

Mapping Twitter

Fira de les Ciències 2018

Escola Mare de Déu de la Salut (Sabadell)

Maria BAÑERAS
Àngel GALLEGO
Elena MAES

¿Qué es Mapping Twitter?

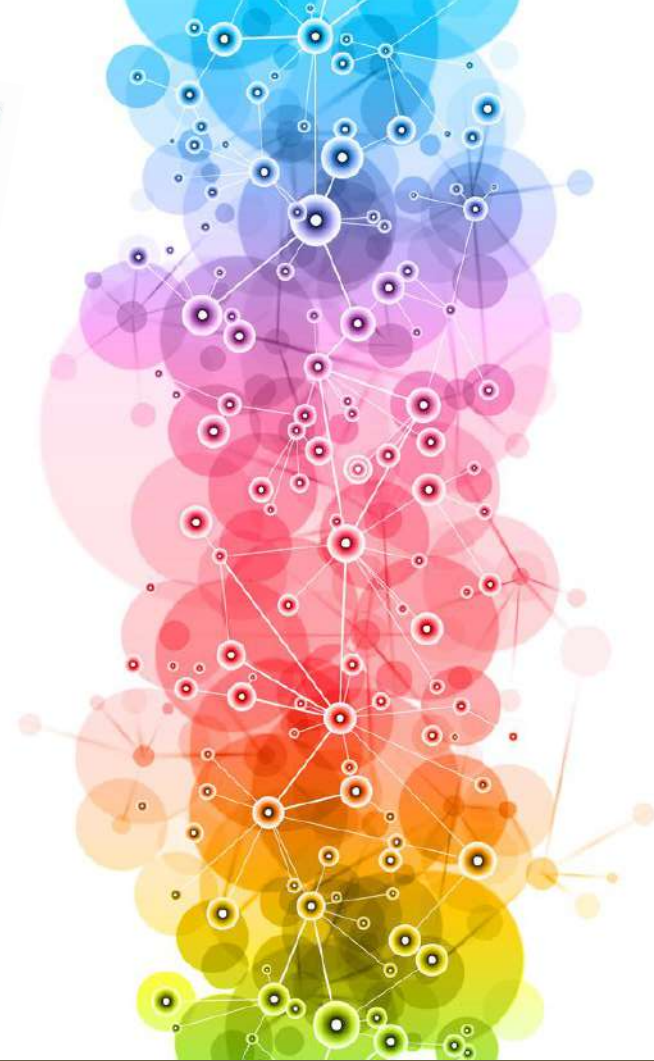
- Proyecto en colaboración con el Departamento de Filología Española de la UAB.
- Combina elementos de diferentes áreas del conocimiento: *big data*, lingüística y ciencia ciudadana.



Objetivos generales



- Entrar en contacto con las nuevas **herramientas digitales** de la geolingüística.
- Acercarse a la **dimensión científica** del estudio de la lengua.



Objetivos específicos



- Estudiar la **diversidad** lingüística y cultural de Cataluña.
- Observar diferentes **fenómenos lingüísticos** (contacto entre lenguas, adquisición del lenguaje, alternancia de código, lenguas heredadas, etc.).



La geolingüística

→ Un atlas lingüístico es una colección de mapas con información lingüística.



Tradicionalmente, se elaboraban viajando a cada localidad y haciendo preguntas a los informantes.

La geolingüística 2.0



→ Un atlas lingüístico se puede generar también a partir de datos de las **redes sociales** o de **crowdsourcing** (ciencia ciudadana).



Se pueden obtener cantidades masivas de datos (*big data*).



La geolinguística 2.0



→ Un ejemplo de crowdsourcing: www.abruzzesemolisano.it



Abruzzo e Molise data crowdsourcing

Partecipa anche tu alla documentazione delle lingue di Abruzzo e Molise! Vai dai tuoi nonni e chiedi loro di parlare per 10 minuti (raccontando la loro giovinezza, o la guerra), e registrati con il tuo telefonino. Poi entra nel sito (vai alla mappa), seleziona il cerchietto che corrisponde al tuo paese, e fai l'upload dell'audio o del video che hai registrato. Grazie!

Clicca sul cerchietto corrispondente alla località i cui dati hai registrato (nella prossima pagina):

Vai alla mappa

La geolingüística 2.0



→ Una de la **redes sociales** más interesantes para este tipo de actividades es **Twitter**.

*~400 millones
de tuits al día*

*usuarios de
diferentes perfiles*

*textos breves
(280 caracteres)*

*espontaneidad y
rasgos de oralidad*

*geolocalización
opcional (0,5-25%)*

*obtención del 1%
gratuitamente*



La geolingüística 2.0



→ Se pueden obtener datos como éstos:



IRATXE @iratxe_96 · 31 de mar.

No puedo con la nieve osea literal mañana es abril y **esta puto nevando**

Abrir

Responder Retwittear Favorito Más



Seguido por **Pupita Gratuita**



Maria Valls @mvalsrock · 29 de abr.

Que alguien me pegue un tiro que **esta puto nevando!**

pic.twitter.com/GHaCyxMGjq

Ver foto

Responder Retwittear Favorito Más



kenk(eva)sión @evittakaos · 29 de abr.

No os miento. Aquí en Colmenar **está puto nevando.**

pic.twitter.com/2wOgl3aafc

Ver foto

Responder Retwittear Favorito Más



{} @justsaracano · 23 de feb. de 2013

Pues no, aquí **no está puto nevando.** Aquí hace un sol que deslumbra.

Abrir

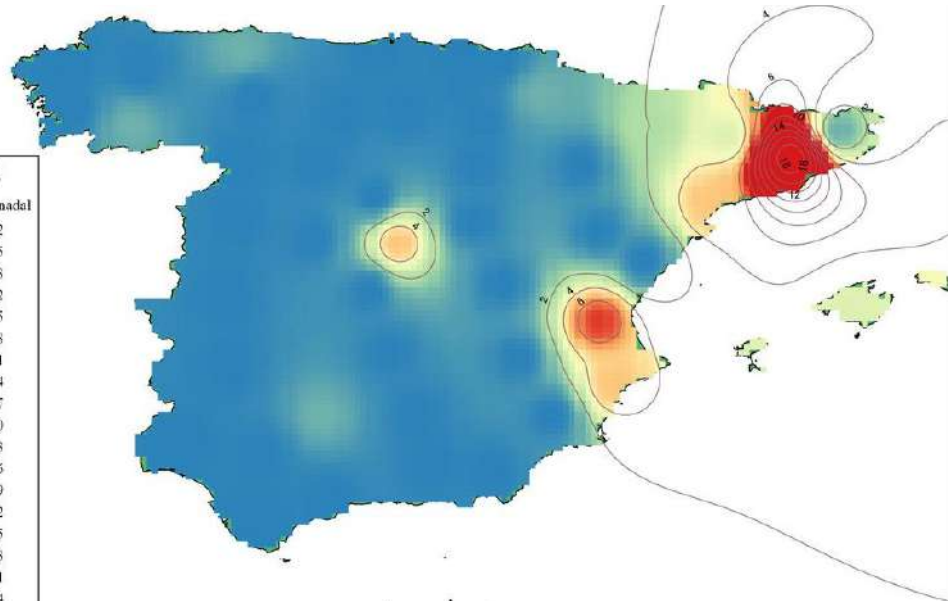
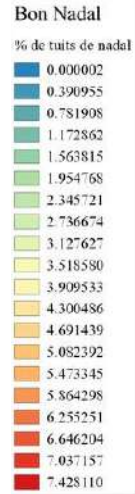
Responder Retwittear Favorito Más



La geolingüística 2.0



→ ...o imágenes como éstas:



@aruiztinoco

La geolingüística 2.0



→ ...o imágenes como éstas:

#Tuits que contienen la palabra "accidente" alrededor de Madrid. ¿Puntos negros en las carreteras?
#kibana



@aruiztinoco

La geolingüística 2.0

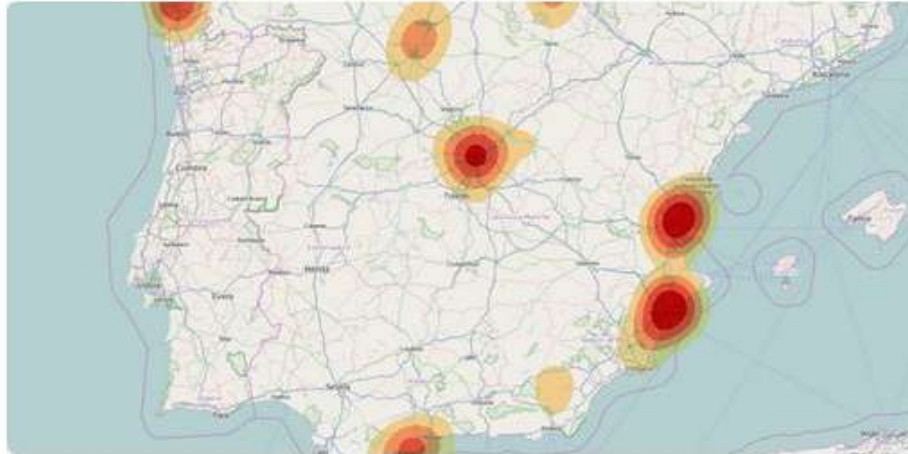


→ ...o imágenes como éstas:

Antonio Ruiz Tinoco @aruiztinoco · Jul 2

Lugares que muestran mayor densidad de uso de "de la hostia".

#lexicalvariation #Spanish #dialect #qgis



83

40

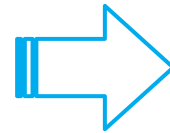
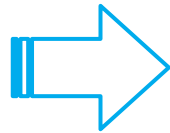


@aruiztinoco

Pero... ¿en qué consiste el proyecto?

→ Elaborar mapas a partir de datos extraídos de redes sociales (Twitter). ¿Cómo?

1. Escoge un tema:



Pero... ¿en qué consiste el proyecto?

2. Busca en la base de datos del ADiBa, que ha recopilado centenares de miles de tuits del Área Metropolitana de Barcelona (~2500 en Sabadell).

www.adiba.cat



Atles Dialectal de Barcelona

Pero... ¿en qué consiste el proyecto?

→ Interfaz de búsqueda:

Text:

Localitat:

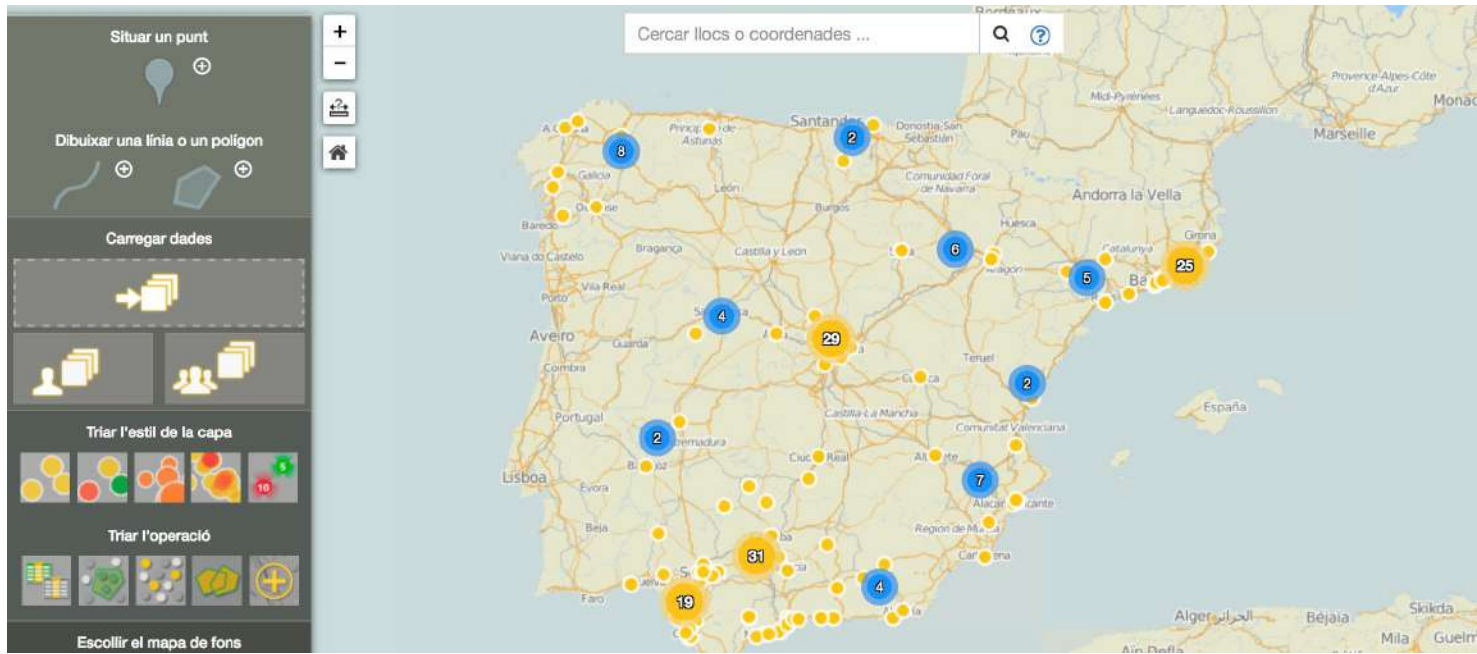
Data inici: **Data final:**

S'han trobat **22006** registres

Text	Lema	Etiqueta	Conjunt	Localitat	Data
'papi' es ahora una tendencia en #Barcelona https://t.co/pG3q5PZteZ https://t.co/cWfDr2IYIW	' papi ' ser ahora uno tendencia en tag	Fe NCMS000 Fe VSIP3S0 RG DI0FS0 NCFS000 SP NCMS000	' (Fe) papi (NCMS000) ' (Fe) es (VSIP3S0) ahora (RG) una (DI0FS0) tendencia (NCFS000) en (SP) tag (NCMS000)	Barcelona	14-02-2018

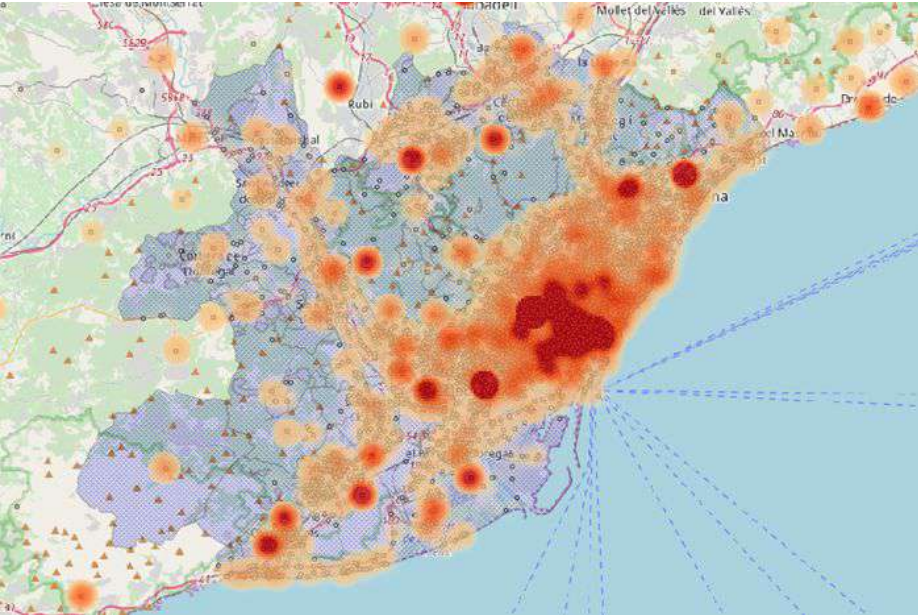
Pero... ¿en qué consiste el proyecto?

3. Mapea los datos obtenidos:



Pero... ¿en qué consiste el proyecto?

3. Mapea los datos obtenidos:



Propuesta de temas



- **Temas lingüísticos**
 - uso de diferentes abreviaturas en función de la localidad (*ferros vs ferrocata*)
 - fenómenos de contacto entre lenguas (*tomar un café vs hacer un café*)
- **Temas sociales**
 - popularidad de personajes públicos en función de la localidad (equipos de fútbol, cantantes)
 - seguimiento de un objeto o evento cultural a través del tiempo (serie, entrega de premios)

...u otros **temas de libre elección.**

Gracias

