



PREVENTING HATE AGAINST REFUGEES AND MIGRANTS

Interface

USER'S GUIDE



This project is funded by the European Union's Rights, Equality and Citizenship Programme (2014-2020), REC-RRAC-RACI-AG-2019, Grant Agreement number 875217

November 2020

<http://pharmproject.usal.es>

This report constitutes Deliverable 19, for Work Package 4 of the PHARM project.

October 2020

© 2020 – PHARM, Preventing Hate Against Refugees and Migrants, – GA number 875217.

General contact: pharm.project.eu@gmail.com

For more information type the e-mail address of the corresponding author

Information may be quoted provided the source is stated accurately and clearly.
This publication is also available via <https://pharmproject.usal.es/>

This project is funded by the European Union's Rights, Equality and Citizenship Programme (2014-2020).
REC-RRAC-RACI-AG-2019. Grant Agreement number 875217

The content of this report represents the views of the author only and is his/her sole responsibility. The European Commission does not accept any responsibility for use that may be made of the information it contains.

Copyright © Pharm Consortium, 2020

All rights reserved. No part of the report may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, and may not be quoted or cited, without prior permission in writing from the project coordinator.

The views expressed during the execution of the project, in whatever form and or by whatever medium, are the sole responsibility of the authors. The European Union is not liable for any use that may be made of the information contained therein.

VERSION CONTROL SHEET

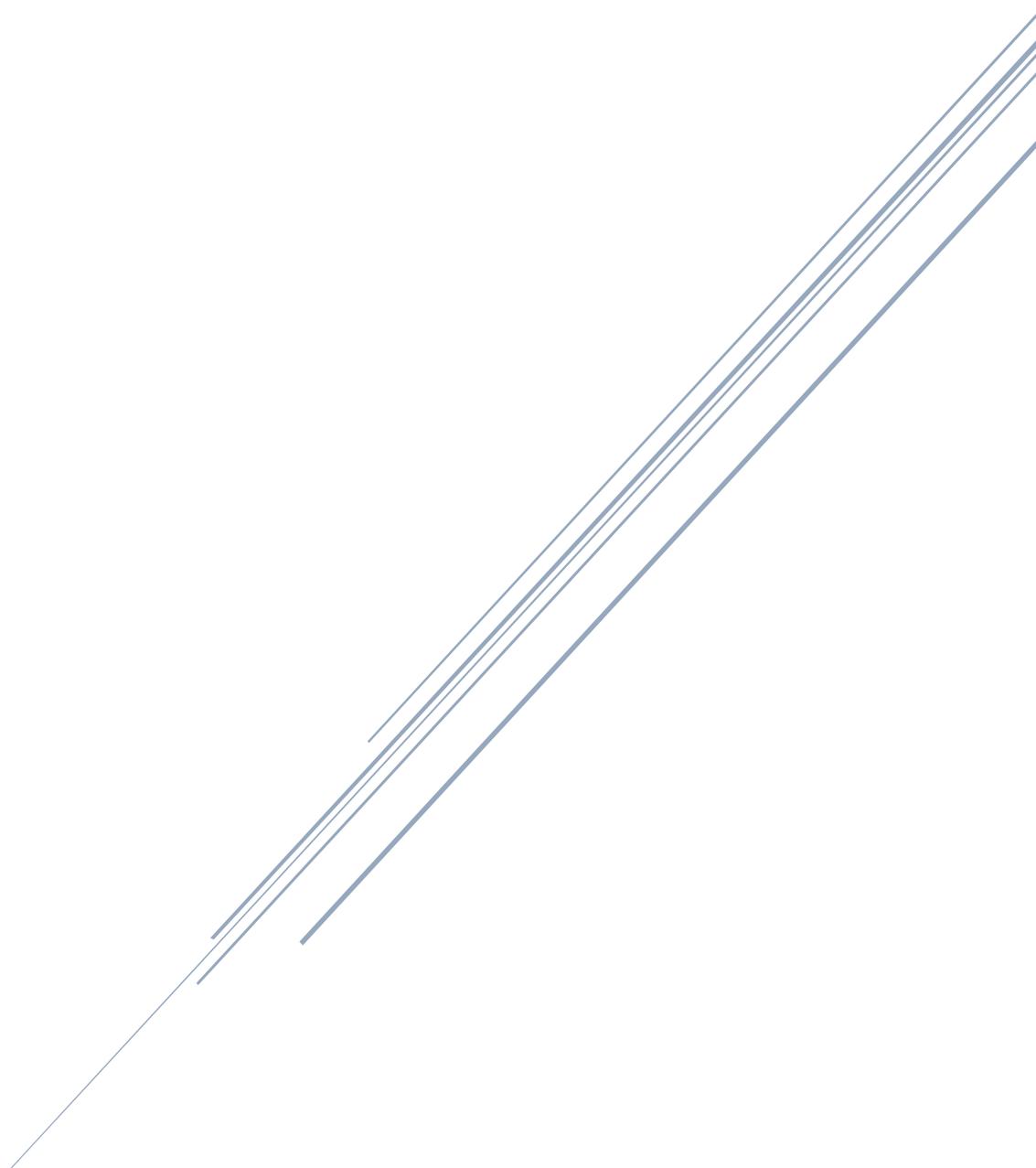
Deliverable number + title	D10. Guide of use of the interface
Prepared by	AUTH TEAM
Work package number	2
Work package leader	Andreas Veglis
Dissemination level (PU, CO)	PU
Delivery date	30/11/2020
Submission date to EC	30/11/2020
Main authors	Lazaros Vrysis / Nikolaos Vryzas
Reviewers	Carlos Arcila Calderón / Patricia Sánchez Holgado

REVISION HISTORY

Version	Date	Summary of changes	Initials	Changes marked
1				

GUIDA TECNICA E PER L'UTENTE

Per l'interfaccia Internet del progetto PHARM



Versione italiana 1.0

Sommarario

1	Informazioni tecniche	2
1.1	L'interfaccia Internet	2
1.2	Informazioni tecniche	3
1.2.1	Attuazione	3
1.3	Utenti	4
1.4	Azioni.....	5
1.4.1	Search (Cerca)	5
1.4.2	Analyze (Analizza).....	5
1.4.3	Scrape.....	6
1.4.4	Annotate (Annota)	7
1.4.5	Submit (Invia)	7
1.5	Formato dei dati.....	7
2	Guida all'interfaccia.....	10
2.1	Introduzione.....	10
2.2	Home.....	10
2.3	Search (Ricerca).....	10
2.4	Analyze (Analizza).....	11
2.5	Scrape.....	12
2.6	Annotate (Annota)	13
2.7	Submit.....	13
2.8	Login.....	15
3	Guida alle annotazioni.....	16
3.1	Generale.....	16
3.2	Annotate (Annota)	16
3.3	Approve (Approva).....	18

1 Informazioni tecniche

PHARM è un progetto europeo finanziato dall'Unione Europea, nell'ambito del programma Rights, Equality e Citizenship REC-RRAC-RACI-AG-2019 (GA n. 875217). L'obiettivo principale di Preventing Hate Against Refugees and Migrants (PHARM) è monitorare e modellare l'incitamento all'odio contro rifugiati e migranti in Grecia, Italia e Spagna per prevedere e combattere i crimini ispirati dall'odio e contrastarne gli effetti utilizzando tecniche all'avanguardia, come data journalism e narrazione persuasiva. Il risultato principale sarà l'identificazione e la riduzione dell'incitamento all'odio online e la previsione di potenziali crimini d'odio.

1.1 L'interfaccia Internet

L'interfaccia Internet PHARM è un'interfaccia web per un formato integrato semistrutturato orientato all'archiviazione e alla ricerca di notizie multi-fonte e contenuti dei social media. L'interfaccia è costituita da un database e da un'interfaccia web front-end per l'esposizione di dati e funzionalità agli utenti. La prima schermata che viene presentata all'utente è la "Schermata iniziale", che è presentata nella Figura 1. Oltre ad alcune informazioni di base sul progetto, questa schermata fornisce un punto di partenza per accedere alle analisi supportate su un testo.

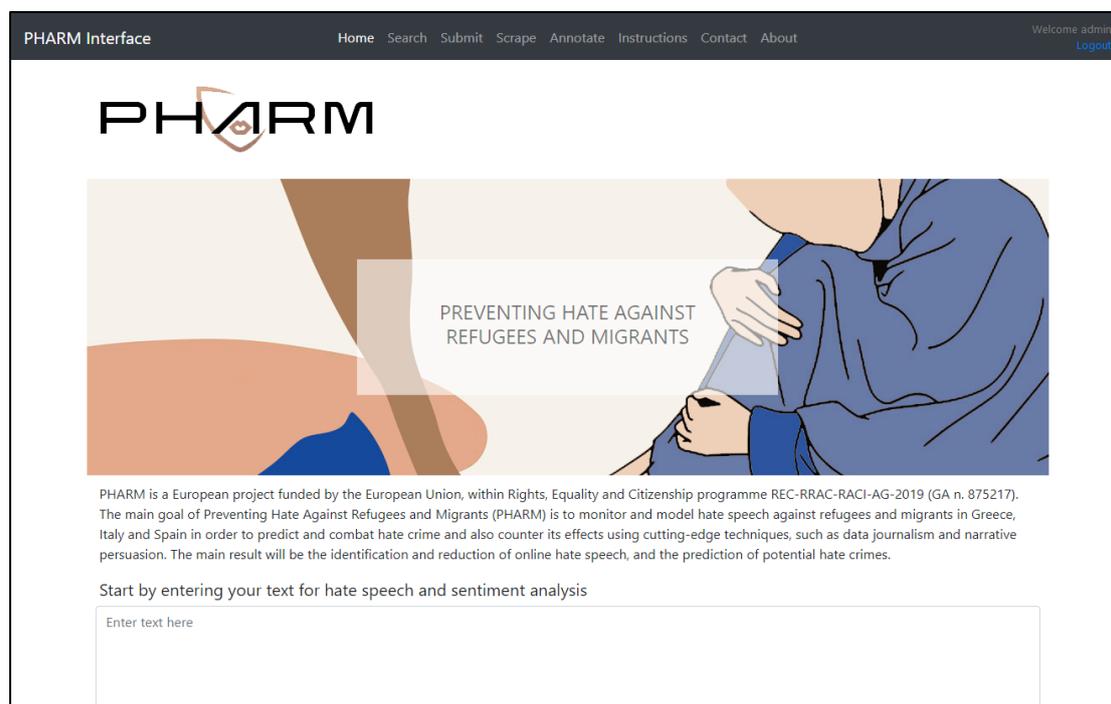


Figura 1 La schermata principale dell'interfaccia Internet

Le azioni principali supportate dall'interfaccia sono:

- **Search (Ricerca):** La ricerca riguarda l'interrogazione del database per il filtraggio e il *download* dei dati rispetto all'incitamento all'odio.
- **Analyze (Analizza):** L'analisi fornisce risultati di vari tipi di analisi del testo, come rilevamento di incitamento all'odio, sentiment analysis e stima della geolocalizzazione.
- **Scrape:** Questa azione consente di estrarre dati di testo dalle piattaforme social di Twitter e YouTube.
- **Annotate (Annota):** Abilita l'annotazione dei dati raccolti.

- **Submit (Invia):** Questa funzionalità consente di aggiungere nuovi estratti di discorsi di incitamento all'odio al database.

Le funzioni di analisi e ricerca sono liberamente esposte a tutti gli utenti che utilizzano l'Interfaccia Internet, mentre per accedere alle azioni *Submit*, *Scrape* e *Annotate* è necessaria l'autenticazione. Se gli utenti vogliono far parte del progetto PHARM per contribuire con dati aggiuntivi, devono contattare il nostro team per richiedere di accedere con tale ruolo attraverso la pagina dei contatti di Inter-face.

1.2 Informazioni tecniche

1.2.1 Attuazione

Per lo sviluppo dell'Interfaccia è stato utilizzato il *framework web Python Flask*. La scelta è motivata in quanto anche le procedure di *Natural Language Processing (NLP)* e di analisi dei dati sono scritte in Python. Con questo approccio, tutte le funzionalità dell'interfaccia possono essere incluse in un progetto comune. Il front-end dell'interfaccia utente grafica è stato progettato principalmente in Bootstrap, una popolare libreria HTML, CSS e JavaScript. Ove necessario, sono stati aggiunti ulteriori blocchi HTML, CSS e JavaScript. Il progetto Flask è stato implementato su una macchina virtuale e viene offerto utilizzando la *Waitress*, un server WSGI Python puro di qualità di produzione con alte prestazioni.

Anche le funzioni di analisi che alimentano l'interfaccia sono implementate nel linguaggio Python. Sono state distribuite varie librerie per implementare i metodi di analisi supportati, insieme ad algoritmi personalizzati che sono stati sviluppati specificamente per il progetto PHARM. Alcuni dei metodi più importanti consentono lo scraping dei dati, la stima di data e ora, la stima della geolocalizzazione, il rilevamento della lingua e il rilevamento di incitamento all'odio. Funzioni aggiuntive, come la sentiment analysis, saranno via via disponibili nell'Interfaccia rispettando le tempistiche del progetto. Di seguito vengono presentate le principali funzionalità e scelte progettuali di questi metodi. Per ulteriori informazioni tecniche, visitare il [PHARM Github Repository](#).

1.2.1.1 Scraping

Il progetto supporta la raccolta di testo da Twitter tramite l'API appropriata e il metodo stream. Tweepy è utilizzato per accedere all'API di Twitter. Sono stati sviluppati quattro dizionari per filtrare i tweet, incluse parole chiave in greco, inglese, spagnolo e italiano. La raccolta di commenti su YouTube è supportata tramite l'API di Google. È richiesta all'utente una query di ricerca pertinente all'argomento (esempio: "Rifugiati migranti") e vengono raccolti i commenti dai risultati principali (video).

1.2.1.2 Stima data e ora

È stato implementato un metodo per rilevare e standardizzare le informazioni su data e ora da metadati e testo. Oltre alla posizione e alla lingua, quando sono disponibili i metadati, PHARM utilizza informazioni aggiuntive pertinenti per l'analisi dei discorsi di incitamento all'odio. Alcune di queste informazioni aggiuntive, come la data o l'ora, potrebbero essere disponibili nella maggior parte dei casi in diversi formati. Ciò introduce la necessità della standardizzazione. Per le esigenze di questo requisito, vengono sfruttati i pacchetti Python *dateparser*, *datefinder* e *parsedatetime*, classificati da una maggiore precisione a una maggiore probabilità di restituire un risultato. Se il metodo più accurato non riesce a rilevare un oggetto *DateTime*, viene chiamato il servizio successivo. Il rilevamento si basa sull'analisi dei metadati, in cui sono comunemente presenti le informazioni sulla data. Se il rilevamento di

data e ora (*DateTime*) non riesce per tutti i servizi attraverso i metadati, lo stesso flusso di lavoro viene applicato ai dati di testo.

1.2.1.3 *Stima geolocalizzazione*

È stato implementato un metodo per rilevare la geolocalizzazione dai dati di testo. È stata implementata la libreria *geopy*, insieme al *geocoder nominatim*. Le entità nominate (caratteristiche linguistiche) sono isolate dai testi, secondo la seguente classificazione: GPE (paesi, città, stati), LOC (montagne, specchi d'acqua), FAC (edifici, aeroporti, autostrade, ecc.), ORG (aziende, agenzie, istituzioni, ecc.).

1.2.1.4 *Rilevamento della lingua*

PHARM analizza testo prodotto in lingua greca, italiana e spagnola, ma molte delle fonti potrebbero avere contenuti in altre lingue straniere o dialetti locali. Per lavorare con queste tre lingue nazionali, è stata codificata una procedura per rilevare la lingua del testo multimediale quando non è stata dichiarata correttamente. Viene adottato un approccio ricorsivo per una maggiore robustezza. Vengono utilizzati i servizi *Textblob*, *googletrans* e *langdetect*. Se uno di questi non raggiunge l'obiettivo, sono utilizzati quelli successivi.

1.2.1.5 *Rilevamento Incitamento all'odio*

Sono stati implementati un paio di metodi per trovare i termini chiave per il rilevamento dell'incitamento all'odio. Questi includono una semplice corrispondenza delle stringhe, una corrispondenza approssimativa delle stringhe con l'uso di metriche appropriate, come *Levenshtein Distance*, *Damerau-Levenshtein Distance*, *Jaro Distance*, *Jaro-Winkler Distance*, *Match Rating Approach Comparison*, *Hamming Distance*. La corrispondenza dei termini mira anche a rendere neutrale i suffissi, adattandoli ai nomi di molte lingue (ad esempio, la lingua greca presenta suffissi diversi per il genere o versioni singolare / plurale). È stato inoltre sviluppato un approccio word-vector, tenendo conto del significato semantico dei termini. È stato implementato ed è in fase di valutazione un approccio ibrido basato sul dizionario con frasi predefinite, insieme a combinazioni di termini dinamici (cioè aggettivi combinati con nomi). Viene caricato un modello linguistico (in base alla lingua del testo), vengono seguite pratiche comuni (minuscolo, lemmatizzazione, stop-words e rimozione della punteggiatura) e vengono cercati i termini mirati nel testo. Se viene trovato un termine (o una combinazione di termini), il testo viene contrassegnato come "incitamento all'odio". Un ulteriore algoritmo di classificazione è in fase di sviluppo e sarà incorporato nell'interfaccia.

1.2.1.6 *Sentiment analysis*

Due approcci per l'analisi dell'incitamento all'odio sono in fase di sviluppo e saranno incorporati nell'interfaccia, secondo il calendario del progetto. Questi includono una tecnologia di apprendimento automatico non supervisionata, basata sul lessico e una supervisionata, che sfrutta i metodi per la sentiment analysis.

1.3 Utenti

L'interfaccia Internet autorizza due tipi di utenti: il visitatore e il contribuente. Un visitatore può accedere ai moduli di analisi e ricerca, mentre il contribuente ha pieno accesso a tutte le azioni disponibili. I flussi di lavoro di base per un visitatore (utente non autenticato) e un collaboratore (utente autenticato) sono definiti come segue:

Collaboratore

- Search -> Analyze
- Scrape -> Annotate -> Submit

Visitatore

- Search -> Analyze

1.4 Azioni

Come accennato in precedenza, le azioni principali supportate dall'interfaccia sono cinque: Analyze (Analizza), Search (Cerca), Scrape, Annotate (Annota) e Submit (Invia).

Di seguito una descrizione più dettagliata di queste funzioni.

1.4.1 Search (Cerca)

Una delle funzionalità principali dell'interfaccia Internet è la navigazione attraverso le registrazioni di incitamento all'odio (testo) del database. L'utente può visualizzare tutti i risultati o applicare una serie di filtri (esempio: Origine, lingua, data). In dettaglio, i filtri disponibili sono:

- Selezione della sorgente (*Twitter, Facebook, Articoli del sito web, commenti del sito web*).
- Selezione della lingua (*inglese, greco, italiano, spagnolo*).
- Selezione data e ora (mostra solo i risultati all'interno di un periodo specifico).
- Filtro per parole chiave (una query di ricerca per trovare occorrenze nei testi).

L'utente può visualizzare in anteprima i risultati, scegliere la visualizzazione "Semplice" o "Scientifica", scaricarli come file CSV o JSON o visualizzare informazioni più dettagliate per ogni record.

1.4.2 Analyze (Analizza)

Quando viene selezionato un record (o per un testo inserito nella schermata principale), viene visualizzata una vista che presenta informazioni dettagliate (Figura 2). La posizione è contrassegnata su una mappa e i risultati di vari algoritmi di analisi del testo vengono presentati con grafici (icone, barre, ecc.). I risultati riguardano il rilevamento dell'incitamento all'odio (sia per metodi di classificazione non supervisionati che per metodi di classificazione supervisionati), sentiment analysis (per metodi di classificazione sia non supervisionati che supervisionati), rilevamento di parole frequenti e raccolta di stralci di incitamento all'odio.

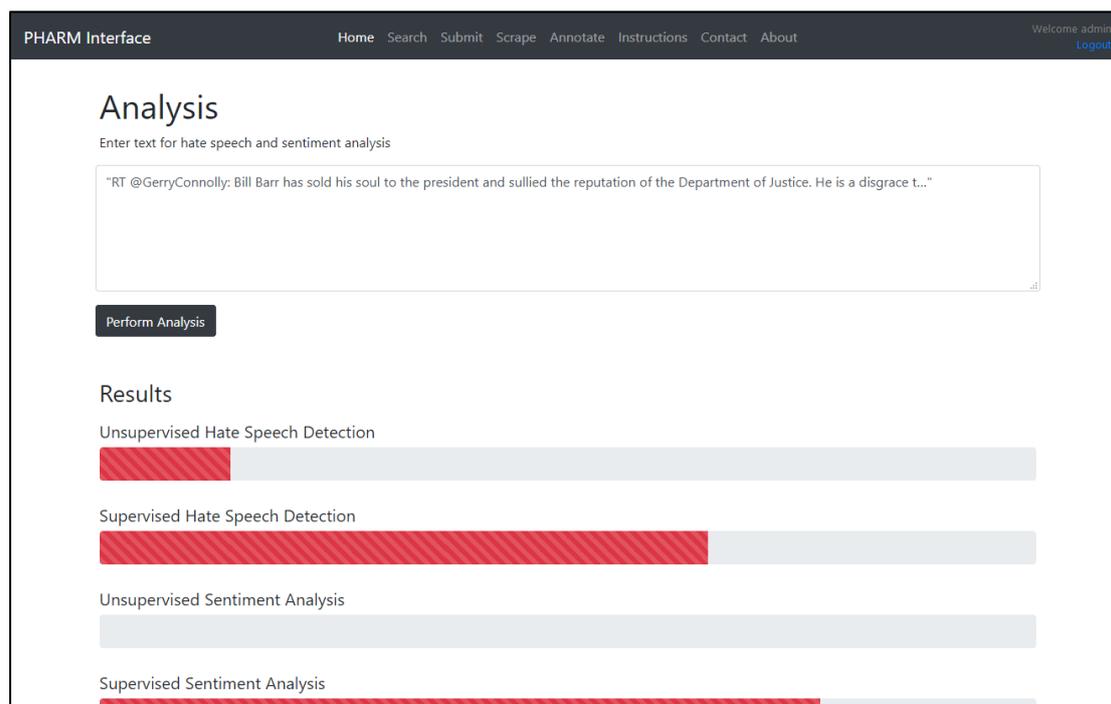


Figura 2 La schermata di analisi dell'interfaccia Internet

1.4.3 Scrape

Questo modulo consente la raccolta di massa di dati di testo da due piattaforme popolari: Twitter e YouTube. Un utente può raccogliere dati sull'incitamento all'odio da Twitter, semplicemente scegliendo la lingua desiderata (greco, inglese, italiano o spagnolo) e attivando il processo premendo il pulsante "Scrape". Lo streaming dei dati inizia e i tweet vengono raccolti sulla base di un lessico specifico delle lingue sviluppate nel contesto del progetto. Il processo si interrompe dopo un intervallo di tempo configurabile dall'utente. Viene fornito un collegamento per scaricare un file JSON che contiene i dati. Questi dati possono essere utilizzati per qualsiasi attività desiderabile di elaborazione del linguaggio naturale (NLP). L'utente può ripetere il processo più volte.

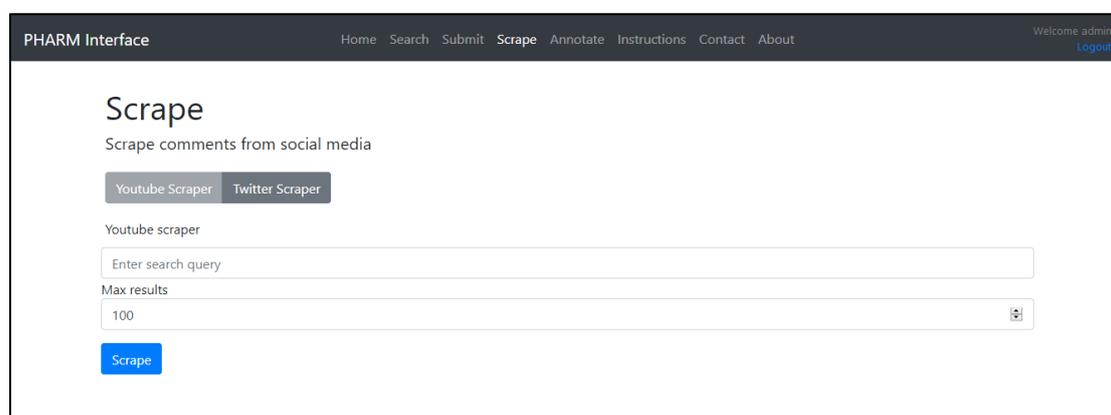


Figura 3 La schermata Scrape dell'interfaccia Internet

Nel caso di YouTube, invece di selezionare la lingua, è necessario impostare una query di ricerca. La query di ricerca può includere singoli termini di ricerca o una combinazione di essi, separati da una virgola. Il processo in questo caso è istantaneo che mostra i dati corrispondenti. Questi dati possono essere scaricati anche come file (CSV o JSON). Per ulteriori

informazioni sui lessici di filtraggio di Twitter e sul formato dei dati basato su JSON, è possibile fare riferimento a [PHARM GitHub repository](#).

1.4.4 Annotate (Annota)

Il processo di annotazione è alimentato dallo strumento doccano. Doccano è un sistema di gestione delle annotazioni per i dati di testo e può essere utilizzato per lo sviluppo di set di dati per facilitare la classificazione, l'analisi del sentiment, l'etichettatura di entità o le attività di traduzione. Nel contesto del progetto PHARM, viene utilizzato per la classificazione del testo e ogni record deve essere etichettato con tag specifici.

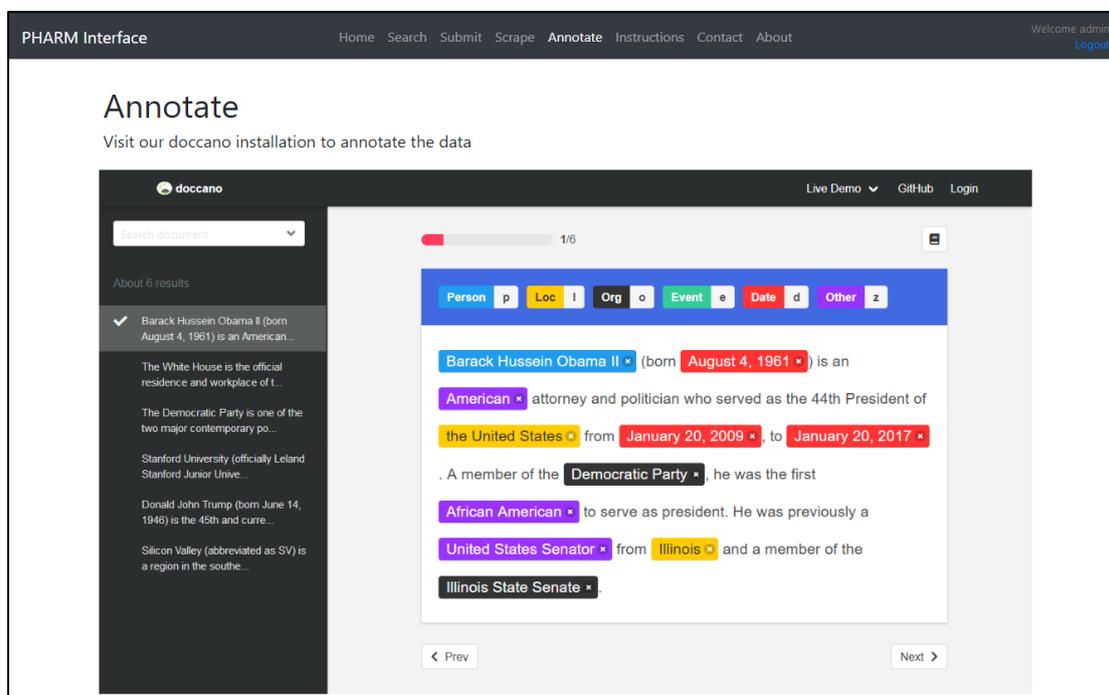


Figura 4 Schermata Annotate dell'interfaccia Internet

1.4.5 Submit (Invia)

Per aggiungere un nuovo record (o un insieme di record), l'utente deve selezionare "Submit" (Invia) dal menu in alto (Figura 5). L'immissione dei dati può essere eseguita manualmente (uno per uno) o assieme. Per quanto riguarda il primo metodo, l'utente dovrebbe impostare tutti i dati (testo) e metadati (fonte, lingua, data, odio, sentimento, ecc.) Tramite i moduli di input corrispondenti (esempio: campi di testo, pulsanti di opzione, caselle di controllo, ecc.). Se i dati sono già formati in modo appropriato (vedi Paragrafo 1.5), possono essere importati anche come file JSON.

1.5 Formato dei dati

Tenendo conto dei requisiti del progetto (ovvero l'uso di alcune informazioni supplementari pertinenti per l'analisi dell'incitamento all'odio), le fonti utilizzate per raccogliere il contenuto pertinente (ad es. Articoli e commenti del sito Web, commenti di YouTube e tweet di Twitter), considerazioni sull'interoperabilità e la compatibilità per l'importazione/esportazione di dati su applicazioni di terze parti che vengono sfruttate (es. piattaforma doccano per annotazioni), sono state impostate alcune specifiche generali per il formato dei dati.

Figura 5 La visualizzazione Invia dell'interfaccia Internet

Il campo principale è il testo (contenuto), accompagnato da id, annotazioni e metacampi. Il meta field è un contenitore che include tutti i metadati aggiuntivi. Un set fondamentale/minimo di metadati viene utilizzato per tutte le piattaforme (ad esempio type, plang, pdate, phate, psent, pterms, ploc). È più probabile che questi campi vengano trovati per tutti i record su origini diverse. La figura seguente fornisce una vista gerarchica dello schema di dati proposto (Tabella 1).

Nei casi di web scraping, i metadati dipendono dai metadati disponibili forniti da ciascun sito e dalla struttura specifica del sito, mentre per i commenti di YouTube e i tweet di Twitter, dove vengono utilizzate le API corrispondenti, sono stati selezionati metadati specifici che vengono raccolti insieme al testo del commento (Tabella 2).

Tabella 1 I campi comuni del formato dati PHARM

Field	Description
id	unique identifier
annotations	hate speech and sentiment annotations
meta	metadata
text	content
type	type of text (tweet, article, comment, etc.)
plang	language detection via PHARM analysis
pdate	datetime estimation via PHARM analysis
phate	hate speech detection via PHARM analysis
psent	sentiment analysis via PHARM analysis
pterms	frequent terms collection via PHARM analysis
ploc	geolocation estimation via PHARM analysis

Tabella 2 I campi aggiuntivi (metadati) per i record di YouTube e Twitter

Twitter	YouTube
tweet_id	comment_id
is_retweet	reply_count
is_quote	like_count
user_id	video_id
username	video_title
scr_name	channel
location	video_description
followers	author_id
friends	author_name
date	date

Per una presentazione più descrittiva, di seguito vengono riportati dati di esempio da diverse origini.

Facebook

```
{
  "id": "1",
  "annotations": [
    { "label": "Hate", "user": 2 },
    { "label": "Negative", "user": 2 }
  ],
  "meta": {
    "id": "80056833",
    "type": "facebook_comment",
    "source": "https://www.facebook.com/provinciale.rivista",
    "plang": "it",
    "pdate": "1571-10-15 00:00:00",
    "phate": "(nero cinesi)",
    "pterm": "",
    "ploc": "Italy, United States of America",
    "text": "andrebbe anche fatta con Venezia (principale flotta a Lepanto). 1571: le mie potentissime galeazze spaccheranno il culo all infedele ottomano. 2020: per favore basta quarantena, devo vendere lasagne sur-gelate in nero ai cinesi a 50 euro"
  }
}
```

Twitter

```
{
  "id": "2",
  "annotations": [
    { "label": "Hate", "user": 2 },
    { "label": "Neutral", "user": 2 }
  ],
  "meta": {
    "type": "twitter_comment",
    "date": "10/13/2020",
    "tweet_id": 1315981643111432192,
    "is_retweet": true,
    "is_quote": false,
    "user_id": 1025701121749340160,
    "username": "Christina Dim",
    "scr_name": "ChristinaDim31",
    "location": "Αθήνα",
    "followers": 3178,
    "friends": 4756,
    "quoted_text": "",
    "pid": 27454420,
    "plang": "el",
    "pdate": "",
    "phate": "(Τουρκία μετανάστες)",
    "pterm": "",
    "ploc": "Turkey",
    "text": "RT @kanekos69: Εν τω μεταξύ αν γίνει στραβή με Τουρκία βλέπω μετανάστες εθελοντές στο μέτωπο και δεξιούς προσφυγες στα Παρίσια"
  }
}
```

Youtube

```
{
  "id": "3",
  "annotations": [
    { "label": "No Hate", "user": 2 },
    { "label": "Neutral", "user": 2 }
  ],
  "meta": {
    "type": "youtube_comment",
    "comment_id": "Ugy-SPKz3HG04OohJnfR4AaABAq",
    "reply_count": 3,
    "like_count": 30,
    "video_id": "Gaz6UvRW0G8",
    "channel": "Μηδέν Ένα Μηδέν 010",
    "video_title": "0 1 0 ~ Πρόσφυγες",
    "video_desc": "Lyrics/Raps - 0 1 0 Beat by Apo (Αισθήσεις) Recorded @ Blackspot Studio Mix/Master by Sativa Cover by SpyOne (Baseline Co.) 0 1 0 IG: ...",
    "author_id": { "value": "UCPP-OugMmE8pNbWRbFv4YCA" },
    "author_name": "GATE21QNZ",
    "rating": "none",
    "date": "2020-09-03T21:16:31Z",
    "plang": "el",
    "pdate": "2020-10-25 00:00:00",
    "phate": "(Αφγανιστάν ισλαμιστές)",
    "pterm": "",
    "ploc": "Afghanistan",
    "text": "Είμαι 25 ετών, ονομάζομαι Μήτσος ο μαλάκας. Μαζεύω ισλαμιστές και λιποτάκτες από μια τούρκο θάρκα, Το παίζω ανοιχτόκαρδος με τις τσέπες του μπαμπάκαα, Και όταν ο Μήτσος πήγε να κάνει φίλους στο Αφγανιστάν, Τον σφάζαν σαν τους προγόνους του, για χάρη του Ισλαμ."
  }
}
```

Website Article

```
{
  "id": "4",
  "annotations": [
    { "label": "No Hate", "user": 2 },
    { "label": "Positive", "user": 2 }
  ],
  "meta": {
    "id": "92860318",
    "type": "article",
    "source": "http://defencereview.gr/gnorizontas-ta-gallika-ploia-oi-fremm-kai-o/",
    "meta": "Άμυνα Ελλάδα 9 Σεπτεμβρίου 2020 18:22",
    "title": "Γνωρίζοντας τα γαλλικά πλοία: Οι FREMM και οι [email protected] (Video)",
    "lang": "el",
    "date": "2020-09-09 00:00:00",
    "hate": "",
    "terms": "",
    "loc": "",
    "text": "Συχνά πυκνά αναφερόμαστε στις γαλλικές ναυπηγικές σχεδιάσεις. Τα παρακάτω βίντεο που βρήκα-με είναι αντιπροσωπευτικά για τις δυνατότητες των πλοίων με πολύ καλά σκηνοθε-τημένα βίντεο και ενδιαφέροντα πλάνα. Αξίζει να τα δείτε:"
  }
}
```

Website Comment

```
{
  "id": "5",
  "annotations": [
    { "label": "No Hate", "user": 2 },
    { "label": "Neutral", "user": 2 }
  ],
  "meta": {
    "id": "92860318",
    "type": "comment",
    "source": "http://defencereview.gr/gnorizontas-ta-gallika-ploia-oi-fremm-kai-o/",
    "lang": "el",
    "date": "",
    "hate": "",
    "terms": "",
    "loc": "",
    "text": "Το μπαραζ επεκτείνεται. Ελπίζω αυριο τετοια ωρα πανω κατω να μην κλαίμε"
  }
}
```

È inoltre possibile ottenere ulteriori informazioni sul formato dei dati e sui campi utilizzati visitando il [PHARM GitHub repository](#).

2 Guida all'interfaccia

2.1 Introduzione

L'interfaccia Internet consente la ricerca, la visualizzazione, l'analisi, l'aggiunta e lo scraping di testi di incitamento all'odio per il progetto PHARM, fornendo un'interfaccia grafica *user-friendly*. L'interfaccia Internet è accessibile visitando il seguente collegamento ipertestuale: pharm-interface.usal.es.

2.2 Home

La "Schermata iniziale" serve come punto di partenza per effettuare analisi di testo su un testo definito dall'utente. Digita (o incolla) il testo desiderato nel file **textbox** e premi il pulsante **"Perform Analysis" button**, come illustrato nella Figura 6.

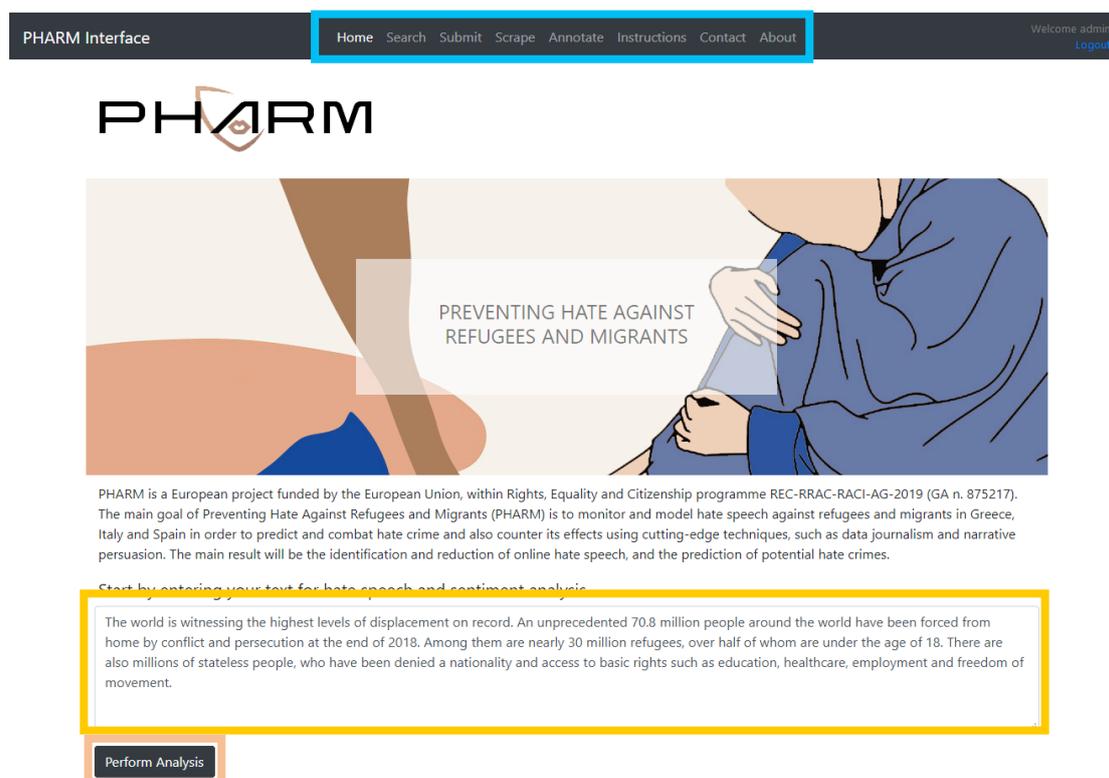


Figura 6 Funzionalità della schermata iniziale

Si apre una nuova scheda nel browser Web e vengono presentati i risultati dell'analisi (Figura 9. Funzionalità della schermata di analisi). Per maggiori informazioni fare riferimento alla Sezione 2.4.

La navigazione a tutte le sottopagine dell'interfaccia Internet è disponibile utilizzando il [top menu](#).

2.3 Search (Ricerca)

Una delle azioni principali dell'interfaccia è la navigazione attraverso i record di incitamento all'odio del database.

Gli utenti possono visitare la "Schermata di ricerca" (Search Screen) selezionando "Cerca" (Search) dal menu in alto. Dopo aver impostato il file **desired filters**, vale a dire origine, lingua, intervallo di date e query di ricerca, devono premere il tasto **"Search" button** (Figura 7

Funzionalità della schermata di ricerca). I filtri di origine e di lingua sono obbligatori, mentre l'impostazione di una data e di una query di ricerca è facoltativa.

Viene caricata la "Schermata dei risultati della ricerca" (Search Results Screen) e viene visualizzato un elenco dei risultati corrispondenti (Figura 8 Funzionalità della schermata Risultati della ricerca). L'utente può visualizzare in anteprima i record, scegliendo il file "Simple" or the "Scientific" view, download them as a CSV or a JSON file, delete a record o display detailed information su di esso.

Va notato che l'azione di eliminazione è disponibile solo per gli utenti autenticati.

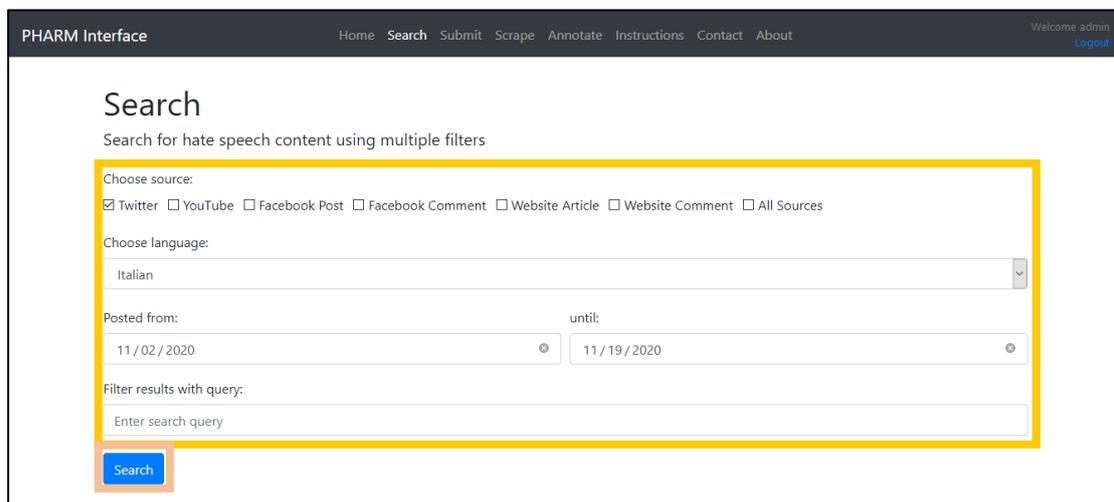


Figura 7 Funzionalità della schermata di ricerca

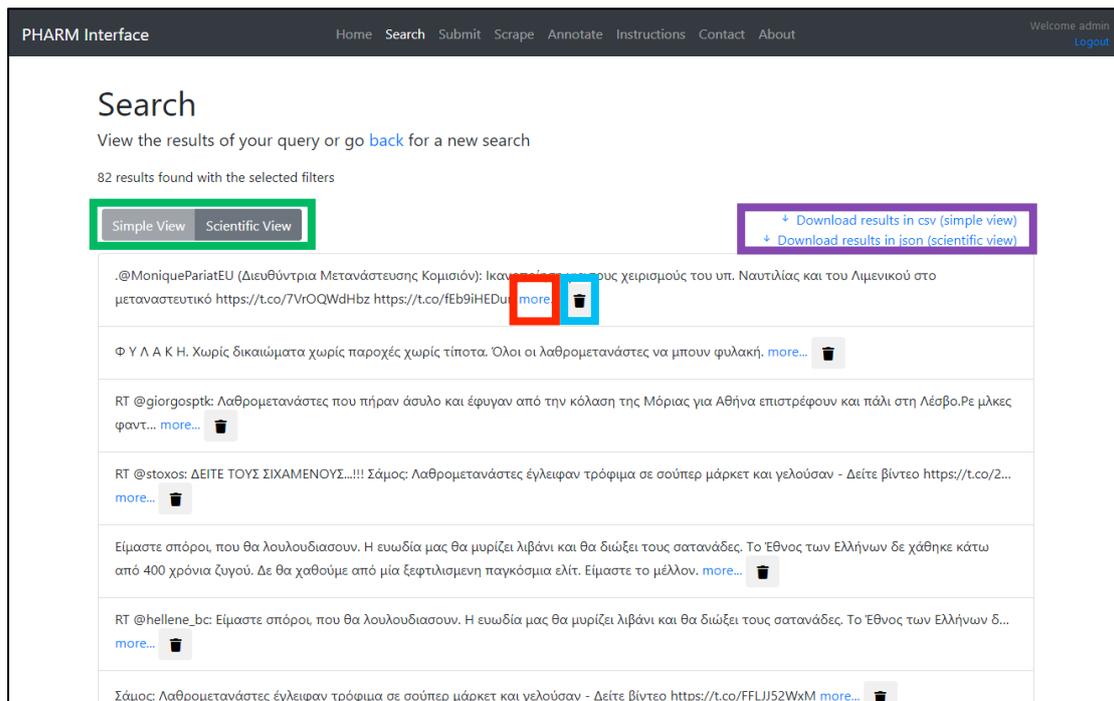


Figura 8 Funzionalità della schermata dei risultati di ricerca

2.4 Analyze (Analizza)

Quando viene selezionato un record (o viene inserito un testo nella "Schermata iniziale"), viene visualizzata una vista che presenta informazioni dettagliate (Figura 9. Funzionalità del-

la schermata). Vengono presentati i risultati dei vari algoritmi di analisi del testo. Gli algoritmi riguardano il rilevamento di incitamento all'odio, l'analisi del sentiment, il rilevamento di parole frequenti, la stima della geolocalizzazione.

L'utente può eseguire una nuova analisi del testo senza lasciare lo schermo, modificando il testo nella casella di testo e premendo il pulsante "Perform Analysis".

Nella parte inferiore della pagina, ci sono scorciatoie per suggerire una correzione al nostro team o fare riferimento al repository PHARM GitHub per trovare ulteriori informazioni sugli algoritmi di analisi del testo distribuiti nell'interfaccia.

The screenshot displays the PHARM Interface with the following elements:

- Navigation Bar:** Home, Search, Submit, Scrape, Annotate, Instructions, Contact, About. User: Welcome admin, Logout.
- Analysis Section:**
 - Header: Analysis
 - Text input: Enter text for hate speech and sentiment analysis. The input text is: ".@MoniquePariatEU (Διευθύντρια Μετανάστευσης Κομισιόν): Ικανοποίηση για τους χειρισμούς του υπ. Ναυτιλίας και του Λιμενικού στο μεταναστευτικό https://t.co/7VrOQWdHbz https://t.co/fEb9iHEDun"
 - Button: Perform Analysis
- Results Section:**
 - Unsupervised Hate Speech Detection: Progress bar (approx. 45% red).
 - Supervised Hate Speech Detection: Progress bar (approx. 5% red).
 - Unsupervised Sentiment Analysis: Progress bar (approx. 10% red).
 - Supervised Sentiment Analysis: Progress bar (approx. 70% red).
 - Geographical estimation of hate speech references: A map showing the location of 'business Consulting' in Lagina, with various landmarks and a search box.
- Footer:** Do you disagree with the hate-speech and sentiment analysis evaluation? Consider contributing your opinion to help us improve. For more information about the analysis algorithms please visit our GitHub page.

Figura 9. Funzionalità della schermata di analisi

2.5 Scrape

Questa schermata fornisce funzionalità per abilitare la raccolta di dati in formato testuale da due piattaforme social, Twitter e YouTube.

Quando è selezionato **YouTube Scraper**, viene visualizzata la schermata illustrata nella Figura 10. L'utente deve digitare **alcuni termini di ricerca** per raccogliere commenti pertinenti da YouTube. È anche possibile **impostare il numero di risultati**. Infine, è necessario premere il pulsante "Scrape" per caricare i risultati, che vengono presentati come mostrato in Figura 8, mentre i dati possono essere scaricati anche in formato CSV o JSON.

Nel caso di Twitter scraper, lo schermo ha una forma simile (Figura 11 Funzionalità di Twitter scraping). È **necessario selezionare una lingua**, insieme **all'intervallo di tempo in cui l'operazione di scraping** sarà attiva. Quando si preme il **pulsante "Scrape"** viene generato un collegamento ipertestuale per scaricare i dati corrispondenti dopo che è trascorso l'intervallo di tempo predefinito.

Figura 10 Funzionalità di scraping di YouTube

Figura 11 Funzionalità di scraping di Twitter

2.6 Annotate (Annota)

Per annotare i dati, seleziona " annotate " (Annota) dal menu in alto per passare alla piattaforma di annotazione doccano. Per ulteriori informazioni fare riferimento alla Sezione 3.

2.7 Submit

Per aggiungere un nuovo record, l'utente deve selezionare "Submit" (Invia) dal menu in alto. Anche l'immissione dei dati può essere eseguita **one-by-one** o **massively**.

Per quanto riguarda il primo metodo, l'utente deve impostare tutti i dati (testo) e i metadati (fonte, lingua, data, sentimento, ecc.) Tramite i moduli di input corrispondenti. I dati vengono aggiunti al database premendo il pulsante "Submit" (Invia) (Figura 12 Funzionalità di invio di testo singolo).

Se i dati sono già formati in modo appropriato (vedere paragrafo 1.5), possono essere importati come file (JSON). Seleziona "Batch File", cerca il file e premi il pulsante "Invia" (Submit button).

Una volta completata l'importazione dei dati, viene visualizzata la schermata Riepilogo importazione (Import Summary Screen) (Figura 14). Questa schermata visualizza informazioni utili sul totale dei record che sono stati elaborati, sul numero di duplicati o sui record che sono già memorizzati nel database.

Inoltre, vengono rilevate voci non valide, prive di campi obbligatori e l'utente viene informato di conseguenza.

The screenshot shows the 'Submit' page of the PHARM Interface. The page title is 'Submit' and the subtitle is 'Submit a single or a batch of records to the PHARM database'. There are two tabs: 'Single Text' (selected) and 'Batch File'. Below the tabs, there is a text input field with the placeholder 'Enter text here'. Underneath, there is a section for 'Insert tags (Optional)' with a tag 'refugees x'. There are two dropdown menus: 'Choose text language:' set to 'English' and 'Choose date (Optional):' set to 'mm / dd / yyyy'. Below these are radio buttons for 'Choose text source:' with options: Twitter, YouTube, Facebook Comment, Facebook Post, Website Article, and Website Comment. At the bottom, there is an input field for 'Insert source url (Optional)' and a blue 'Submit' button.

Figura 12 Funzionalità di invio di testo singolo

The screenshot shows the 'Submit' page of the PHARM Interface. The page title is 'Submit' and the subtitle is 'Submit a single or a batch of records to the PHARM database'. There are two tabs: 'Single Text' and 'Batch File' (selected). Below the tabs, there is a text input field with the placeholder 'Submit batch hate-speech texts to the PHARM database. The records should be in a JSON format. For the JSON data format configuration, please refer to the [Instructions](#) page.' Below this is a file upload section with a 'Browse...' button and the text 'No file selected.'. At the bottom, there is a blue 'Submit' button.

Figura 13 Funzionalità di invio dei Batch records

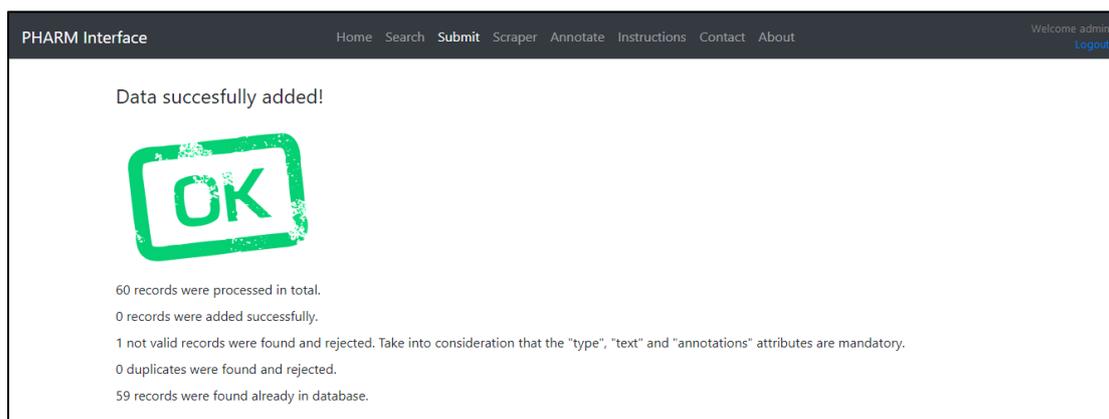


Figura 14 Schermata di riepilogo dell'importazione (Import Summary Screen) dell'interfaccia Internet

2.8 Login

Quando l'utente seleziona una pagina che richiede l'autenticazione (ad esempio Cerca, Scrape e Annota), viene visualizzata la schermata di accesso (Figura 15 la schermata di accesso - Login Screen- dell'interfaccia Internet), in cui è necessario inserire il nome utente e la password per procedere.

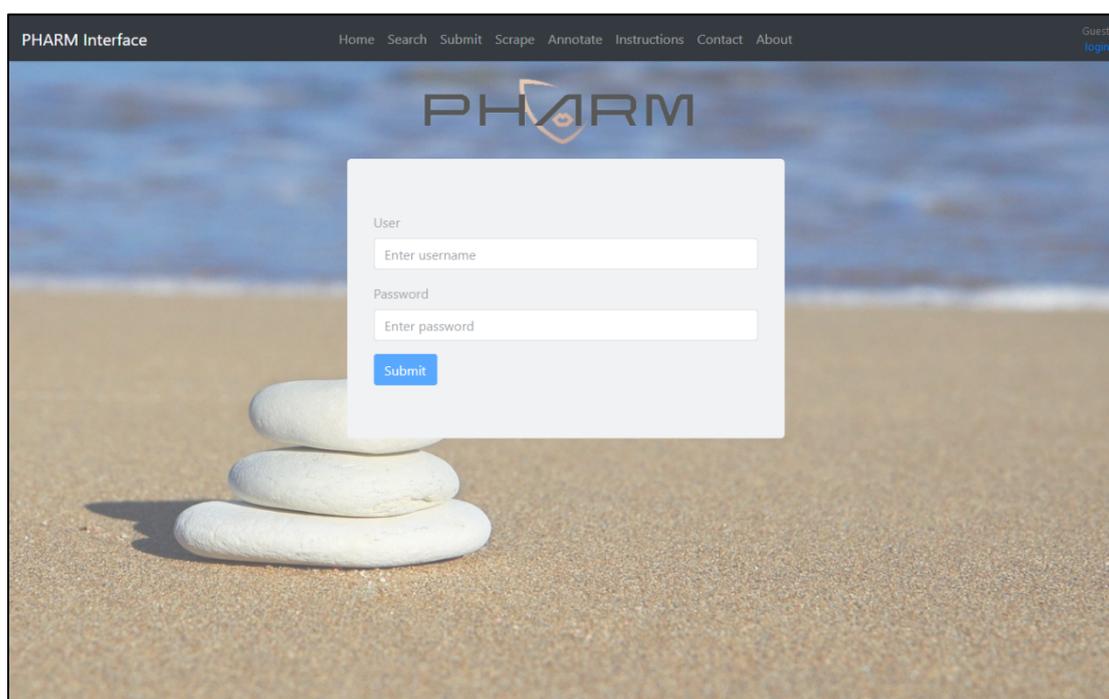


Figura 15 Schermata di accesso dell'interfaccia Internet

Per ottenere un account autenticato, visita la "Schermata di contatto" (Contact Screen) e contatta il nostro team.

3 Guida alle annotazioni

L'interfaccia Internet utilizza lo strumento doccano per supportare le attività di annotazione richieste.

3.1 Generale

Doccano è un sistema di gestione delle annotazioni per i dati di testo. Può essere utilizzato per sviluppare set di dati per la classificazione, l'analisi del sentiment, l'etichettatura di entità o la traduzione. Nel contesto del progetto PHARM, viene utilizzato per scopi di classificazione del testo. Pertanto, ogni voce dovrebbe essere etichettata con tag.

Innanzitutto, l'utente può dare un'occhiata alla demo online per l'analisi del sentiment (la stessa procedura può essere seguita anche per la classificazione): <http://doccano.herokuapp.com/demo/text-classification>. L'utente può accedere all'installazione di doccano visitando <http://pharm-doccano.usal.es>. Lo strumento è accessibile anche dalla schermata corrispondente dell'interfaccia Internet di PHARM (Figura 4 Schermata Annotate dell'interfaccia Internet).

In poche parole, l'utente/annotatore dovrebbe assegnare etichette ("Odio" o "Nessun odio", "Positivo", "Neutro" o "Negativo") a ciascuna voce. Le annotazioni vengono memorizzate automaticamente.

3.2 Annotate (Annota)

Accedi utilizzando le tue credenziali come annotatore (Figura 16 Schermata di accesso del doccano).

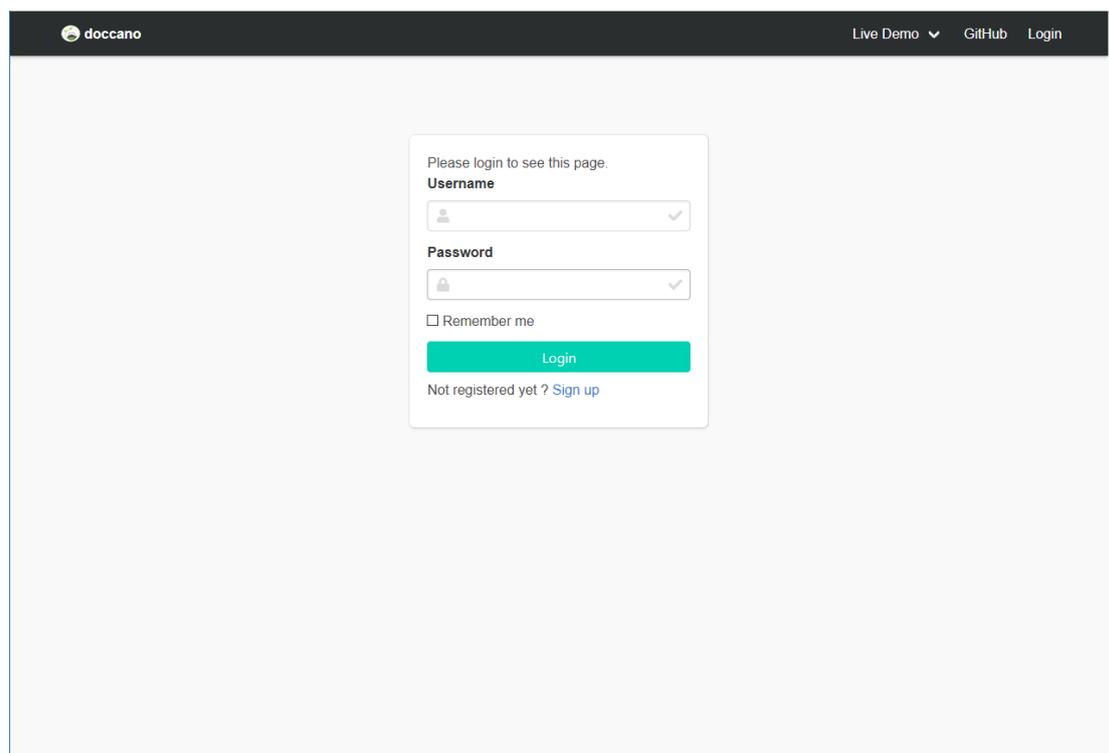


Figura 16 Schermata di accesso del doccano

Scegli il progetto a cui ti è stato chiesto di collaborare (Figura 17 Schermata di selezione del progetto). Quindi, leggi il testo (per la voce corrente), scegli "Odio" o "Nessun odio" (se il testo contiene contenuti contro rifugiati o migranti o no) e contrassegnalo come "Positivo",

"Neutro" o "Negativo" (secondo il sentimento del testo). È obbligatorio scegliere un'etichetta "blu" e una "viola" per ciascuna voce, poiché l'annotazione viene eseguita per due schemi di classificazione individuali simultaneamente (Figura 18 Schermata di annotazione e funzionalità di selezione dell'etichetta del doccano). Le etichette "blu" servono per il rilevamento dell'incitamento all'odio, mentre le etichette "viola" sono necessarie per l'analisi del sentimento. Le voci annotate sono evidenziate con un'icona di segno di spunta. Infine, scegli "Avanti" (next) per passare alla voce successiva.

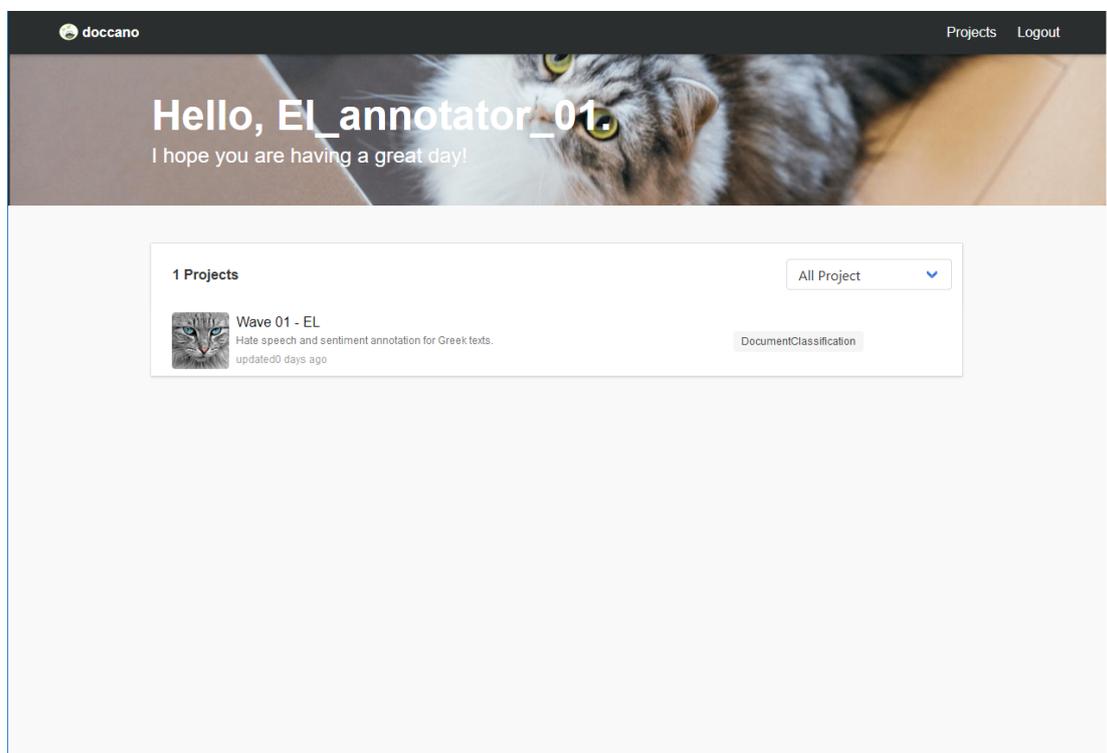


Figura 17 Schermata di selezione del progetto del doccano

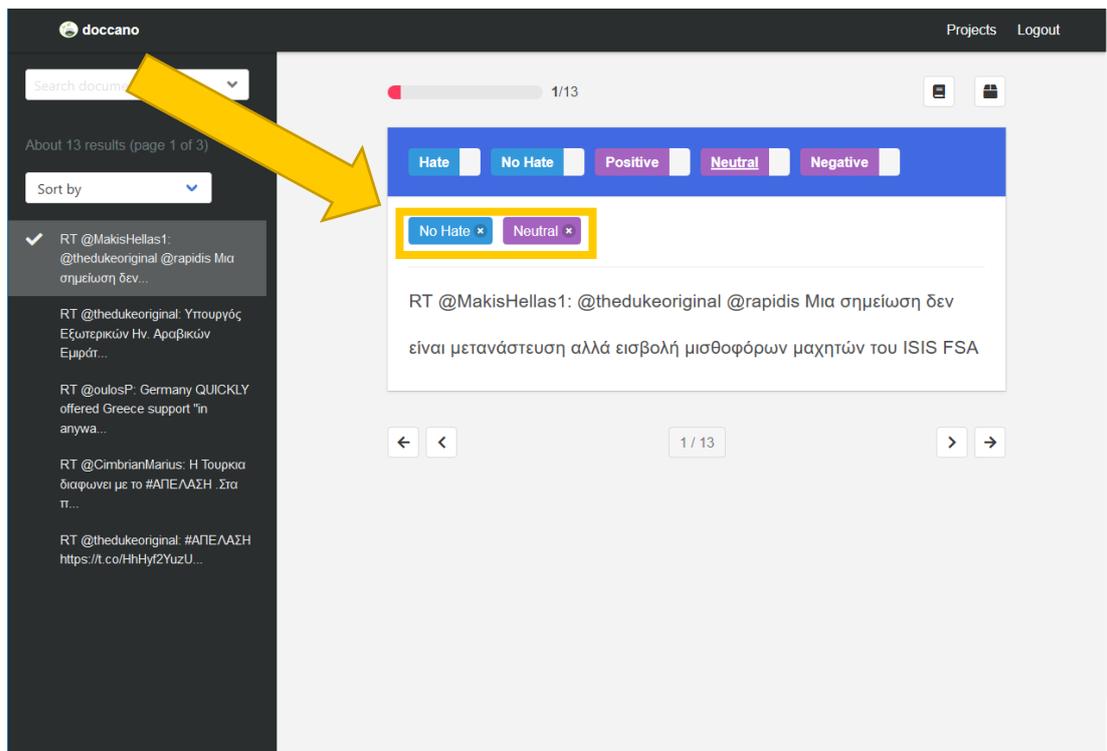


Figure 1 Annotation Screen and label selection functionality of the doccano

3.3 Approve (Approva)

Accedi a doccano e seleziona il progetto appropriato (segui gli stessi passaggi descritti in precedenza). Approvare un'annotazione utilizzando il pulsante mostrato nella Figura 19. Le voci approvate vengono visualizzate con un'icona con il thumbs up icon.

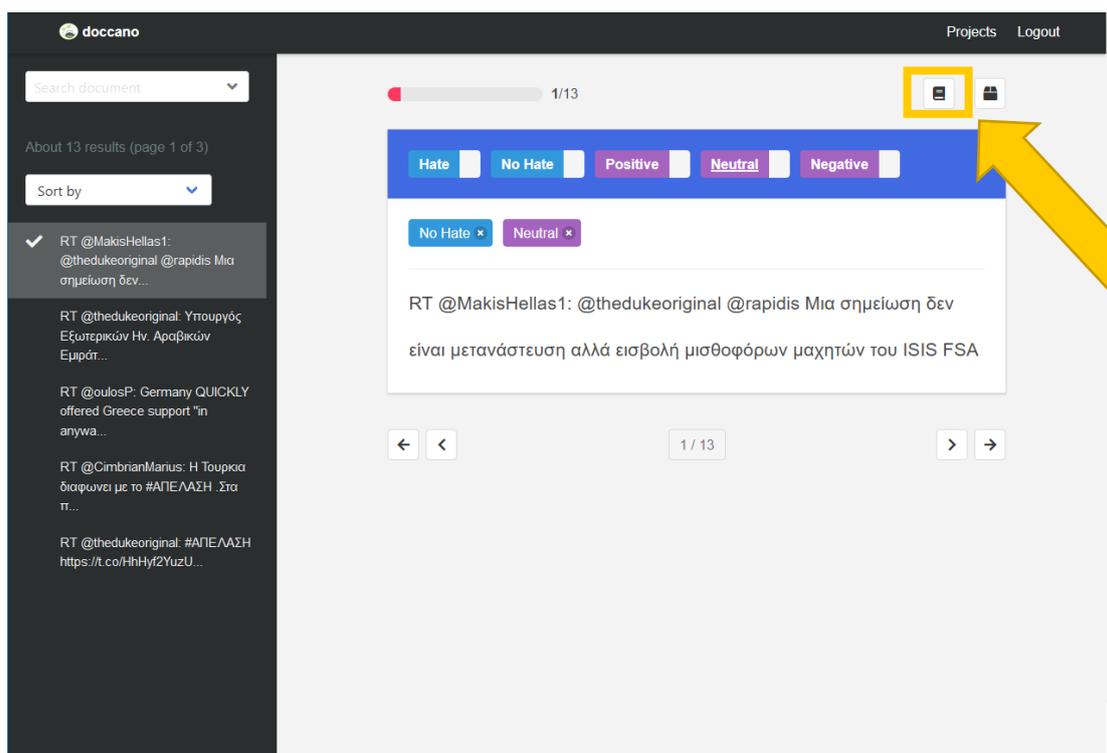


Figura 19 Schermata di annotazione e funzionalità di approvazione del doccano



PREVENTING HATE AGAINST REFUGEES AND MIGRANTS

ABOUT PHARM (2020-2022)

Preventing Hate against Refugees and Migrants (PHARM)

Migration to Europe has grown in the last years in scale and complexity. The so called 'refugee crisis' and the migratory pressure is particularly acute in southern EU countries as the main entrance to the EU.

The main goal of PHARM project is to monitor and model hate speech against refugees and migrants in Greece, Italy and Spain in order to predict and combat hate crime and also counter its effects using cutting-edge techniques, such as data journalism and narrative persuasion. The activities distributed in 5 coordinated work packages include:

(i) Implementation of a conceptual and methodological common framework for large-scale analysis and detection of hate speech; (ii) Implementation and evaluation of machine learning approaches to model and predict hate crimes against refugees and migrants based on hate speech features; (iii) Survey journalists to understand how they inform and raise awareness about hate speech and how they can help building and disseminating counter-narratives based in data-driven news pieces; (iv) Creation, evaluation and dissemination of counter-narrative fictional stories adapted to different characteristics of citizens using large-scale narrative persuasion.

COORDINATOR

University of Salamanca, Faculty of Social Sciences (ES)

PARTNERS

University of Milan, (IT)

Aristotle University of Thessaloniki (GR)

VISIT: <http://pharmproject.usal.es>

CONTACT US: pharm.project.eu@gmail.com

FOLLOW US:

@Pharm_project

Pharm_project

Pharm Project

Pharm Project