

## 10. Media, Big data e Intelligenza Artificiale

La prima parte di questo modulo tratta dei dati e dell'Intelligenza Artificiale e del valore aggiunto della MIL.

La seconda parte, il MILAB, offre un approccio pratico al tema dei dati e dell'Intelligenza Artificiale.

### Alcune definizioni (si veda il glossario)

- Data
- Data mining
- Algoritmo
- Intelligenza artificiale (IA)
- Società democratica

### Sintesi

#### PARTE I Costruzione delle conoscenze

1. Prima dei dati e dell'IA
2. I dati e l'IA nel mondo digitale
3. Caratteristiche dell'IA
4. Opportunità e rischi
5. Obiettivi didattici, competenze e valore aggiunto della MIL
6. Valutazione
7. Materiali formativi

#### PARTE II MILAB

- Attività fase 1
- Attività fase 2
- Attività fase 3

**PARTE I (4 sessioni da 45 MIN)****1. Prima dell'Intelligenza Artificiale (IA) (10 MIN)**

Le statistiche venivano applicate a piccoli insiemi di dati.

Le statistiche venivano già usate per la profilazione delle persone. Nielsen, negli Stati Uniti, monitorava campioni di telespettatori per osservarne le abitudini ed indicare le preferenze generali del pubblico. Tali informazioni venivano utilizzate per la creazione di pubblicità per prodotti radio-televisivi.

**2. I dati e l'IA nel mondo digitale (35 MIN)****Cosa cambia per i dati e l'IA?**

- ❖ La novità è rappresentata dalla potenza di calcolo dei computer e dalla quantità di informazioni che produciamo, mediante cui essi possono estrarre più informazioni! Vengono estratti degli insiemi di dati molto più grandi per ottenere una profilazione, individuazione e previsione più precise.

Il termine big data si riferisce all'immensa quantità di informazioni codificate che viene analizzata mediante gli algoritmi in modo da scoprire modelli, tendenze e correlazioni. Essi possono essere elaborati con diversi tipi di analisi statistiche per la profilazione delle persone, per prevedere comportamenti, dati analitici sull'apprendimento. Spesso vengono descritti mediante le "5 V": Volume (grandi quantità di registri archiviati), Velocità (rapidità nella creazione e nel trasferimento dei dati), Varietà (dati strutturati e non strutturati), Veridicità (precisione e correttezza) e Valore (infrastrutture informatiche costose ed estrazione).

(si veda "Una Breve Introduzione alle 5 V dei Big Data e alla Tecnologia Hadoop", Ishwarappa Kalbandi e J Anuradha (2015)

[https://www.researchgate.net/publication/282536587\\_A\\_Brief\\_Introduction\\_on\\_Big\\_Data\\_5Vs\\_Characteristics\\_and\\_Hadoop\\_Technology](https://www.researchgate.net/publication/282536587_A_Brief_Introduction_on_Big_Data_5Vs_Characteristics_and_Hadoop_Technology)).

**Suggerimento per l'insegnante:**

Non dimenticare mai che sono necessari degli esseri umani per creare degli algoritmi.

Visitare un sito di e-learning, come quello della Khan Academy, per saperne di più sugli algoritmi e su come funzionano.

<https://www.khanacademy.org/computing/computer-science/algorithms/intro-to-algorithms/a/route-finding>

- ❖ Usi degli algoritmi (informazione e previsione):
  - ⇒ profilazione per pubblicità / informazioni su misura
  - ⇒ medicina e trattamenti online
  - ⇒ guida e automobili autonome
  - ⇒ navigazione e geolocalizzazione
  - ⇒ inchieste di polizia (sorveglianza, tracciamento)
  - ⇒ ambiente (monitoraggio, previsione)

...

**Suggerimento per l'insegnante:**

Mostrare agli studenti i video di Hans Rosling sulla demografia e la visualizzazione, della serie "Il Bello delle Statistiche".

<https://www.youtube.com/watch?v=jbkSRLYSojo>

- ❖ Come viene influenzata dai dati la **produzione mediatica** delle informazioni?
  - Visualizzazione dei dati
  - Ricerca sui dati (fuga di notizie)

- Giornalisti robotici (IA)
- Droni

#### Suggerimento per l'insegnante:

È possibile concentrarsi su esempi pratici di giornalismo prodotto dai robot.

Ad es. nel 2015 il quotidiano francese *Le Monde* ha utilizzato un robot per la copertura delle elezioni locali. Il risultato fu la pubblicazione di 36.000 articoli.

- ❖ Che effetto hanno i dati sul **consumo mediatico** delle informazioni?

Due tipi principali di parametri per il pubblico e le comunità:

- **dati sulla sensibilizzazione**, come le impressioni, le visualizzazioni, i click, i collegamenti
- **dati sulla partecipazione**, come i mi piace, le condivisioni, i commenti

La logica sottostante le piattaforme basate su algoritmi vale anche per i motori di ricerca ed i servizi di streaming video...tutti luoghi in cui si veicolano informazioni e intrattenimento. Per quanto riguarda il settore commerciale, in base all'attenzione posta sull'economia (il tempo passato online dall'utente), diventa fondamentale la capacità di filtrare, classificare e misurare le informazioni, poiché in tal modo è possibile trasformarle in denaro attraverso pubblicità e sponsor.

I feed delle notizie (o *newsfeed*) e gli indici dei social media si basano su quattro tipi di algoritmi, che curano l'informazione con delle procedure di editorializzazione più o meno evidenti (D. Cardon, *Che cosa sognano gli algoritmi: Le nostre vite nell'era dei big data*. Seuil, 2016.):

- ⇒ **popolarità** (rilevazione del pubblico, es. numero di click, numero di visualizzazioni su YouTube)
- ⇒ **autorevolezza** (numero di link e hyperlink. l'algoritmo PageRank di Google)
- ⇒ **reputazione** o affinità (numero di mi piace e commenti che mostrano il rapporto e le conversazioni degli utenti su Facebook, Twitter, ecc.)
- ⇒ **previsione** (raccolta di impronte digitali e dati online)

Questi modelli coesistono e sono in concorrenza sul web. Tuttavia, si tende sempre di più ad osservare un passaggio dall'informazione sotto forma di notizia e documenti, all'informazione sotto forma di dati e grafici sociali (basati sull'identità e la comunità).

#### Suggerimento per l'insegnante:

Google, Yahoo, Bing, Qwant...: i motori di ricerca non sono neutrali. Hanno una linea editoriale come gli altri mezzi di comunicazione, attuata tramite i rispettivi algoritmi! Google, ad esempio, associa fortemente la "popolarità" delle informazioni a quelle preferite, non alla qualità dell'informazione.

Sessione 2: 45 MIN

### 3. Caratteristiche dell'IA

- ❖ **Discussione: Caratteristiche dei dati e dell'IA secondo te (25 MIN)**

#### Suggerimenti:

- Micro-profilazione
- Manipolazione
- Connessione di oggetti
- Interazione con bot e assistenti digitali (es. Siri)

- Creazione di bolle di filtraggio
- Sostituzione degli esseri umani con i robot
- Controllare le persone
- Prevedere il futuro
- Monitorare gusti e comportamenti
- ...

#### Suggerimento per l'insegnante:

L'IA richiede l'esercizio del pensiero collettivo. Ciò che potrebbe essere buono per un singolo individuo (personalizzazione dei servizi), potrebbe rappresentare un problema per la società in generale. È importante porre il dibattito sull'IA a livello sociale. Soprattutto in virtù del fatto che le culture digitali tendono ad accrescere l'individualismo e il comunitarismo.

#### ❖ Forme ed esempi di dati in relazione alla cittadinanza e alla sensibilizzazione del consumatore (20 MIN)

- ⇒ Nudge
- ⇒ Robot formati (apprendimento automatico)
- ⇒ Deepfake
- ⇒ Catturatori (RFID)

#### Suggerimento per l'insegnante:

L'identificazione a radiofrequenza (RFID) è una tecnologia che registra la presenza di un oggetto collegato utilizzando i segnali radio. Viene utilizzato per l'identificazione (persone, pacchetti...), inventari (merce, beni...)

#### Riepilogo

- Impronta digitale
- L'IA non è né davvero "umana", né è un'"intelligenza": (al momento) si tratta per lo più di statistiche applicate ad insiemi di big data e di modelli di apprendimento automatico dei comportamenti online.
- Tra i rischi e le questioni etiche sollevate dall'IA, vi sono:
  - Il nudge ed il sistema di credito sociale (o Classificazione del Credito Sociale in Cina)
  - Bolle di filtraggio e camere dell'eco
  - Disordini informativi: fake news, deepfake
- Questioni etiche:
  - libertà di coscienza
  - libertà di parola
  - trasparenza degli algoritmi
  - supervisione dei robot

#### Suggerimento per l'insegnante:

Le bolle di filtraggio e le camere dell'eco tendono a ridurre la portata delle informazioni ed a polarizzare l'opinione, mentre i sistemi di credito sociale vengono utilizzati per esortare le persone a comportarsi secondo alcune norme o aspettative.

*Bolle di filtraggio:* ad esempio il PageRank di Google e il Newsfeed di Facebook dispongono di algoritmi che creano una "bolla di filtraggio" che spinge gli utenti verso le informazioni che si avvicinano al loro punto di vista (ed eliminano le informazioni in contrasto con la propria cultura o comunità online). Eli Pariser, che ha coniato il termine, sottolinea come tale situazione potrebbe portare ad un "determinismo dell'informazione", in cui le nostre precedenti abitudini di navigazione online modellano il nostro accesso a nuove informazioni.

*Sistema di credito sociale*: l'idea è di monitorare i comportamenti sociali del cittadino (abitudini di acquisto, valutazione del credito, regolarità lavorativa, interazioni sociali) ed assegnargli un punteggio per valutare la sua affidabilità. La Cina prevede di mettere in atto la Classificazione del Credito Sociale entro il 2020.

Per un punto di vista distopico sul sistema di credito sociale, gli studenti possono guardare e commentare il primo episodio della terza stagione della serie inglese Black Mirror, dal titolo "Caduta Libera".

Sessione 3: 45 MIN

#### 4. Opportunità e rischi dei Big data e dell'IA

❖ **Discussione**: Come viene usata l'IA al giorno d'oggi? (15 MIN)

Chiedere agli studenti di cercare online degli articoli che mostrino i principali 10 algoritmi che dominano il nostro mondo. Eventualmente, indirizzarli alla lista di Gizmodo:

Google Page Rank, Newsfeed di Facebook, "Trova appuntamenti" di OKCupid, Crittografia NSA, gli algoritmi di raccomandazione "Potrebbe interessarti anche", Google AdWords, Mercato azionario ad alta frequenza, Compressione degli MP3 per ridurre la dimensione dei dati, IBM CRUSH (Criminal Reduction Utilizing Statistical History - Riduzione della criminalità mediante analisi statistiche) per l'analisi predittiva e Auto-tune.

Link: <https://io9.gizmodo.com/the-10-algorithms-that-dominate-our-world-1580110464>

Quali sono le principali funzioni/obiettivi? (suggerimenti: raccomandazioni, accoppiamento, predizione, creazione di conoscenze, sorveglianza, processo di elaborazione delle decisioni). Discutere dei pro e dei contro.

##### Suggerimento per l'insegnante:

In alternativa, chiedere agli studenti di accedere al sito web dell'autorità garante del proprio Paese e di osservare come descrive i propri obiettivi in materia di dati e IA.

Gli studenti possono dare un'occhiata anche alle linee guida etiche UE sull'affidabilità dell'IA:

<https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>

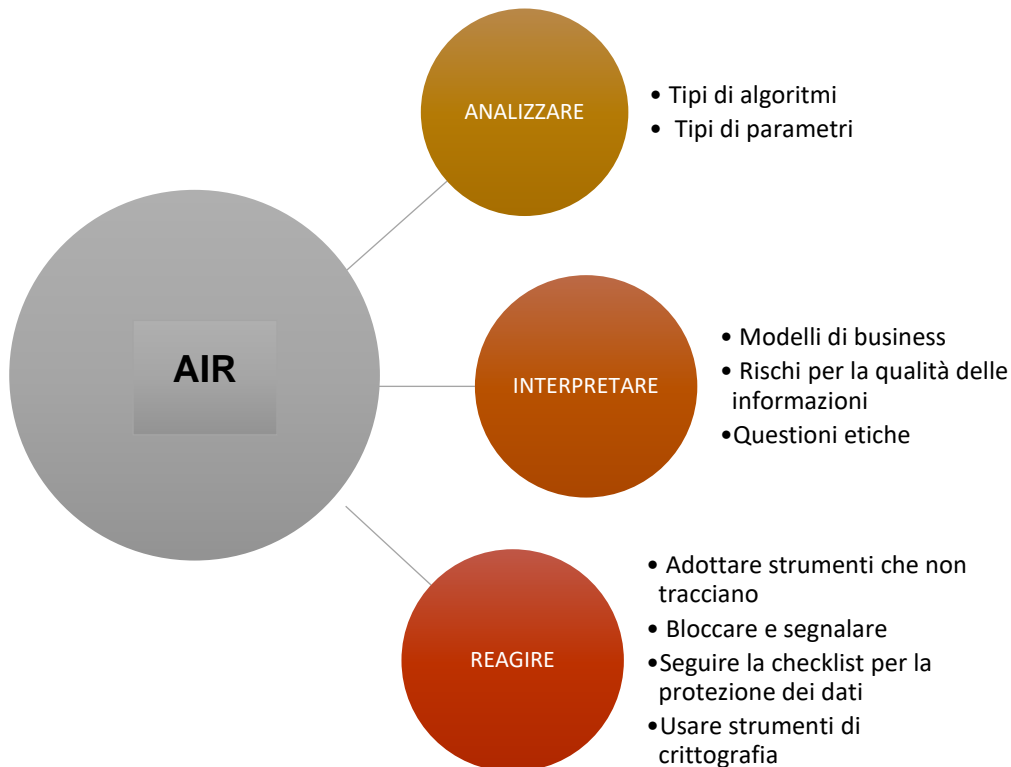
Qualora gli studenti siano interessati all'etica dell'IA, ricapitolare i "6 tipi di questioni etiche" sollevate dagli algoritmi, che mostrano come questi ultimi siano carichi di valori. Si veda Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, Luciano Floridi « L'etica degli algoritmi: Mappatura del dibattito », Big data & society, 1 dicembre, 2016

<https://journals.sagepub.com/doi/full/10.1177/2053951716679679>

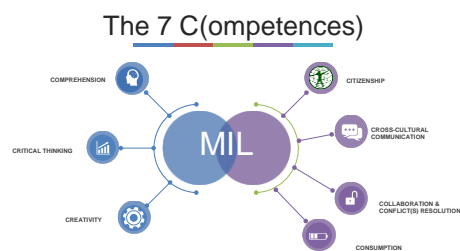
❖ **Opportunità e rischi (15 MIN)**



**Costruzione delle capacità di pensiero critico degli studenti: AIR**



## 5. Focus sui dati e l'IA in relazione alle **COMPETENZE DI CITTADINANZA E DI SENSIBILIZZAZIONE DEL CONSUMATORE (7C) (15 MIN)**

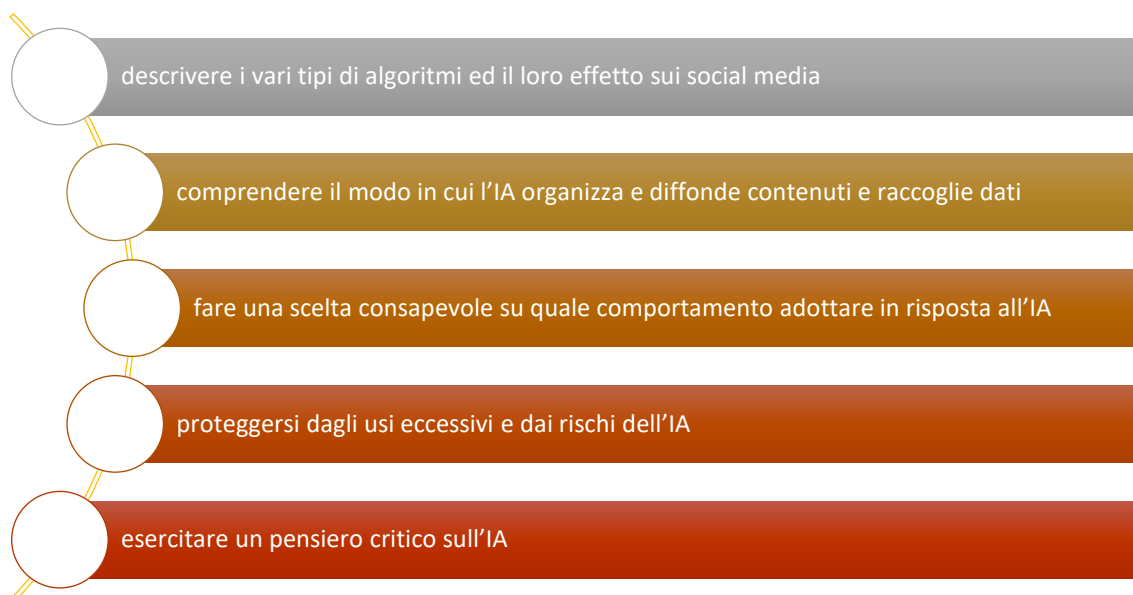


Le competenze di sensibilizzazione del consumatore e di cittadinanza si uniscono per garantire che i cittadini comprendano l'ambiente digitale creato dai media, prodotti e servizi basati sull'IA. Capire le implicazioni della realtà commerciale dell'IA e le loro conseguenze etiche sull'autonomia e la privacy è fondamentale per il futuro delle società democratiche e per il benessere online.

Si vedano l'introduzione e il glossario

## Divisione delle competenze MIL in categorie/indicatori delle **COMPETENZE** di CITTADINANZA E SENSIBILIZZAZIONE DEL CONSUMATORE

Gli studenti dovrebbero essere in grado di:



### Valore aggiunto della MIL per i diritti umani

- Privacy
- Liberta d'espressione
- Dignità (questioni etiche sull'estrazione e l'interpretazione dei dati)

Sessione 4: 45 MIN

## 6. Valutazione

Saggio: cosa significa il termine "Intelligenza" all'interno dell'espressione "Intelligenza Artificiale"?

NB: ricordare che il termine inglese "*intelligence*" significa anche "spionaggio"

*oppure*

Chiedere agli studenti di guardare il video “*Project Debater*” che mostra come un computer abbia vinto un dibattito sul finanziamento degli asili negli Stati Uniti

(<https://www.youtube.com/watch?v=FmGNwMyFCqo>).

Chiedere di scrivere un post su cosa pensano del video. Come si sentono.

#### **7. Materiali di supporto didattico (si veda la sezione aggiuntiva ai Moduli Didattici)**

- ⇒ Riferimenti ad altri materiali e risorse
- ⇒ Link utili per animazione pedagogica
- ⇒ Glossario
- ⇒ Software per l'integrazione della MIL nei risultati didattici (Risorse online per Paese)  
<https://docs.google.com/spreadsheets/d/10wxqYEe9O8GiSKo8kTjv8uQqpkOeHJp5k0ytBcZsdU/edit?usp=sharing>



## PARTE II - MILAB (4 sessioni da 45 MIN, a seconda del tempo assegnato alla MIL)

*Le attività MILAB sono strutturate in tre fasi (1-avvio; 2-consolidamento; e 3-approfondimento). Le tre fasi sono indicative: possono essere seguite come suggerito o utilizzate secondo un approccio di "collegamento" o modulare a seconda del tempo assegnato alla MIL e/o al livello di risultato desiderato.*

*Le fasi prevedono varie attività pedagogiche, come ad esempio: workshop, giochi di ruolo, esercizi scritti e giochi. Sperimentano vari formati multimediali, come i blog, i video e gli articoli web, al fine di formare le capacità di pensiero critico degli studenti (AIR).*

### ESPLORAZIONE DELLA CITTADINANZA e della SENSIBILIZZAZIONE DEL CONSUMATORE:

*Contributi ai social media tramite tweet, post, storie...*

---

#### FASE 1 (15-16):

##### COMPNDERE GLI ALGORITMI

---

###### ❖ Ricerca di informazioni (45 MIN)

Dividere gli studenti in tre gruppi: chiedere a ciascun gruppo di cercare le stesse informazioni conflittuali/controverse su vari servizi (Google, Wikipedia, Twitter). Far confrontare i risultati delle ricerche.

*Cosa indicano le differenze in termini di logica dell'attenzione? e di modello di business? Si vedano i 4 tipi di algoritmi (popolarità, autorevolezza, reputazione, predizione).*

##### Suggerimento per l'insegnante:

Scegliere una personalità politica polarizzante, o un evento attuale controverso come il cambiamento climatico o la crisi dei migranti.

Oppure,

Chiedere agli studenti di osservare i risultati del giornalismo d'inchiesta eseguito tramite i big data, come il consorzio internazionale che portò alle rivelazioni dei *Panama Papers* (link : <https://www.icij.org/>)

Quali sono i vantaggi del trattamento combinato e decentrato dei big data? Questo tipo di giornalismo quanto ha cambiato le politiche in materia di trasparenza? Come vengono usate queste informazioni dai politici per: modificare le leggi, multare e sanzionare le frodi, recuperare denaro nelle dichiarazioni dei redditi?

###### ❖ Ricerca di informazioni (continuazione) (45 MIN)

Proseguire con l'analisi dei risultati delle ricerche.

- ✓ Cosa succede agli altri tipi di rappresentazioni?
- ✓ Cosa succede alla qualità dell'informazione?
- ✓ Che tipo di conoscenza si può costruire sulla base di tali risultati?
- ✓ Quali sono i rischi dell'eventuale comprensione o fraintendimento della situazione attuale? Ad esempio per prendere decisioni (es. votare alle elezioni)?

### ❖ Dibattito (45 MIN)

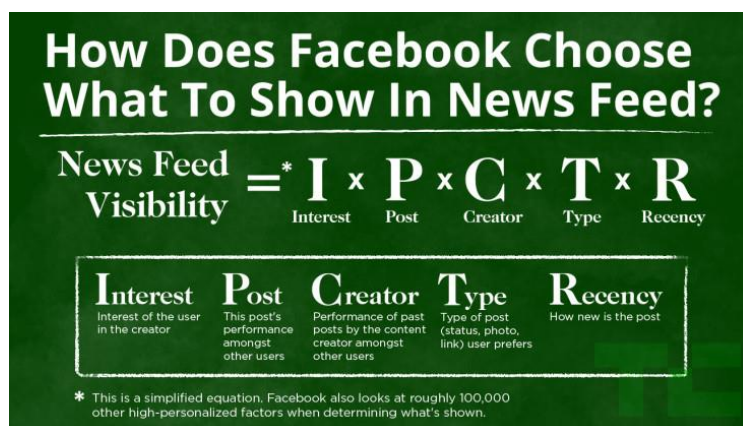
Sulla base dei risultati, i tre gruppi elaborano delle soluzioni etiche su come affrontare le bolle di filtraggio e le camere dell'eco create dagli algoritmi e dall'IA.

### ❖ VALUTAZIONE E COMMENTO (45 MIN)

Ogni studente scrive un **post sul blog** sui propri sentimenti e reazioni in base alla propria esperienza.

#### Suggerimento per l'insegnante:

In alternativa, è possibile mostrare agli studenti l'algoritmo Newsfeed per aiutarli a comprendere il modo in cui vengono create le bolle di filtraggio.



<https://techcrunch.com/2014/04/03/the-filtered-feed-problem/>

## FASE 2 (16-17)

### APPROFONDIMENTO SUI DATI E L'IA

---

#### ❖ SIMULAZIONE (45 MIN)

Chiedere agli studenti di prestare attenzione alla propria cronologia per una settimana. Possono tenere un registro o accedere alla sezione cronologia tramite la barra degli strumenti.

#### ❖ SIMULAZIONE (continuazione) (45 MIN)

Chiedere agli studenti di visualizzare la propria "bolla di filtraggio" rispetto al loro utilizzo di Google.

Far utilizzare agli utenti l'estensione sviluppata da Mozilla, Lightbeam, per aiutarsi a visualizzare il tracciamento in base alla propria cronologia di navigazione. Cosa notano?

#### ❖ SIMULAZIONE (continuazione) (45 MIN)

Dividere gli studenti in 2 o 3 gruppi. Chiedergli di provare dei motori di ricerca che non eseguono il tracciamento, come DuckDuckGo o Qqwant. Qual è la loro politica in materia di IA? Come si adattano al RGPD?

❖ **VALUTAZIONE E COMMENTO (45 MINS)**

*Chiedere agli studenti di confrontare una specifica ricerca sul tema prescelto eseguita tramite più motori di ricerca, assicurandosi che almeno uno tracci i dati personali e che almeno uno non li tracci. Riflettere sulle conseguenze e su cosa c'è in ballo.*

**FASE 3 (18-19):**

**COSTRUIRE UN RAPPORTO ETICO SOSTENIBILE CON I DATI E L'IA**

---

❖ **Scenario DISASTRO INFORMATIVO (45 MIN)**

Ricerca: Autovetture autonome. Prendere in esame il caso di Waymo, il progetto di autovettura autonoma di Google. Ci sono state delle vittime? di che tipo? Qual è stato il risultato a livello giuridico? chi è stato giudicato responsabile?

❖ **Scenario DISASTRO INFORMATIVO (45 MIN)**

Ricerca: Riconoscimento del volto e sorveglianza. Mettere a confronto il caso del riconoscimento facciale da parte dello Stato in Cina con la decisione della città di San Francisco di rendere tale pratica illegale. Argomentazioni a favore e contrarie.

❖ **Scenario DISASTRO INFORMATIVO (45 MIN)**

Ricerca: Deepfake. Guardare il video del Presidente Obama realizzato dall'Università di Washington (Washington State University). Cosa sarebbe accaduto se le persone avessero creduto alle sue parole?

Link: <https://www.youtube.com/watch?v=UCwbJxW-ZRg>

❖ **VALUTAZIONE E COMMENTO (45 MIN)**

**Saggio:** *Alla luce dei tre scenari catastrofici analizzati, i robot potrebbero dominare il mondo? Come possono i cittadini stabilire dei confini all'interno delle società democratiche?*