

“IL FUTURO PARTE DA QUI” – REGIONE CALABRIA

Intelligenza artificiale al servizio delle nuove emergenze globali.



“AI-based *CONTENT DISCRIMINATOR* US *MISINFORMATION*”

Andrea Simeri



SUSTAINABLE DEVELOPMENT GOALS

1 NO POVERTY 	2 ZERO HUNGER 	3 GOOD HEALTH AND WELL-BEING 	4 QUALITY EDUCATION 	5 GENDER EQUALITY 	6 CLEAN WATER AND SANITATION 
7 AFFORDABLE AND CLEAN ENERGY 	8 DECENT WORK AND ECONOMIC GROWTH 	9 INDUSTRY, INNOVATION AND INFRASTRUCTURE 	10 REDUCED INEQUALITIES 	11 SUSTAINABLE CITIES AND COMMUNITIES 	12 RESPONSIBLE CONSUMPTION AND PRODUCTION 
13 CLIMATE ACTION 	14 LIFE BELOW WATER 	15 LIFE ON LAND 	16 PEACE, JUSTICE AND STRONG INSTITUTIONS 	17 PARTNERSHIPS FOR THE GOALS 	



AGENDA 2030



16 PEACE, JUSTICE AND STRONG INSTITUTIONS



Conflitti, poca sicurezza, istituzioni deboli ed accesso limitato alla giustizia rappresentano una grande minaccia allo sviluppo sostenibile.

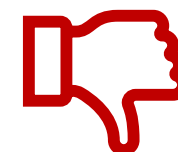
Tra gli obiettivi del sedicesimo SGD:

- **(16.10) Assicurare il pubblico accesso alle informazioni e proteggere le libertà fondamentali, in accord alle leggi nazionali ed agli accordi internazionali.**

Il mondo delle informazioni è cambiato in seguito all'introduzione del WEB 2.0.
Chiunque può produrre informazioni. Chiunque può generare informazioni.



Questo rappresenta una buona notizia fin tanto che queste informazioni sono reali.

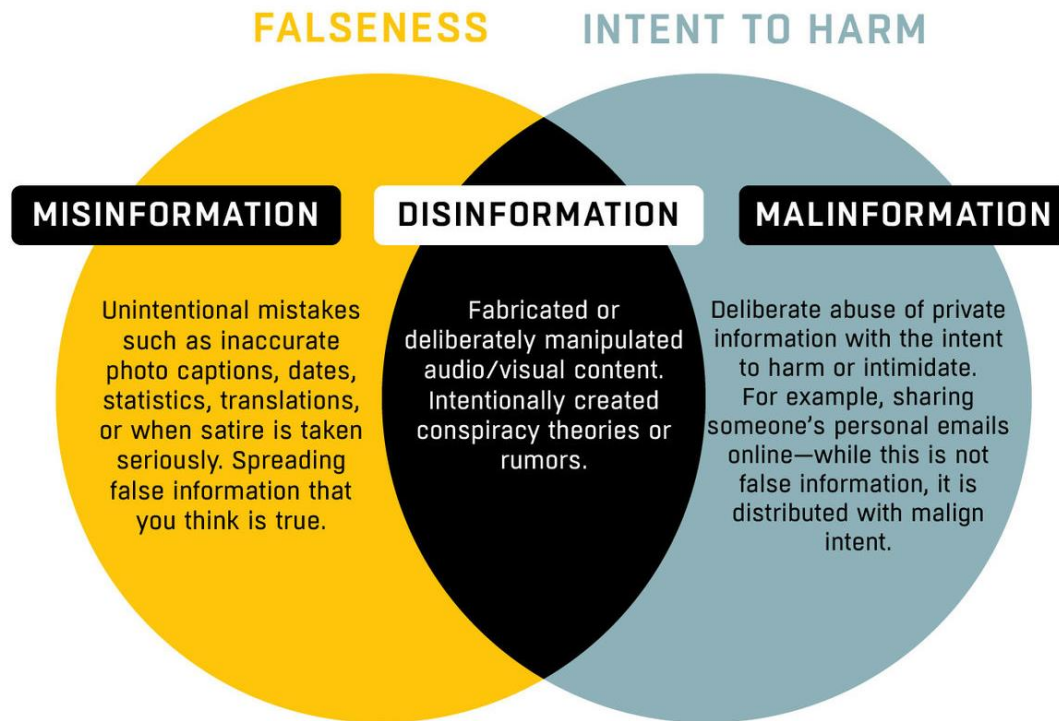


Ma può essere rischioso invece se sono:

- Ingannevoli;
- Poco accurate;
- False.



PROBLEMA



MISINFORMATION:

Informazioni false che sono diffuse senza l'intento di fare del male. Esse non sono attribuite ad una particolare fonte e quindi spesso sono inaffidabili e non verificate.

DISINFORMATION:

Informazioni volutamente ingannevoli e con contenuti falsi; racconti modificati di avvenimenti accaduti; propaganda.

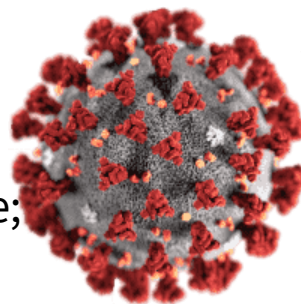
MALINFORMATION:

Diffusione volontaria di informazioni private per scopi personali o di interesse pubblico, come il revenge porn. Manipolazioni volontaria di date, luogo rispetto ai contenuti reali.



PROBLEMA

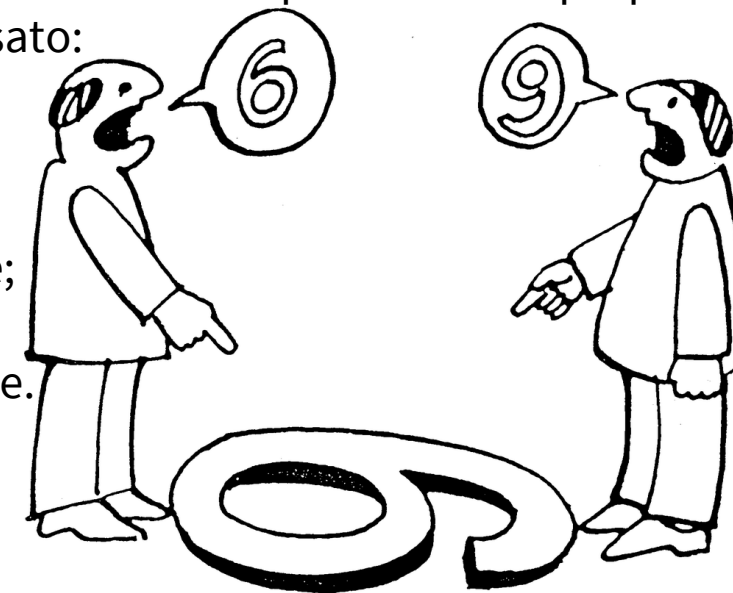
- Informazioni errate rispetto alla crisi pandemica del COVID hanno causato:



- Morti;
- Confusione;
- Panico;
- Scarsa fiducia nelle istituzioni.

- Informazioni manipolate al fine di promuovere i propri interessi hanno causato:

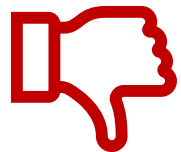
- Conflitti;
- Decisioni errate;
- Radicalizzazione.





SOLUZIONE: CASO D'USO

MISINFORMATION RILEVATA

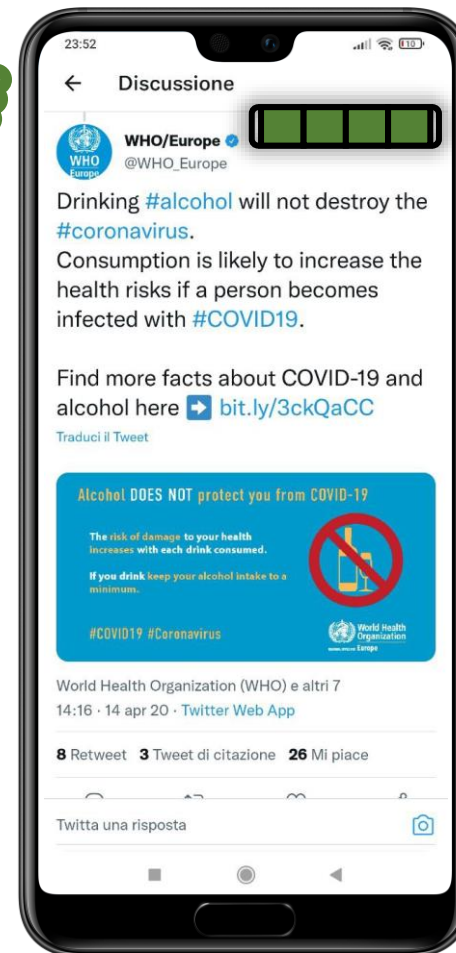


- Non si tratta di una sorgente verificata;
- Il suo significato semantico non è in linea con quello di articoli simili pubblicati da sorgenti verificate;



Basso livello di affidabilità

INFORMAZIONE AFFIDABILE



- E' una sorgente verificata.



Alto livello di affidabilità

SOLUZIONE

Contenuto informativo da verificare

Pagine Gialle di sorgenti verificate



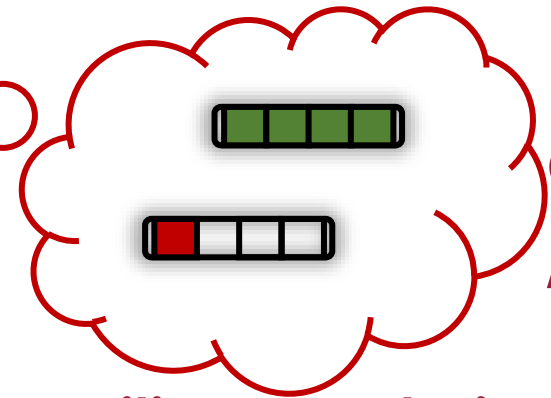
Quanto sono semanticamente equivalenti



?



AI-based CONTENT DISCRIMINATOR



% di AFFIDABILITA'

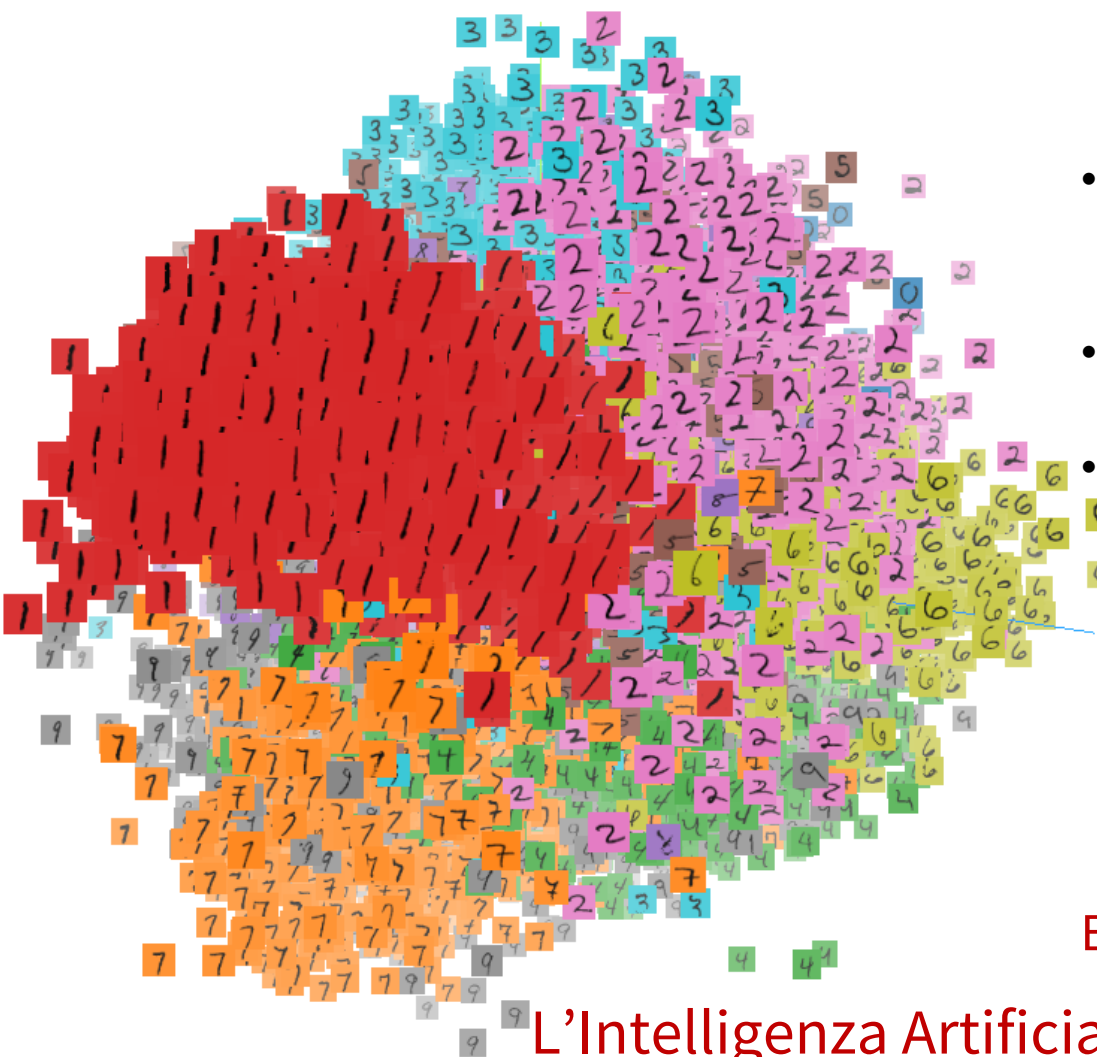
CONTENT DISCRIMINATOR:

- Non è basato su equivalenza lessicografica;
- Non è basato sulla corrispondenza di keyword;

Non viene utilizzata una logica a due valori(SI o NO) perché condanniamo la censura.



Ma cosa si intende per “semanticamente equivalente”?

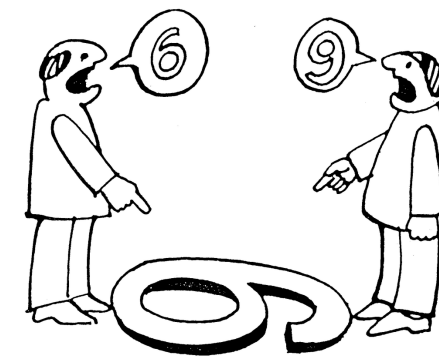
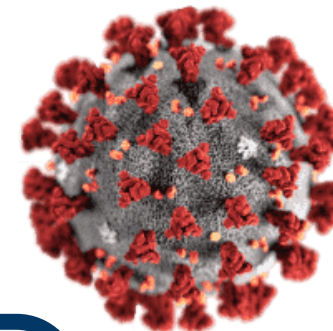


- Ogni riquadro numerato rappresenta un contenuto informativo nello spazio;
- Esso lavora come un sistema di coordinate spaziali;
- Questa rappresentazione gode delle seguenti proprietà:
 - Due punti vicini rappresentano due contenuti informativi semanticamente equivalenti;
 - Due punti lontani rappresentano due contenuti informativi che hanno un significato diverso.

E' quindi possibile calcolare la distanza da una notizia verificata!

L'Intelligenza Artificiale ed in particolare l'NLP possono realizzare tutto ciò!

AI-based Content Discriminator vs MISINFORMATION



- Ingegnere informatico
- Studente di dottorato in ICT
- AI, DL and NLP

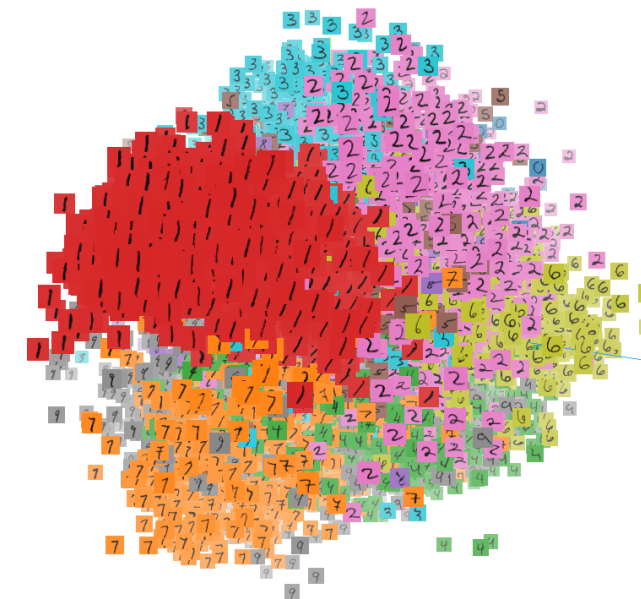
 UNIVERSITÀ DELLA CALABRIA
DIPARTIMENTO DI
INGEGNERIA INFORMATICA,
MODELLISTICA, ELETTRONICA
E SISTEMISTICA
DIMES



andrea.simeri.kr@gmail.com



<https://www.linkedin.com/in/andrea-simeri>





GRAZIE PER L'ATTENZIONE.