Análisis de Opiniones en Monumentos Culturales

IX JORNADA R

UNIVERSIDAD DE GRANADA

Ana Valdivia
Granada, 16 de Noviembre del 2017

Ana Valdivia

Mathematician (UPC)

PhD in **Natural Language Processing** and **Sentiment Analysis** (UGR)

#DataScience4SocialGood

UNIVERSIDAD
DE GRANADA

De la Torre achaca el retraso del Museo de Málaga a la crisis económica y destaca su apuesta cultural

La crisis económica obliga a cerrar el museo Chillida-Leku

Crisis económica afecta a museos

La Junta reduce a 400.000 euros la partida del patrimonio histórico

☐ Las inversiones en bienes culturales y monumentos cae en picado desde 2012 en Sevilla

☐ El PP critica los «recortes» y que 168 millones no se hayan ejecutado

El PSOE denuncia un recorte de 44% en patrimonio cultural con 137 bienes en peligro

Los recortes económicos ponen en peligro a los monumentos romanos

UNIVERSIDAD DE GRANADA

La crisis económica ralentiza la aplicación de nuevas tecnologías en el ámbito de los museos

Javier Belloso, gerente de la empresa Miramultimedia, defiende que el uso de las TIC fideliza a los visitantes y atrae a los más jóvenes hacia las colecciones museográficas

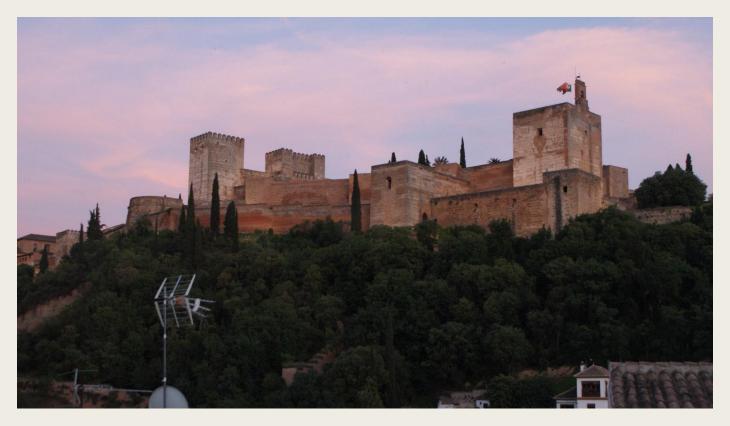Crisis y museos. Una oportunidad para hacer las cosas mejor

Escrito en diciembre de 2014 por Fco Javier Zubiaur Carreño

Museos de EEUU se reinventan para afrontar crisis económica

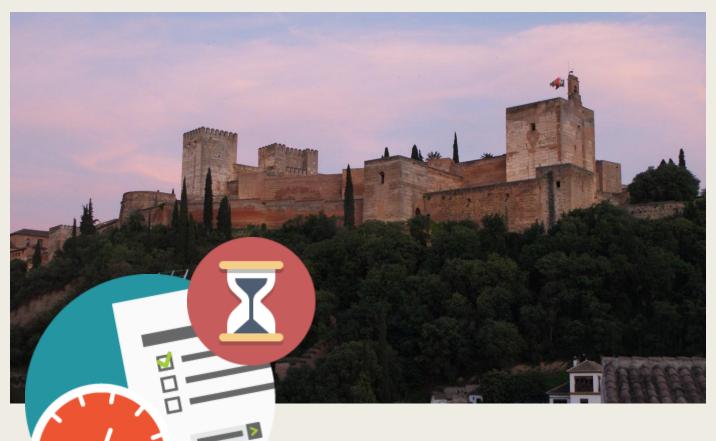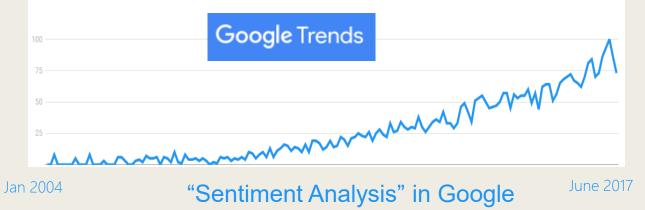## Llamando a las puertas del saber: los museos se resisten a sucumbir a la crisis económica
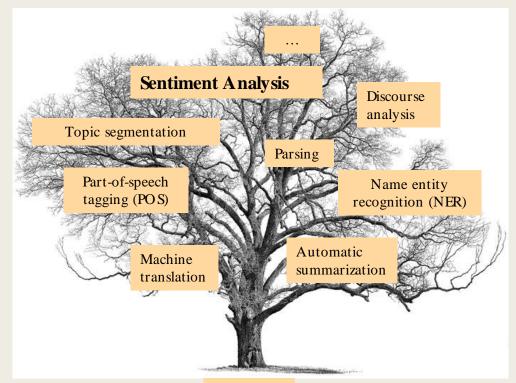
UNIVERSIDAD
DE GRANADA

Sentiment Analysis

It is the field of knowledge that **analyses people`s opinions**, reviews or thoughts about products, companies or experiences identifying its sentiment.

Google Trends

100

75

50

25

Jan 2004

"Sentiment Analysis" in Google

June 2017

UNIVERSIDAD DE GRANADA

...

Sentiment Analysis

Discourse analysis

Topic segmentation

Parsing

Part-of-speech tagging (POS)

Name entity recognition (NER)

Machine translation

Automatic summarization

NLP

Sentiment Analysis

It is the field of knowledge that **analyses people`s opinions**, reviews or thoughts about products, companies or experiences identifying its sentiment.

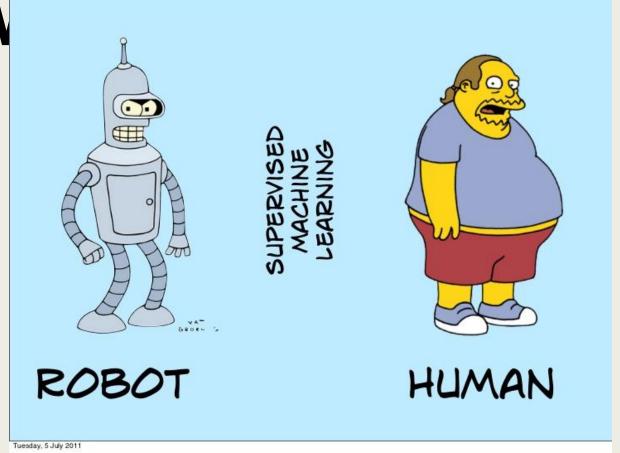SA has experienced an **important growth** due to the development of **Web 2.0**.

SA can be divided in **different tasks**:

1. Sentiment Classification

2. Subjectivity Classification

3. Opinion Summarization

4. Sarcasm and Irony

5. …
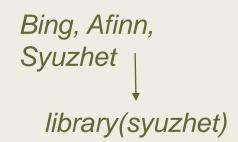
UNIVERSIDAD
DE GRANADA

Sentiment Analysis Methods (SAM

UNIVERSIDAD
DE GRANADA

Sentiment Analysis Methods (SAMs)

1. Bag of Words

Bing, Afinn, Syuzhet

library(syuzhet)

Sentiment Analysis Methods (SAMs)

1. Bag of Words
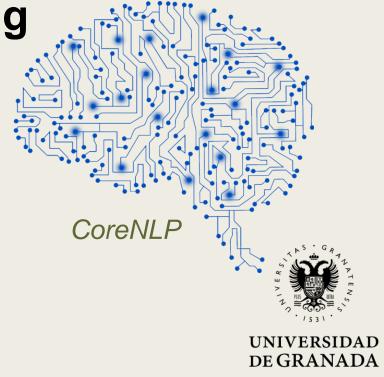
2. Machine Learning

Bing, Afinn, Syuzhet

Microsoft Azure

UNIVERSIDAD DE GRANADA

CoreNLP

UNIVERSIDAD DE GRANADA

Brief **#dataviz**
of
Alhambra's dataset

UNIVERSIDAD
DE GRANADA

Some nice plots :)

Monthly Total Number of Reviews

Some nice plots :)

Monthly Rating Evolution

Some nice plots :)

Some nice wordclouds :)

Some nice wordclouds :)

Some nice wordclouds :)

coreNLP

To download the R library and corresponding CoreNLP java library, run the following in R:

```
devtools::install_github("statsmaths/coreNLP")
coreNLP::downloadCoreNLP()
```

Then, in order to run the package, the following code initializes rJava correctly (if run from the same directory as the above):

```
library(coreNLP)
initCoreNLP()
```

As a simple example of usage, this how to annotate the opening lines to "The Cat in the Hat":

```
catInHat = c("the sun did not shine.", "it was too wet to play.",
             "so we sat in the house all that cold, cold, wet day.")
output = annotateString(catInHat)
```

UNIVERSIDAD
DE GRANADA

Which gives the following information via `getToken`, `getDependency` and `getSentiment`:

```
> getToken(output)[,c(1:3,6:7)]
   sentence  word lemma POS     NER
1         1   the   the  DT       O
2         1   sun   sun  NN       O
3         1   did    do VBD       O
4         1   not   not  RB       O
5         1 shine shine  VB       O
6         1     .     .   .       O
7         2    it    it PRP       O
8         2   was    be VBD       O
9         2   too   too  RB       O
10        2   wet   wet  JJ       O
11        2    to    to  TO       O
12        2  play  play  VB       O
13        2     .     .   .       O
14        3    so    so  IN       O
15        3    we    we PRP       O
16        3   sat   sit VBD       O
17        3    in    in  IN       O
18        3   the   the  DT       O
19        3 house house  NN       O
20        3   all   all  DT       O
```

getToken()

Example:

```
catInHat = c("the sun did not shine.", "it was too wet to play.",
             "so we sat in the house all that cold, cold, wet day.")
output = annotateString(catInHat)
```

UNIVERSIDAD
DE GRANADA

```
> getDependency(output)
   sentence governor dependent    type governorIdx dependentIdx
1         1     ROOT     shine     root          0            5
2         1      sun       the      det          2            1
3         1    shine       sun    nsubj          5            2
4         1    shine       did      aux          5            3
5         1    shine       not      neg          5            4
6         2     ROOT       wet     root          0            4
7         2      wet        it    nsubj          4            1
8         2      wet       was      cop          4            2
9         2      wet       too   advmod          4            3
10        2     play        to      aux          6            5
11        2      wet      play    xcomp          4            6
12        3     ROOT       sat     root          0            3
13        3      sat        so      dep          3            1
14        3      sat        we    nsubj          3            2
15        3      sat        in     prep          3            4
16        3    house       the      det          6            5
17        3       in     house     pobj          4            6
18        3      sat       all     tmod          3            7
19        3      day      that      det         14            8
20        3      day      cold     amod         14            9
21        3      day      cold     amod         14           11
22        3      day       wet     amod         14           13
23        3      all       day      dep          7           14
```
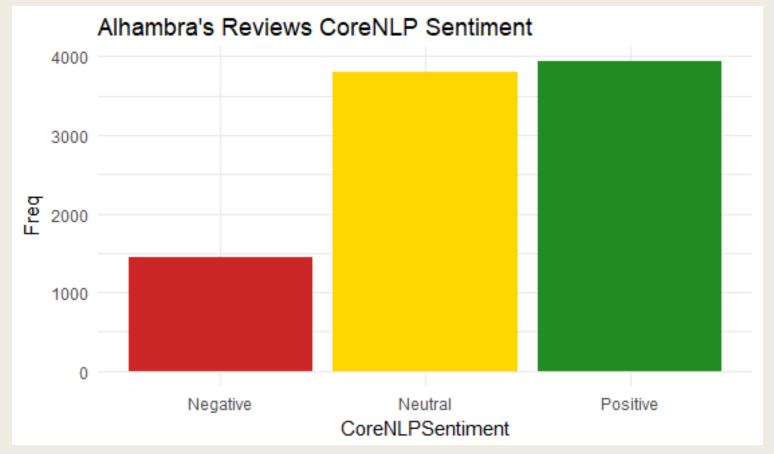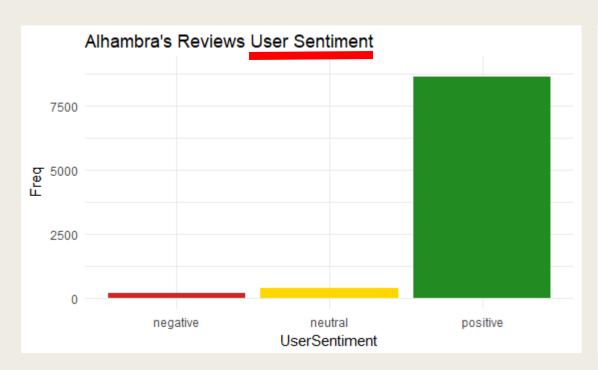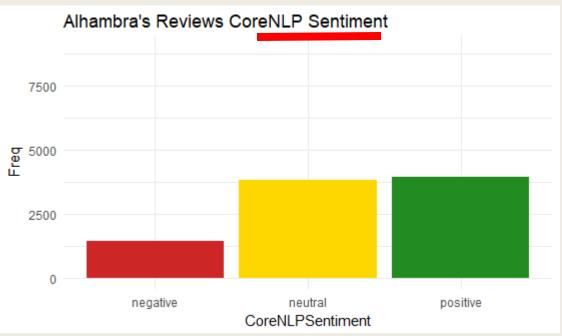
getDependency()

```
> getSentiment(output)
  id sentimentValue sentiment
1  1              1  Negative
2  2              1  Negative
3  3              1  Negative
```

getSentiment()

Brief **#dataviz**
of
Alhambra's dataset
with
CoreNLP

UNIVERSIDAD
DE GRANADA

Some nice plots :)

Análisis de Opiniones en Monumentos Culturales
IX JORNADAⓇ

Ooopps…..

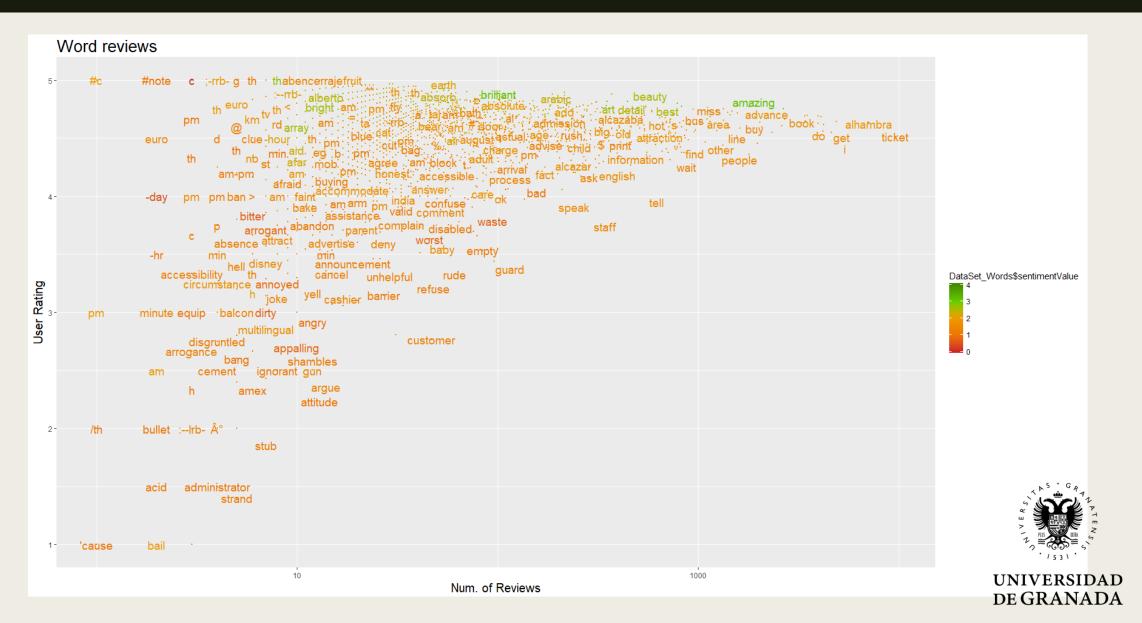Word reviews

Word reviews

Análisis de Opiniones en Monumentos Culturales

IX JORNADA Ⓡ
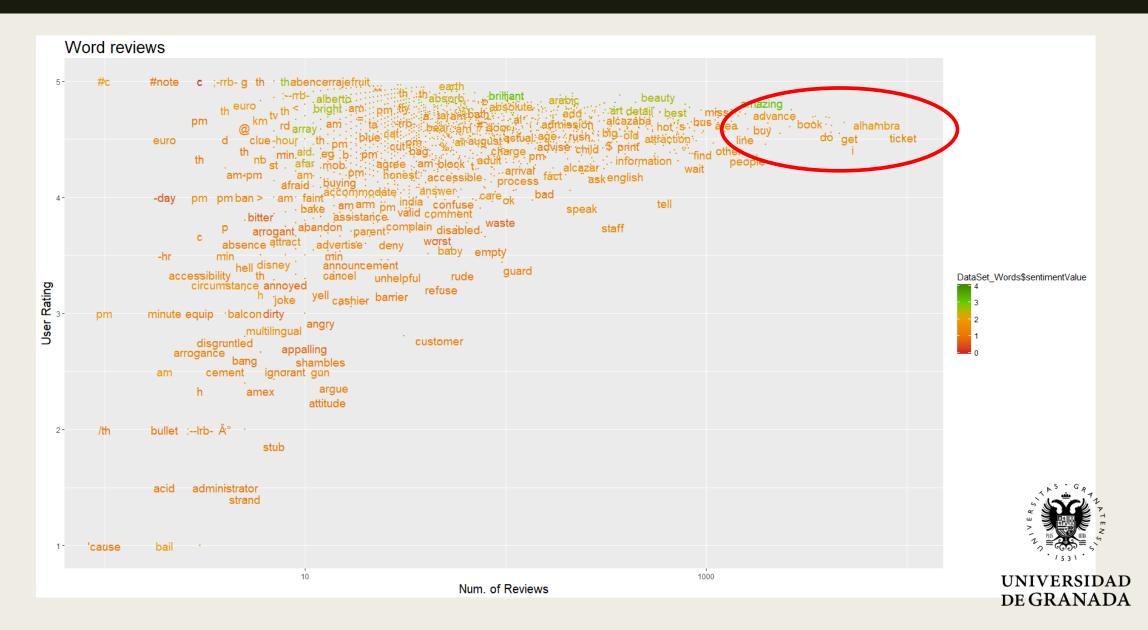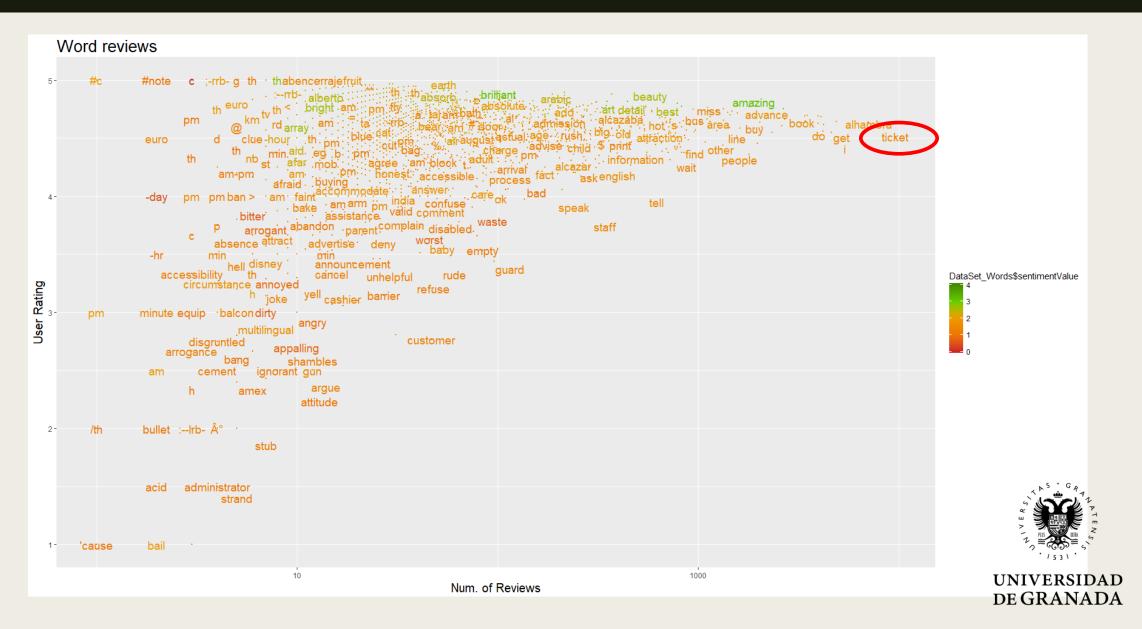
Conclusions

1. Sentiment Analysis can be a good tool to help **digitize** cultural monuments.

2. We can **replace the survey methodology** with **Sentiment Analysis.**

3. We may **lose information** using the User Rating of TripAdvisor.

UNIVERSIDAD
DE GRANADA

avaldivia@ugr.es

@ana_valdi

#9JRes