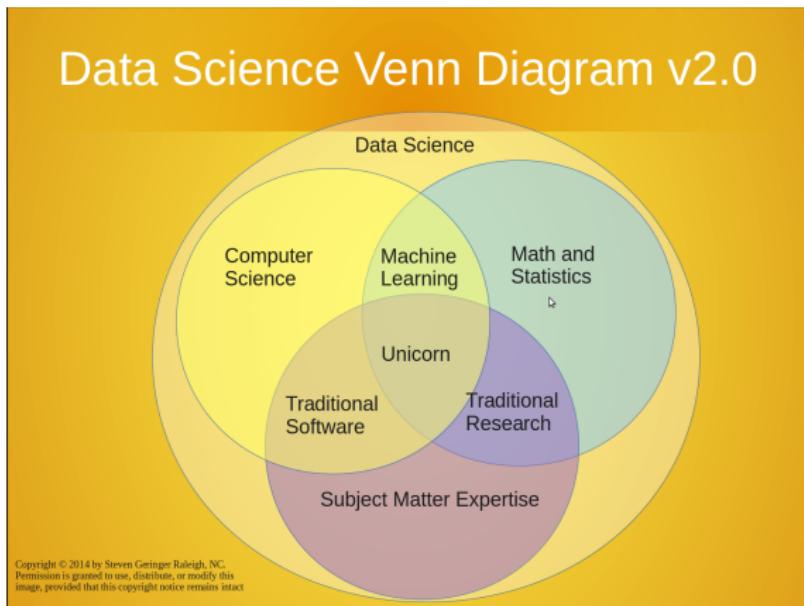
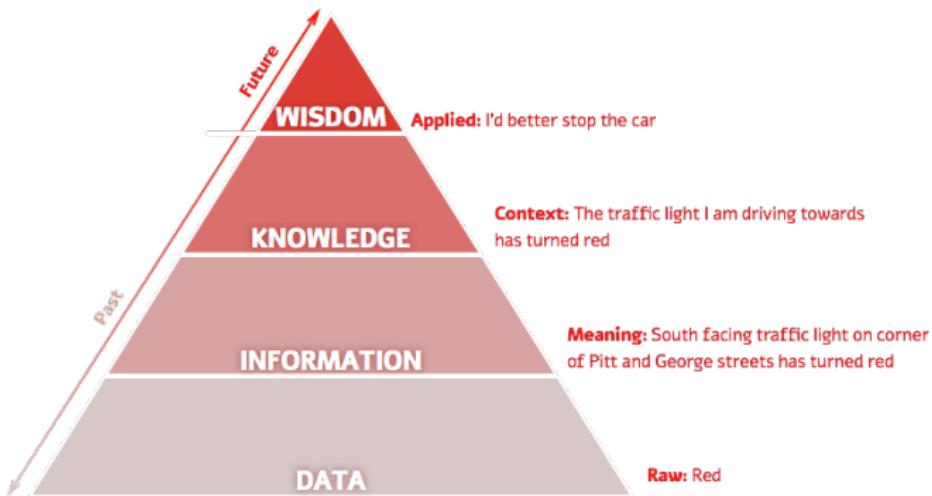


Cientistas & Unicórnios

- “Data Scientist” is a Data Analyst who lives in California.
- A data scientist is someone *who is better at statistics than any software engineer and better at software engineering than any statistician.* (Wills, Cloudera)



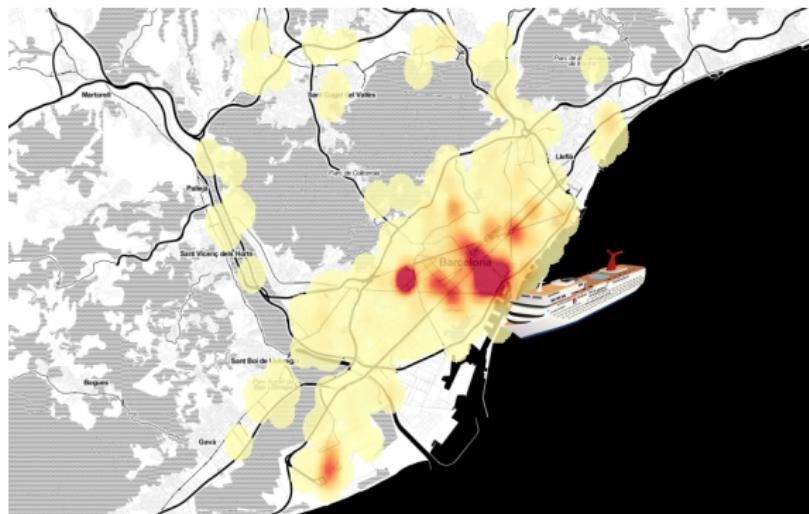
Dados, Informação e Conhecimento



Algo espacial...

Lei de Tobler:

Everything is related to everything else, but near things are more related to each other.

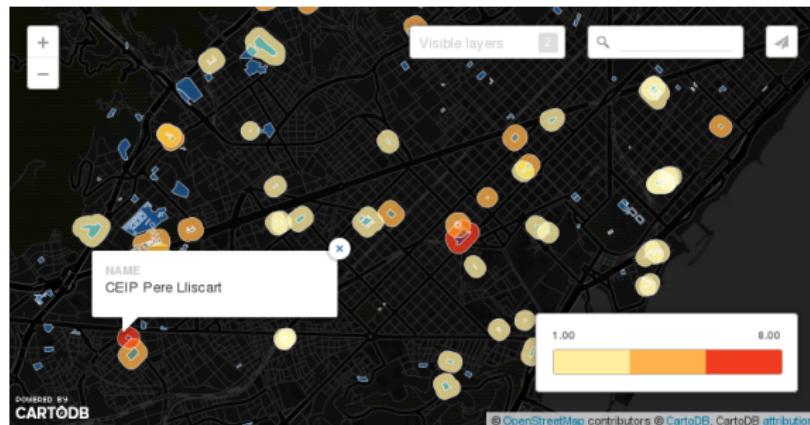


Heatmap de Tweets perto de um Cruzeiro (Eurecat, unpublished).

Algo espacial...

Lei de Tobler:

Everything is related to everything else, but near things are more related to each other.

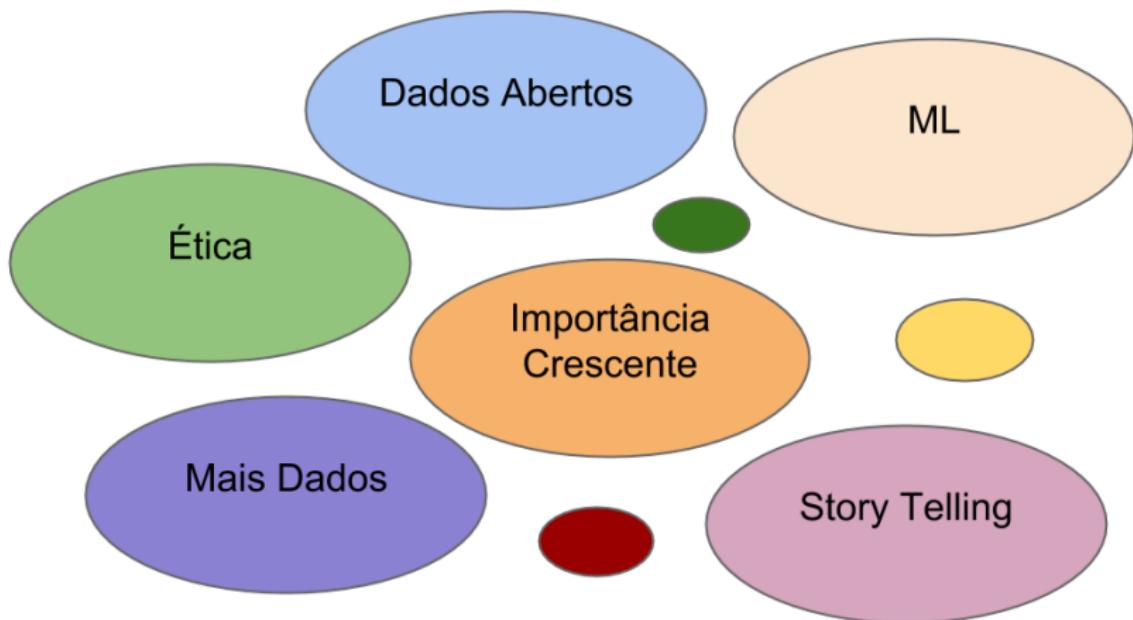


Buffers de acidentes a volta das escolas (Eurecat, unpublished).

Onde Vamos?



(Algumas) Tendências



Importância Crescente

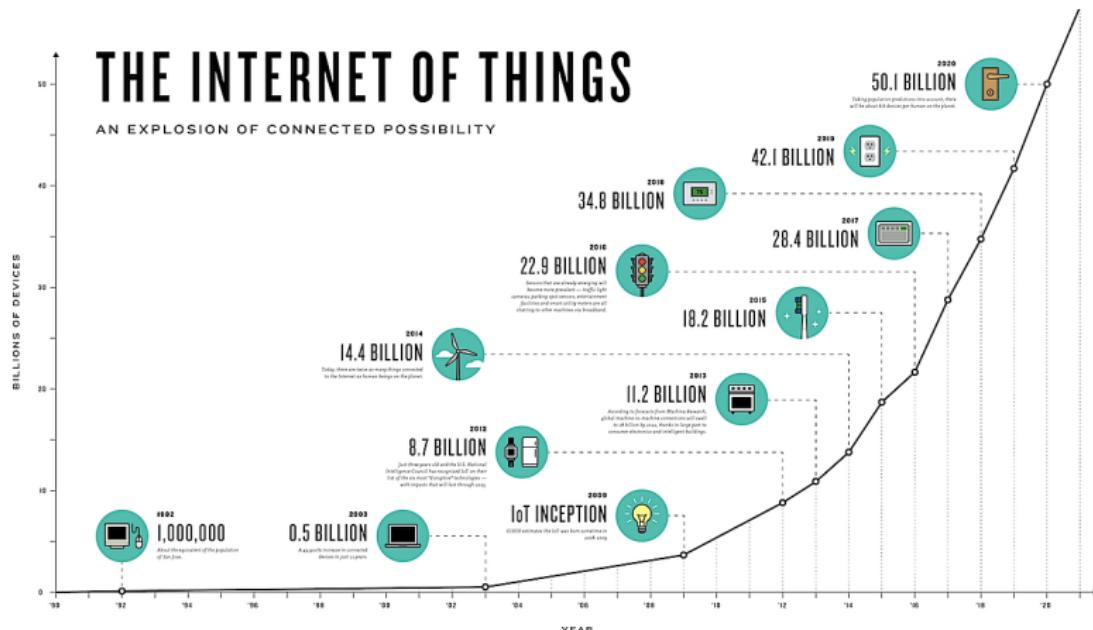


Runkeeper: fitness track app <https://runkeeper.com/>



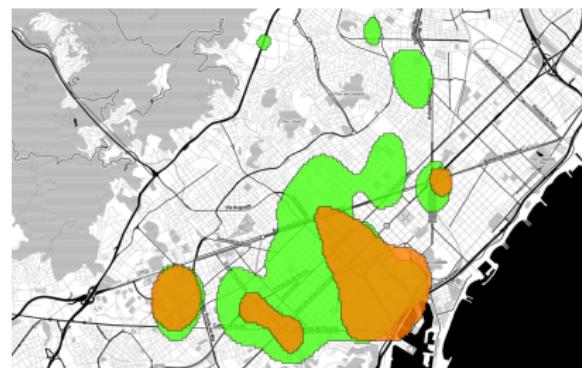
Withings: balança "inteligente" <http://www2.withings.com/eu/en>

Mais Dados



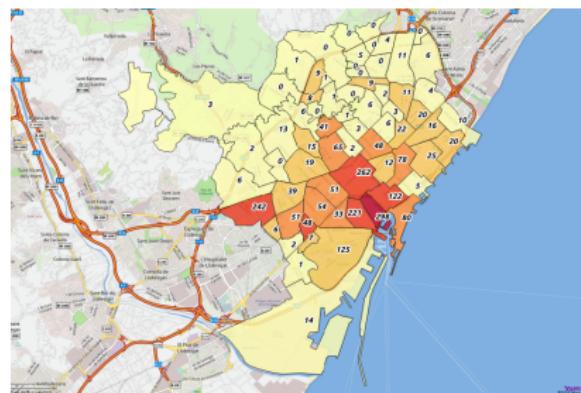
Infografia baseada en dados da Cisco <http://www.i-scoop.eu/internet-of-things/>

E Ainda Mais Dados



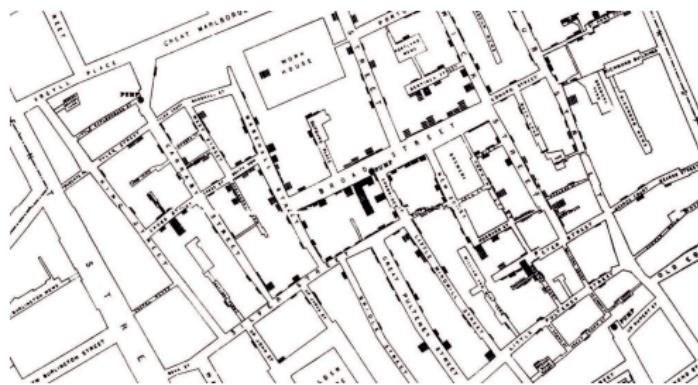
Clusters de Tweets enviados por locais e estrangeiros (Eurecat, unpublished).

E Ainda Mais Dados



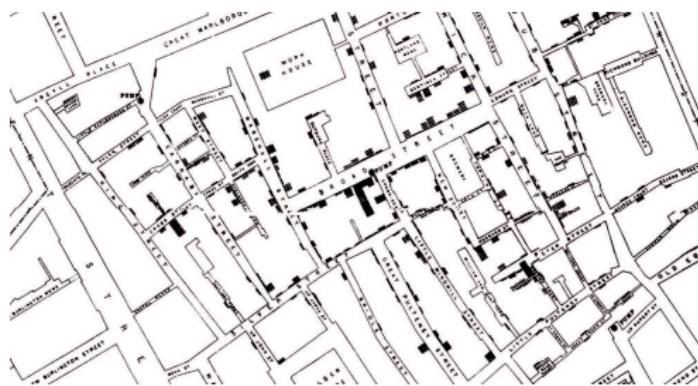
Distribuição de densidades de Tweets enviados por estrangeiros (Eurecat, unpublished).

“Arqueologia” de Dados



Mapa de casos de colera em Londres, produzido por John Snow (1864).

“Arqueologia” de Dados



Mapa de casos de colera em Londres, produzido por John Snow (1864).

```

Path list file is : [/home/joana/projects/no/test1/arg2', '/home/joana/projects/no/test2/arg1']

Rootdir list is : ['lmg2', 'mrg1']

Output file is : /tmp/out.html

File duplicates is : 0

077fcf2677116ed42fa2b269702346 ['./home/joana/projects/no/test1/arg2/2000408_ang2/warnits02.arg', './home/joana/projects/no/test2/arg1/2000408_ang2/warnits02.arg']

238975376916022778249546202d1 ['./home/joana/projects/no/test1/arg2/2000408_ang2/warnmp20.arg', './home/joana/projects/no/test1/arg2/2000408_ang2/Par1_Pel_1/sancmp20.arg', './home/joana/projects/no/test2/arg1/2000408_ang2/warnmp20.arg']

1515c15e0d5025268c842784888751 ['./home/joana/projects/no/test1/arg2/2000408_ang2/warnmp20.arg', './home/joana/projects/no/test2/arg1/2000408_ang2/warnmp20.arg']

59b8b21ef5d49e8cd47407d254135 ['./home/joana/projects/no/test1/arg2/2000408_ang2/warnmp30.arg', './home/joana/projects/no/test1/arg2/2000408_ang2/warnmp30.arg', './home/joana/projects/no/test1/arg2/2000408_ang2/sancos02.arg', './home/joana/projects/no/test1/arg2/2000408_ang2/sancos02.arg']

542424242424242424242424242424 ['./home/joana/projects/no/test1/arg2/2000408_ang2/warnmp30.arg', './home/joana/projects/no/test1/arg2/2000408_ang2/sancos02.arg', './home/joana/projects/no/test2/arg1/2000408_ang2/warnmp30.arg', './home/joana/projects/no/test2/arg1/2000408_ang2/sancos02.arg']

1622175b313957ca202395141b8 ['./home/joana/projects/no/test2/arg1/2008_Pelagic_Gabon_Gabon/sanxx86.arg', './home/joana/projects/no/test2/arg1/2008_Pelagic_Gabon_Gabon/sanxx86.arg']

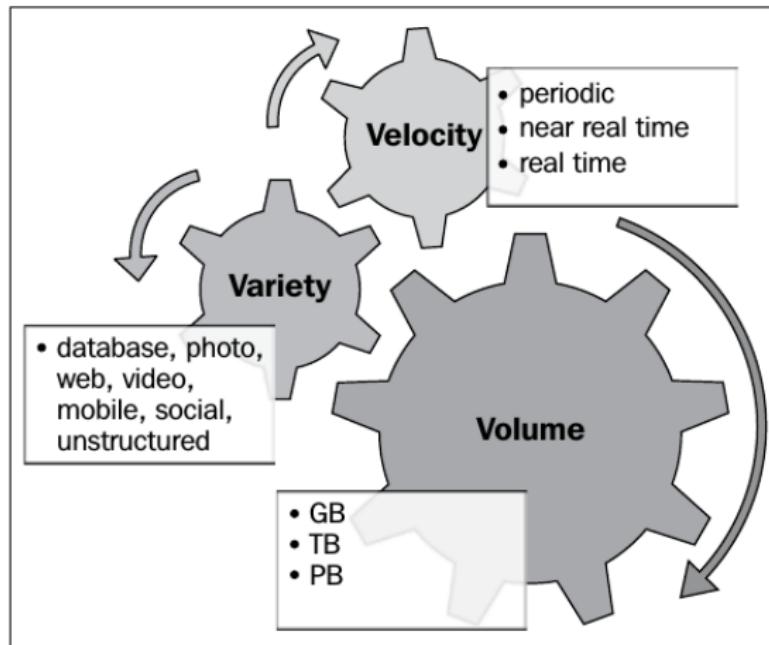
23900593835856ba59a37e70f53c1 ['./home/joana/projects/no/test1/arg2/2000408_ang2/AMNTS02.MRG', './home/joana/projects/no/test2/arg1/2000408_ang2/AMNTS02.MRG']

False. Applications is : 0

File duplicates is : 0
```

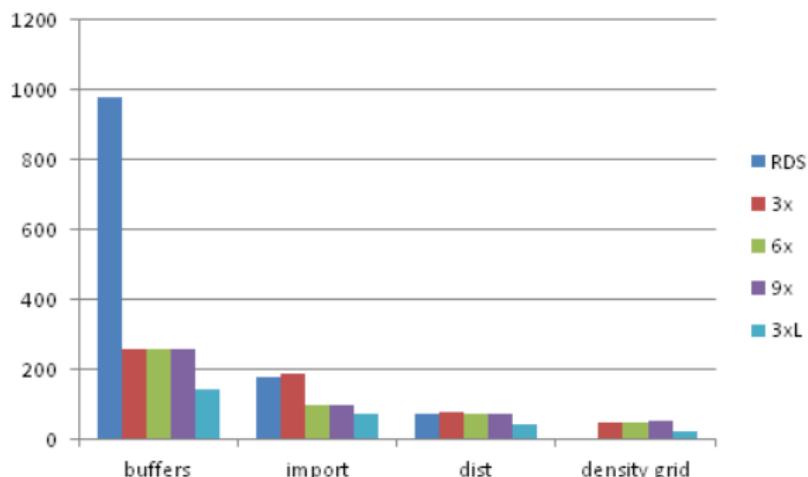
Descodificacao de metadados atraves da estrutura de directorios (Institute of Marine Research, unpublished).

Tecnologias de *Big Data*



Os 3 Vs de Big Data (Cuesta, H., 2013).

Tecnologias de *Big Data*



Benchmarking de bases de dados espaciais na cloud (Simoes, 2015).

Uso Cada vez Mais Generalizado de ML

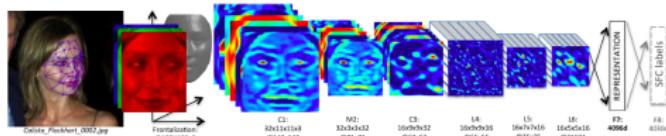
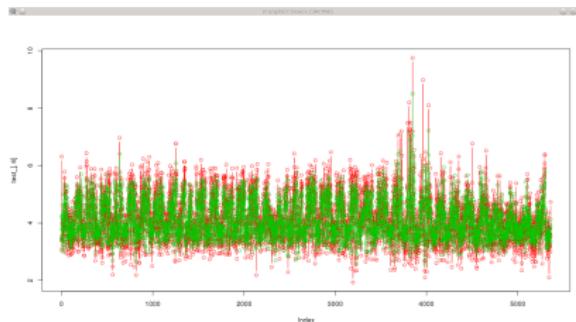


Figure 2. Outline of the DeepFace architecture. A front-end of a single convolution-pooling-convolution filtering on the rectified input, followed by three locally-connected layers and two fully-connected layers. Colors illustrate outputs for each layer. The net includes more than 120 million parameters, where more than 95% come from the local and fully connected layers.

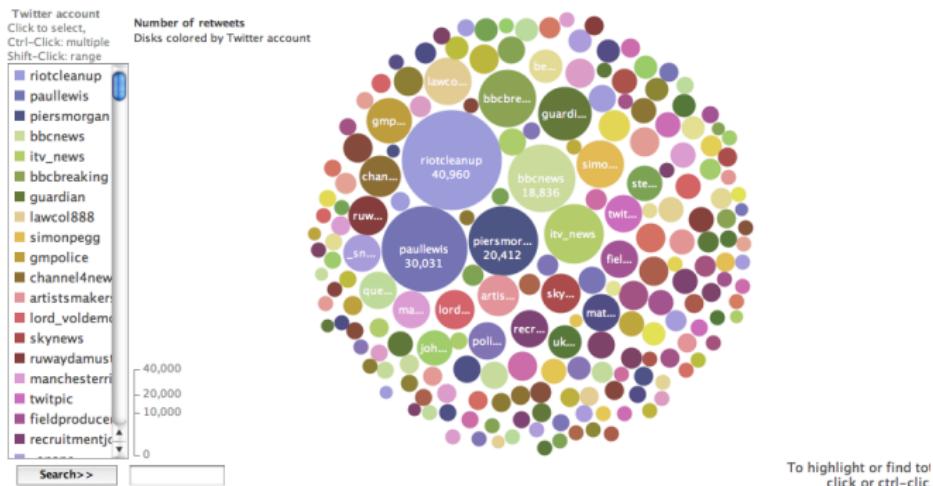
Arquitectura de uma red de Deep Learning para reconhecimento facial

<https://gigaom.com/2015/03/06/how-paypal-uses-deep-learning-and-detective-work-to-fight-fraud/>.



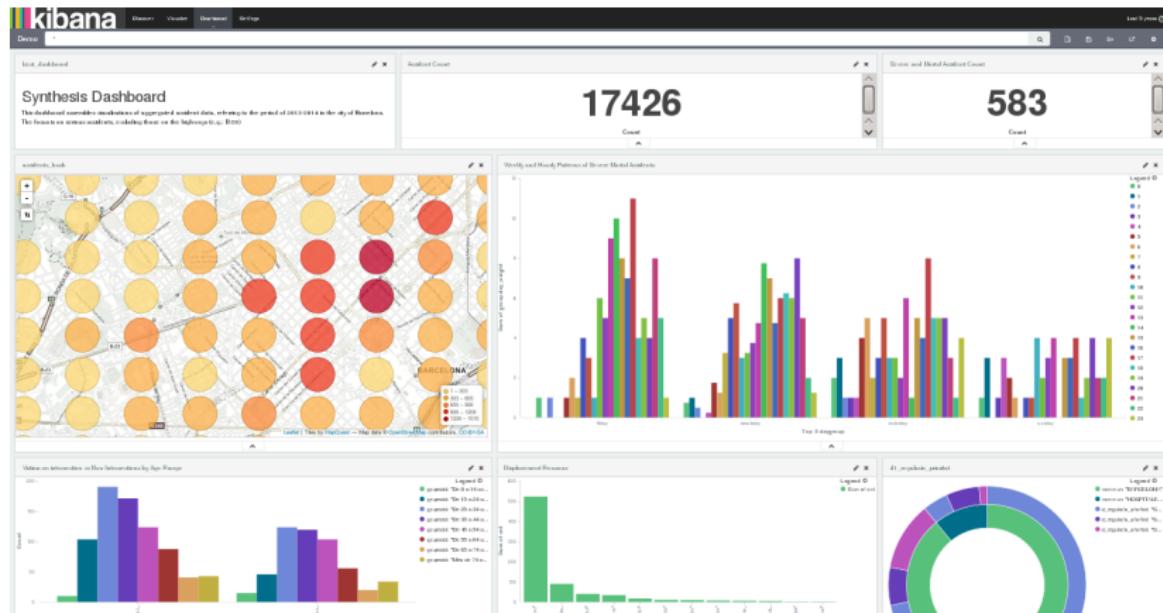
Tempos de viagem na cidade de Barcelona: ajuste entre as previsões SVM (verde) e os valores observados (vermelho) (Eurecat, unpublished).

Jornalismo de Dados & *Story Telling*



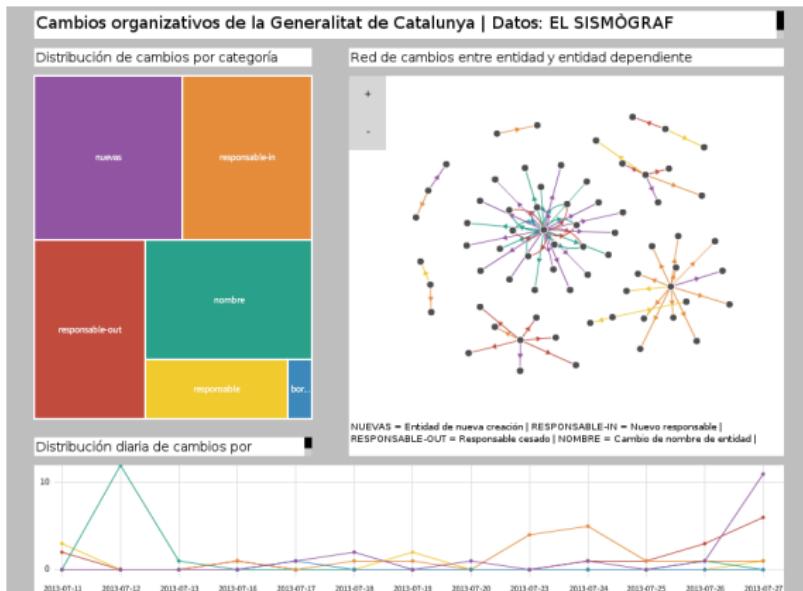
How riot rumours spread on Twitter (Guardian Interactive team, 2011).

Jornalismo de Dados & Story Telling



Dashboard sobre acidentes graves e mortais na cidade de Barcelona (Eurecat, unpublished).

Mais (e Melhores) Dados Abertos



Sismógrafo da Catalunya, @OpenGovCat <http://opengov.cat/en/>

Ética

Ofuscacao espacial:

Ética

Ofuscacao espacial:

- Annonimity set

Ética

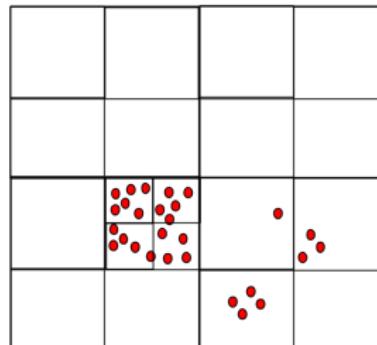
Ofuscacão espacial:

- Anonymity set
- MSV

Ética

Ofuscacão espacial:

- Anonymity set
- MSV



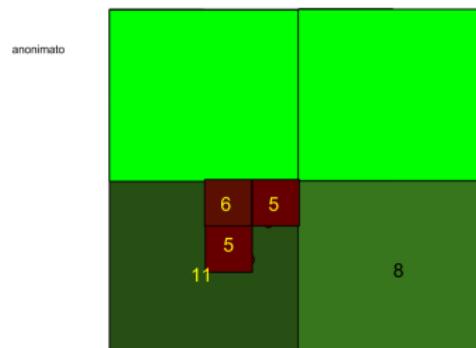
Pontos originais (Eurecat, unpublished).

Ética

Ofuscaccao espacial:

- Annonimity set
- MSV

N=35; MSV=5



Pontos ofuscados (Eurecat, unpublished).

SIG Open-Source?



Importância destas Tendências para o FOSS4G

Mais aplicações + mais quantidades de dados = maior comunidade

Importância destas Tendências para o FOSS4G

Mais aplicações + mais quantidades de dados = maior comunidade

- Infraestruturas colaborativas de dados.

Importância destas Tendências para o FOSS4G

Mais aplicações + mais quantidades de dados = maior comunidade

- Infraestruturas colaborativas de dados.
- Bibliotecas de ML com capacidades espaciais.

Importância destas Tendências para o FOSS4G

Mais aplicações + mais quantidades de dados = maior comunidade

- Infraestruturas colaborativas de dados.
- Bibliotecas de ML com capacidades espaciais.
- Visualização (interativa) de dados espaciais (3D).

Importância destas Tendências para o FOSS4G

Mais aplicações + mais quantidades de dados = maior comunidade

- Infraestruturas colaborativas de dados.
- Bibliotecas de ML com capacidades espaciais.
- Visualização (interativa) de dados espaciais (3D).
- Infraestrutura e processamento de Big Spatial Data.

Importância destas Tendências para o FOSS4G

Mais aplicações + mais quantidades de dados = maior comunidade

- Infraestruturas colaborativas de dados.
- Bibliotecas de ML com capacidades espaciais.
- Visualização (interativa) de dados espaciais (3D).
- Infraestrutura e processamento de Big Spatial Data.
- Ética e privacidade.

Obrigada pela vossa Atenção

Esta apresentação encontra-se disponível em:

<http://tinyurl.com/nfbrhvl>

