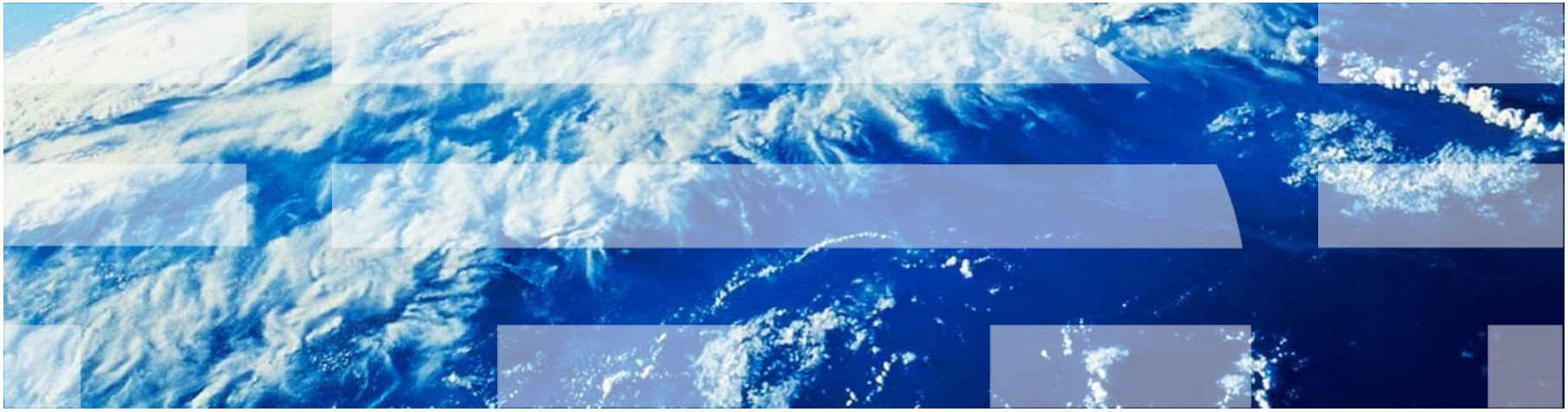
Social Media Sentiment Analysis



Agenda

- What is sentiment analysis?
- Implementation overview
- What did we use?
- Examples:
  - Euro 2012
  - London 2012
  - Mobile World Congress 2013
  - CeBIT 2013
- Summary

© 2013 IBM Corporation

What is sentiment analysis

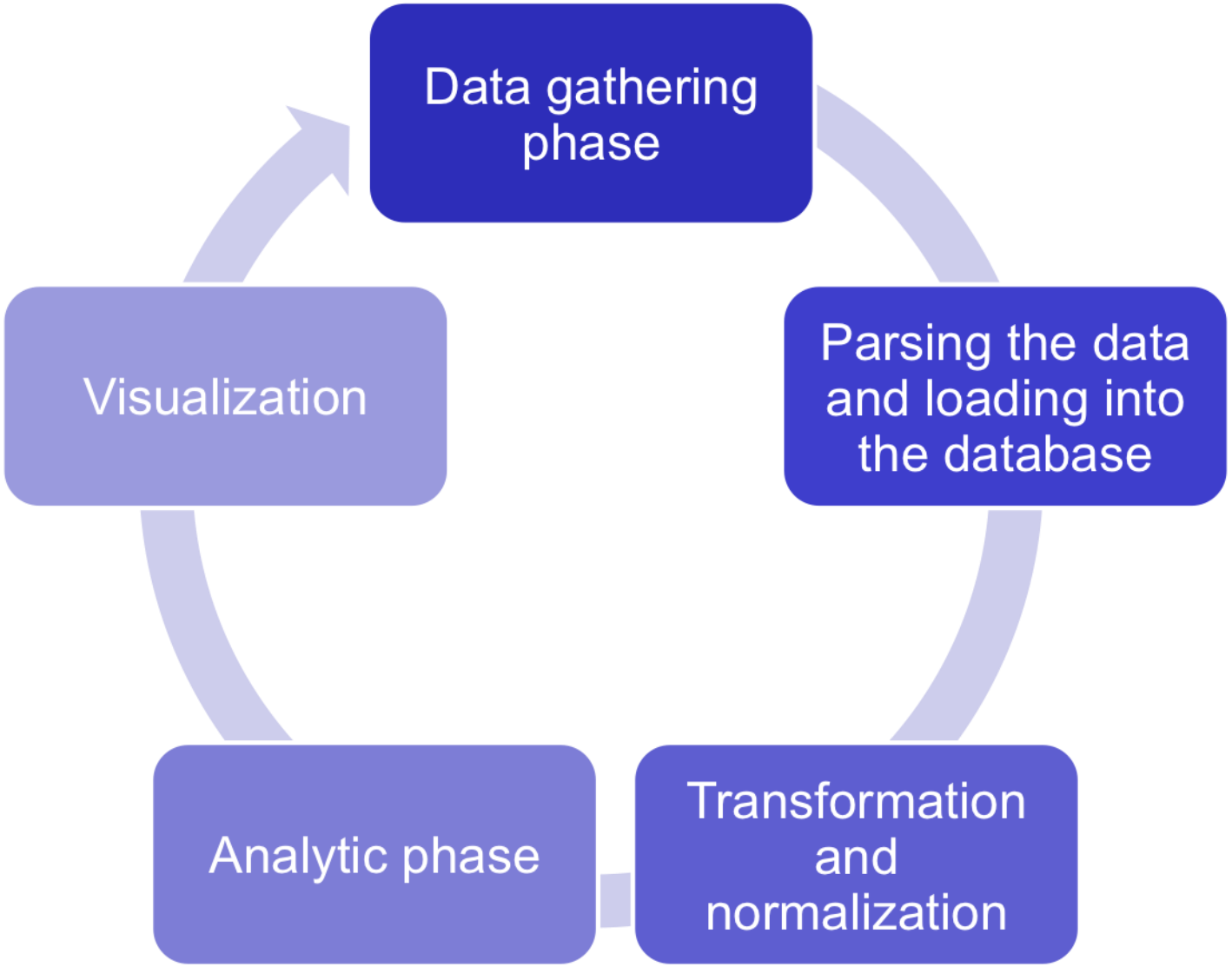
- About 12+ TB of new messages each day
- 25 TB of data in total per day

20 000 000 of tweets x 140 characters = 1 400 000 pages
Let's assume 300 pages per book...
... about 4700 books

- In a nutshell, sentiment analysis (opinion mining) involves analyzing comments (feelings, emotions, opinions) left on social portals.
- Sophisticated methods were used to identify and extract subjective information from the material.

© 2013 IBM Corporation

Implementation overview



© 2013 IBM Corporation

What did we use?

IBM PureData System for Analytics with Map-Reduce

- Data transformation and normalization
- Splitting sentences into words
- Explanation of acronyms
- Removing stop words
- Discovering the language
- Processing queries and analytics

IBM Spss Modeler

- Discovering relationships in the data
- Presentation layer

Revolution R Enterprise

- Analytical computations
- Presentation layer

Tweeter API

- Tweeter Streaming and Search API

© 2013 IBM Corporation

IBM PureData System for Analytics
The Simple Appliance for Serious Analytics

Built-in Expertise

– **No indexes or tuning**

– **Data model agnostic**

– **Hardware accelerated, fully parallel, optimized, In Database Analytics**

Integration by Design

– **Server, Storage, Database in one easy to use package**

– **Automatic parallelization and resource optimization to scale efficiently and economically**

– **Enterprise-class security and platform management**

Simplified Experience

– **Up and running in hours**

– **Minimal up front design and tuning**

– **Minimal ongoing administration**

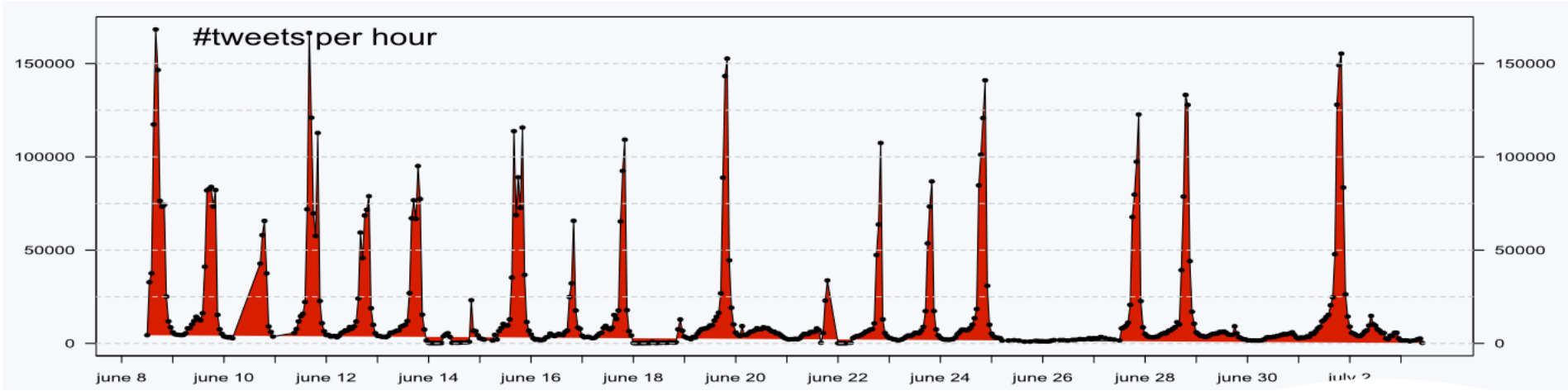
– **Standard interfaces to best of breed Analytics, Business Intelligence, and data integration tools**

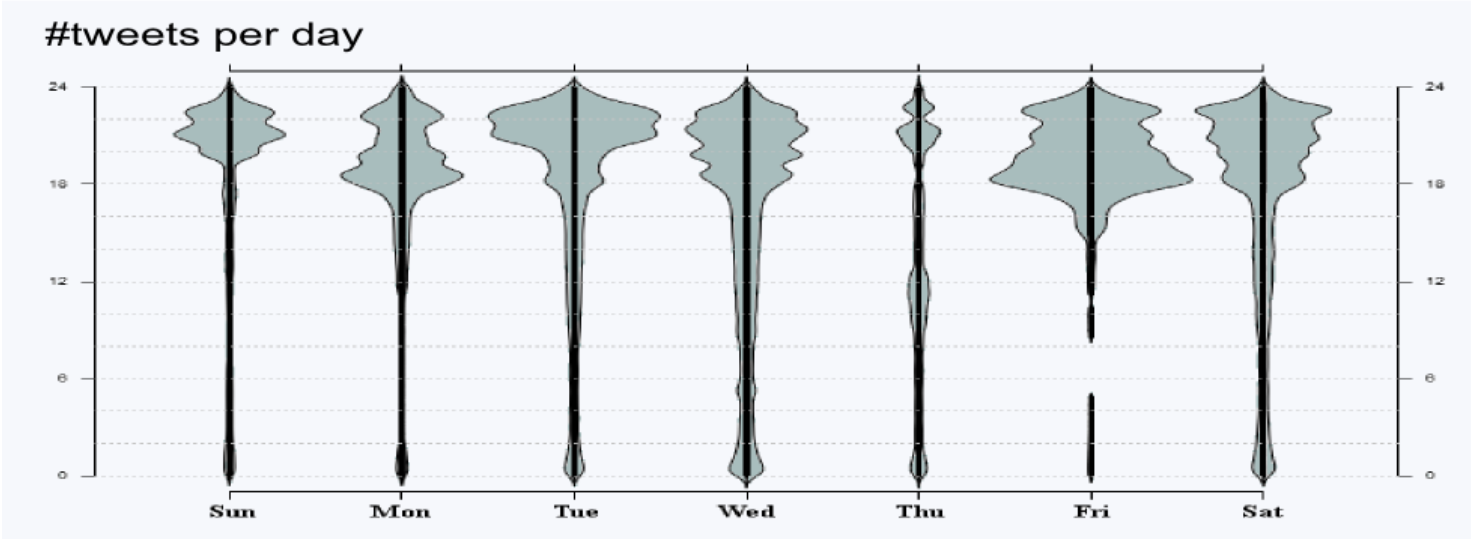
– **Built-in, complex analytical capabilities allow users to derive insight from their data quickly**

– **Easy connectivity to other Big Data Platform components**

© 2013 IBM Corporation

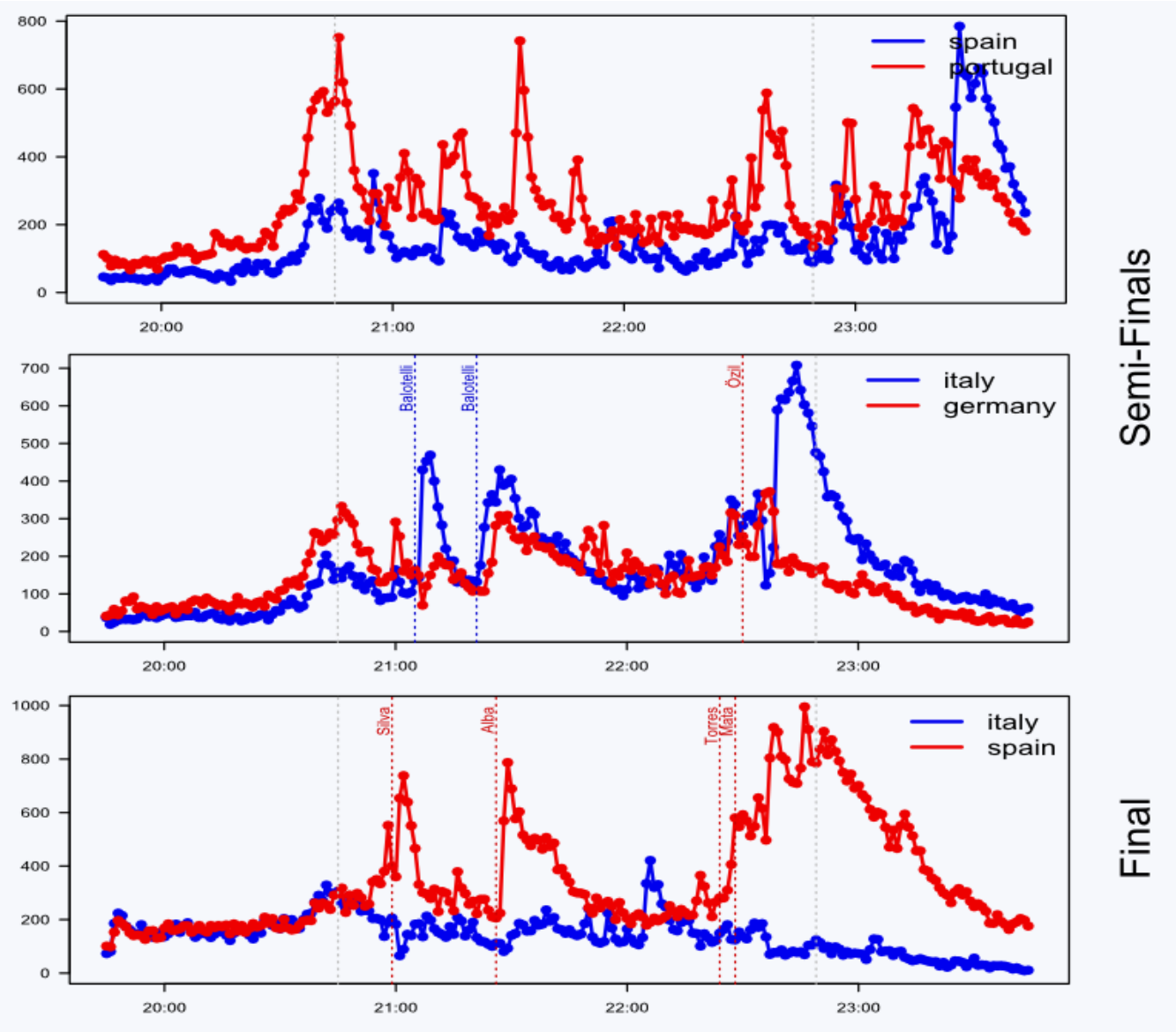
Euro 2012 – posts per hour and per day





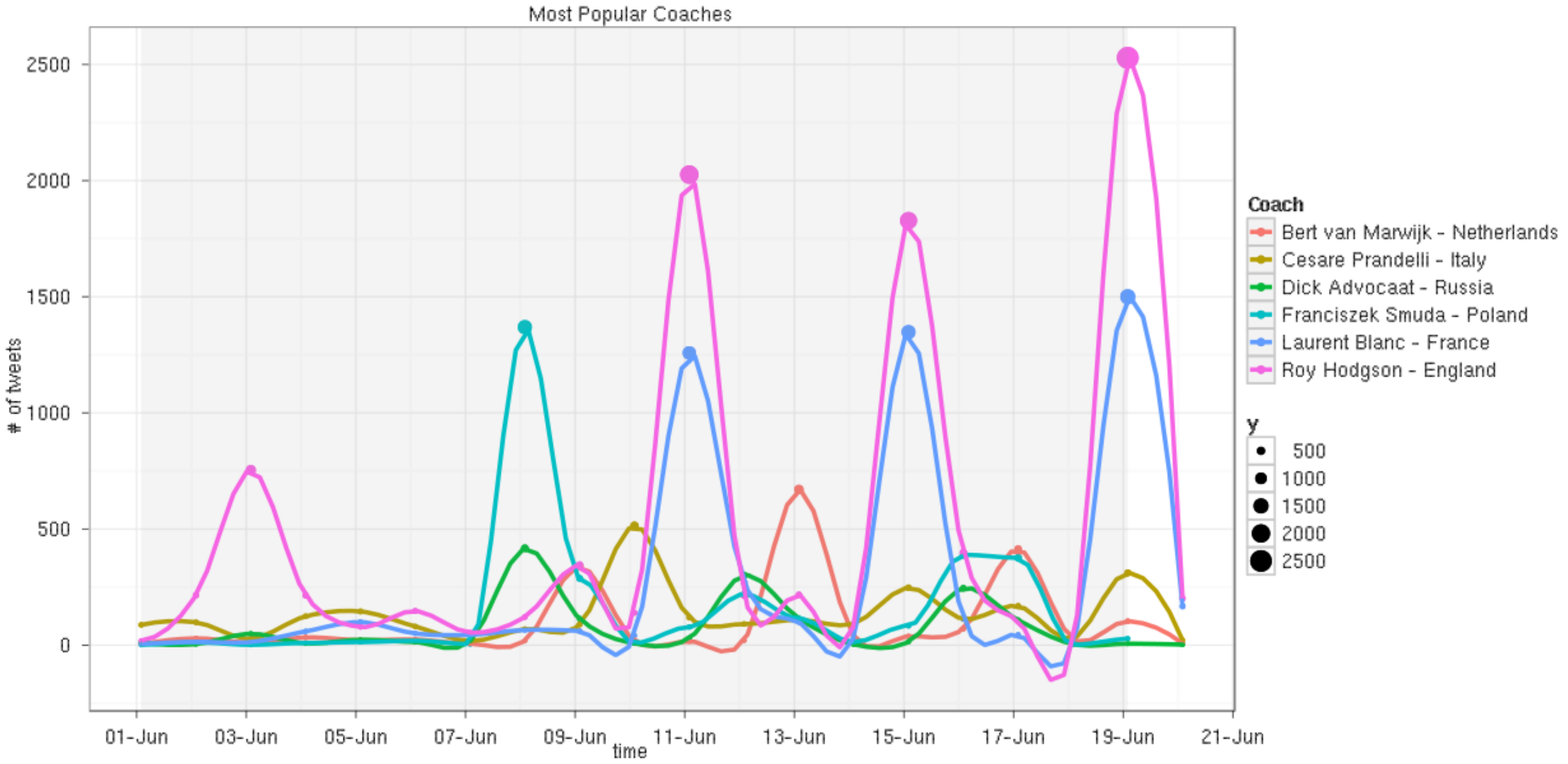
© 2013 IBM Corporation

Euro 2012 – semi finals and final sentiment



© 2013 IBM Corporation

Euro 2012 – most popular coaches in group phase



© 2013 IBM Corporation

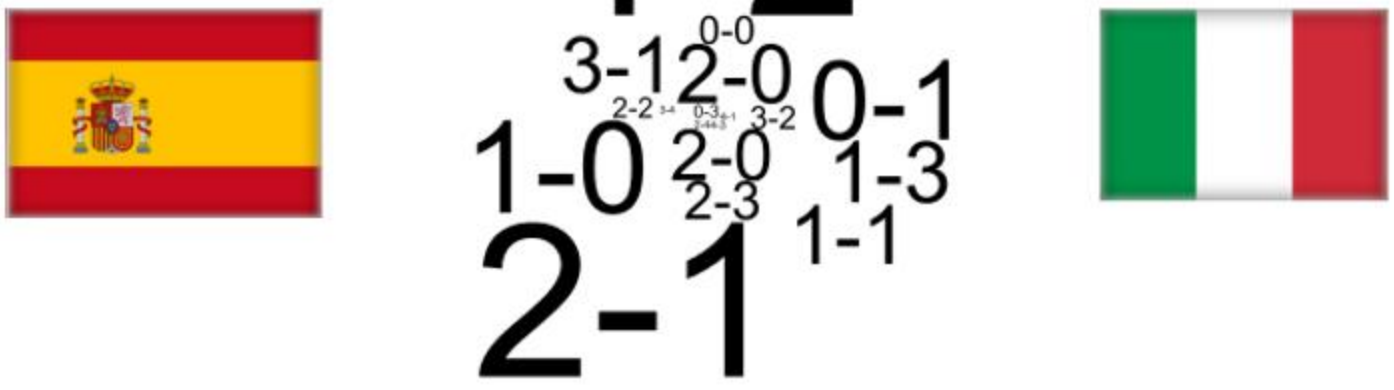
Euro 2012 – users predictions of games' scores

Team 1	Score	Hits	Team 2	Team1 wins / draw / Team 2 wins	Predictions			
Poland	1-1	11%	Greece	68.9% / 18.3% / 12.9%	2-1	21%	2-0	21%
Russia	4-1	1%	Czech	69.8% / 22.2% / 8%	2-1	21%	2-0	20%
Netherlands	0-1	0%	Denmark	91.9% / 6.1% / 2%	3-1	24%	2-0	24%
Germany	1-0	3%	Portugal	77.2% / 15.6% / 7.2%	2-1	29%	3-1	21%
Spain	1-1	12%	Italy	70.8% / 12.5% / 16.7%	2-0	42%	1-0	21%
Ireland	1-3	2%	Croatia	25.8% / 33.9% / 40.3%	1-1	27%	1-2	23%
France	1-1	18%	England	40.9% / 24.3% / 34.8%	1-1	18%	2-1	17%
Ukraine	2-1	6%	Sweden	16.8% / 28.2% / 55%	1-2	19%	1-1	16%
Greece	1-2	22%	Czech	21.1% / 32.4% / 46.5%	1-1	23%	1-2	22%
Poland	1-1	5%	Russia	8.4% / 16% / 75.6%	1-3	25%	1-2	24%
Denmark	2-3	0%	Portugal	11.8% / 27% / 61.2%	1-2	21%	1-1	20%
Netherlands	1-2	22%	Germany	21.9% / 34.7% / 43.4%	1-2	22%	1-1	17%
Italy	1-1	33%	Croatia	42.4% / 42.4% / 15.2%	1-1	33%	2-1	21%
Spain	4-0	21%	Ireland	87.5% / 4.2% / 8.3%	3-0	46%	4-0	21%
Ukraine	0-2	25%	France	6.1% / 12.4% / 81.4%	1-2	26%	0-2	25%
Sweden	2-3	1%	England	3.8% / 15.1% / 81.1%	1-2	25%	0-2	23%
Czech	1-0	2%	Poland	18.3% / 28.2% / 53.4%	1-2	38%	1-1	21%
Greece	1-0	1%	Russia	1.9% / 3.9% / 94.2%	0-3	29%	0-2	27%
Denmark	1-2	16%	Germany	3.9% / 10.3% / 85.8%	0-2	33%	1-2	16%

© 2013 IBM Corporation

Euro 2012 finals - users' predictions

Spain	4-0	0%	Italy	48.9% /	11.7% /	39.4%	2-1	20%	1-2	19%

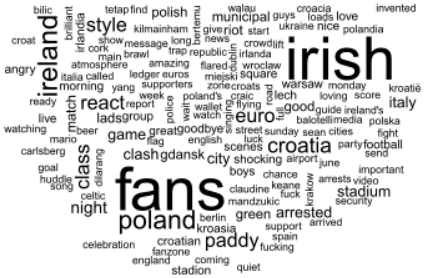


© 2013 IBM Corporation

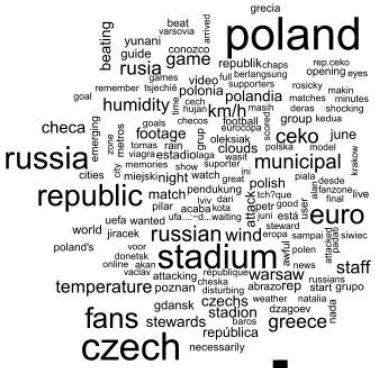
Euro 2012 – most popular topics in host cities

gdansk

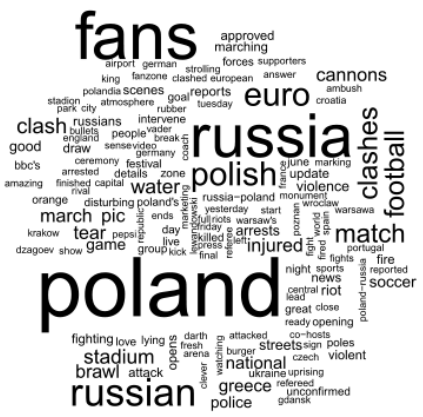
poznan



wroclaw

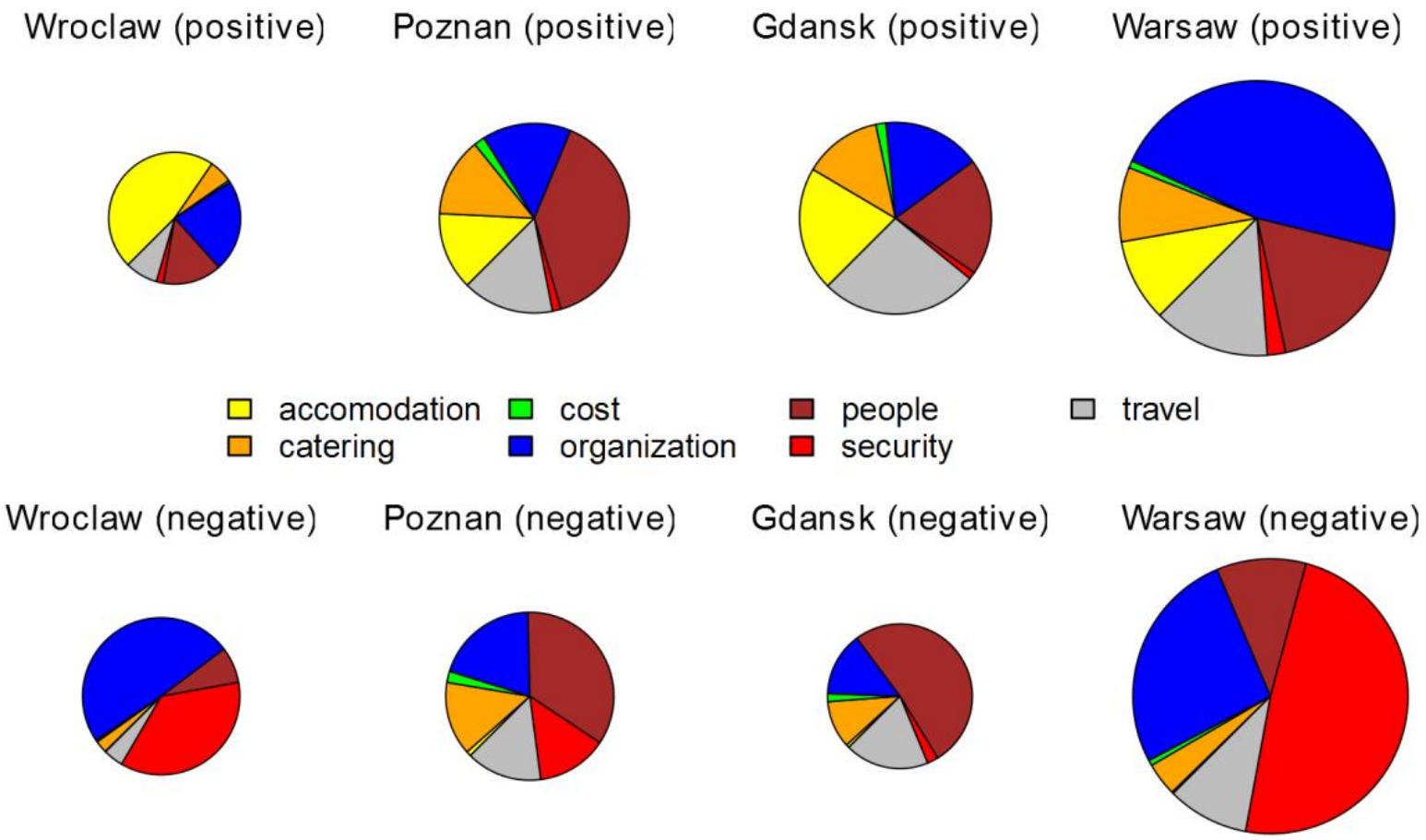


warsaw



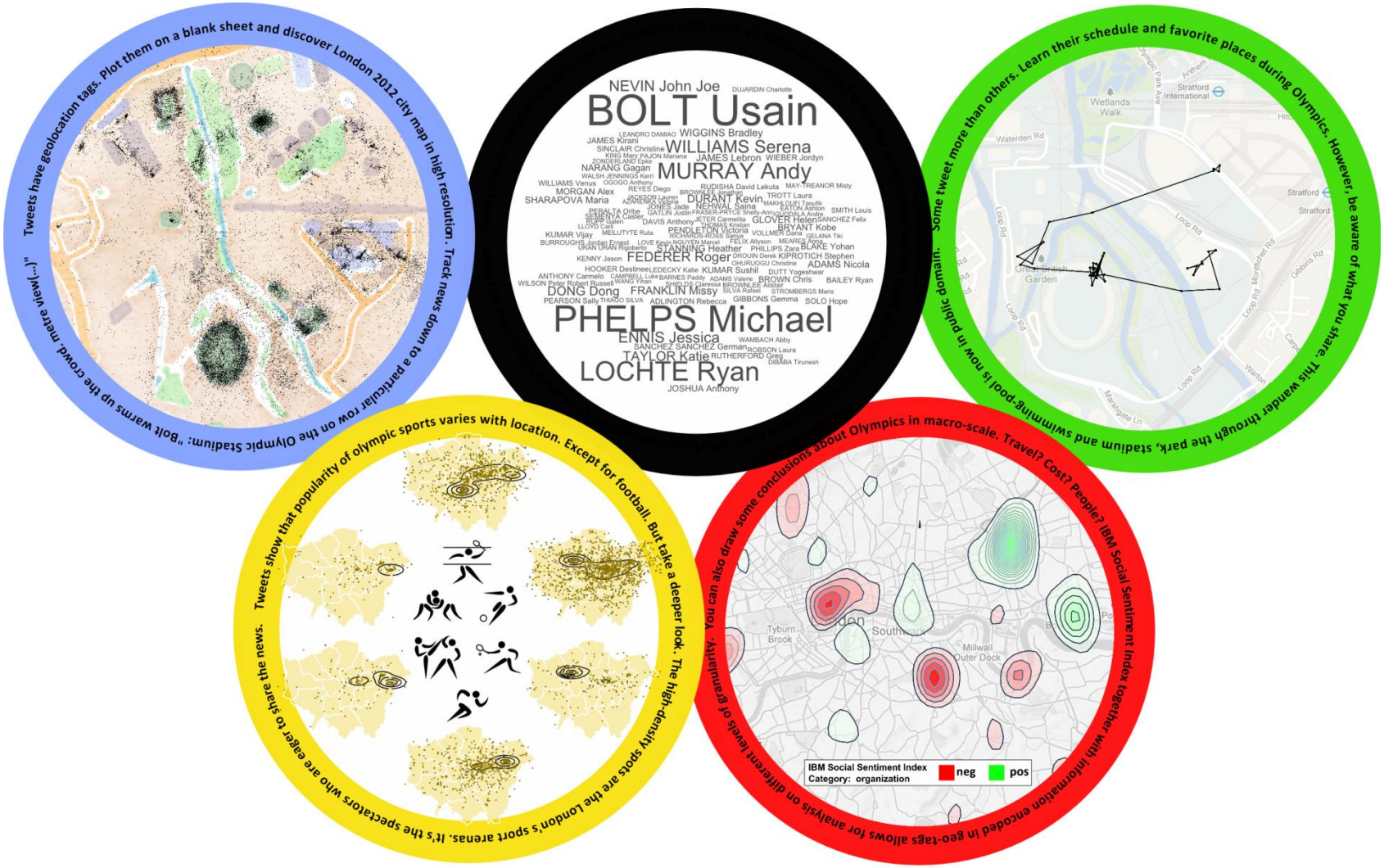
© 2013 IBM Corporation

Euro 2012 – opinions about the organisation of the event



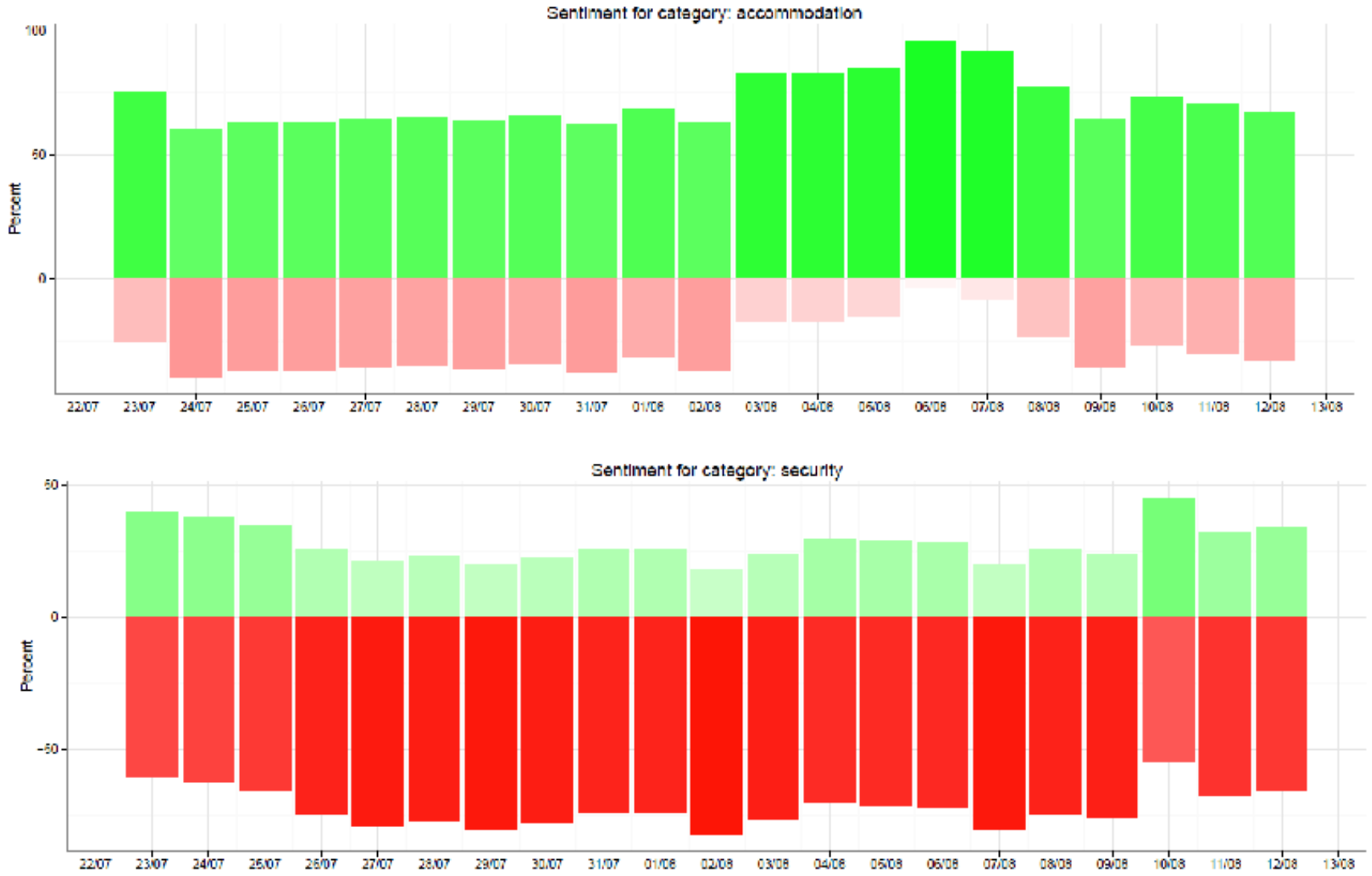
© 2013 IBM Corporation

London 2012



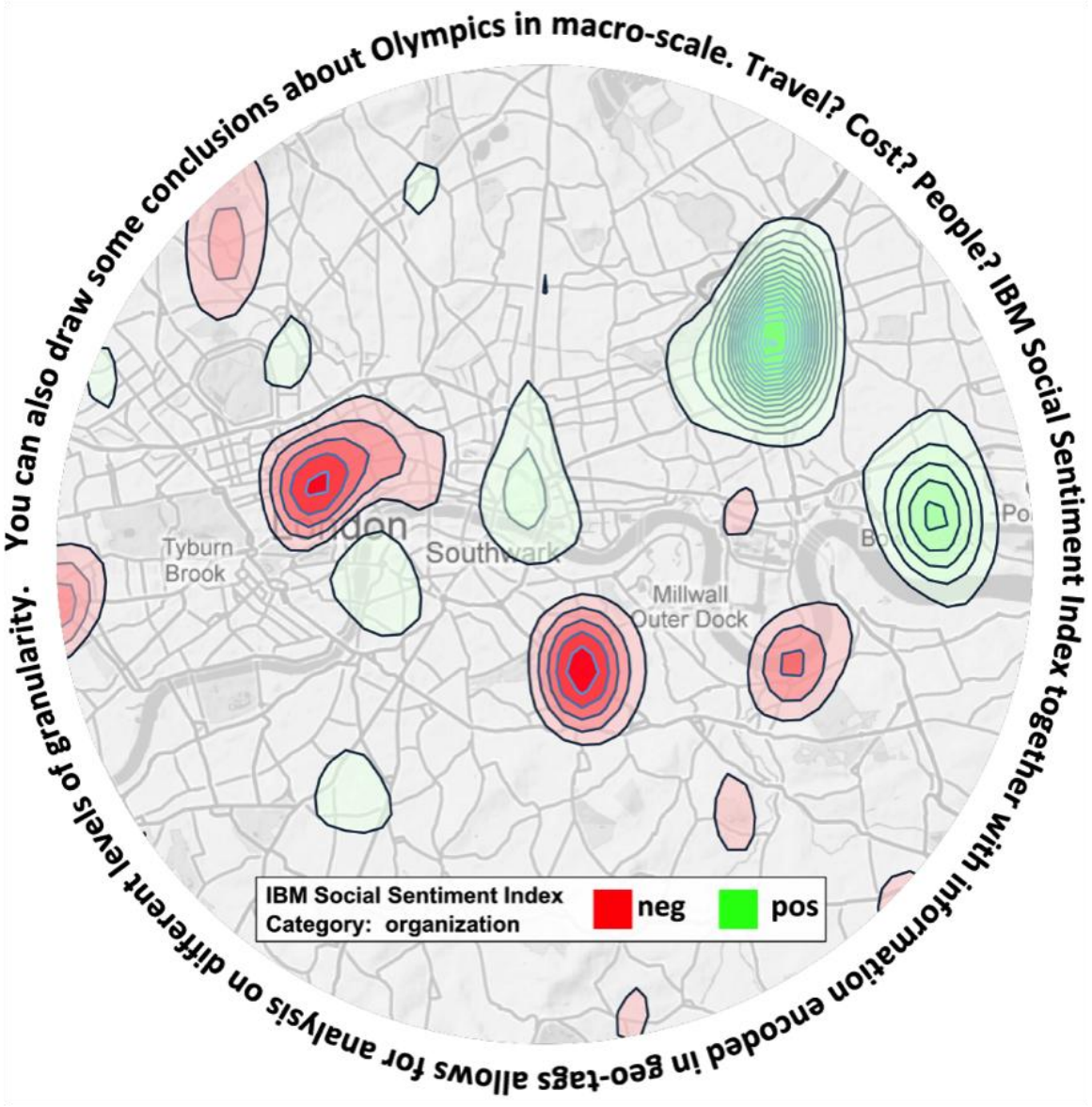
© 2013 IBM Corporation

London 2012



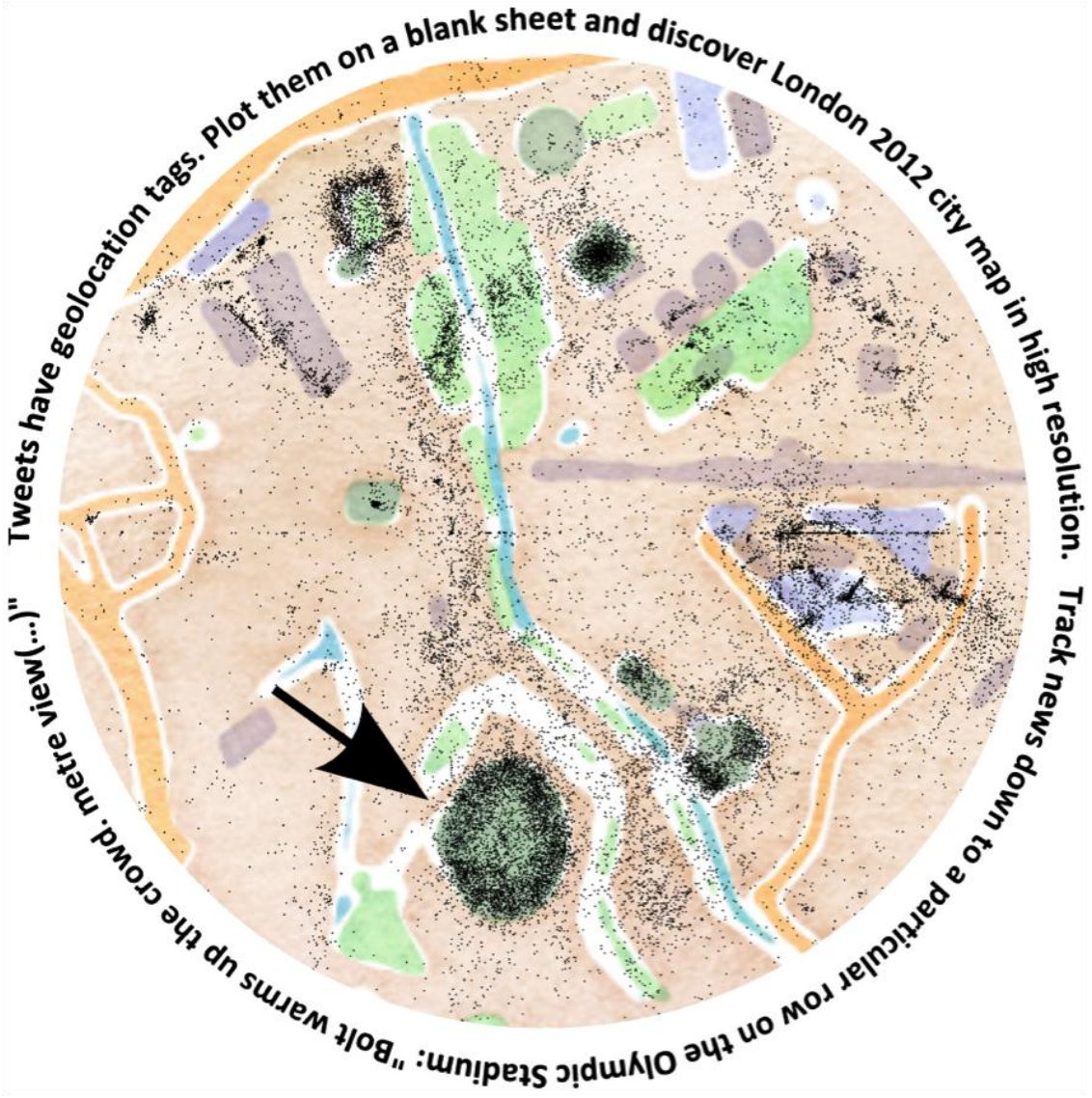
© 2013 IBM Corporation

London 2012 – sentiment index and geo tags used together



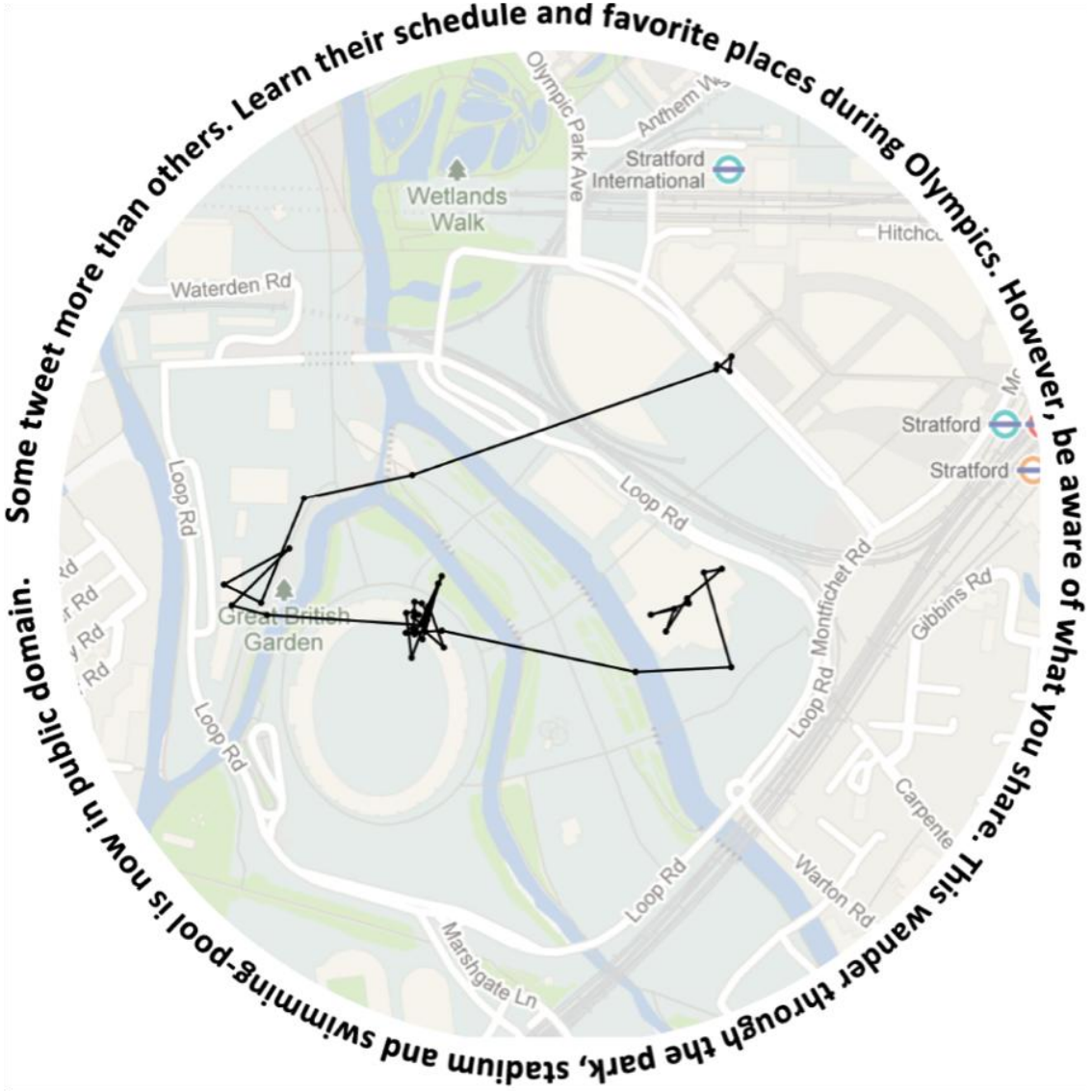
© 2013 IBM Corporation

London 2012 – usage geolocation information



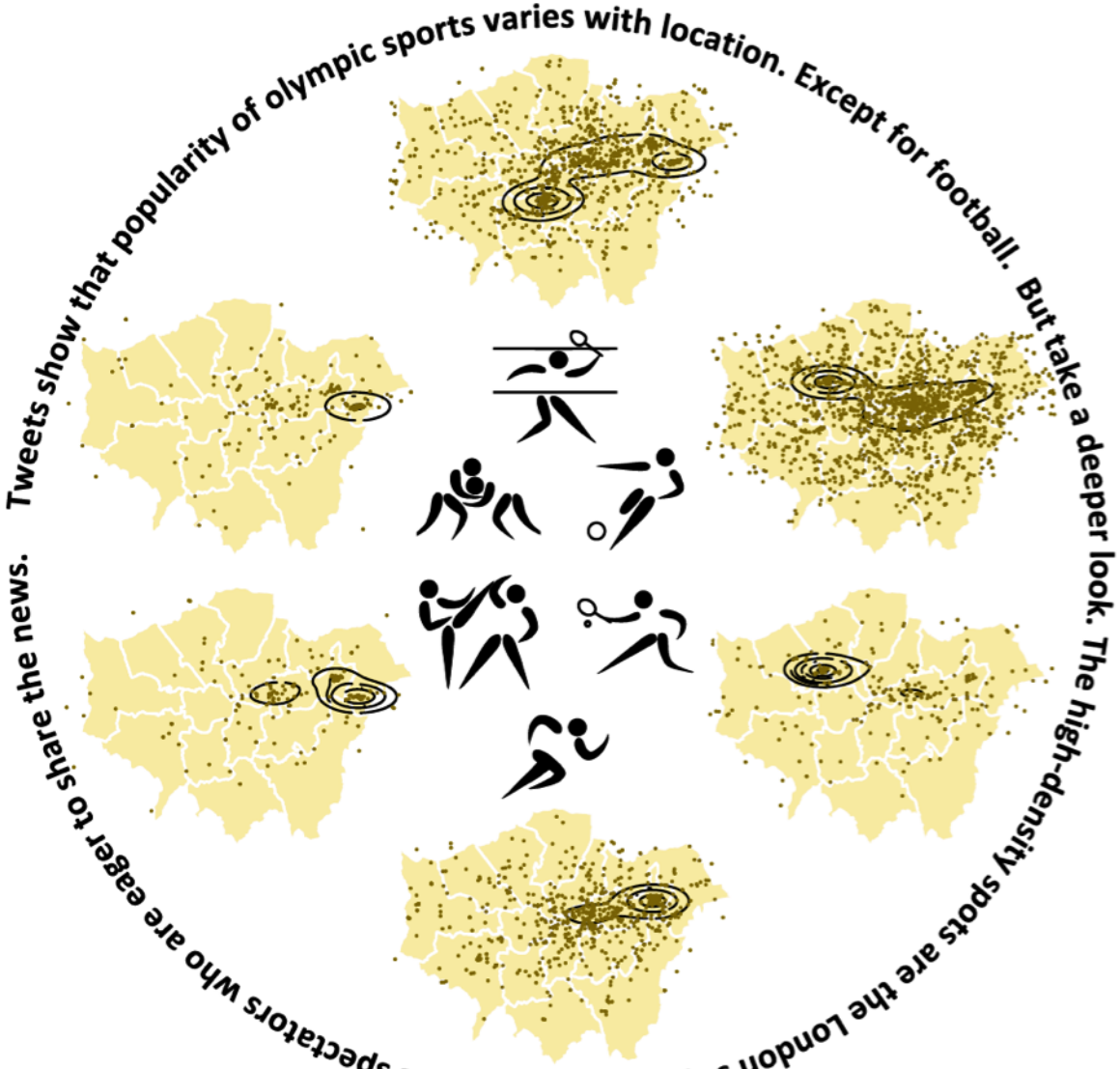
© 2013 IBM Corporation

London 2012 – tracking comments on a very detailed level



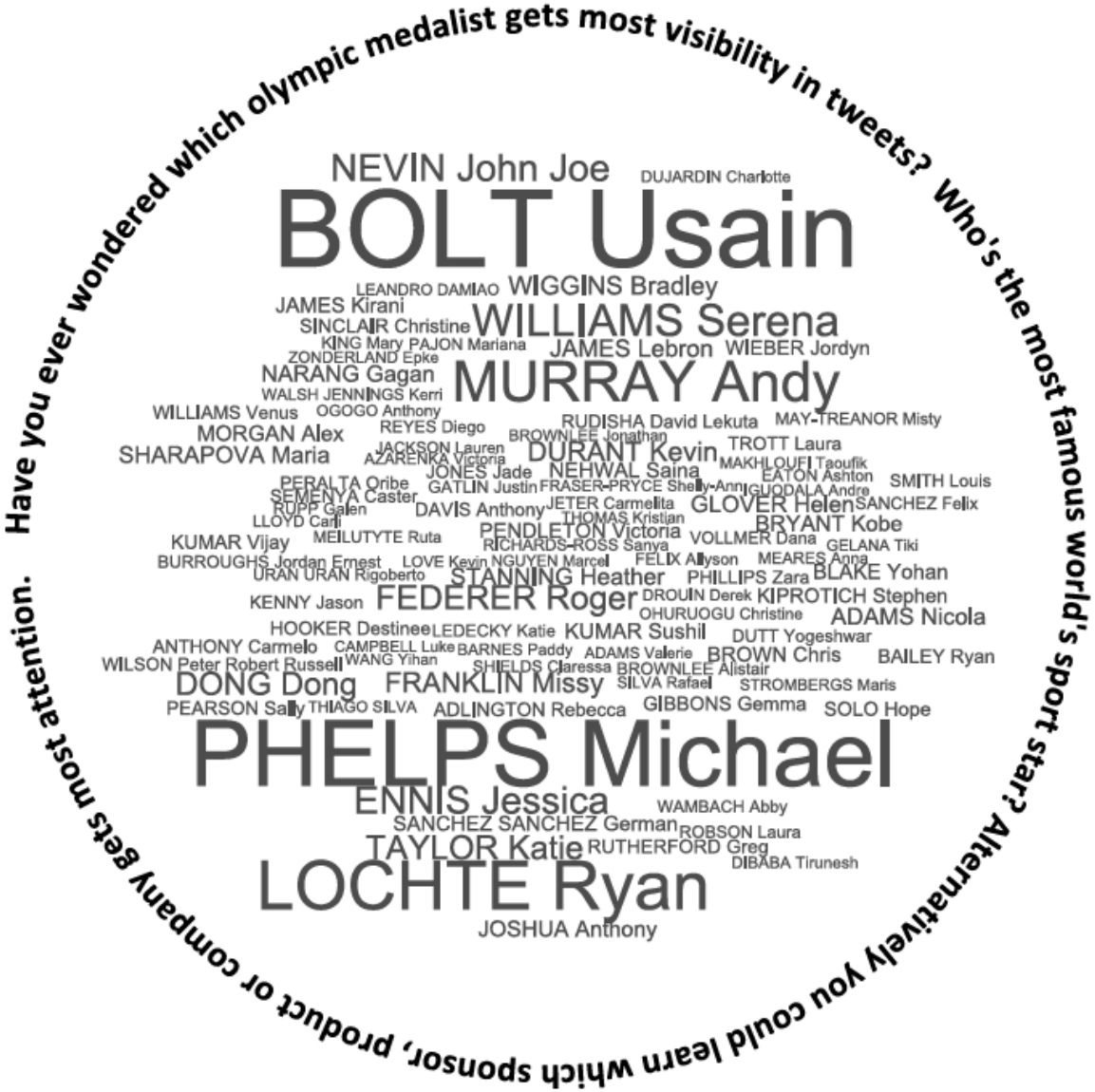
© 2013 IBM Corporation

London 2012 – the popularity of Olympic sports depends on location



© 2013 IBM Corporation

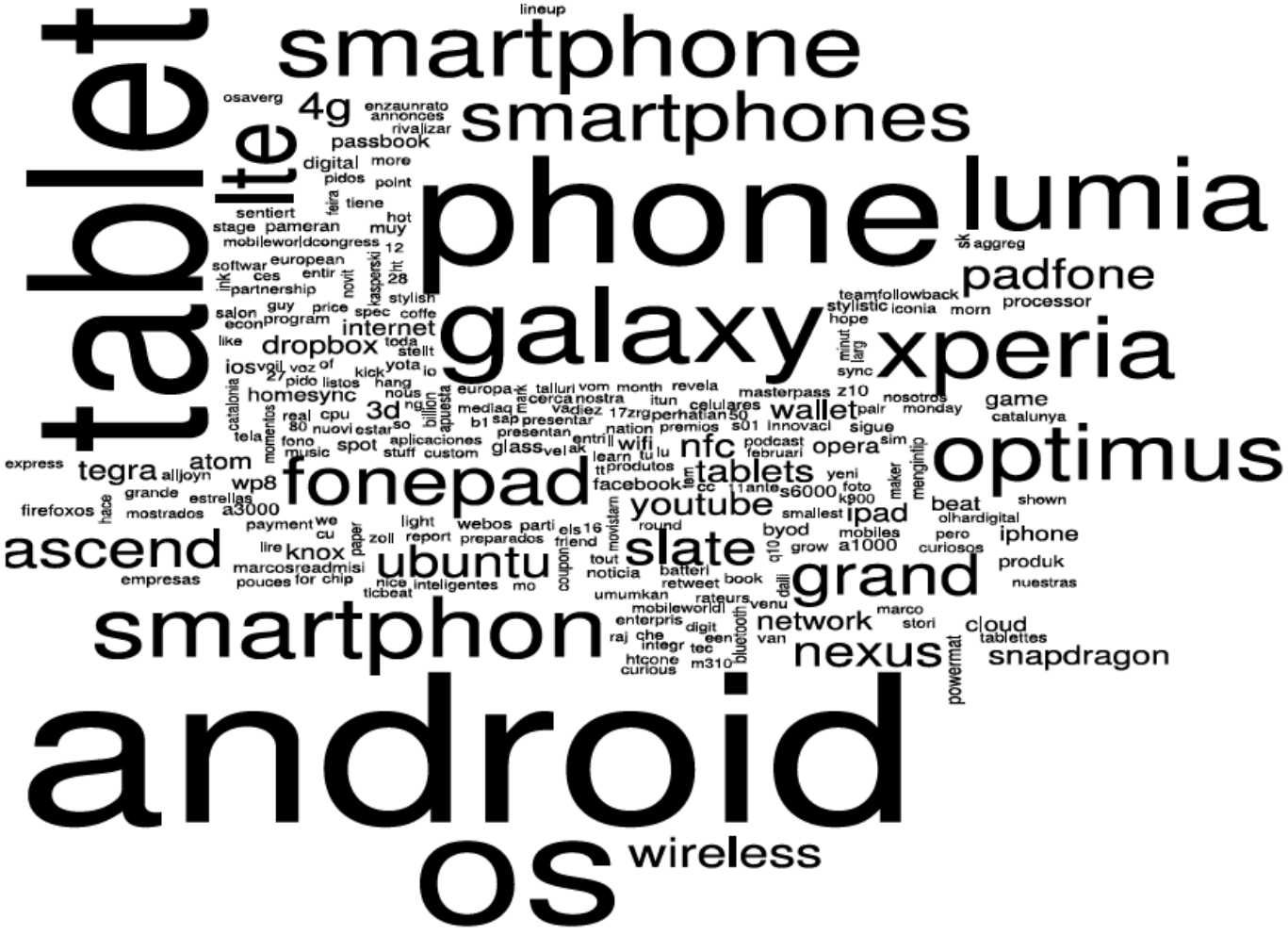
London 2012 – the most popular medalists



© 2013 IBM Corporation

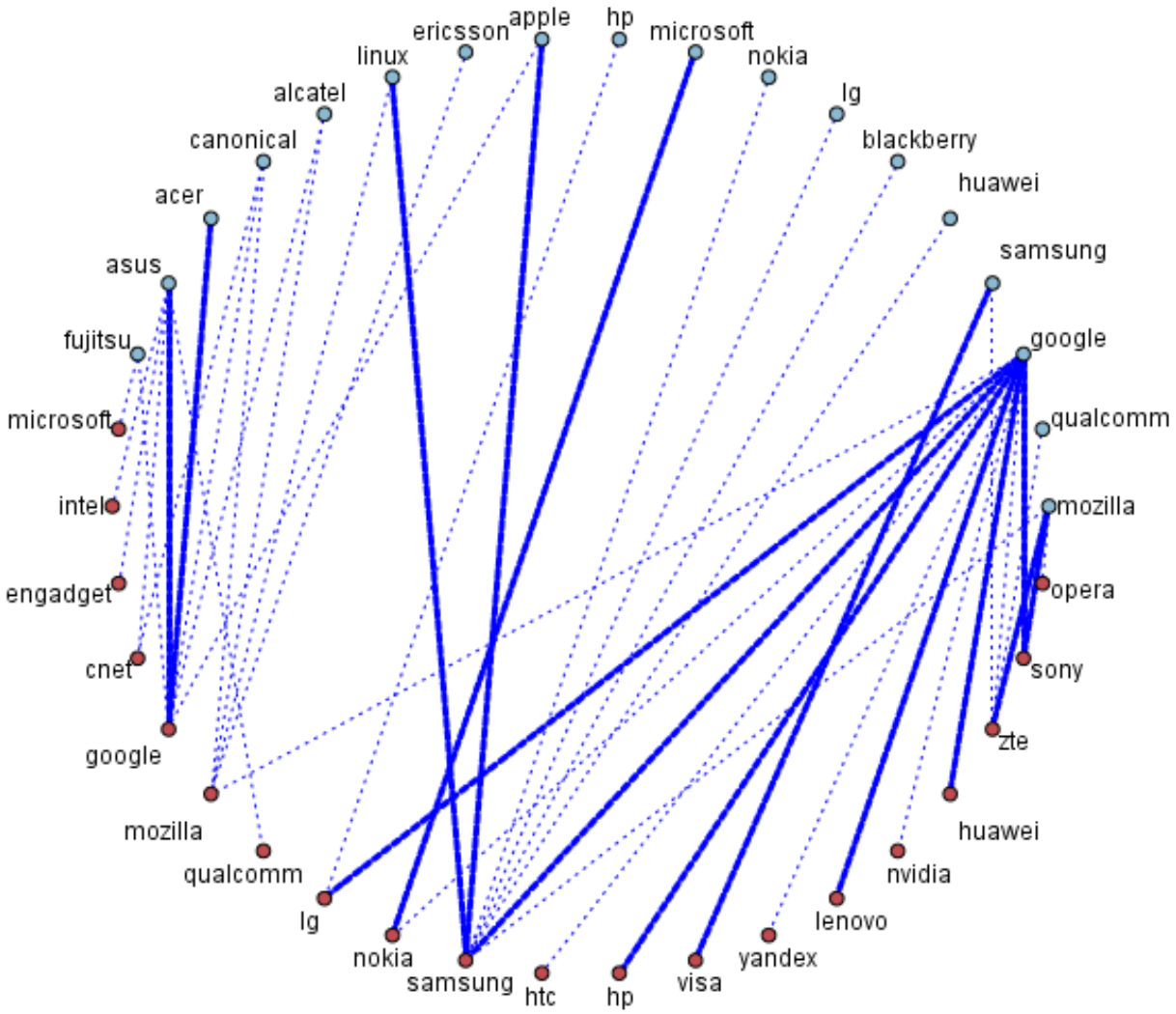
MOBILE WORLD CONGRESS
Barcelona | 25 – 28 February 2013

Main topics


tablet
smartphone
smartphones
phone
lumia
galaxy
xperia
padfone
optimus
fonepad
android
os
wireless
ascend
slate
grand
smartphon
ubuntu
nexus
youtube
tablets
wallet
nfc
dropbox
internet
homesync
tegra
atom
wp8
ipad
iphone
beat
network
cloud
snapdragon
lte
4g
passbook
digital
facebook
opera
wifi
glass
knox
3d
lineup
osaverg
enzaunrato
annonces
rivalizar
more
pidos
point
sentiert
tiene
stage
pameran
hot
muy
mobileworldcongress
european
partnership
stylish
price
spec
coffe
guy
program
like
toda
stellt
yota
kick
hang
nous
europa
mediaq
vadiez
17zrg
perhatian
50
nation
premios
s01
innovacl
sigue
podcast
februari
sim
aplicaciones
spot
stuff
custom
music
express
alljoyn
grande
estrellas
firefoxos
mostrados
a3000
payment
we
cu
light
webos
parti
els16
friend
preparados
report
zoll
paper
lire
marcos
readmisi
empresas
pouces
for
chip
nice
inteligentes
mo
coupon
tout
noticia
batteri
retweet
book
umumkan
rateurs
venu
mobileworldl
enterpris
digit
raj
che
integr
tec
htcone
curious
m310
bluetooth
van
een
marco
stori
tablettes
powermat
dalil
byod
mobiles
a1000
grow
q10
pero
curiosos
produk
nuestras
olhardigital
shown
smallest
k900
foto
yeni
s6000
cc
11ante
maker
menginlip
teamfollowback
stylistic
iconia
morn
processor
hope
minut
larg
sync
aggreg
masterpass
z10
nosotros
monday
pair
game
catalunya
s01
innovacl
sigue
podcast
sim
februari
opera
produtos
learn
tu
lu
ak
vel
entri
nation
premios
presentan
presentar
b1
sap
billion
apuesta
mark
mecerca
nostra
itun
celulares
vom
month
revela
talluri
catalonia
momentos
tela
fono
music
real
cpu
80
nuovi
estar
so
ng
hace
grande
mostrados
a3000
estrellas
wp8
alljoyn
express
ios
vgil
voz
of
pido
listos
hang
kick
yota
io
27
like
econ
program
salon
guy
price
spec
coffe
stylish
kaspersky
novit
ink
softwar
ces
entir
partnership
12
28
htt
mobileworldcongress
stage
pameran
muy
sentiert
tiene
tetra
digital
more
pidos
point
passbook
4g
enzaunrato
annonces
rivalizar
osaverg
lineup
sk
aggreg
teamfollowback
stylistic
iconia
morn
processor
hope
minut
larg
sync
masterpass
z10
nosotros
monday
pair
game
catalunya
s01
innovacl
sigue
podcast
sim
februari
opera
produtos
learn
tu
lu
ak
vel
entri
nation
premios
presentan
presentar
b1
sap
billion
apuesta
mark
mecerca
nostra
itun
celulares
vom
month
revela
talluri
catalonia
momentos
tela
fono
music
real
cpu
80
nuovi
estar
so
ng
spot
stuff
custom
aplicaciones
glass
wifi
nfc
wallet
3d
homesync
internet
dropbox

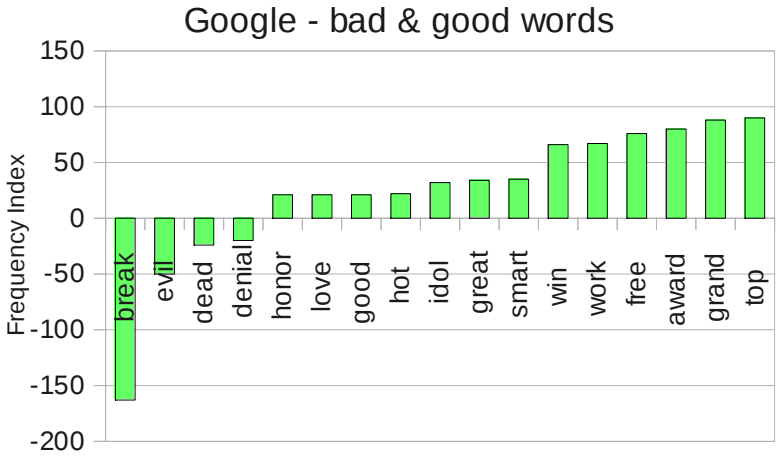
Pair of companies mentioned in a single comment

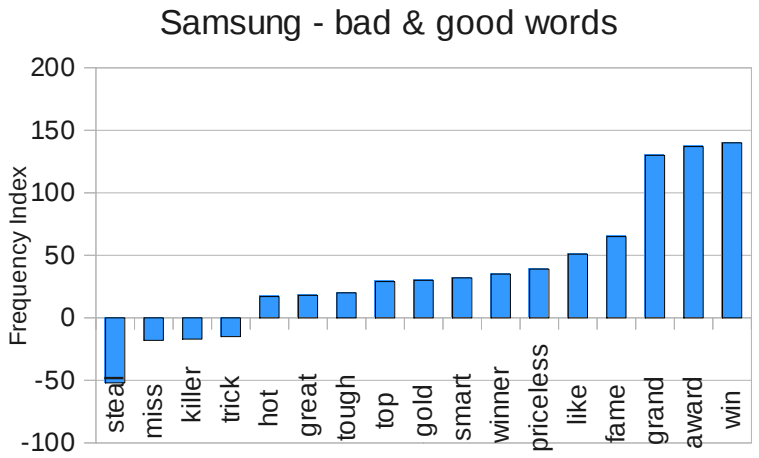
MOBILE WORLD CONGRESS
Barcelona | 25 – 28 February 2013

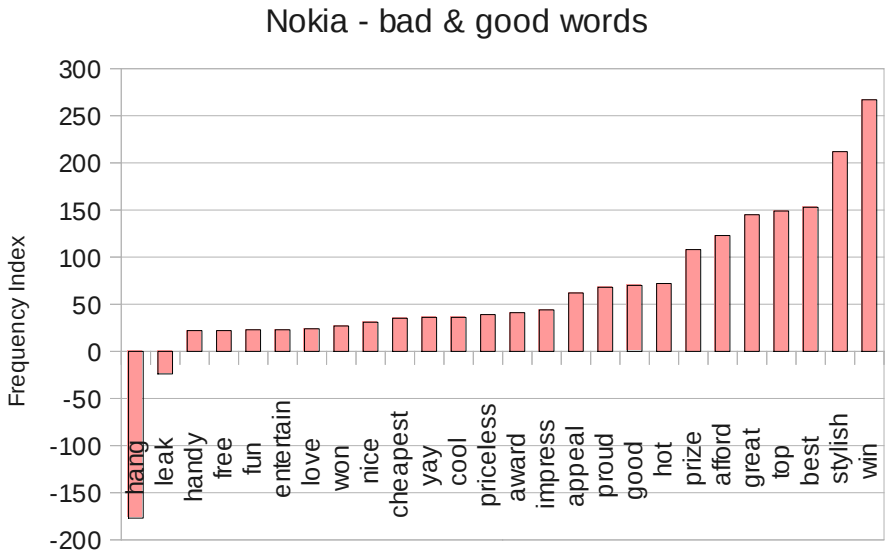


© 2013 IBM Corporation

MWC 2013 – what people say about the most popular companies





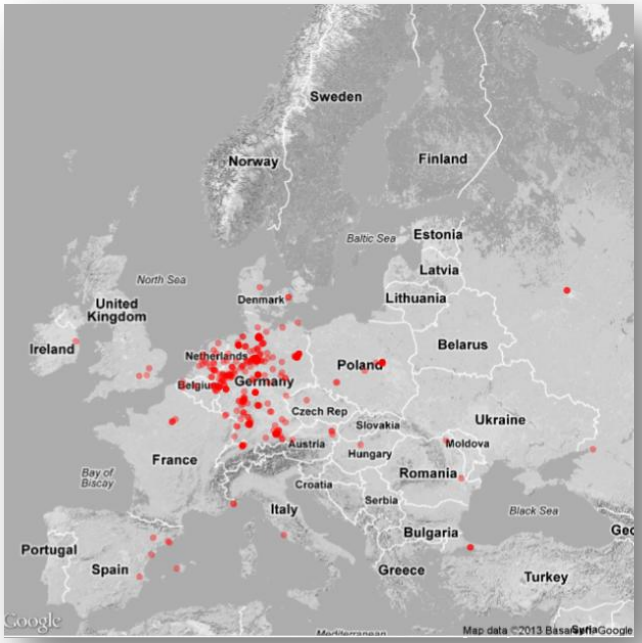


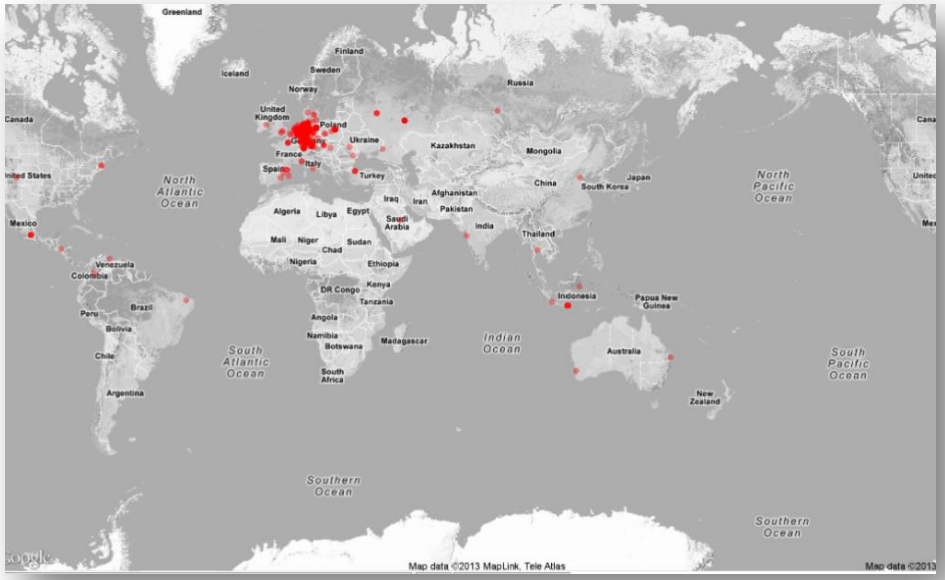
© 2013 IBM Corporation

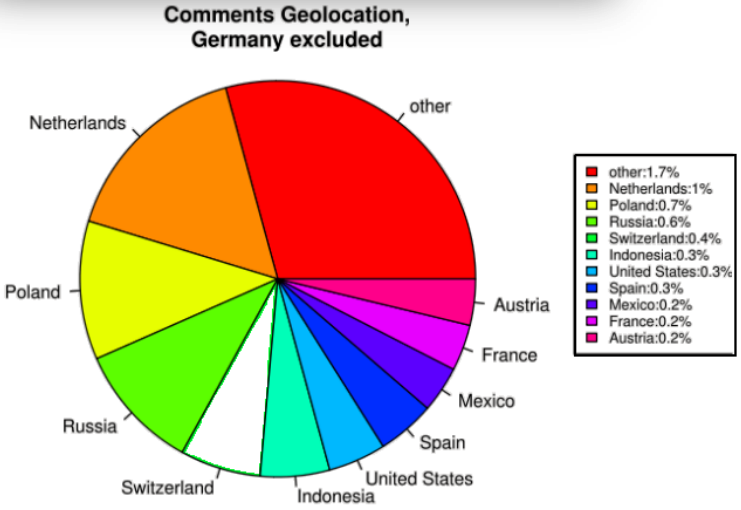
Most popular topics

© 2013 IBM Corporation

CeBIT 2013 – sources of comments

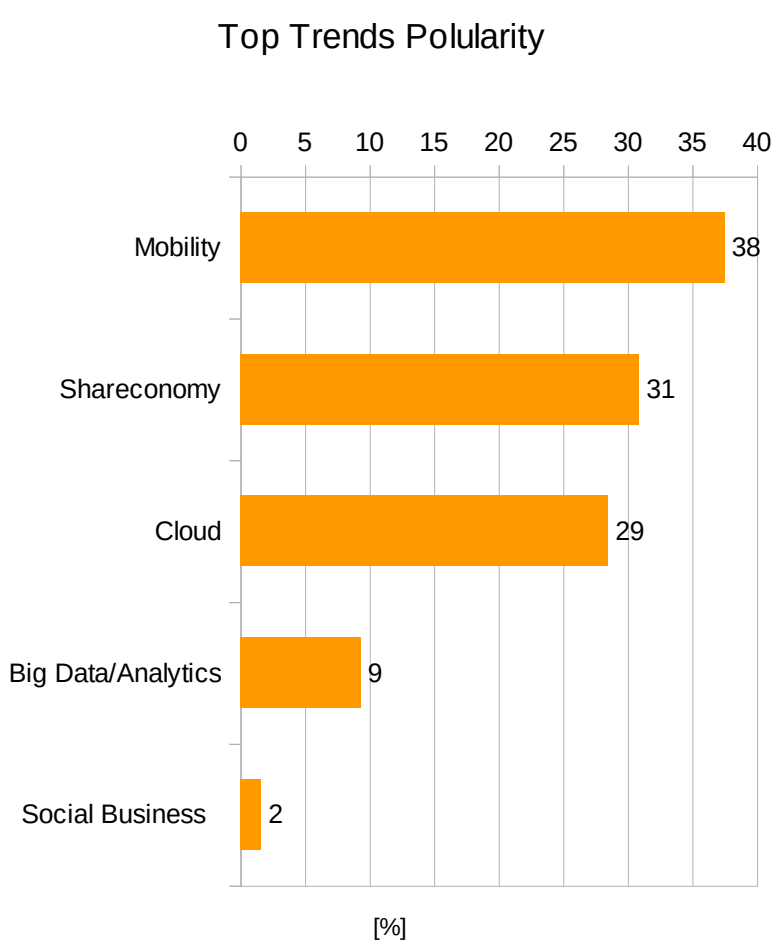


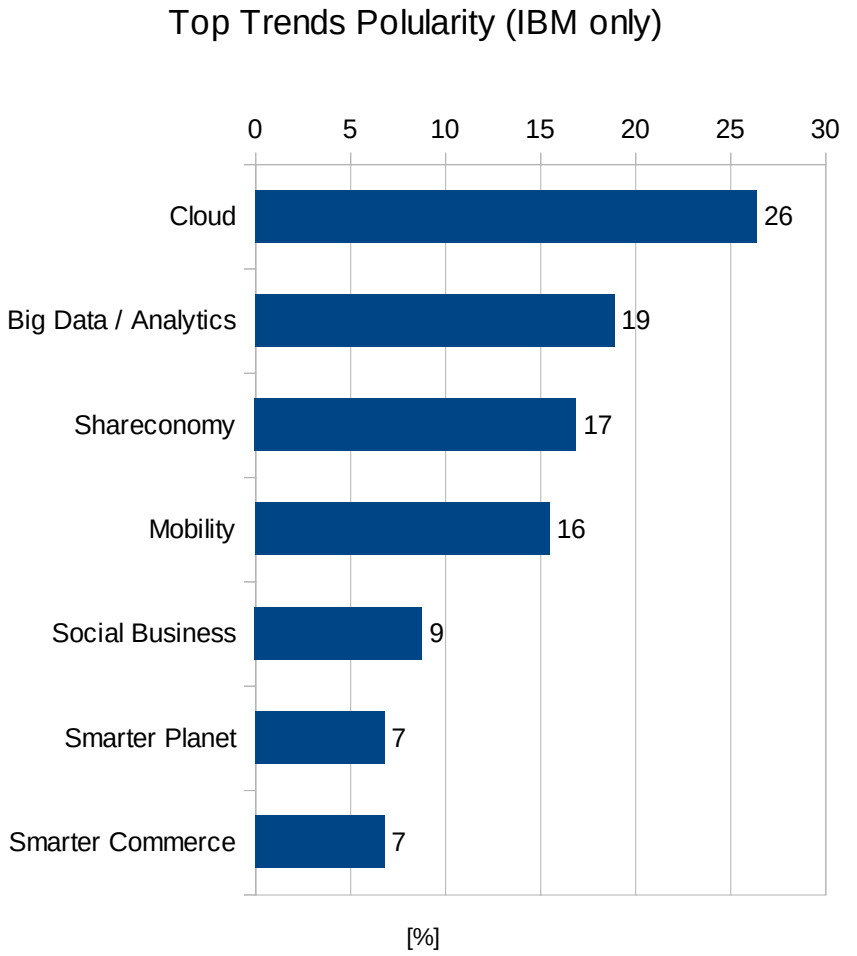




© 2013 IBM Corporation

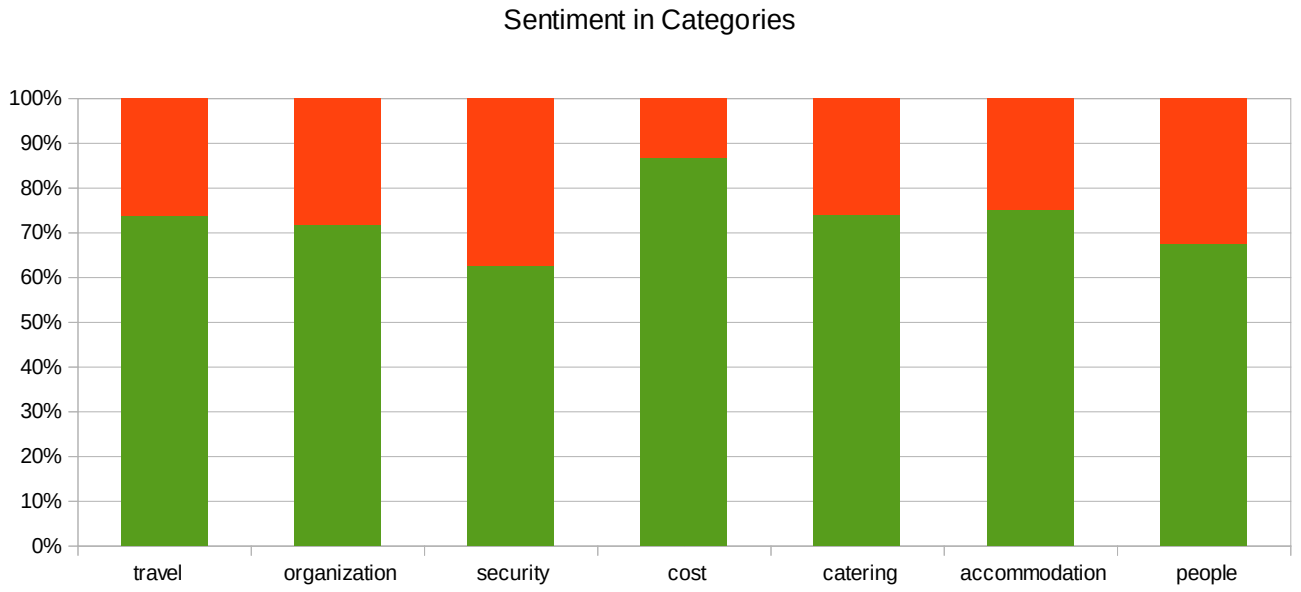
CeBIT 2013 – the popularity of top trends in technology

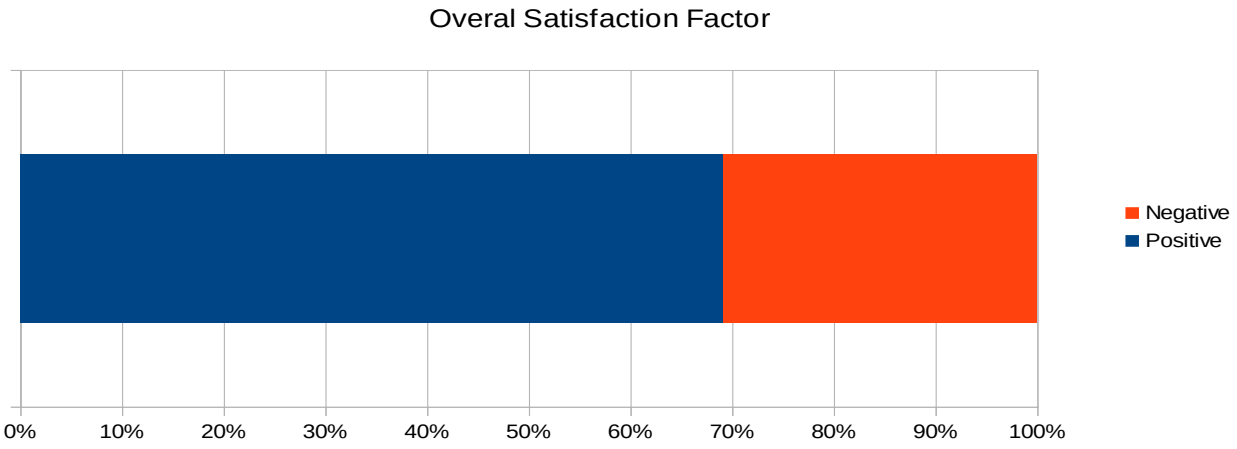




© 2013 IBM Corporation

CeBIT 2013 – sentiment in categories





© 2013 IBM Corporation

Sentiment analysis allows to...

- Measure customers' reactions after a product lunch
- Analyze what people hear and say about a company or a product
- Answer the questions: Which features of a product are mostly commented on? What are the pros and cons?
- Track the activities of the competition. What do they say and do to gain customers? What do people think about the competitors?
- Measure the success of advertising campaigns

© 2013 IBM Corporation

Thank you !

tomasz.stradomski@pl.ibm.com

© 2013 IBM Corporation