



Corso di Laurea Magistrale in Informatica

Machine Learning Fairness

Analisi empirica dello stato della pratica

Prof. Fabio Palomba
Dott.ssa. Giulia Sellitto

Carmine Ferrara
Mat: 0522500990



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

sesa^{lab}
SOFTWARE ENGINEERING
SALERNO



*When you think of A.I., it's forward-looking.
But A.I. is based on data, and data is a
reflection of our history*

Joy Buolamwini



*Le soluzioni di Machine Learning sono progettate e sviluppate da umani,
e naturalmente gli uomini riproducono il loro giudizio all'interno delle soluzioni di Machine Learning*



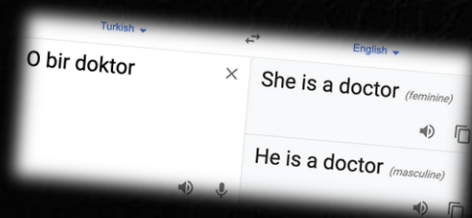
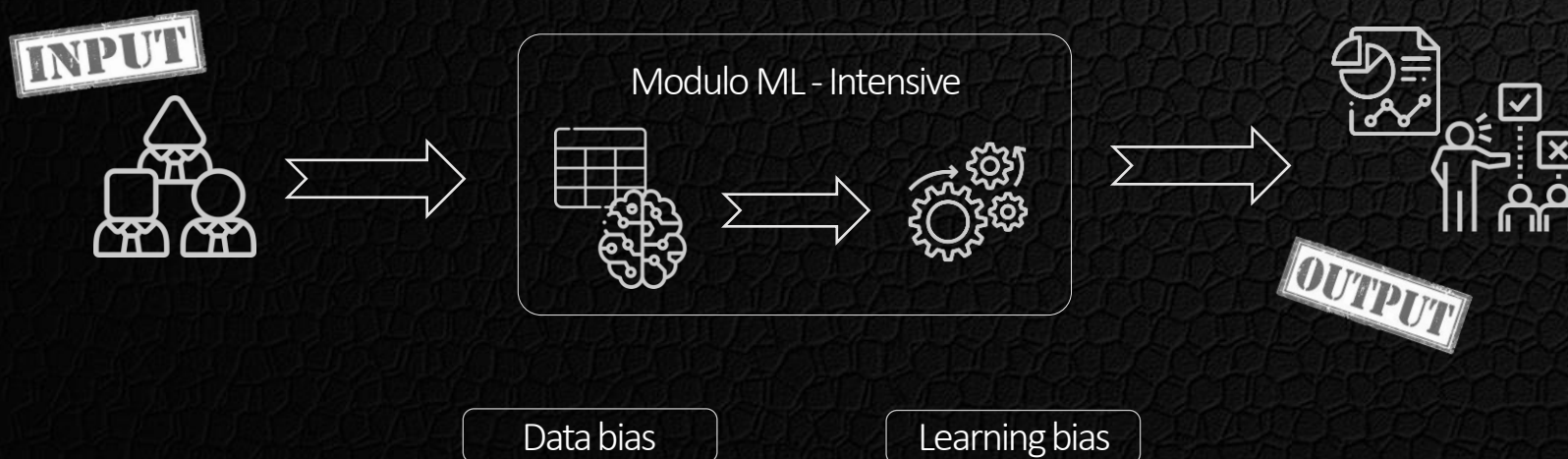
✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Introduzione e Background

Nuove tipologie di vulnerabilità nello sviluppo ML-Intensive



Per alcune professioni quali “Dottore o Dottoressa”, la traduzione dal turco di Google Translate non tengono conto del Gender del soggetto.



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Introduzione e Background

Machine Learning Fairness

Un vero e proprio *Hot Topic* di ricerca

Con più di 20 differenti *definizioni e metriche*

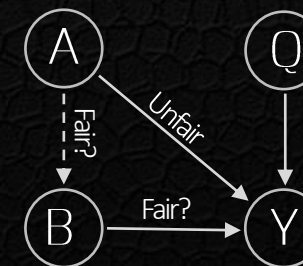
Fairness: Set di requisiti, metodi e tecniche che rendono una soluzione di intelligenza artificiale eticamente corretta.

Approcci Probabilistici



Approcci di Similarità

Approcci Causali



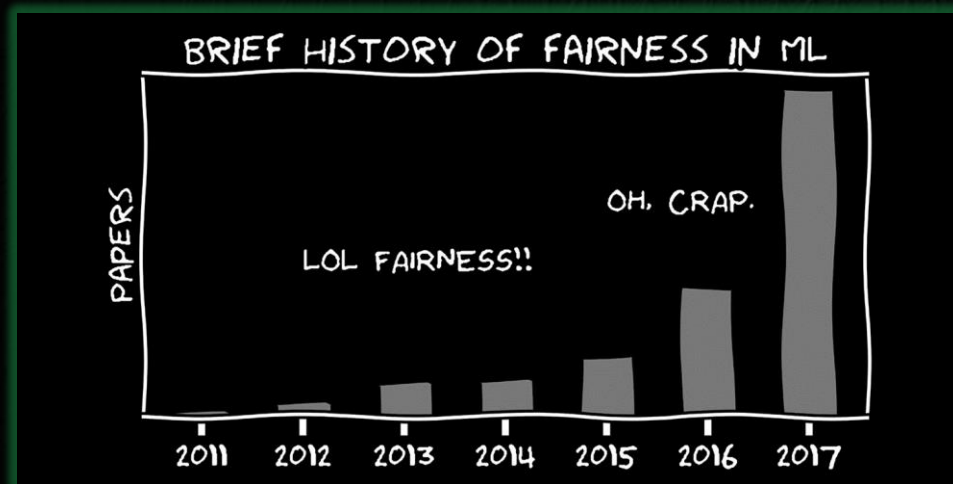
✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Problematica di Ricerca

È osservabile come negli ultimi anni Fairness sia cresciuta considerevolmente come argomento di ricerca...



Dai numerosi studi di intelligenza artificiale atti a migliorare i livelli di fairness di un sistema ML-Intensive...



... Fino a quelli ingegneristici che trattano fairness come un vero aspetto di qualità nello sviluppo ML-Intensive



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Problematica di Ricerca

Ma come applicare fairness in casi pratici di sviluppo?

Secondo i più recenti studi di ricerca, Fairness
è un'aspetto di qualità...

Astratto

Context-Specific

*Difficile da trattare
praticamente*



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Obiettivi di Ricerca

... si rende quindi necessario indagare empiricamente circa le pratiche lavorative nel trattamento di Fairness, cercando di capire...

Quali sono le figure necessarie in un Team Fair-Oriented?



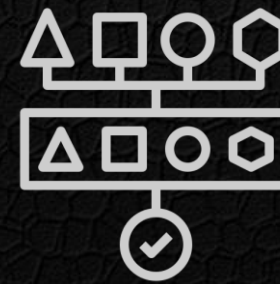
Quali sono le definizioni o gli approcci maggiormente utilizzati?



Quanto Fairness è rilevante rispetto altre specifiche?



In che fasi di sviluppo ML viene trattata Fairness?



Quanto mature sono le compagnie Fair-Oriented?



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Metodologia di Ricerca

Come rispondere ai quesiti?



Redazione di un **survey** atto a coinvolgere esperti del settore



Manager
Data Scientists
Ingegneri del Software

*Perché un **survey**?*

Coinvolgere in modo mirato uno specifico gruppo di candidati

Integrare differenti culture e background

Ottenere rapidamente dati omogenei facili da analizzare

....



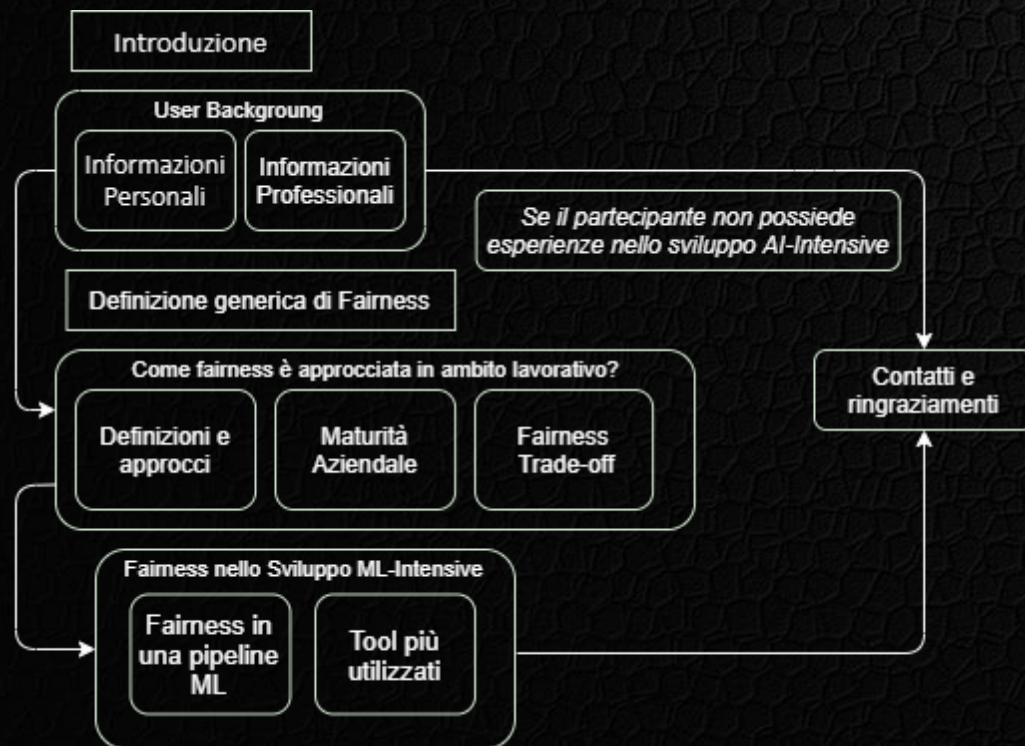
✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Metodologia di Ricerca

Dettagli sulla fase di studio Empirico



Test pilota con studenti di SE4AI




Piattaforma nota nel mondo della ricerca
per questa tipologia di studi la quale offre sistemi
di filtering automatici per i partecipanti


Pulizia dei dati


È stata condotta un'attività di data quality prescreening per rimuovere risposte non rilevanti e/o ottenute da figure non adeguate ai nostri obiettivi

203 risposte collezionate e 116 usate in analisi



 c.ferrara49@studenti.unisa.it

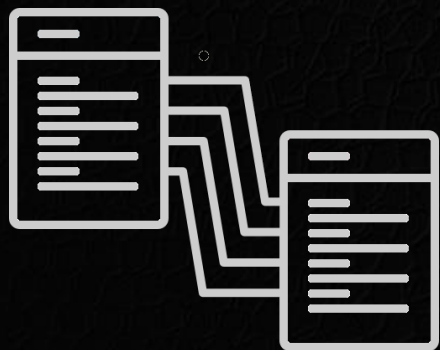
 <https://cferrara98.github.io/>

 <https://tinyurl.com/27zbux8z>

Analisi dei Dati



Statistica descrittiva e Grafici



In fase di analisi è stata ricostruita la mappatura tra quesiti del Survey e gli obiettivi di ricerca

Attività di pre-processing come trasformazioni di scala o uso di acronimi utili all'elaborazione dei dati



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Risultati di Ricerca

Composizione del Campione



Circa il 90% dei partecipanti all'indagine possiede un titolo di studi triennale o superiori

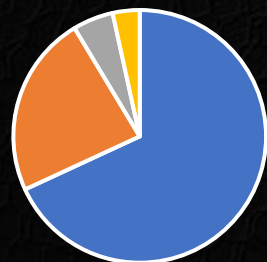


Tutti i 116 partecipanti hanno esperienza lavorativa con Machine Learning e Fairness

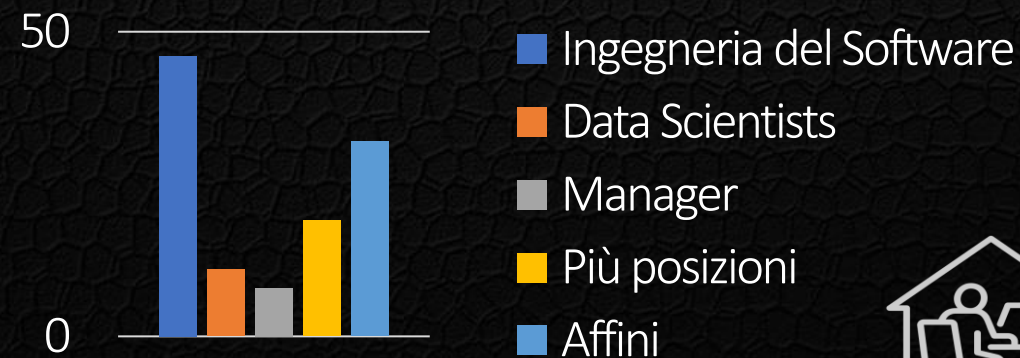


Provenienza dei partecipanti

- Europa
- Africa
- Nord America
- Altro



Occupazione dei partecipanti



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Risultati di Ricerca

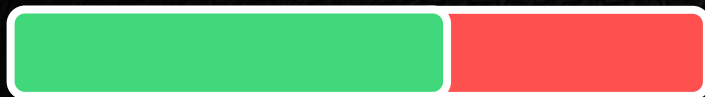
Trattare Fairness in ambito aziendale

Definizioni e Metriche Formali

Approcci Pratici

Similarità degli Individui

Analisi delle dipendenze



Probabilità e
Casual Reasoning

Configurazioni
di Learning

Altrimenti

Metodologie empiriche o Correlazione statistica



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

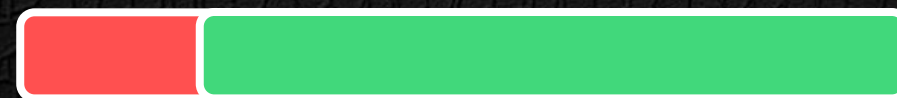
Risultati di Ricerca

Trattare Fairness in ambito aziendale

Usability & Scalability



Accuracy & Sicurezza



Fairness

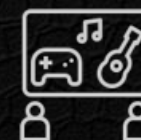
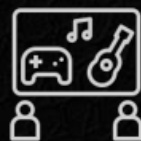
Fairness

Intrattenimento

Sicurezza



Fairness



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Risultati di Ricerca

Trattare Fairness in ambito aziendale

Il trattamento di fairness è ancora poco maturo e standardizzato nelle pratiche aziendali ma questo aspetto è da considerarsi in costante miglioramento



Il 26% degli intervistati tratta Fairness in maniera sporadica senza standard precisi

Un ulteriore 24% afferma che la propria compagnia adotti standard di sviluppo Fair Oriented



Della restante parte oltre il 30% dichiara di trattare le specifiche etiche con standard definiti e volti all'ottimizzazione



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Risultati di ricerca

Implicazioni e Studi Futuri

... L'immatunità di Fairness fa sì che studi di ricerca mirati siano necessari



Sistematizzazione di Definizioni e Approcci di Sviluppo
per il trattamento di Fairness secondo le esigenze professionali e di dominio

Capire Fairness tramite l'ausilio di casi d'uso reali in
ambito aziendale e accademico



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Risultati di Ricerca

Team Ottimale per il trattamento di Fairness



Figure come **Manager e Esperti Specifici**
sono di estrema rilevanza



Data Scientists e Ingegneri del Software

Figure trasversali con responsabilità equiparabili



Implicazioni e Studi Futuri



Standard di Responsabilità in un Team Fair Oriented



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

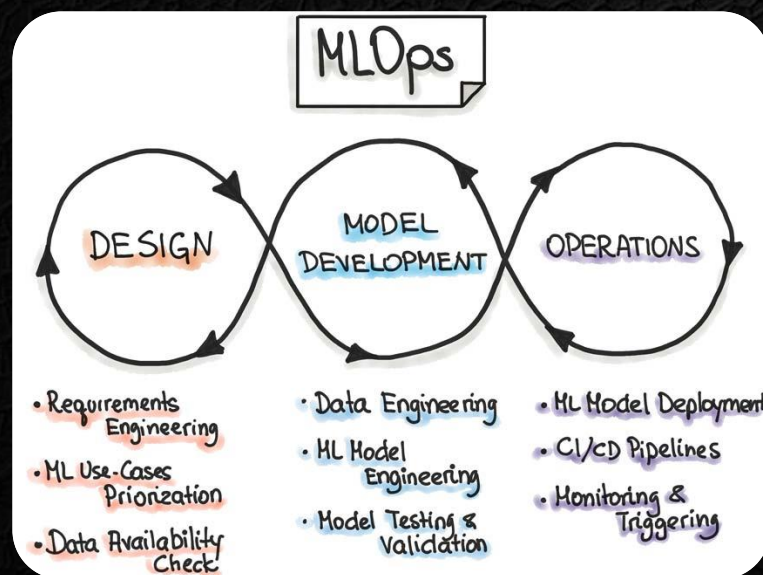
Risultati di Ricerca

Fairness nello Sviluppo di Soluzioni ML-Intensive

Fairness matura man mano che il modello evolve e matura



Fairlearn



What If...

you could inspect a machine learning model, with minimal coding required?



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Risultati di Ricerca

Fairness nello Sviluppo di Soluzioni ML-Intensive

Fairness matura man mano che il modello evolve e matura



Analisi di Bad & Best
Practices Specifiche

Implicazioni e Studi Futuri

Protocolli Gestionali
Evolutivi Fair Oriented



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>



*Grazie
per l'attenzione!*

Crediti: Flaticon (icone), Storyset (illustrazioni)

Machine Learning Fairness
Analisi empirica dello stato della pratica



lab
sesa
SOFTWARE ENGINEERING
SALERNO



c.ferrara49@studenti.unisa.it



<https://cferrara98.github.io/>



<https://tinyurl.com/27zbux8z>

