

Corso di Laurea (Triennale) in Informatica

# Equità in Sistemi di Machine Learning: Stato dell'arte, Problemi e Sfide

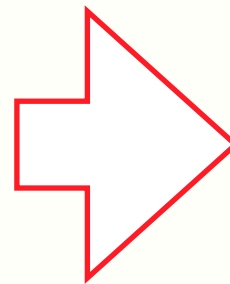
**Prof. Fabio Palomba**  
**Dott. Giammaria Giordano**

**Alessandra Parziale**  
**Mat.: 0512108069**



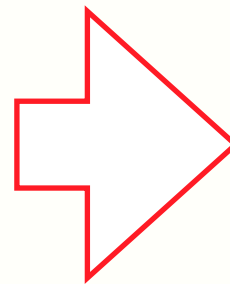
L'**equità (Fairness)** consiste nel trattare i soggetti in modo analogo indipendentemente dal sesso, dalla razza o dallo stato sociale.

L'**equità (Fairness)** consiste nel trattare i soggetti in modo analogo indipendentemente dal sesso, dalla razza o dallo stato sociale.

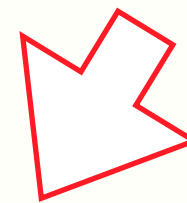


Il concetto di equità è oggi associato al **giudizio** in quanto consiste in una valutazione di qualcuno oppure di qualcosa

L'**equità (Fairness)** consiste nel trattare i soggetti in modo analogo indipendentemente dal sesso, dalla razza o dallo stato sociale.



Il concetto di equità è oggi associato al **giudizio** in quanto consiste in una valutazione di qualcuno oppure di qualcosa

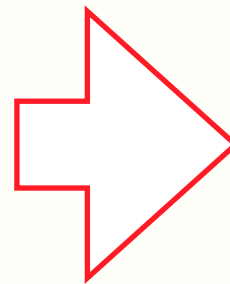


**Giudizi soggettivi**

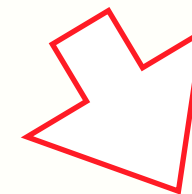
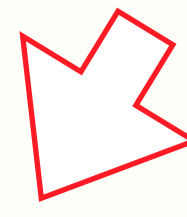


Facilmente  
condizionabili da  
fattori esterni

L'**equità (Fairness)** consiste nel trattare i soggetti in modo analogo indipendentemente dal sesso, dalla razza o dallo stato sociale.



Il concetto di equità è oggi associato al **giudizio** in quanto consiste in una valutazione di qualcuno oppure di qualcosa

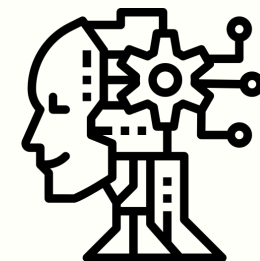


**Giudizi soggettivi**



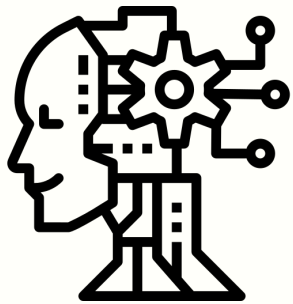
Facilmente  
condizionabili da  
fattori esterni

**Giudizi oggettivi**

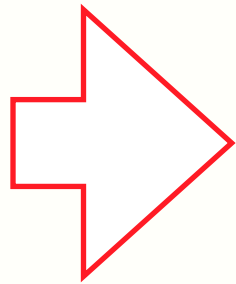
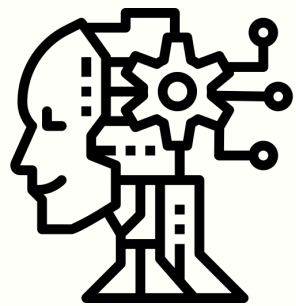


Non sono  
influenzate da  
fattori esterni

# Introduzione e Background

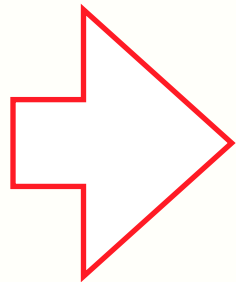
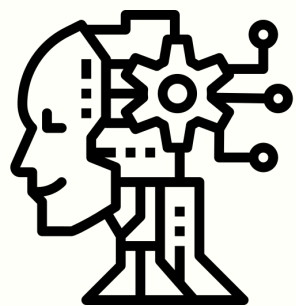


# Introduzione e Background

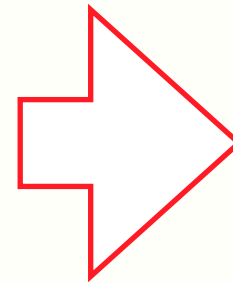


Provocano iniquità

# Introduzione e Background



Provocano iniquità



Non hanno una conoscenza incorporata ma utilizzano degli algoritmi che apprendono secondo delle impostazioni fissate da persone che potrebbero generare delle discriminazione





L'algoritmo utilizzato da Google per associare determinate parole alle immagini era poco cortese con il genere femminile rispetto al genere maschile, infatti associava alle ricerche di genere femminile solo immagini riguardanti l'estetica, la sensualità o la forma fisica mentre agli uomini immagini riguardanti l'ambito lavorativo o professionale

# Introduzione e Background



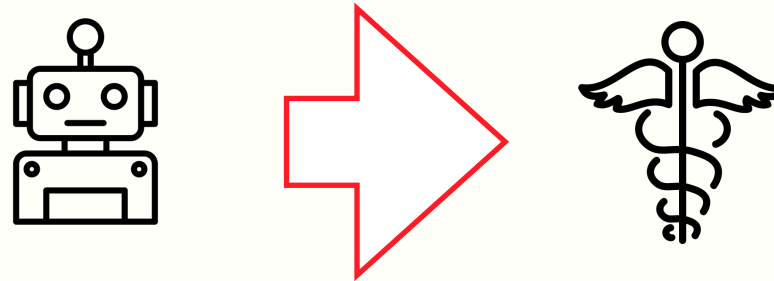
L'algoritmo utilizzato da Google per associare determinate parole alle immagini era poco cortese con il genere femminile rispetto al genere maschile, infatti associava alle ricerche di genere femminile solo immagini riguardanti l'estetica, la sensualità o la forma fisica mentre agli uomini immagini riguardanti l'ambito lavorativo o professionale

Nel 2015, il sistema intelligente capace di valutare i candidati per un lavoro non considerava tutti in modo equo per lavori di sviluppo software o incarichi tecnici. Questo perché gli algoritmi erano stati addestrati osservando dati raccolti in un periodo nel quale nell'industria tecnologica vi era un predominio del genere maschile.

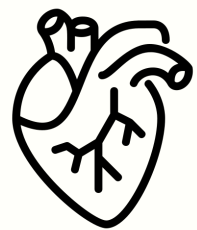
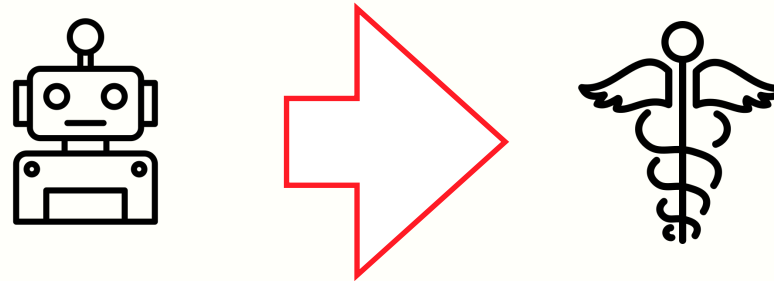


In molti casi l'intelligenza artificiale viene associata alla “**superIntelligenza**” di una macchina, cioè la capacità di superare le prestazioni degli esseri umani.

In molti casi l'intelligenza artificiale viene associata alla “**superIntelligenza**” di una macchina, cioè la capacità di superare le prestazioni degli esseri umani.



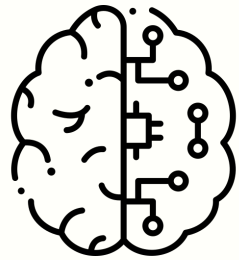
In molti casi l'intelligenza artificiale viene associata alla “**superIntelligenza**” di una macchina, cioè la capacità di superare le prestazioni degli esseri umani.



## Cardiologia:

Dei dispositivi indossabili in grado di **tracciare** e di elaborare i dati del paziente per generare un allarme sul suo stato di salute attraverso lo smartphone

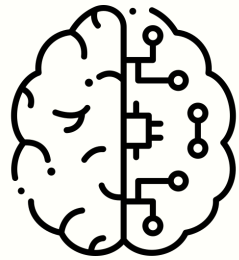




## Neurologia:

Dei **sensori indossabili** per diagnosticare precocemente il morbo di Parkinson e per verificare l'andamento della malattia.





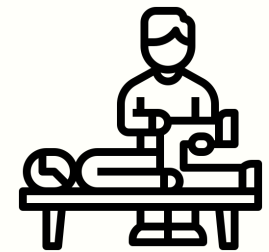
## Neurologia:

Dei **sensori indossabili** per diagnosticare precocemente il morbo di Parkinson e per verificare l'andamento della malattia.



## Fisioterapia:

Dei robot umanoidi in grado di parlare, emettere suoni e luci per aumentare l'attenzione del paziente durante la terapia.





## Agevolazioni:



## Agevolazioni:

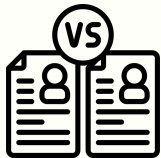


Diagnosi accurata e veloce

## Agevolazioni:



Diagnosi accurata e veloce

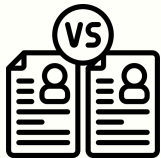


Confronto della storia clinica con  
altri dati raccolti

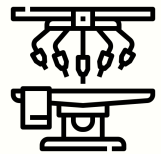
## Agevolazioni:



Diagnosi accurata e veloce



Confronto della storia clinica con  
altri dati raccolti

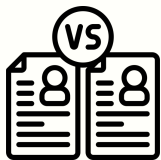


Strumenti all'avanguardia

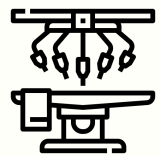
## Agevolazioni:



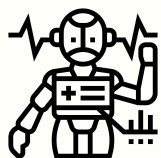
Diagnosi accurata e veloce



Confronto della storia clinica con  
altri dati raccolti



Strumenti all'avanguardia

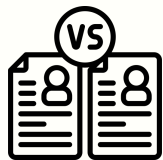


Totale automazione di fasi cliniche  
fondamentali come la prognosi, la diagnosi  
e il trattamento

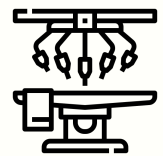
## Agevolazioni:



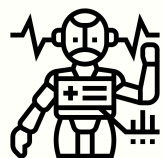
Diagnosi accurata e veloce



Confronto della storia clinica con altri dati raccolti



Strumenti all'avanguardia



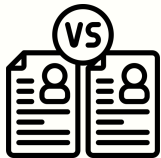
Totale automazione di fasi cliniche fondamentali come la prognosi, la diagnosi e il trattamento

## Criticità:

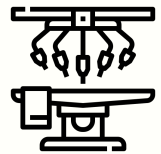
## Agevolazioni:



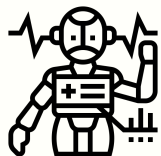
Diagnosi accurata e veloce



Confronto della storia clinica con altri dati raccolti

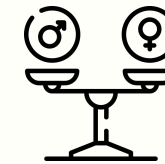


Strumenti all'avanguardia



Totale automazione di fasi cliniche fondamentali come la prognosi, la diagnosi e il trattamento

## Criticità:

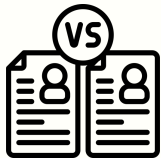


Problemi di equità

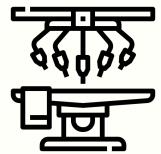
## Agevolazioni:



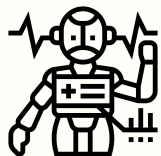
Diagnosi accurata e veloce



Confronto della storia clinica con altri dati raccolti

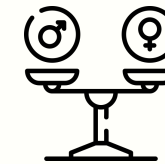


Strumenti all'avanguardia



Totale automazione di fasi cliniche fondamentali come la prognosi, la diagnosi e il trattamento

## Criticità:



Problemi di equità

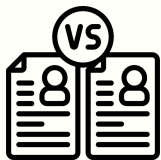


Privacy e sicurezza

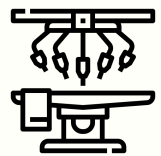
## Agevolazioni:



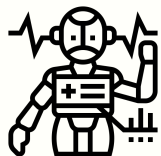
Diagnosi accurata e veloce



Confronto della storia clinica con altri dati raccolti

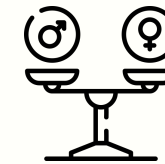


Strumenti all'avanguardia



Totale automazione di fasi cliniche fondamentali come la prognosi, la diagnosi e il trattamento

## Criticità:



Problemi di equità



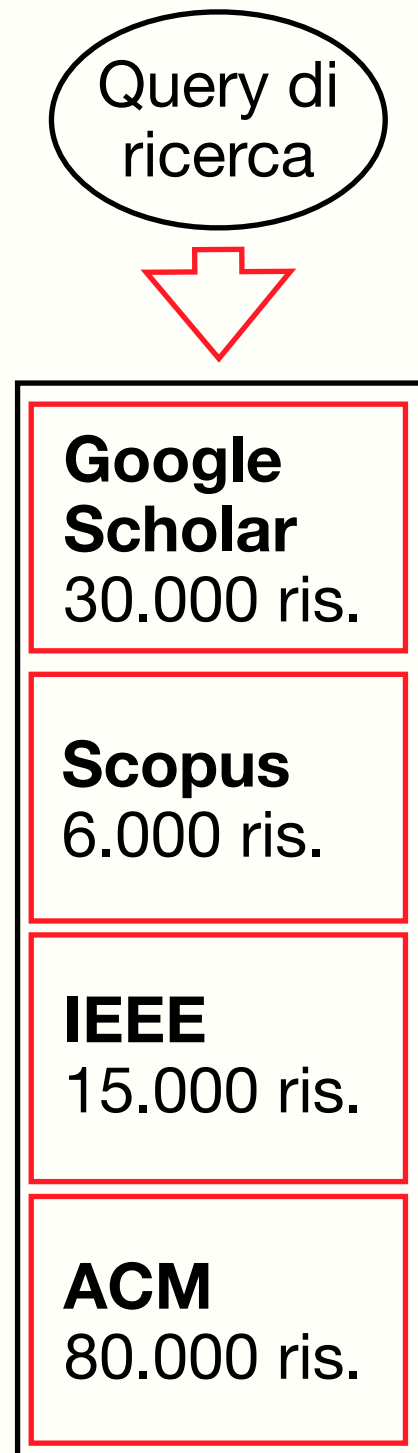
Privacy e sicurezza

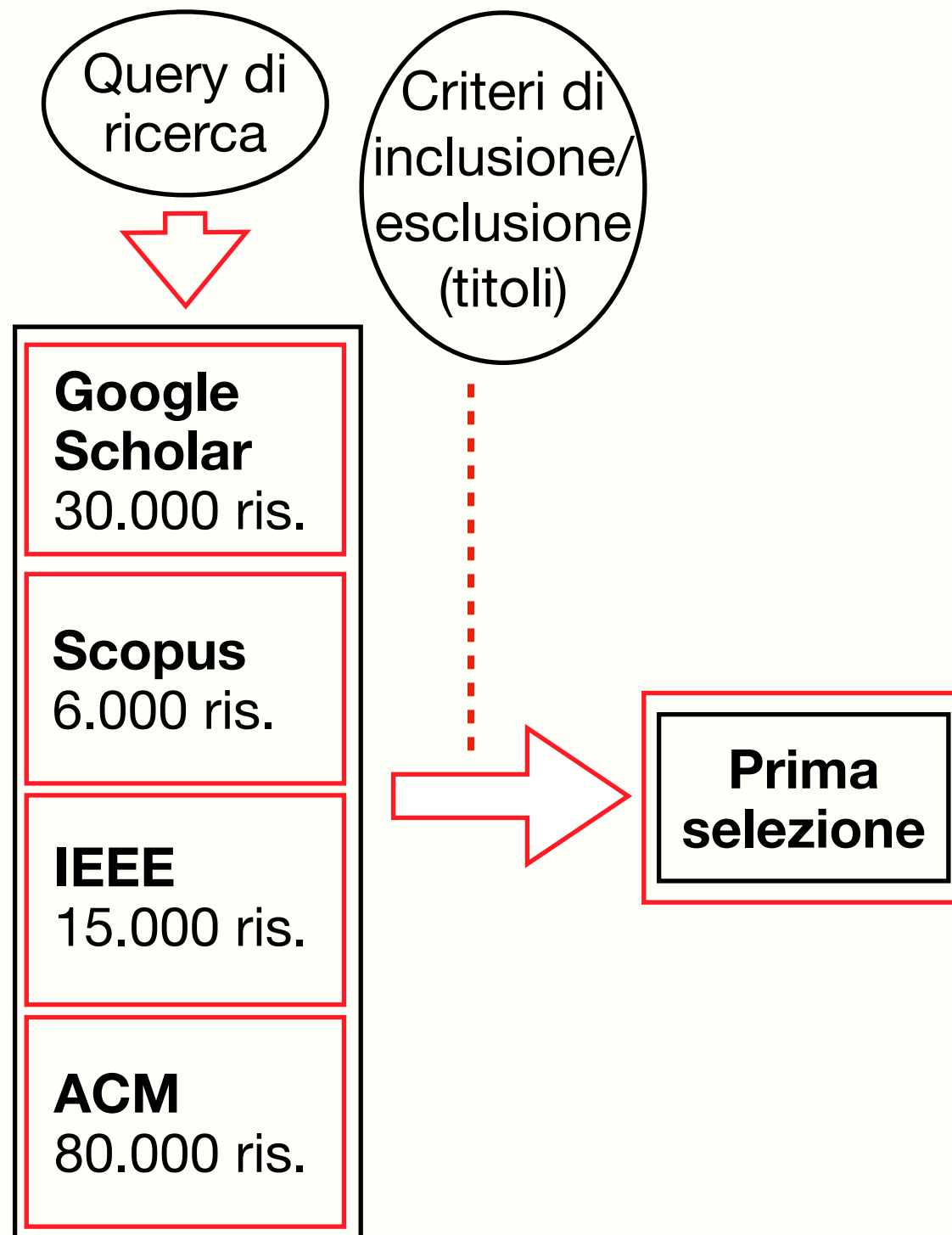


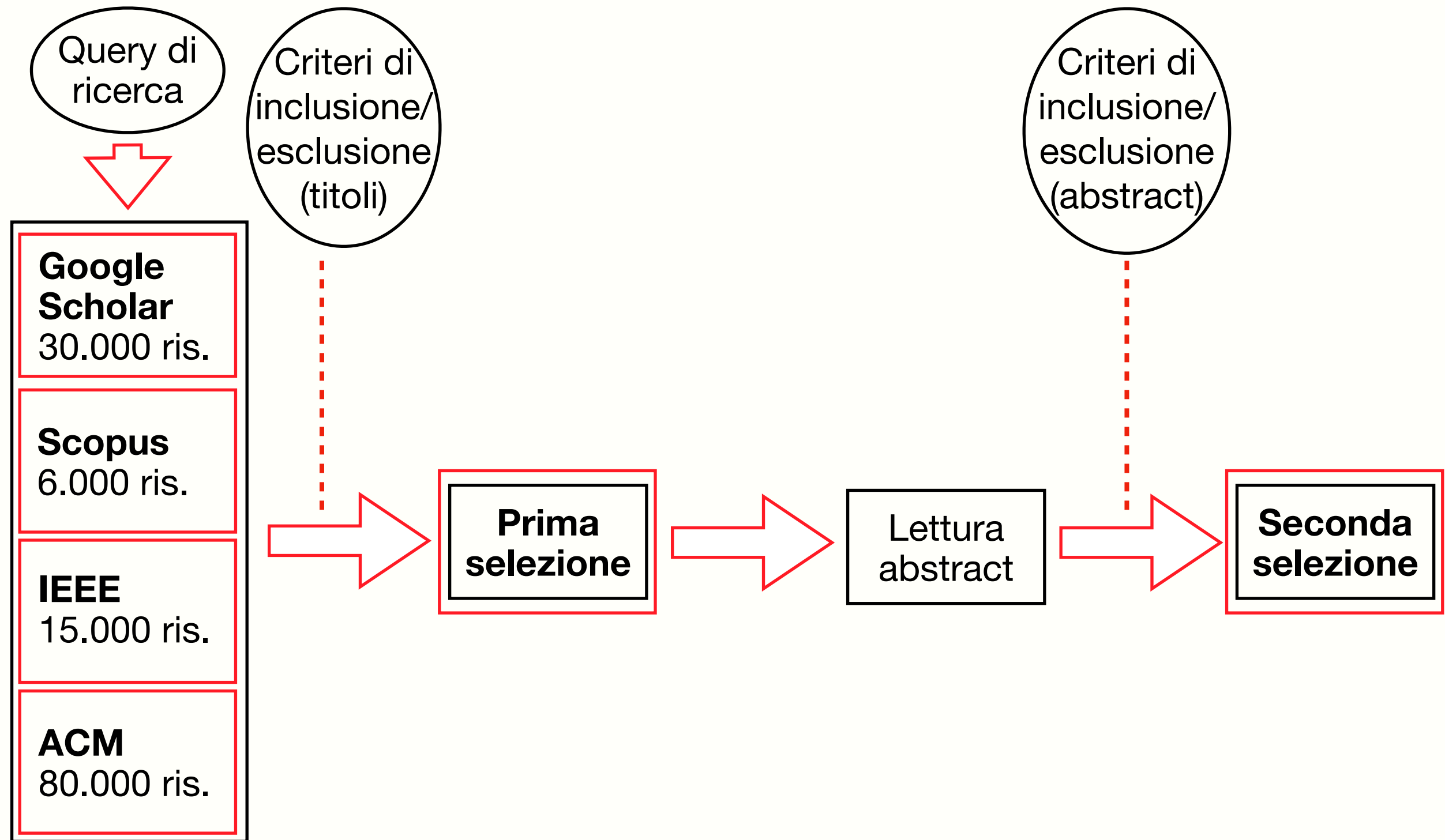
Disumanizzazione

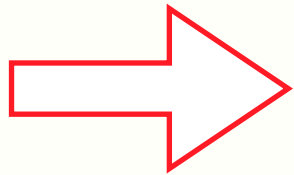


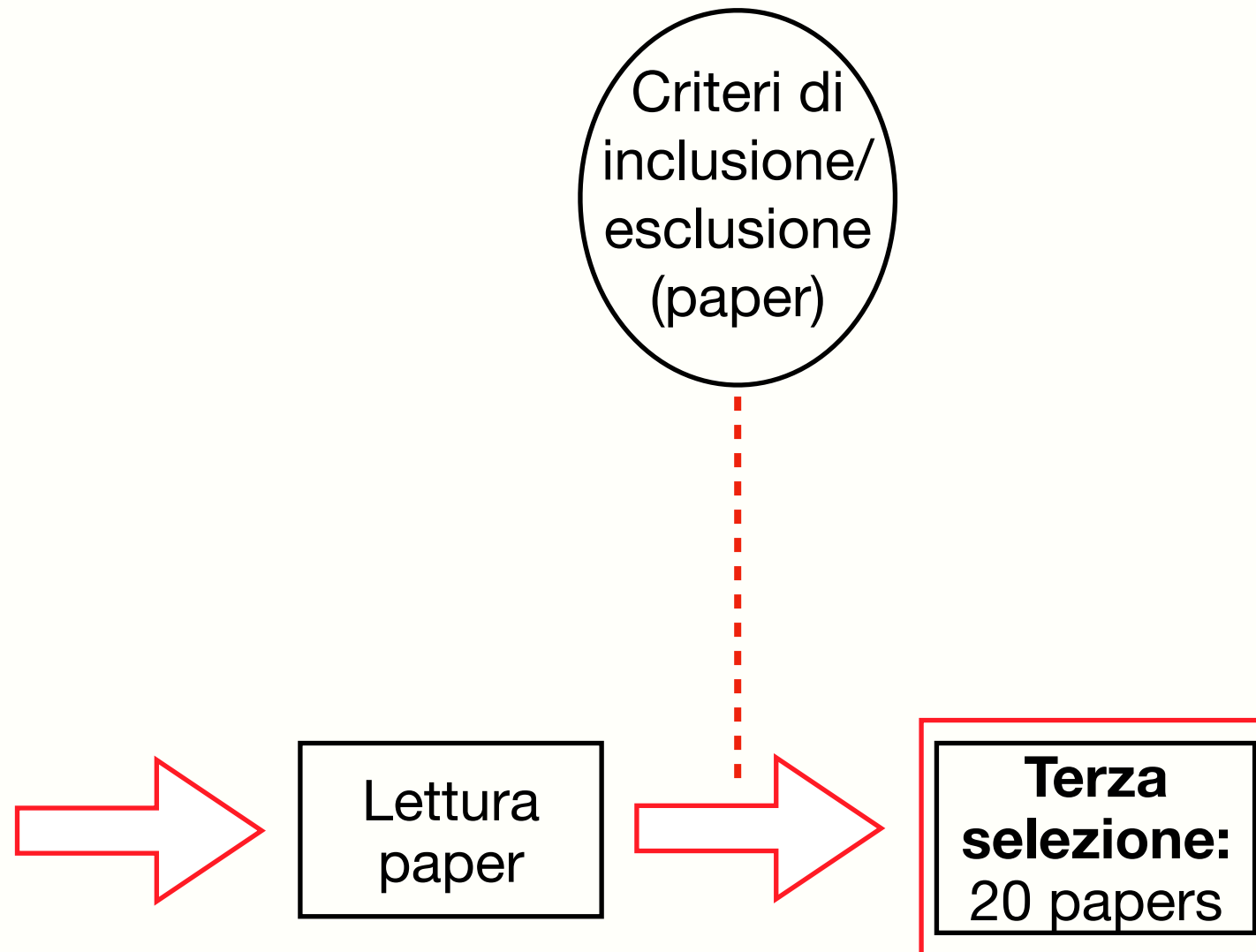
Query di  
ricerca

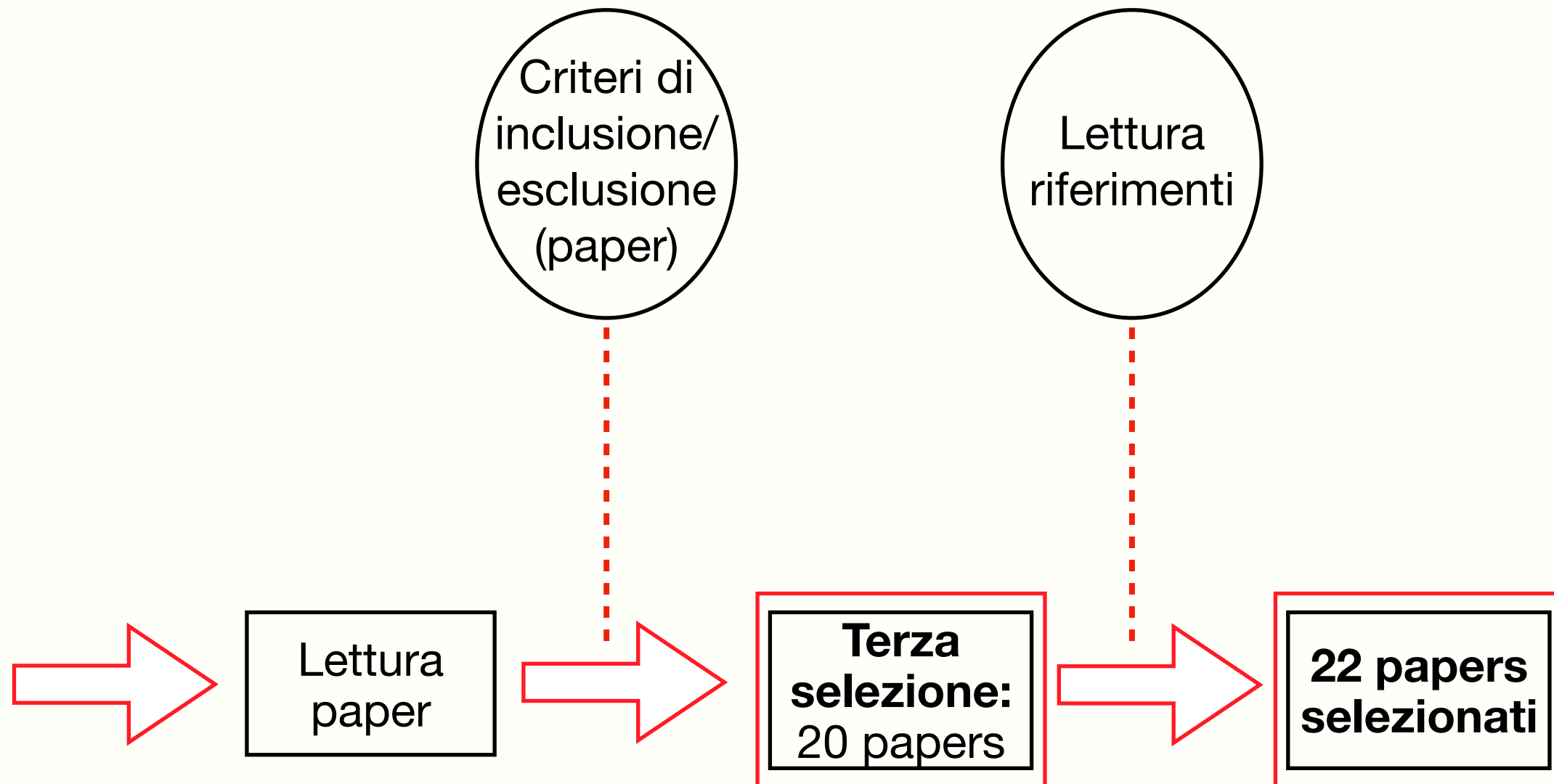


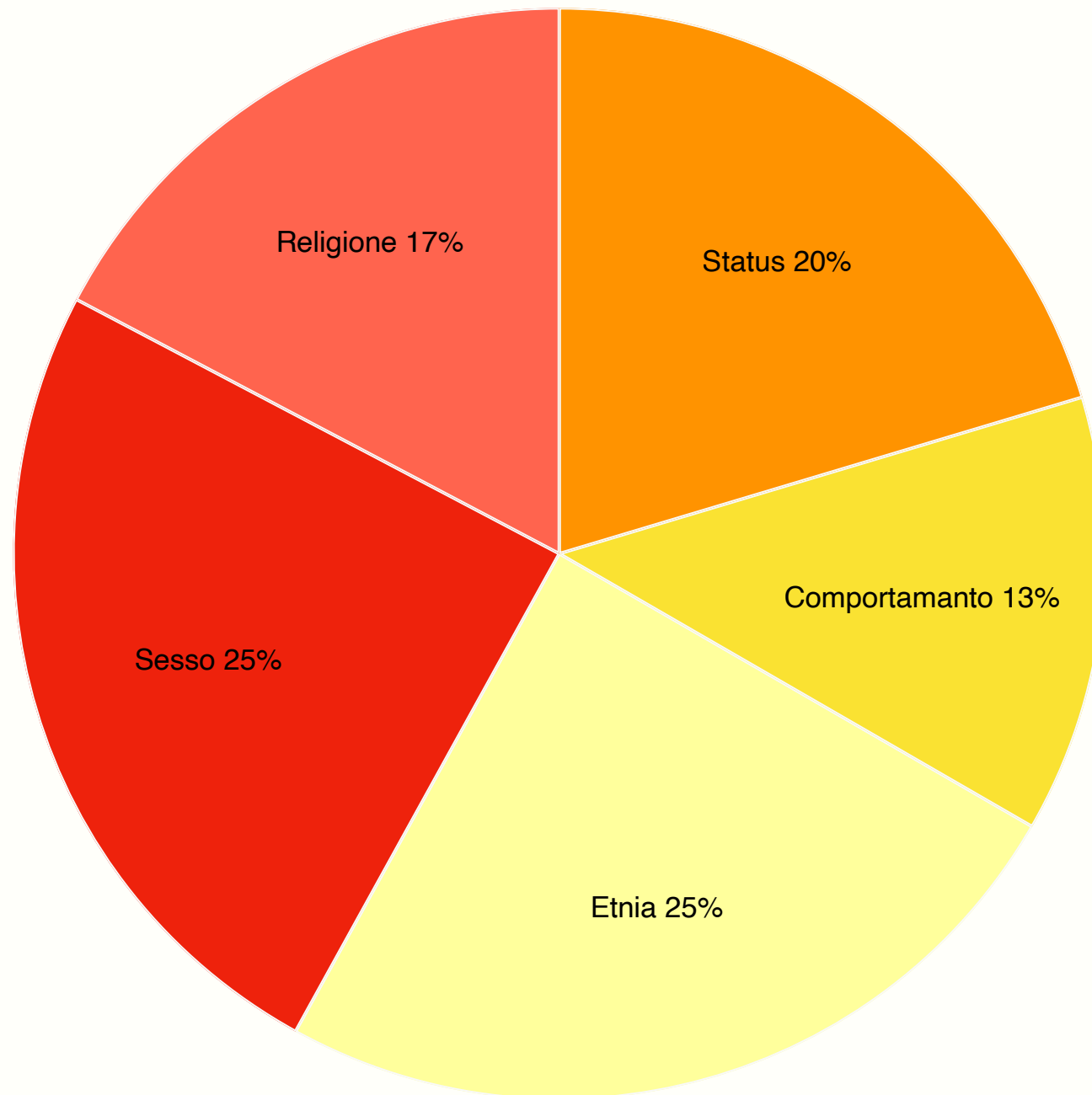




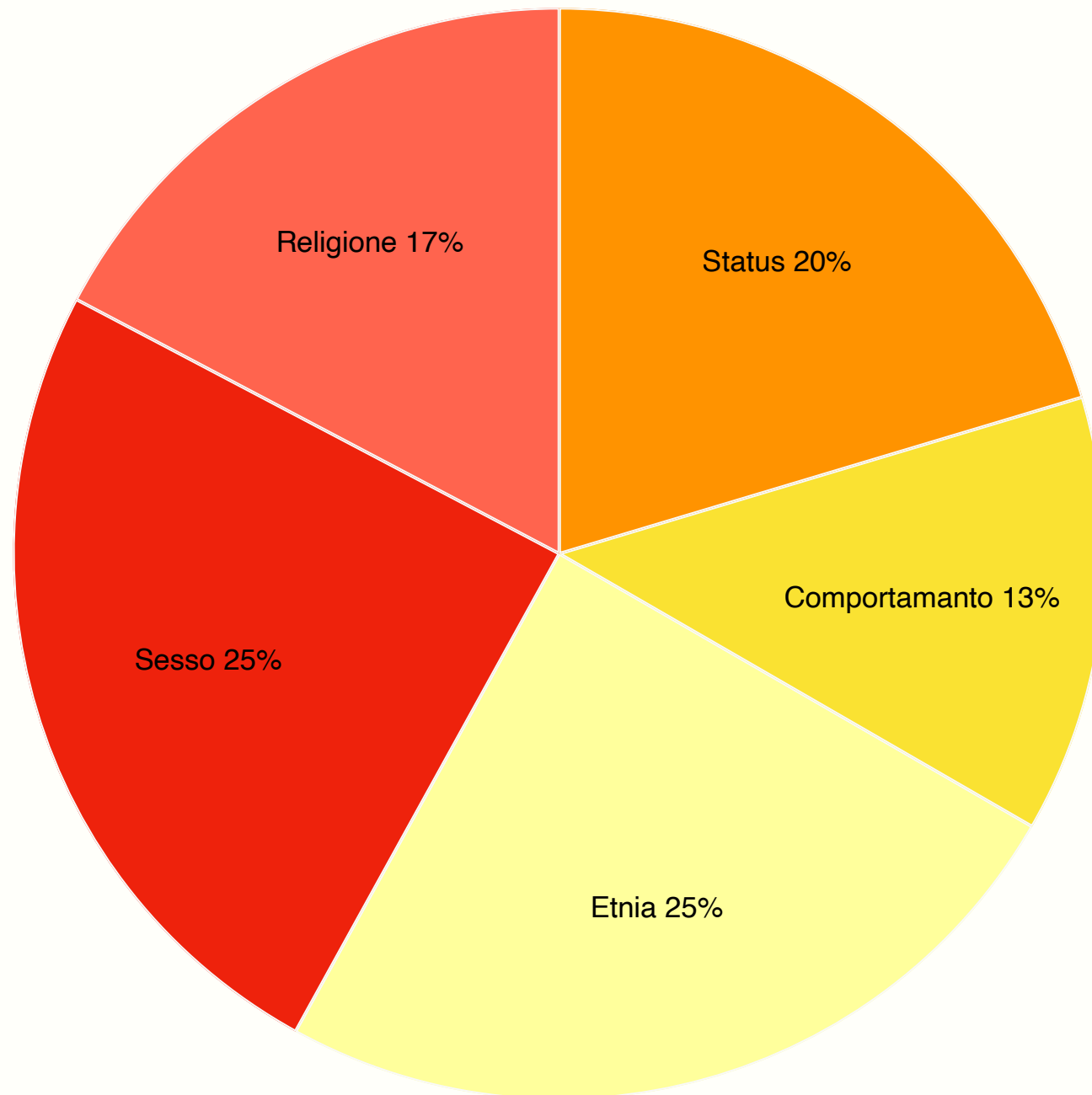






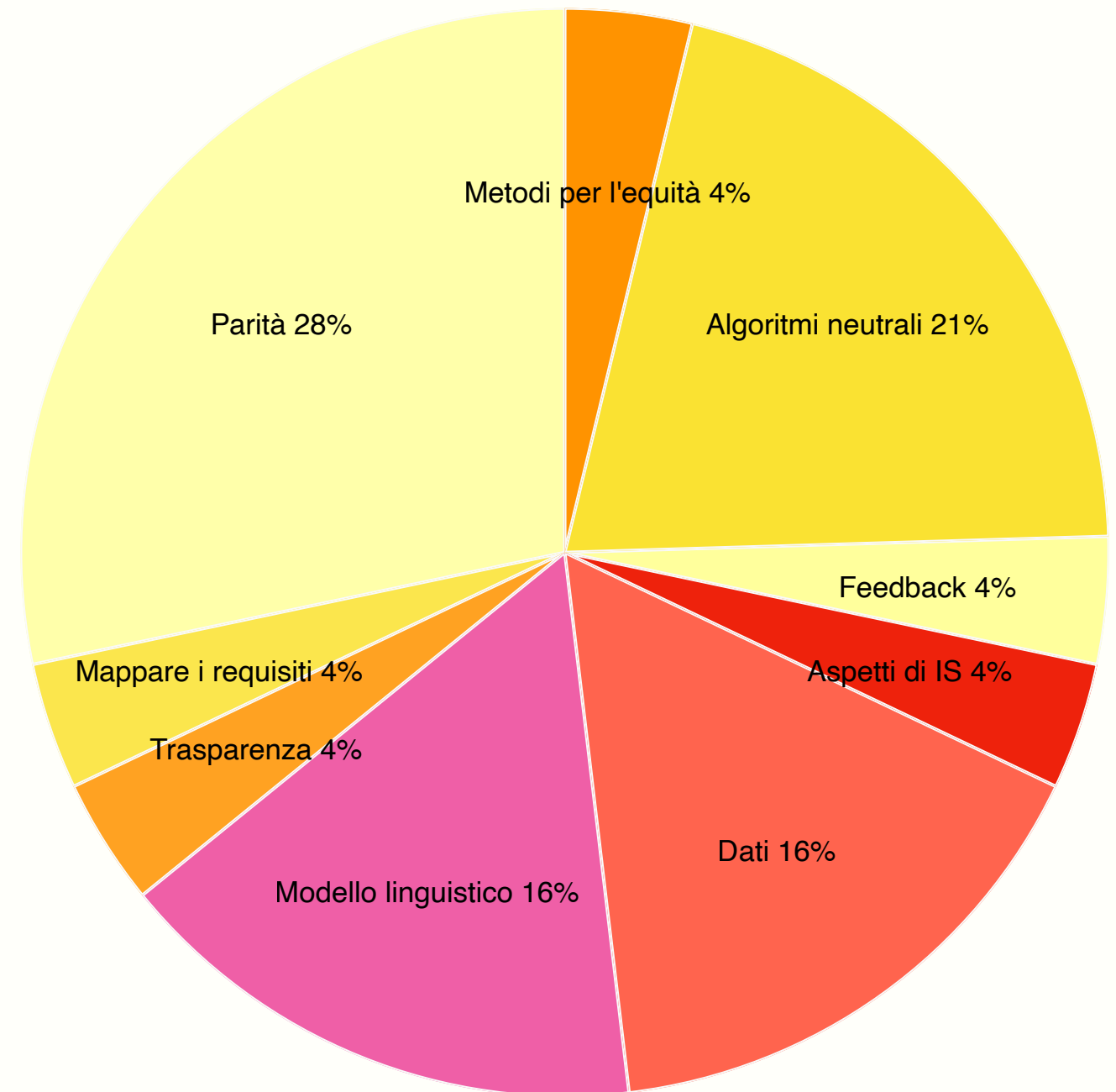






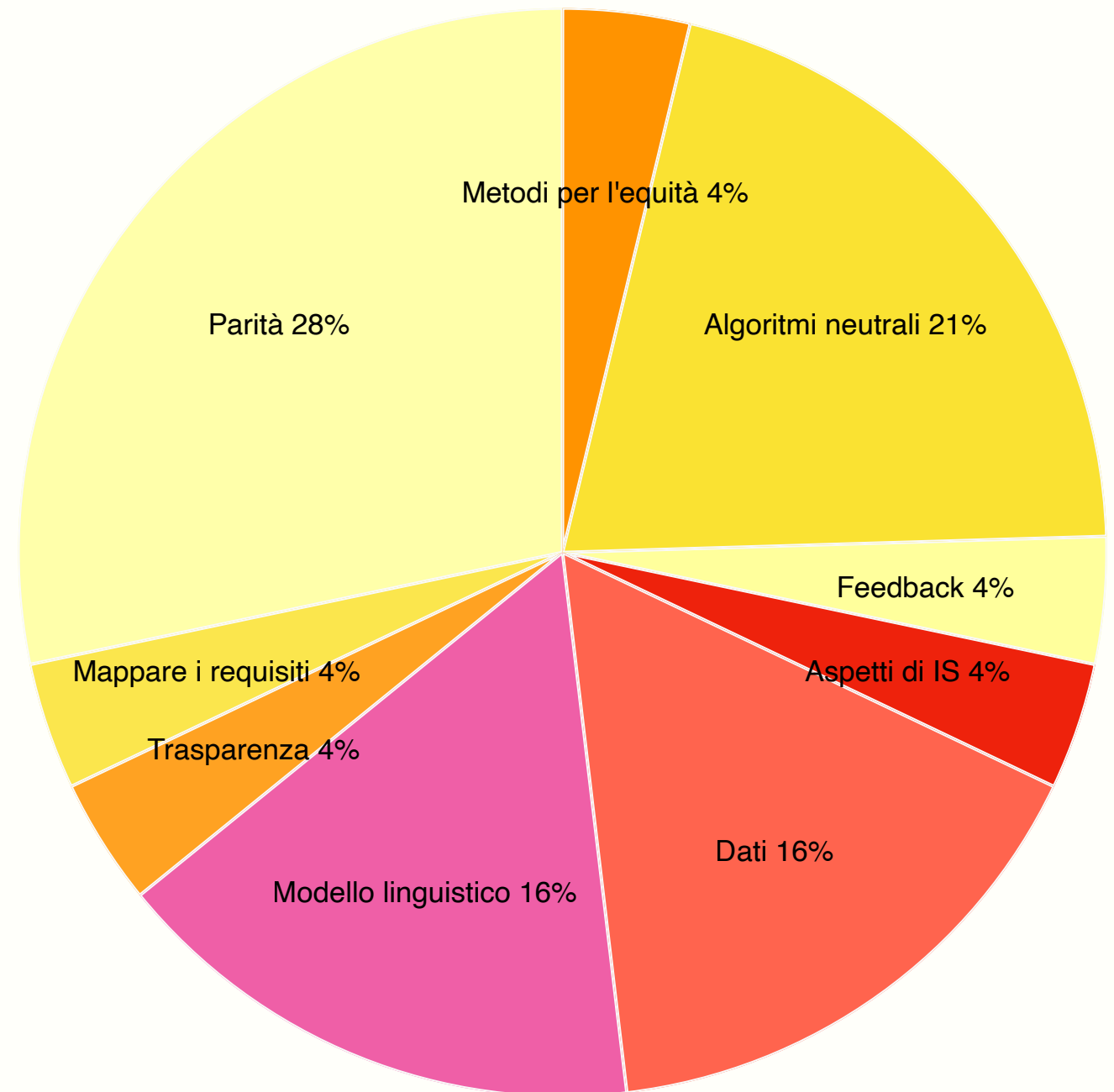
Dalla ricerca è emerso che tra tutti gli attributi considerati, quelli nei quali si evidenzia una maggiore discriminazione nei modelli di apprendimento automatico sono principalmente il sesso e l'etnia, a seguire lo status sociale, la religione e la modalità di comportamento del singolo individuo.

Questo implica che al giorno d'oggi la forma maggiore di discriminazione risiede ancora nel sesso e nell'etnia di un singolo individuo.



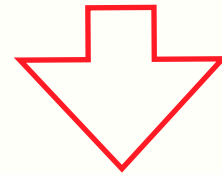
Invece per quanto riguarda i metodi per risolvere la disparità circa il 28% delle ricerche svolte affermano che sia necessaria una parità demografica, individuale, di gruppo e una parità di prestazioni dei modelli, Circa il 20% dichiara che la base per ottenere l'equità sia progettare l'algoritmo specifico in modo da far incorporare esplicitamente o implicitamente l'etnia quindi ottenere comunque algoritmi neutri.

Circa il 16% sostiene che bisognerebbe concentrarsi sul modello linguistico oppure propone un ri-campionamento di dati imparziali.



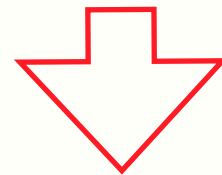
Studiare come l'equità e il pregiudizio varia a seconda della comunità considerata

Studiare come l'equità e il pregiudizio varia a seconda della comunità considerata

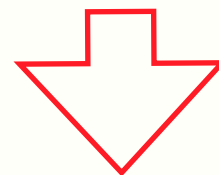


Per dare una visione completa e accurata dei problemi causati dai sistemi intelligenti

Studiare come l'equità e il pregiudizio varia a seconda della comunità considerata



Per dare una visione completa e accurata dei problemi causati dai sistemi intelligenti



Approfondire l'utilizzo di alcune tecnologie avanzate su pazienti con malattie terminali per verificare quanto e come la nuova tecnologia e l'intelligenza artificiale possano intervenire sul ciclo di vita umano

## Introduzione e Background



L'**equità (Fairness)** consiste nel trattare i soggetti in modo analogo indipendentemente dal sesso, dalla razza o dallo stato sociale.



Il concetto di equità è oggi associato al **giudizio** in quanto consiste in una valutazione di qualcuno oppure di qualcosa

Giudizi soggettivi



Facilmente condizionabili da fattori esterni

Giudizi oggettivi



Non sono influenzate da fattori esterni

✉ a.parziale8@studenti.unisa.it  
🌐 <https://github.com/AlessandraParziale>  
🌐 [www.linkedin.com/in/alessandra-parziale](https://www.linkedin.com/in/alessandra-parziale)

Equità in Sistemi di Machine Learning:  
Stato dell'arte, Problemi e Sfide  
Alessandra Parziale  
Università degli Studi di Salerno

## Contesto applicativo



In molti casi l'intelligenza artificiale viene associata alla "**superintelligenza**" di una macchina, cioè la capacità di superare le prestazioni degli esseri umani.



**Cardiologia:**

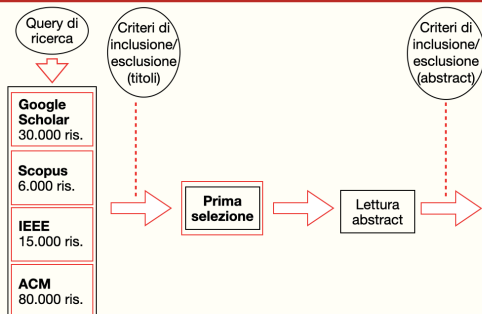
Dei dispositivi indossabili in grado di **tracciare** e di elaborare i dati del paziente per generare un allarme sul suo stato di salute attraverso lo smartphone



✉ a.parziale8@studenti.unisa.it  
🌐 <https://github.com/AlessandraParziale>  
🌐 [www.linkedin.com/in/alessandra-parziale](https://www.linkedin.com/in/alessandra-parziale)

Equità in Sistemi di Machine Learning:  
Stato dell'arte, Problemi e Sfide  
Alessandra Parziale  
Università degli Studi di Salerno

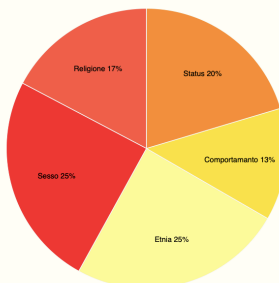
## Metodologia



✉ a.parziale8@studenti.unisa.it  
🌐 <https://github.com/AlessandraParziale>  
🌐 [www.linkedin.com/in/alessandra-parziale](https://www.linkedin.com/in/alessandra-parziale)

Equità in Sistemi di Machine Learning:  
Stato dell'arte, Problemi e Sfide  
Alessandra Parziale  
Università degli Studi di Salerno

## Risultati



Dalla ricerca è emerso che tra tutti gli attributi considerati, quelli nei quali si evidenzia una maggiore discriminazione nei modelli di apprendimento automatico sono principalmente il sesso e l'etnia, a seguire lo status sociale, la religione e la modalità di comportamento del singolo individuo.

Questo implica che al giorno d'oggi la forma maggiore di discriminazione risiede ancora nel sesso e nell'etnia di un singolo individuo.

✉ a.parziale8@studenti.unisa.it  
🌐 <https://github.com/AlessandraParziale>  
🌐 [www.linkedin.com/in/alessandra-parziale](https://www.linkedin.com/in/alessandra-parziale)

Equità in Sistemi di Machine Learning:  
Stato dell'arte, Problemi e Sfide  
Alessandra Parziale  
Università degli Studi di Salerno

## Sviluppi futuri



Studiare come l'equità e il pregiudizio varia a seconda della comunità considerata



Per dare una visione completa e accurata dei problemi causati dai sistemi intelligenti



Approfondire l'utilizzo di alcune tecnologie avanzate su pazienti con malattie terminali per verificare quanto e come la nuova tecnologia e l'intelligenza artificiale possano intervenire sul ciclo di vita umano

✉ a.parziale8@studenti.unisa.it  
🌐 <https://github.com/AlessandraParziale>  
🌐 [www.linkedin.com/in/alessandra-parziale](https://www.linkedin.com/in/alessandra-parziale)

Equità in Sistemi di Machine Learning:  
Stato dell'arte, Problemi e Sfide  
Alessandra Parziale  
Università degli Studi di Salerno



Questa tesi ha contribuito a piantare un albero in \_\_\_\_\_



# Equità in Sistemi di Machine Learning: Stato dell'arte, Problemi e Sfide

Grazie!

Alessandra Parziale

[a.parziale8@studenti.unisa.it](mailto:a.parziale8@studenti.unisa.it)

<https://github.com/AlessandraParziale>

[www.linkedin.com/in/alessandra-parziale](https://www.linkedin.com/in/alessandra-parziale)

