

08 Titolo Progetto

“Le parole sono importanti, hanno un potere enorme. Troppo spesso sono utilizzate in modo improprio, offensivo, sleale, impreciso, maleducato, diseducativo. In una parola, duro (senza cuore). E spesso inconsapevole delle conseguenze.”^[1]

alessia valgimigli

Un progetto per rendere visibile l’odio e l’aggressività che riempiono il mondo online, un luogo dove passiamo parte della nostra vita, spesso senza conoscerne le regole e senza la consapevolezza delle proprie responsabilità.



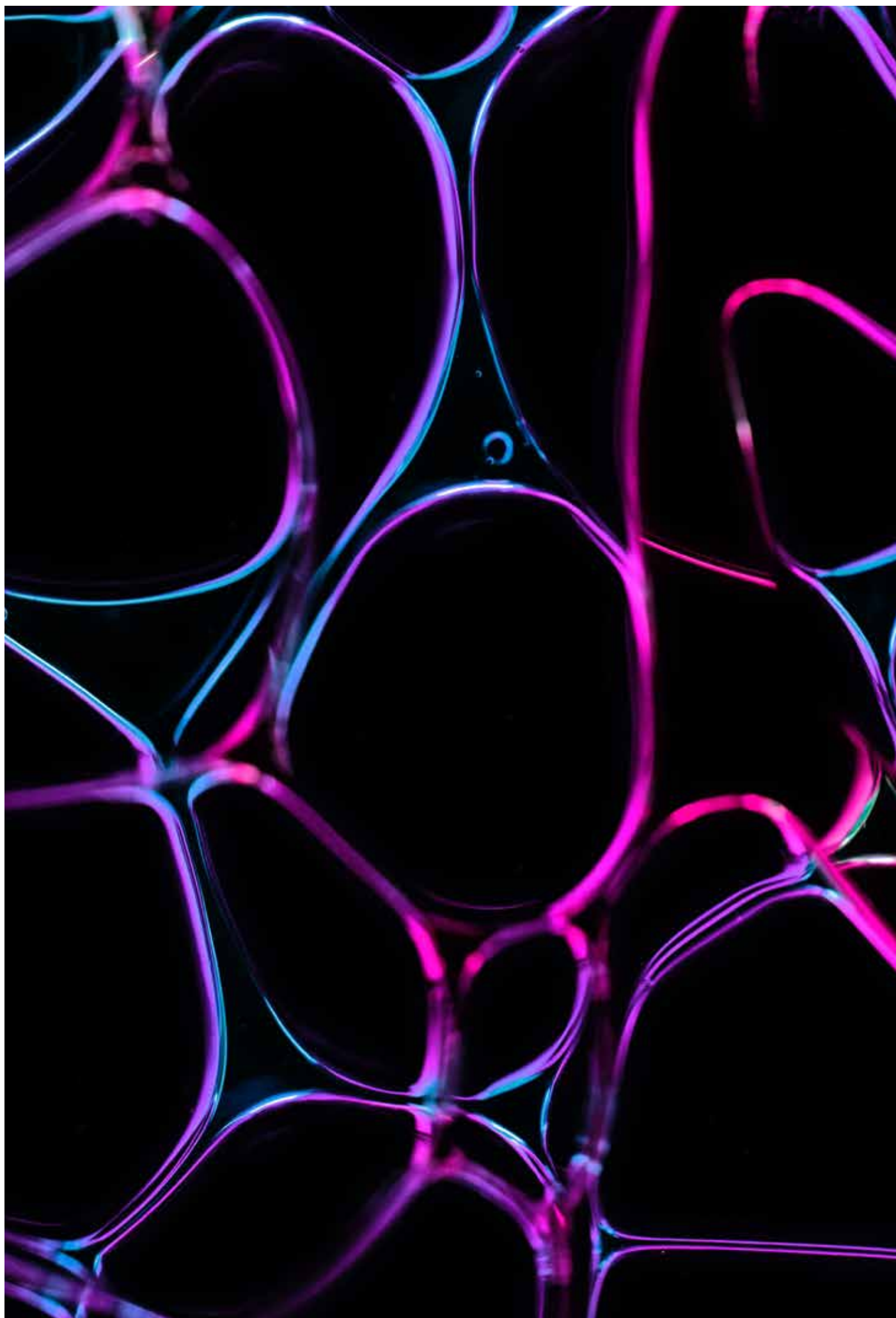
#hatespeech
#parole
#discriminazione
#odio
#Twitter

github.com/ds-2021-unirms
github.com/alessiavalgimigli

[1] descrizione di *Parole Ostili, 10 racconti*, estrapolata dal sito Parole Ostili https://paroleostili.it/i-libri-di-parole-o_stili/libro-10-racconti/

a destra

Immagine rappresentativa del concept e della metafora visiva utilizzata



La violenza dell'hate speech

Il progetto affronta la tematica dell'hate speech, un fenomeno altamente diffuso nella nostra società e nel mondo online. Si tratta di espressioni violente e discriminatorie rivolte ad altre persone con l'intento di aggredire e ferire e che diffondono e giustificano l'odio razziale, la xenofobia, la misoginia e altre forme d'odio basate sull'intolleranza verso gli altri ^[2].

I social network rappresentano un ambiente molto fertile dal punto di vista delle discriminazioni verbali, in quanto risulta più facile giudicare, aggredire e molestare quando si è nascosti e protetti da uno schermo. Ed è proprio attraverso il mondo online che il progetto prende forma. Grazie alla raccolta e all'analisi in tempo reale di tweet contenenti parole discriminatorie, si viene a creare una rappresentazione spaziale che rende visibile come l'hate speech si manifesta su Twitter. Questa visualizzazione nasce come tool di supporto e rafforzamento di piattaforme^[3] volte a denunciare il fenomeno e a creare attorno ad esso un'educazione digitale. La rappresentazione dei dati rende visibile agli utenti due aspetti fondamentali dei discorsi d'odio: la frequenza con la quale essi si verificano in rete e la tipologia di discriminazione che viene fatta.

Riferimenti progettuali

- Love Will Conquer

Si tratta di un'applicazione che raccoglie in tempo reale i tweet contenenti le parole "amore" e "odio" e li geolocalizza su un globo tridimensionale.

I tweet sono animati attraverso una struttura con proprietà variabili a seconda di quando i tweet vengono postati, della loro lunghezza e se questi menzionano amore o odio. Attraverso la rappresentazione è possibile vedere la quantità di amore e odio presente nei vari paesi della Terra.

- Parole tossiche

Parole tossiche è un progetto accademico che prevede il rilevamento di tweet con contenuti d'odio, i quali vanno poi a comporre una nube tossica digitale. Si tratta di un ambiente

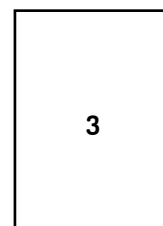
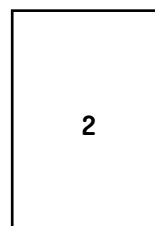
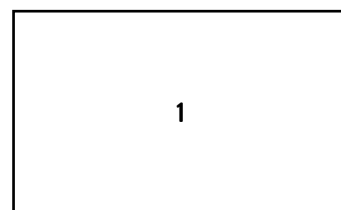
[2] Secondo la Raccomandazione del 1997 del Comitato dei ministri n. 20 del 1997 del Consiglio d'Europa, adottata il 30 ottobre 1997 *discorso d'odio (hate speech) deve essere inteso come l'insieme di tutte le forme di espressione che si diffondono, incitano, sviluppano o giustificano l'odio razziale, la xenofobia, l'antisemitismo e altre forme di odio basate sull'intolleranza e che comprendono l'intolleranza espressa attraverso un aggressivo nazionalismo ed etnocentrismo, la discriminazione e l'ostilità contro le minoranze, i migranti e i popoli che traggono origine dai flussi migratori.*

[3] alcuni esempi di piattaforme che trattano l'argomento:
<http://www.voxdiritti.it/>
<http://www.voxdiritti.it/paroleostili.it>
<https://paroleostili.it/>
<https://www.odiareticosta.it/>

1
Love Will Conquer, Imperial
Leisure for Experiments
with Google, 2012

2
Parole Tossiche, Maria
Chiara Sotgiu, 2016

3
Personal Knowledge Database,
Santiago Ortiz, 2012

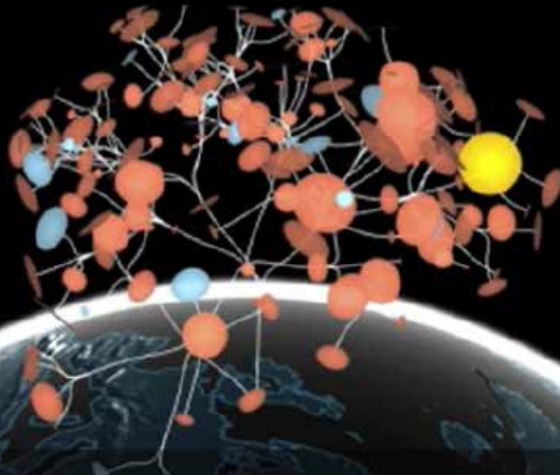


CHOOSE A COUNTRY

United Kingdom
United States
Ireland
Canada
Spain
Brazil
Chile
Germany
Spain
Greece
Indonesia
Italy
Japan
Mexico
Malaysia
Holland
Russia
Sweden
Turkey

The countries above are where the majority of "love"/"hate" tweets come from.

LOVE/HATE WILL CONQUER UNITED KINGDOM



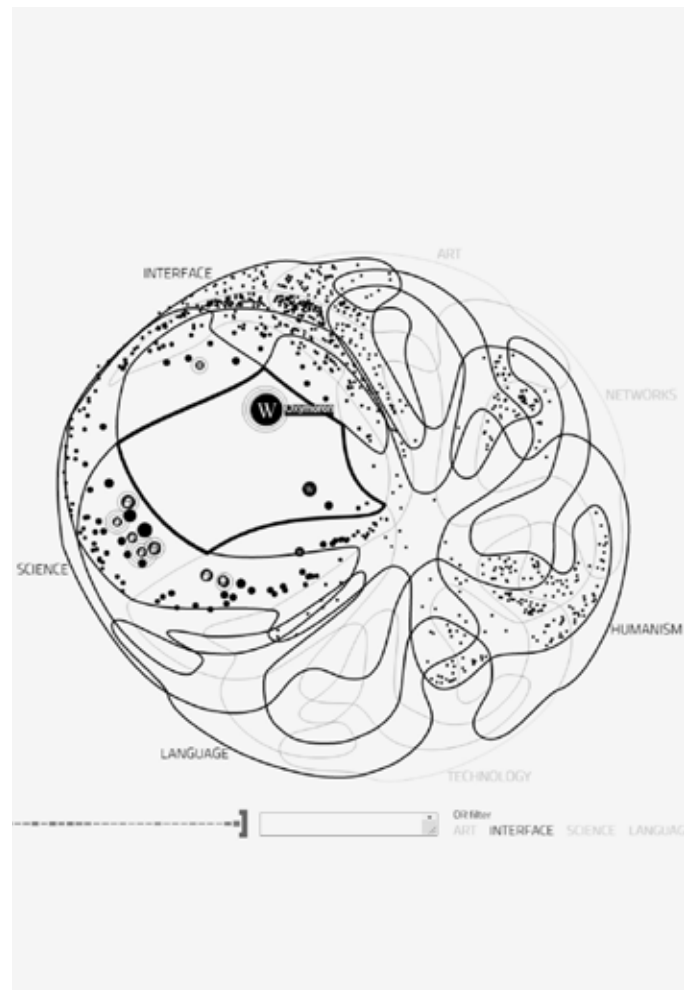
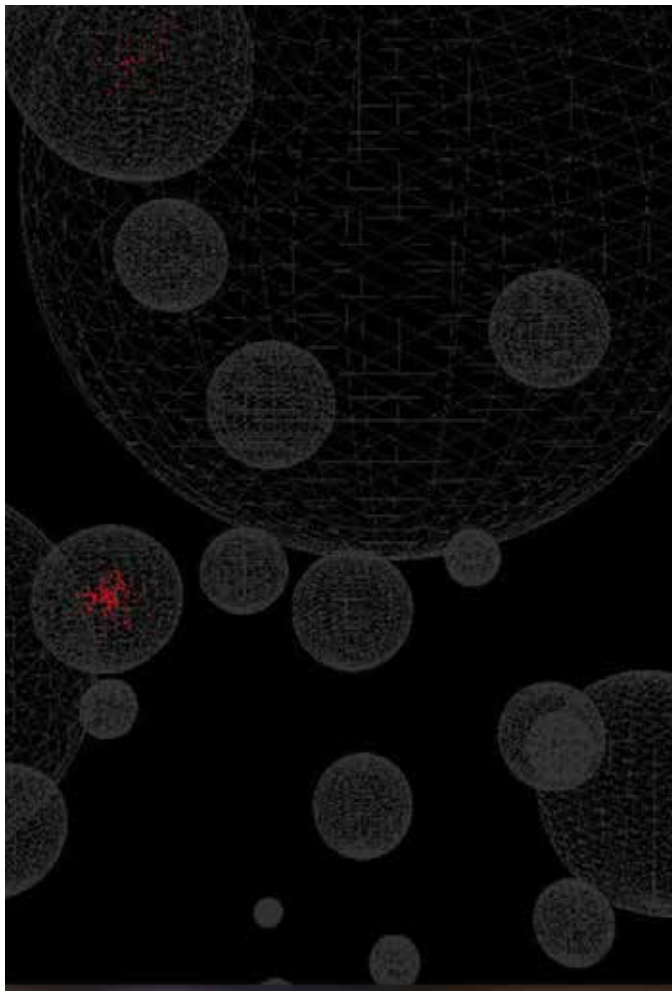
SIGNTASTICA



@MakatonCharity day1 beginners workshop today, always feel more excited than nervous,love meeting people at start of their Makaton journey)
Kingston upon Thames, London | Fri, 17 May 2013 05:58:55 GMT

What is Love Will Conquer? Contact Us

Imperial Leisure | Social Media, Viral Marketing, Digital



tridimensionale navigabile, dove ogni parola crea una nube, ovvero una sfera tridimensionale.

- Personal Knowledge database

Il progetto visualizza e permette l'esplorazione all'interno di un archivio personale, dove sono stati raccolti tutti i riferimenti internet utilizzati dall'autore negli ultimi dieci anni. L'archivio è organizzato attraverso categorie e sottoinsiemi ed è proprio la distribuzione delle informazioni che rende interessante il progetto, insieme alla tipologia di rappresentazione utilizzata.

Il progetto

Un algoritmo interroga Twitter in tempo reale attraverso le API^[4], raccogliendo così gli ultimi tweet pubblicati contenenti parole discriminatorie.

Attraverso una lista di termini contenuta nell'algoritmo, i tweet rilevati vengono raggruppati in quattro categorie diverse, a seconda della forma d'odio che rappresentano^[5].

Questo fenomeno viene visualizzato attraverso la metafora del petrolio: i tweet analizzati diventano delle piccole gocce che, a seconda della tipologia di discriminazione alla quale sono associati, vanno a unirsi a una delle macchie di petrolio, che rappresentano le quattro categorie.

Le gocce all'interno delle macchie rimangono distinte l'una dall'altra, fino a quando non verrà rilevato un tweet contenente una parola d'odio già individuata in precedenza. In quel caso la nuova goccia andrà ad unirsi alla sua omonima, facendola aumentare di dimensione.

In questo modo, all'interno di ogni macchia di petrolio ci saranno delle gocce più o meno grandi, a seconda della frequenza con cui le parole d'odio vengono rilevate su Twitter.

Il petrolio come metafora

Il progetto vuole sottolineare come le parole abbiano un enorme potere comunicativo e come quindi sia importante dargli il giusto peso e il giusto valore. Il modo in cui ci esprimiamo per

[4] l'acronimo API sta per Application Programming Interface, ed è uno strumento che permette agli sviluppatori di lavorare con i dati di Twitter.

[5] le categorie prese in considerazione sono il razzismo, con le parole *musogiallo*, *zingaro*, *clandestino*, *terrone*, *negro*; la misoginia, con le parole *vacca*, *puttana*, *cagna*, *mignotta*, *zoccola*, *troia*; la disabilità, con le parole *mongoloide*, *spastico*, *handicappato*, *storpio*; e l'omofobia, con le parole *frocio*, *lesbicon*, *checca*, *bocchinaro*, *rottinculo*.

in alto

didascalia foto gino che dice cosa sia, dettagli anno, misure, ...

in basso

didascalia foto gino che dice cosa sia, dettagli anno, misure, ...

Rappresentazione goccia

Rappresentazione isole con gocce

rivolgerci alle altre persone può avere un grande impatto su chi ci ascolta e, nel caso dell'hate speech, può provocare conseguenze molto gravi nei confronti delle persone che lo subiscono. Per rappresentare e visualizzare la pesantezza della discriminazione verbale ho deciso di utilizzare la metafora del petrolio, sostanza che inquina e distrugge il nostro ecosistema marino. La macchia oleosa che si crea in seguito alla sua dispersione in mare si allarga rapidamente fino ad occupare grandi porzioni di superficie, diventando così difficile da smaltire e limitando il processo di degradazione. E proprio come il petrolio, l'hate speech soffoca, distrugge, rovina, degrada vite e relazioni. Si allarga e si espande sulle debolezze e vulnerabilità, diventando nocivo per la salute mentale e fisica.

Prototipi

La fase di prototipazione ha avuto inizio con l'utilizzo di Temboo e Processing, strumenti che mi hanno permesso di fare un primo passo verso l'interazione con Twitter e i suoi dati. In questa prima fase, grazie all'utilizzo di Temboo, ho realizzato un algoritmo in grado di interrogare Twitter attraverso una query contenente una parola d'odio. Questa parola viene ricercata tra gli ultimi tweet, restituendo un file JSON, che in seguito viene elaborato visivamente attraverso Processing. In un secondo momento questo meccanismo di analisi dei tweet è stato effettuato attraverso p5js, tralasciando quindi i due strumenti precedenti. La seconda fase della prototipazione è dedicata alla visualizzazione, per la quale ho realizzato un algoritmo attraverso p5js, grazie al quale ogni volta che viene individuato un tweet d'odio viene generata una piccola goccia al centro dello schermo. Questa goccia si muove autonomamente verso la propria macchia di petrolio di riferimento, in base alla categoria discriminatoria a cui appartiene. Parallelamente ai primi due prototipi ho avuto modo di sperimentare con RiTa^[6], libreria che mi ha permesso di individuare parole d'odio all'interno di un file JSON locale e di sostituirle con parole

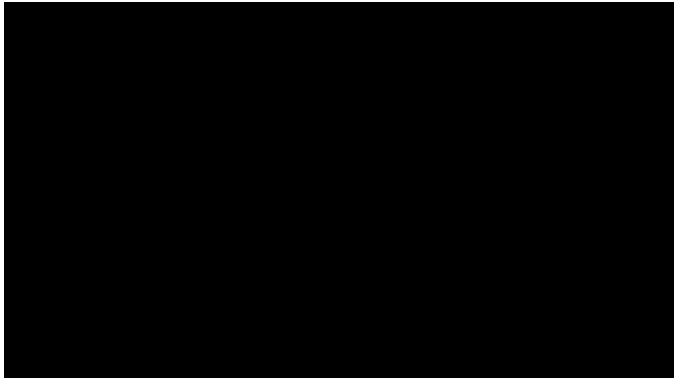
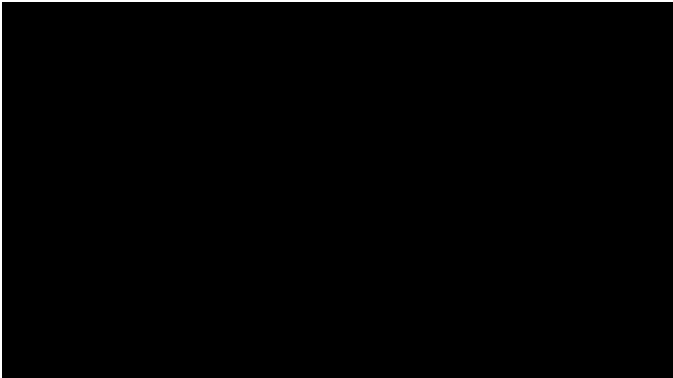
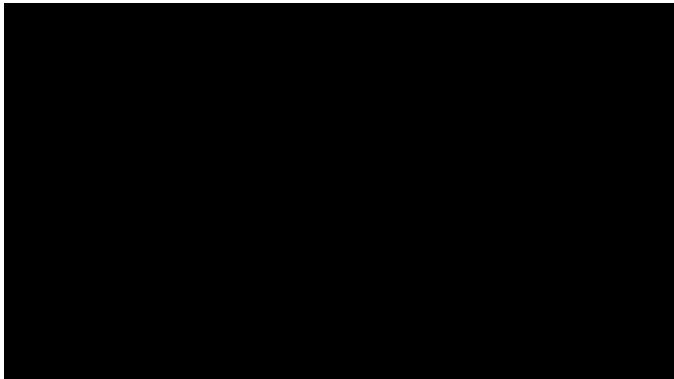
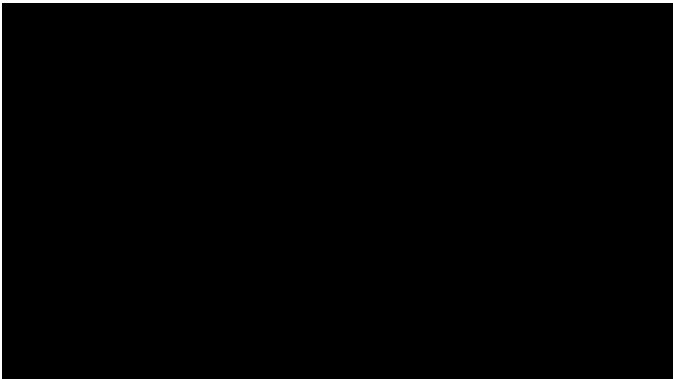
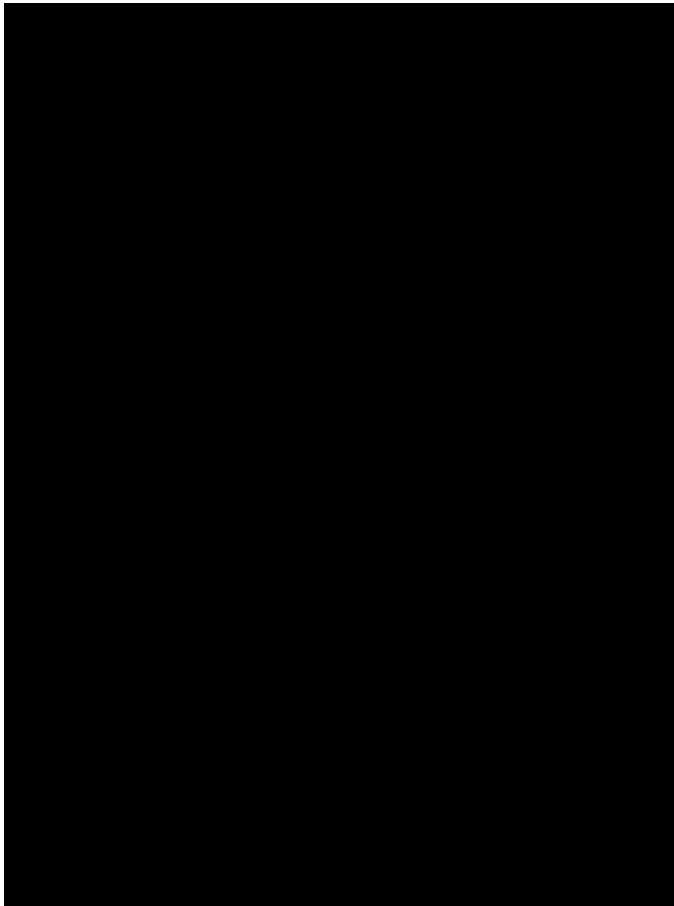
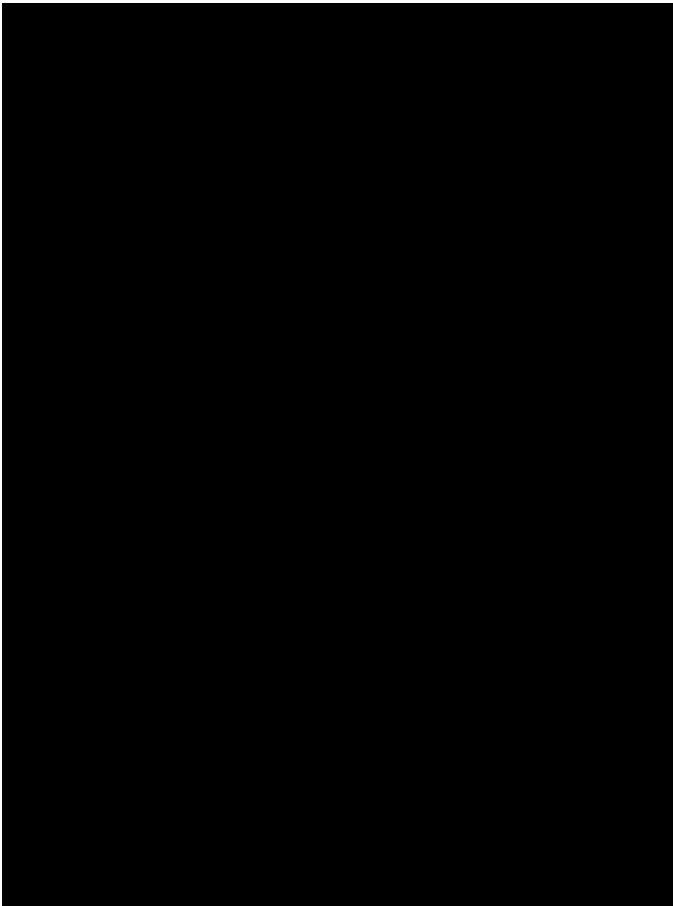
[6] libreria javascript che propone strumenti per generare letteratura computazionale

in alto

didascalia foto gino che dice cosa sia, dettagli anno, misure, ...

in basso

didascalia foto gino che dice cosa sia, dettagli anno, misure, ...



in rima. Sebbene questo concept non rientri nel mio progetto, potrebbe tornare utile per sviluppi futuri o approfondimenti del concept stesso.

Considerazioni e limiti riscontrati

Una delle problematiche principali che si incontrano quando si parla di hate speech è la mancanza di una definizione ufficiale che riesca a definire in modo corretto ed esaustivo il fenomeno. Questa assenza porta con sé non solo problematiche a livello sociale e giuridico, ma anche e soprattutto a livello di riconoscimento del fenomeno stesso. Entrando nel merito del progetto, mi chiedo dunque se sia davvero possibile riconoscere ed etichettare una frase, un post o un tweet come contenuto d'odio. Le parole, come già sottolineato in precedenza hanno un proprio peso, ma molto dipende anche dal contesto in cui queste vengono collocate. Bisognerebbe quindi affrontare l'analisi servendosi di strumenti più complessi, che analizzino la semantica del testo, il contesto, l'emozione dominante e il livello di aggressività del contenuto, come viene fatto nel progetto "Mappa dell'intolleranza" di Vox Diritti. Il secondo limite che ho riscontrato risiede nell'interazione con le API di Twitter. Le application che creo attraverso il profilo da sviluppatore vengono ripetutamente bloccate, impedendomi di poter estrarre dati in tempo reale attraverso le query. Per ovviare al problema, nel momento in cui l'application risultava attiva, ho memorizzato in un unico file tutti i file JSON necessari affinché i prototipi riuscissero a funzionare senza essere collegati direttamente a Twitter.

Sviluppi Futuri

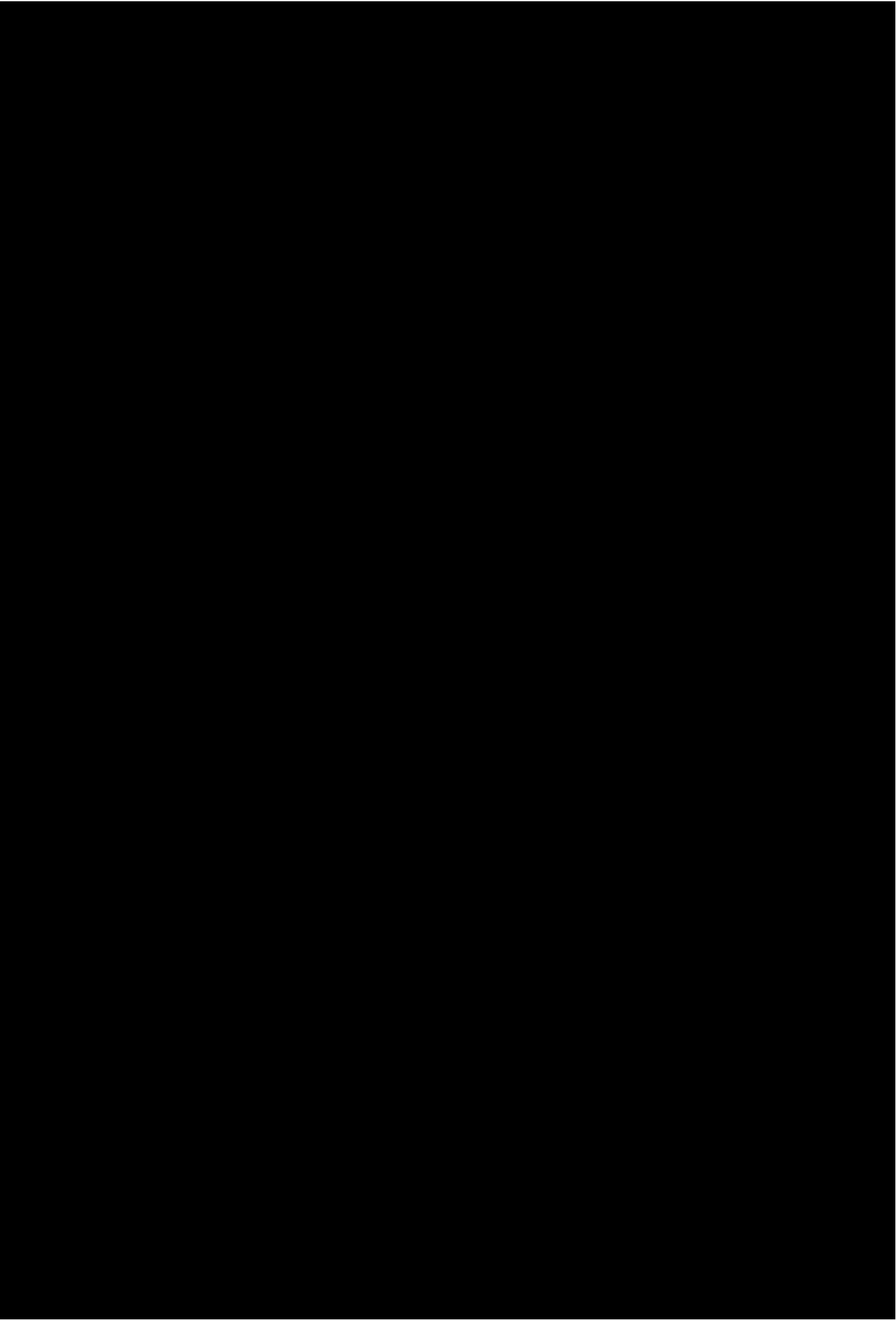
Un primo passo verso uno sviluppo ulteriore del progetto potrebbe essere quello di implementare sistemi più avanzati per analizzare il contenuto dei tweet pubblicati, come nel caso della "Mappa dell'intolleranza" di Vox Diritti, in modo da giudicarne in maniera più fondata il contenuto. Anche la variazione del contesto di utilizzo potrebbe essere uno sviluppo futuro interessante, soprattutto se la visualizzazione di questi dati venisse inserita all'interno del social network stesso. Potrebbe diventare una funzione aggiuntiva di Twitter, volta a sensibilizzare gli utenti al tema e per invitarli a

in alto

didascalia foto gino che dice cosa
sia, dettagli anno, misure, ...

in basso

didascalia foto gino che dice cosa
sia, dettagli anno, misure, ...



segnalare contenuti inappropriati. A sua volta potrebbe essere un sistema di data visualization che compare all'utente in seguito ad una segnalazione eseguita. Un ultimo sviluppo futuro potrebbe prevedere l'implemento della libreria RiTa nell'algoritmo, creando un sistema di censura personale che sostituisca le parole d'odio presenti su Twitter con altre parole assonanti o in rima, in modo che l'utente che sceglie di applicare questo filtro non sia turbato dall'odio presente in rete.

Sitografia

Vox osservatorio italiano sui diritti,
<http://www.voxdiritti.it/>

Parole Ostili
<https://paroleostili.it/>

Temboo
<https://temboo.com/processing/getting-started>

Twitter's Developer Platform
<https://developer.twitter.com/en>

RiTa
<https://rednoise.org/rita/>

Bibliografia

Faloppa F., *Odio. Manuale di resistenza alla violenza delle parole*, Milano, Utet (2020)

Che cos'è l'hate speech, Ep. 2 Stagione 3,
«AntiCorpi» a cura di Jennifer Guerra, novembre
2019, 33:43, https://open.spotify.com/episode/2mi9OfRt6AExItlzXNZ5pJ?si=z5RqYBzMSAe_2i-WogiTDw&dl_branch=1 (consultato 08/06/2021)

