MADRID
inter.noise 2019
June 16 - 19
NOISE CONTROL FOR A BETTER ENVIRONMENT

Noytext: A Web platform to annotate social media documents on noise perception to their use in opinion mining research

Luis Gascó[1,2], César Asensio[1], Guillermo de Arcas[1], Chloé Clavel[3]

1 POLITÉCNICA

1

2 eit

3 TELECOM Paris

The problem

New technologies

Barrier to entry

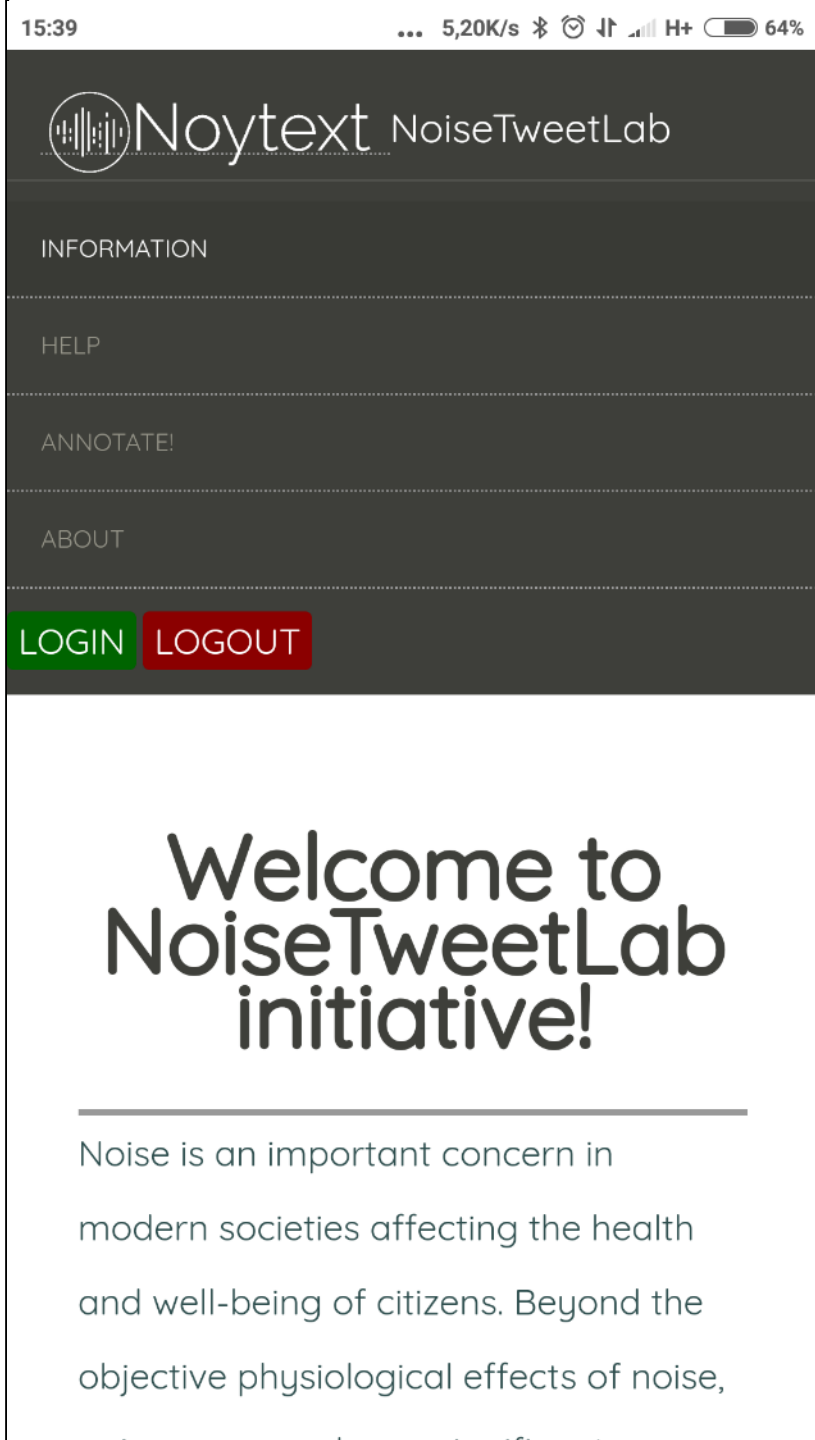

Lack of easy-to-use tools to apply text analytics technologies in acoustics

Noytext

- Simplify the annotation task
- Easy to install in your server or PC
- Cross platform

Structure & customization options

- 4-page schema:
  - Project info.
  - Help
  - **Annotation**
  - About
- Consensous among annotators
- User information



Requirements

Execution environment

- Local

- Remote

Info. & Installation guide

https://www.noytext.com/

https://github.com/luisgasco/noytext

The experiment

NoiseTweetLab

- Collaborative short-text annotation
- **Goal:** To build the first tweets dataset to be used as a baseline to measure the performance of future detection models

Welcome to NoiseTweetLab initiative!

Noise is an important concern in modern societies affecting the health and well-being of citizens. Beyond the objective physiological effects of noise, noise response has a significant subjective component, which is reflected as a community response and has been traditionally evaluated through surveys. These surveys are often costly, invasive and people do not usually take part in them, whether you use one-to-one interview, phone-based polls or web-based forms. But the big boost of online social networks has demonstrated that some people are willing to share their views and feelings about everyday problems, including noise.

We have recently demostrated that it is possible to apply Natural Language Processing and Machine Learning technologies to detect the opinions about noisy activities from online social networks. However, we have detected the lack of databases that allow us to train systems for this purpose. For that reason, we have launched the NoiseTweetLab initiave, in which we request the collaboration of Internet users to help us annotate some short texts from Twitter, with the final goal of building the first database to be used as a baseline to measure the performance of future detection models in this field.

We claim your collaboration by labeling some short texts in the four categories below.

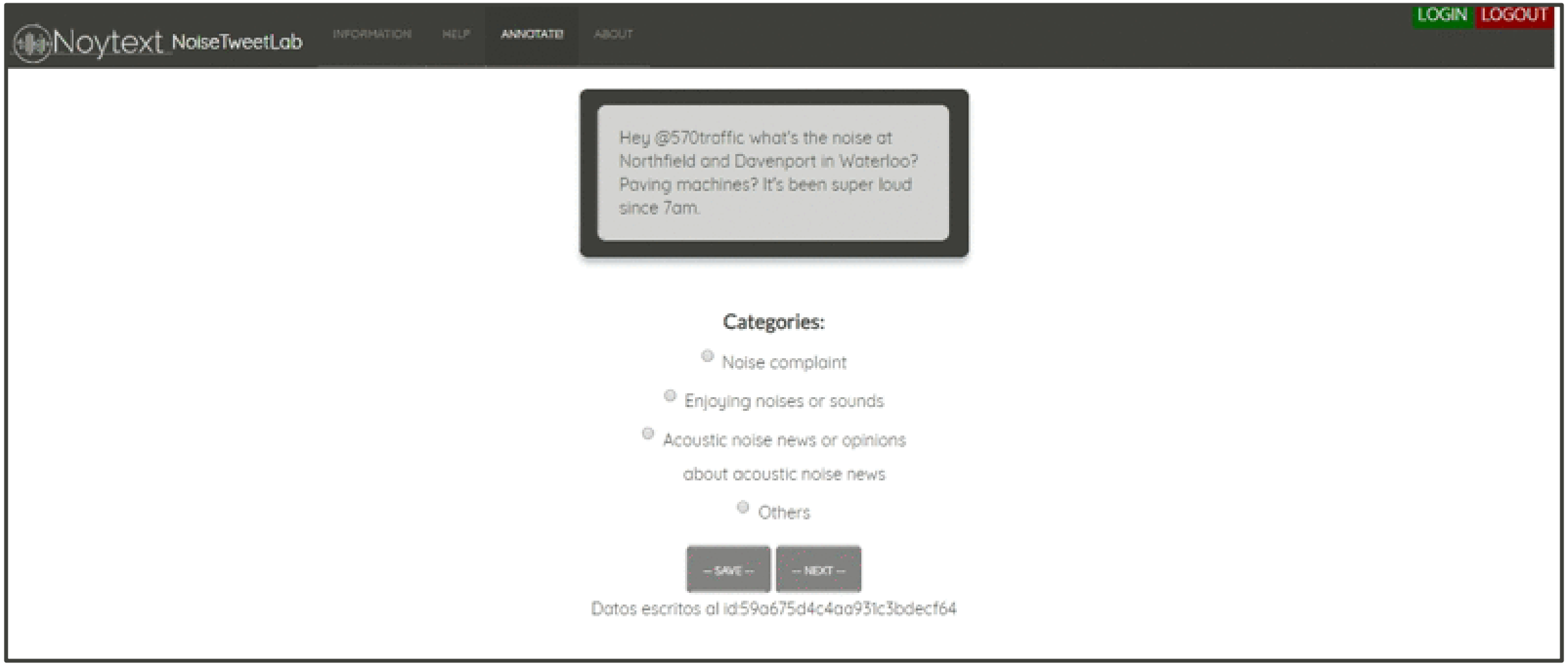
Data classes

The data can be labeled in 4 categories:

- **Noise complaint : The person is complaining about a noise source, or sound, such as neighbors, traffic, aircrafts...**

> **@user198653654**
> Lots of Trucks with noise pipes headed NB on 35w. Also with police lives matter stickers. Whats happening?

> **@user49986414165**
> honestly what the fuck does my upstairs neighbour do all day that makes so much noise?!


Noytext NoiseTweetLab
INFORMATION
HELP
ANNOTATE!
ABOUT
LOGIN
LOGOUT
Hey @570traffic what's the noise at Northfield and Davenport in Waterloo? Paving machines? It's been super loud since 7am.
Categories:
Noise complaint
Enjoying noises or sounds
Acoustic noise news or opinions
about acoustic noise news
Others
-- SAVE --
-- NEXT --
Datos escritos al id59a675d4c4aa931c3bdecf64

Conclusions

If any researcher were interested in starting to apply these technologies, he would need to have the right tools to accelerate his research.

- A open-source tool to annotate their text documents.

- An experiment to generate the first tweet dataset that can be used to intercompare models made by different researchers.


thank you
gracias
danke
謝謝
شكراً جزيلاً
ngiyabonga
teşekkür ederim
рахмат
Баярлалаа
спасибо
faafetai lava
mersi
kia ora
barka
welalin
tack
spas
dank je
misaotra
matondo
paldies
grazzi
mahalo
tapadh leat
vinaka
спасибі
blagodaram
kiitos
dankie
nanni
nandri
dhanyavad
mauruuru
hvala
köszönöm
bedankt
bayarlalaa
gràçie
enkosi
dziękuję
aciu
dankon
akun
хвала
asante
manana
obrigada
murakoze
tenki
djiere dieuf
tau
mochchakkeram
дякую
mamnun
chokrane
chnorakaloutioun
gratias ago
gràcies
sulpáy
taiku
go raibh maith agat
sobodi
dēkuji
obrigado
sağolun
sukriya
kop khun krap
arigatô
takk
dakujem
trugarez
mèsi
didi madloba
ありがとう
najis tuke
terima kasih
tanemirt
grazie
rahmet
ευχαριστώ
diolch
dhanyavadagalu
shukriya
merce
мерси
kam sah hamnida
rahmat
감사합니다
xièxie
merci

http://noisetweetlab.noytext.com/

MADRID
inter.noise 2019
June 16 - 19

Noytext: A Web platform to annotate social media documents on noise perception to their use in opinion mining research

POLITÉCNICA
"Ingeniamos el futuro"
CAMPUS DE EXCELENCIA INTERNACIONAL

INSTRUMENTATION & APPLIED ACOUSTICS RESEARCH GROUP

eit Digital DOCTORAL SCHOOL

Luis Gascó Sánchez

luis.gasco@i2a2.upm.es

https://www.luisgasco.es

https://www.linkedin.com/in/luisgascosanchez/

https://www.researchgate.net/profile/Luis_Gasco_Sanchez

@luisgasco