

#### Architecture of the Triposo travel guide

Douwe Osinga (@dosinga)



#### The Company



#### In Sydney

#### Financed in California

TRACE

**BUDY** 

#### Headquartered in Berlin

THEFT

.......

TREATOR DEPTERTS TO DESCRIPTION OF THE SECOND PROCESSION

ILUMAT.

Thursday, 17 October 13

AND IT FREEMAN



Leba

Thursday, 17 October 13

Tell Atlas

#### Quarterly Jamborees

SALIDA DE SOCORRO

# Last one on a bus in Spain

#### Algorithms are King

### our Guides No Human Editing in NO HUMA LOWED



#### The Product

#### **Our Mission: Build the best travel guide**





#### Mobile





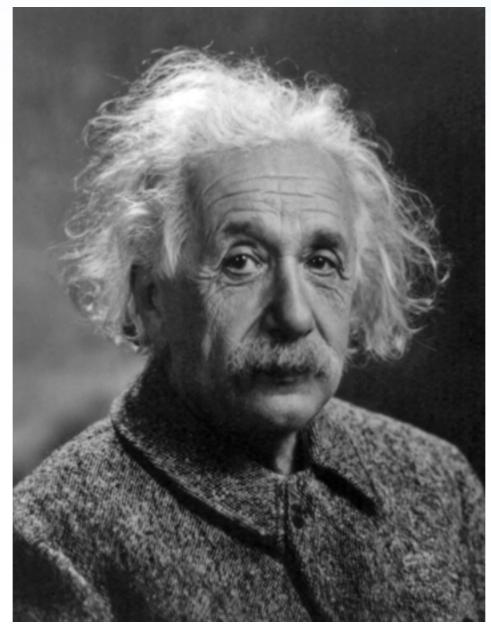
#### Works offline





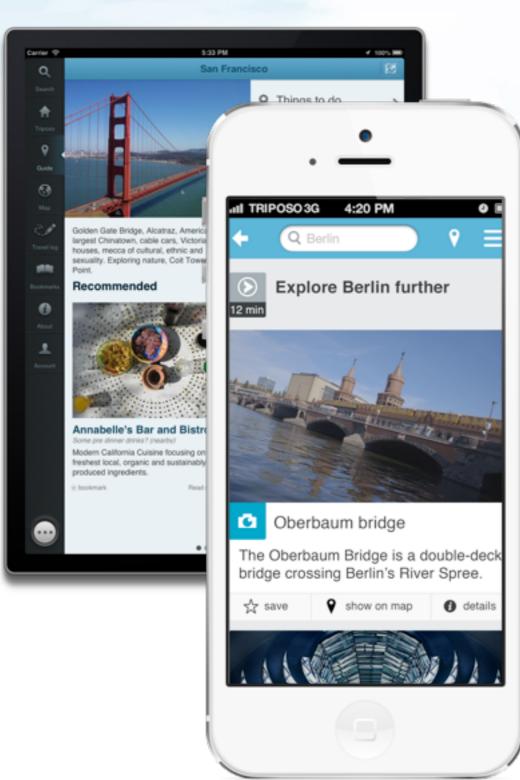
#### Covers the World













#### 25,000 destinations around the world

#### 500,000 points of interest

~5,000,000 downloads



#### The Platform

010101010010101¢ Physical PhysicaPhysicaPhysicaPhysicaPhysicaPhysicaPhysicaPhysicaPhysicaPhysicaPhysicaPhysicaPhysicaPhysicaPhysicaPhysic 

#### **Big data**

QC001010101010001010

(01010100101001010

(010010101010101010)

0)10101001010101010101

9)110101010101010101010

## **facebook** Companies with Big data amazon.com. Walmart > 5. GOQLE

# Do you have big data? LOCK SAMSUNG 6498

#### We didn't.

## Nice, but pricey amazon web services<sup>m</sup>

#### Our current server room

#### Coders are more expensive than Python code is slow

00001\$

mmmmm

000012 DJ 36397805 A

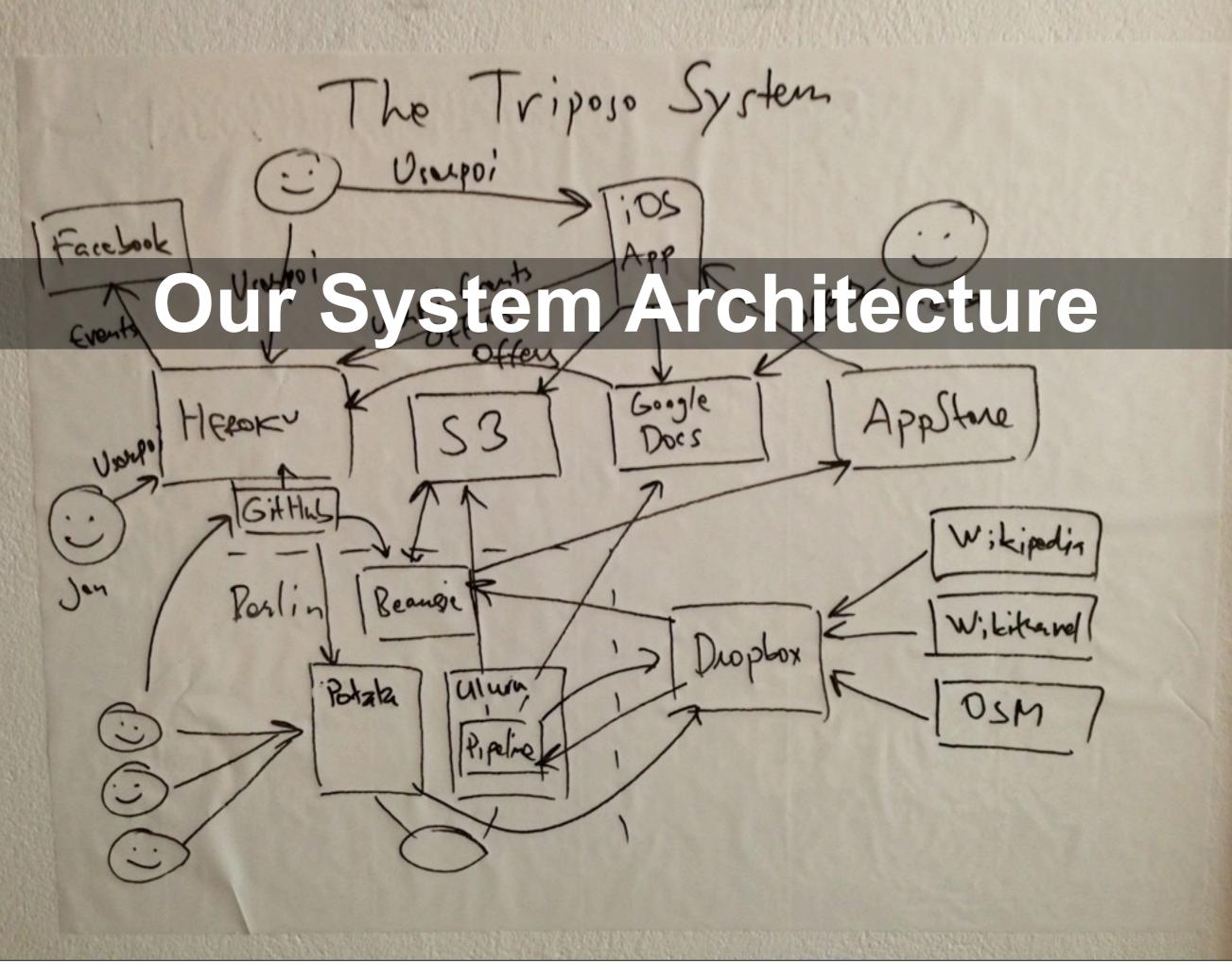
\$10000

\$10000

0000.

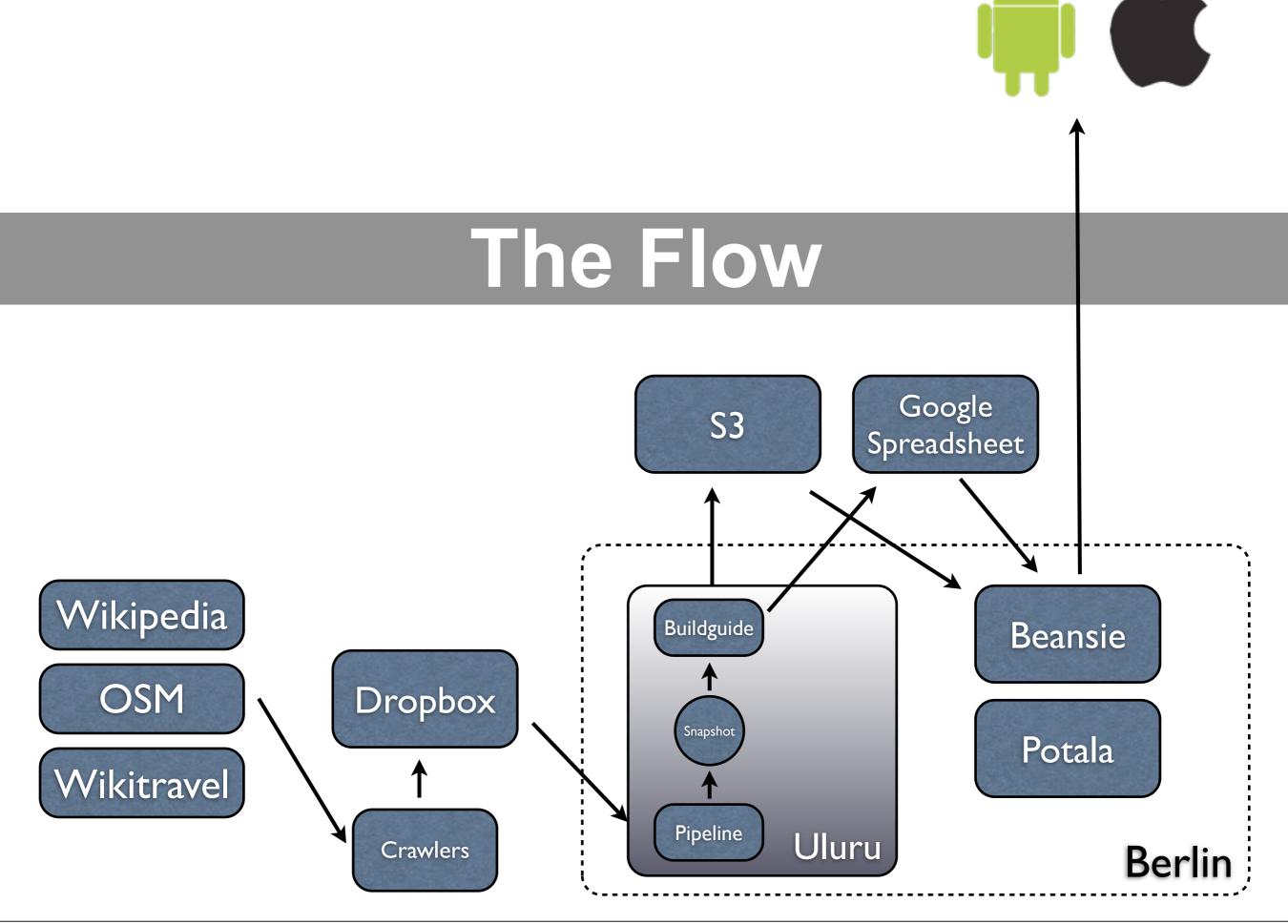
Thursday, 17 October 13

730





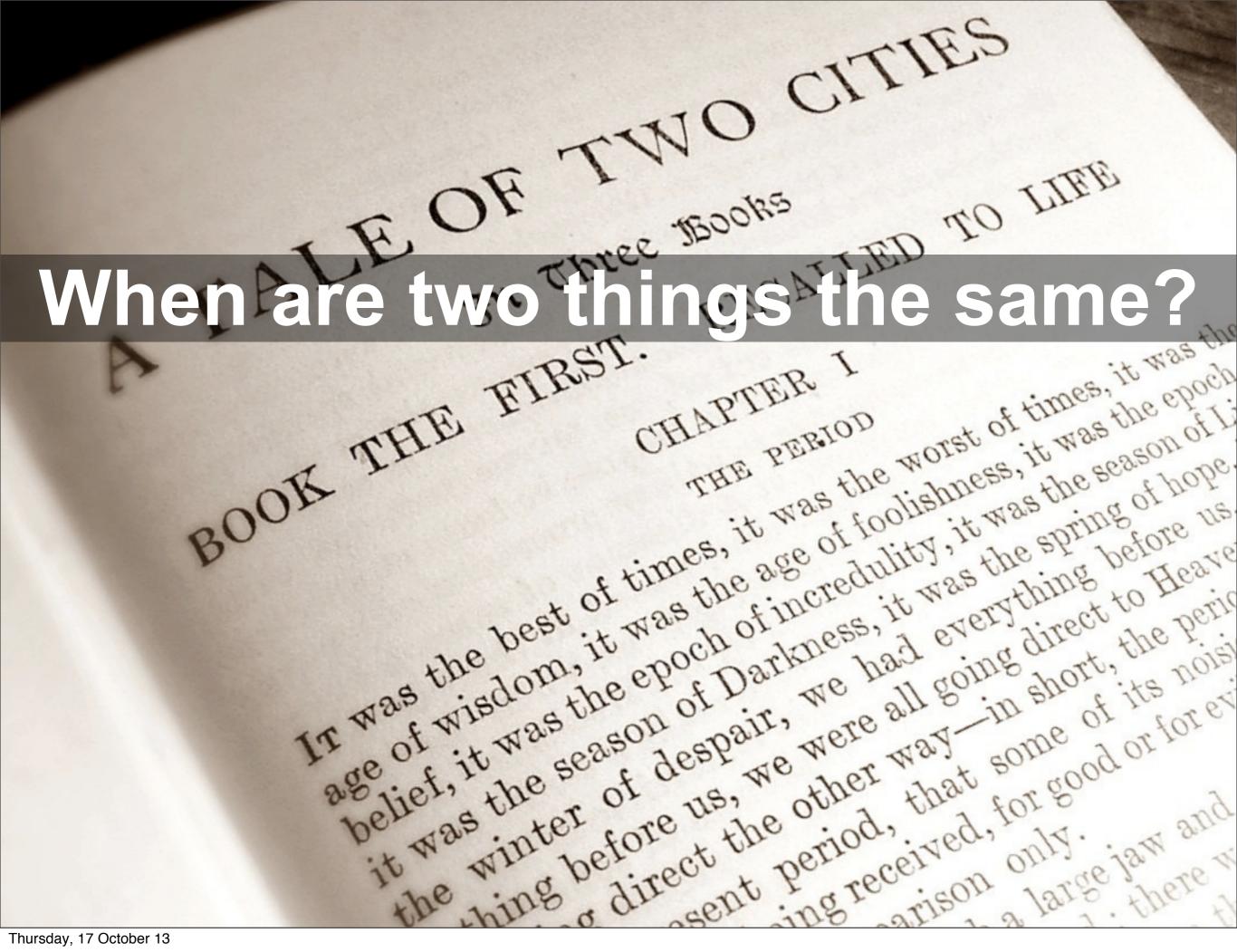
#### Building A database of the world



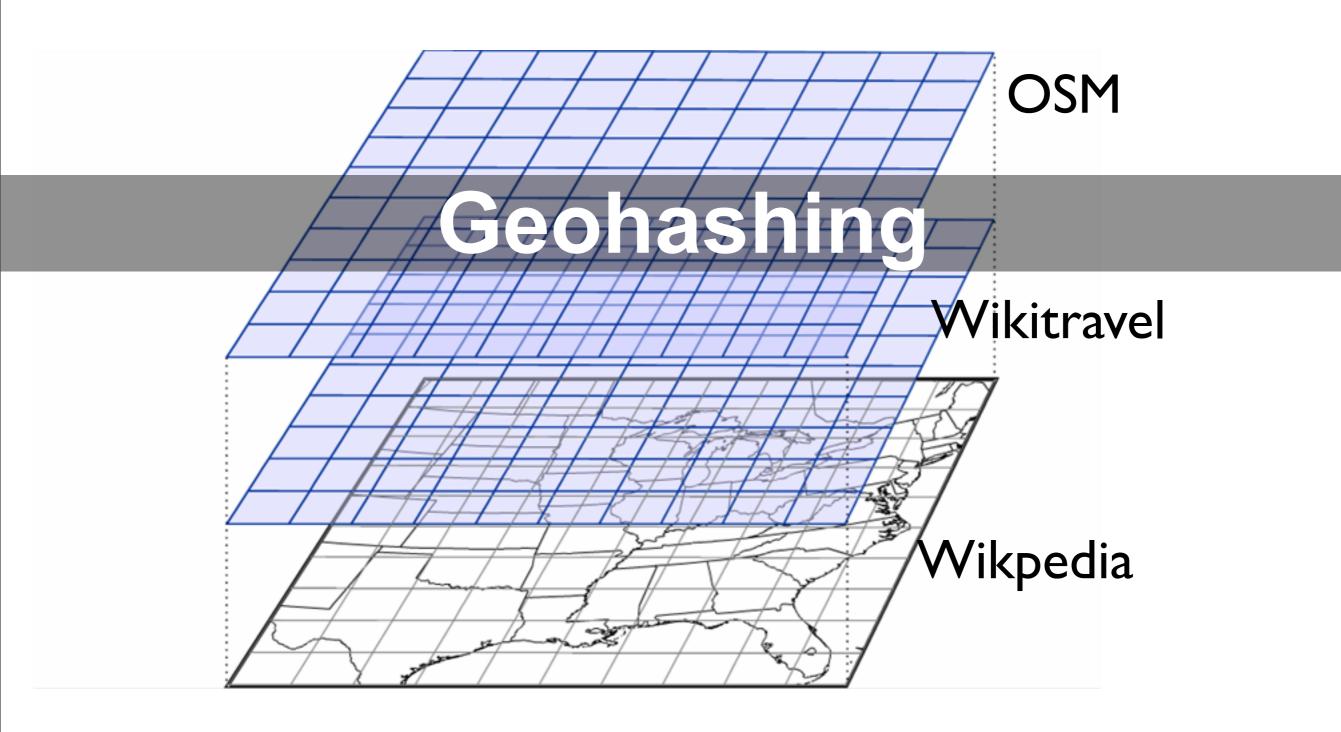
#### Crawl 20 Sources

#### Split everything in Pois & Locs

#### Put every thing back together



#### When they are: Similar in location and name



#### Suroit Camping vs Camping le Suroite

#### Shingling!

#### Suroit Camping vs Camping le Suroite

#### ampi, camp, itca, mpin, oitc, ping, roit, suro, tcam, uroi VS ampi, camp, esur, gles, ingl, lesu, mpin, ngle, oite, ping, roit, suro, uroi

#### 60% Overlap

# John's Bar and Grill vs Restaurant John



### Haarlem Library vs Haarlem City Hall



#### Van Gogh Museum vs Van Gogh Hotel

#### Types from names

Zapf, the most famous artist

The

she the whet termine with the field in the selection of the line in the selection of the se

### Cafe Sydney Opera House vs Sydney Opera House



# 

#### triposo

- 100 - 100									
	TRIPOSO	STORY	CONTACT	HOW WE DO IT	TEAM	LABS			
Home	Gues	Guess the cuisine							
PSI Destinations	obvious from is probably	n the name w	el ( uide , 1 / I a his on the web, s hat it is. "Mike's	eto in on iti ometimes we of t. B Burger Bar" is not an I	ut often when ndian Restau	we can't i urant, while	eyo tisp "Or		
Bookmarks	at four gram	ns has been ra	ther productive t	omatically do the gue o determine language check out Jon's blog	s, so we app		-		
				of a restaurant below a for scores > 3, there i		-			
Travel logs	cafe di mar	rio	Go						
i	1	Italian				8.04			
About	2	Pizza				2.31			
Thursday, 17 October 13	3	American				1.64			



# Making data useful





Thursday, 17 October 13

#### Create account 🔒 Log i

Q



The Free Encyclopedia

#### Mona Lisa

From Wikipedia, the free encyclopedia

	This solids is should be soliding. For other uses, see Mercy Mer		
Main page	This article is about the painting. For other uses, see Mona Lisa (disambiguation).		
Contents	The Mona Lisa (La Gioconda or La Joconde) is a half-length portrait of a woman by the Italian artist Leonardo da Vinci, which has be	en	Mona Lisa
Featured content	acclaimed as "the best known, the most visited, the most written about, the most sung about, the most parodied work of art in the	Italian	La Gioconda, French: La Joconde
Current events	world."[1]	1 1 10	
Random article			
Donate to Wikipedia	and being of hare being int it twein 107 ind 15 is a hour 12 mildo lay lave of hus woring in as at as a 17	4087	
<ul> <li>Interaction</li> </ul>	The monometry of the property of the French Republic, on permanent display at The Low remuseum	in Internet	
Help	Paris since 1797. <sup>[2]</sup> The ambiguity of the subject's expression, which is frequently described as enigmatic, <sup>[3]</sup> the monumentality of the	1000	LAP INT
About Wikipedia	composition, the subtle modeling of forms and the atmospheric illusionism were novel qualities that have contributed to the continuing		
Community portal	fascination and study of the work. <sup>[4]</sup>		
Recent changes			
Contact page	Contents [hide]	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
	1 Title and subject		
Toolbox	2 History	Sec.	
Print/export	2.1 Theft and vandalism	a second	TIME TAD
	3 Aesthetics		
- Languages 🔅	4 Conservation	11	
Afrikaans	4.1 Poplar panel	10	
ስማርኛ	4.2 Frame		
العربية	4.3 Cleaning and touch-up		Contraction of the second
Azərbaycanca याः ना	4.4 Display	a second second	A CONTRACT OF A
Беларуская	5 Fame		and the second
Беларуская	6 Speculation		
(тарашкевіца)	7 Legacy	Artist	Leonardo da Vinci
Български	8 See also 9 Notes	Year	c. 1503–1506, perhaps continuing until c. 1517
র্মন্থিশ	10 References	Туре	Oil on poplar
Bosanski	11 External links		s 77 cm x 53 cm (30 in x 21 in)
Brezhoneg	TT EXIGINAL INKS	Location	Musée du Louvre, Paris
* Català	Title and aubiast	Location	Musee du Louvre, Paris
Česky	Title and subject		

Main article: Lisa del Giocondo

The title of the painting that is known in English as Mona Lisa stems from a description by Renaissance art historian Giorgio Vasari, who wrote "Leonardo undertook to paint, for Francesco del Giocondo, the portrait of Mona Lisa, his wife.\*[5][6] Mona in Italian is a polite form of address originating as ma donna - similar to Ma'am, Madam, or my lady in English. This became madonna, and its contraction mona. The title of the painting, though traditionally spelled "Mona" (as used by Vasari<sup>[5]</sup>), is also commonly spelled in modern Italian as Monna Lisa, but this is rare in English.

Chavacano de

Zamboanga

Cymraeg

Dansk

Eesti

Deutsch

### Wikipedia article distribution

#### Learning from ratings?

### No Soup for you

#### **Opinion Mining**

### Natural Language Processing

#### erns(patterns): e(pattern, re.MULTILINE) for pattern in patterns] if

precompile\_patterns([r'^best\b', r'\bthe best\b', r' r'= (to die for|a must)\b', r'\b(perfect[oa]|delic r'\byumm+y\b', r'\byum+!'. r'\bmm+!', r'\bsuperb[c r'\ble[ck/eu, count-Keg <> (os) precompile\_patterns([r'^(must )?(eat|drink|try|orde r'(loved?|enjoye?d?) (a|the|some|)\s?THEME', r'^] r'^(a\s)?(safe|fine)\s(choice\s)?for THEME']) precompile\_patterns([r'^avoid[!\.\$]', r'^the worst\k r'\b(rude|rudest|sucks|shitty|shittiest|crappy|cra recompile\_patterns([r'\b(not?|keine?|without|ohne|se

```
ON = re.compile(r'\Bn\'t\b')
WESOME = re.compile(r'\b(perfe[ck]t|excellent|aweson
e(r'\b(is|was|were|are|ha[svd]e?\sbeen)\b')
G = re.compile(r'\b(really|truly|definitely|very|tota
FARTER = re.compile(r'^([au]nd|\+|also)\s*')
Fursday, 17 October 13
```



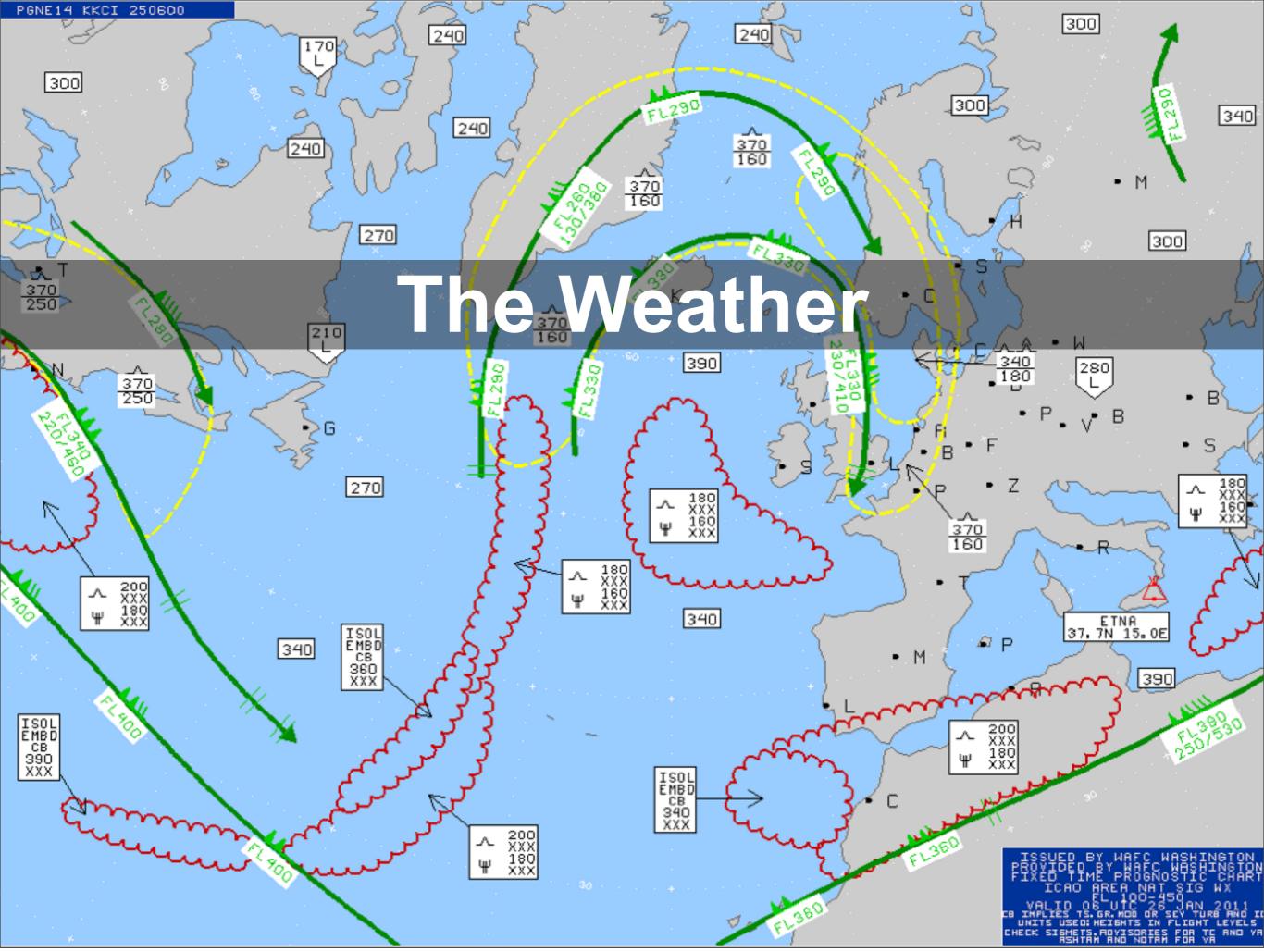
#### **Further experiments**



#### http://labs.triposo.com



# Can I suggest something?



#### YOU ARE HERE

#### Your location

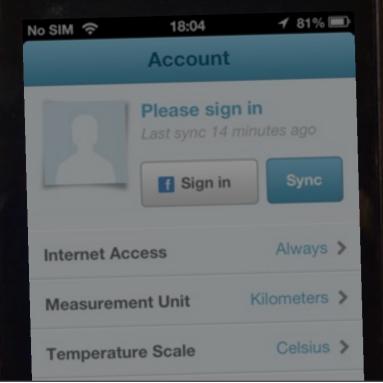




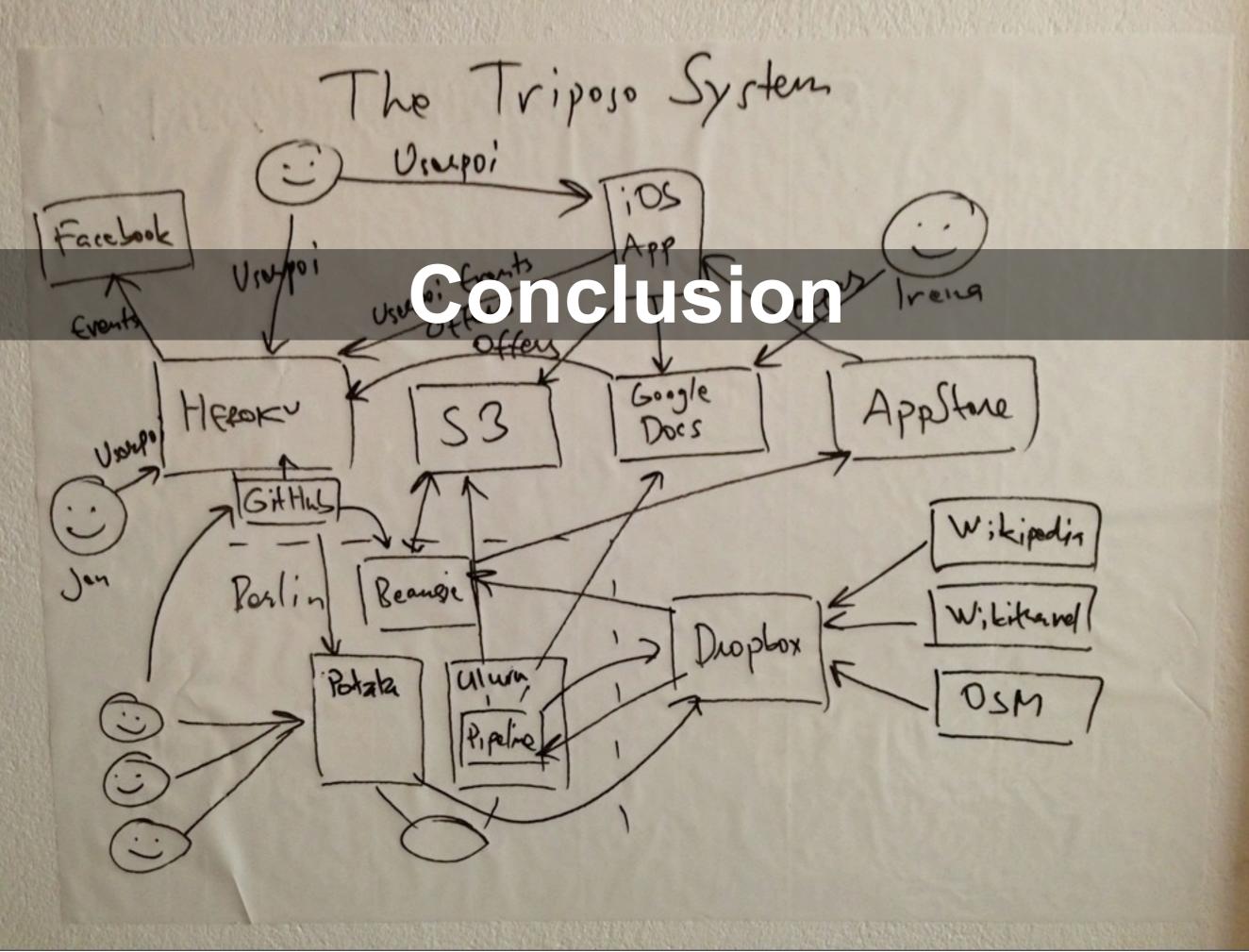
#### Personalization

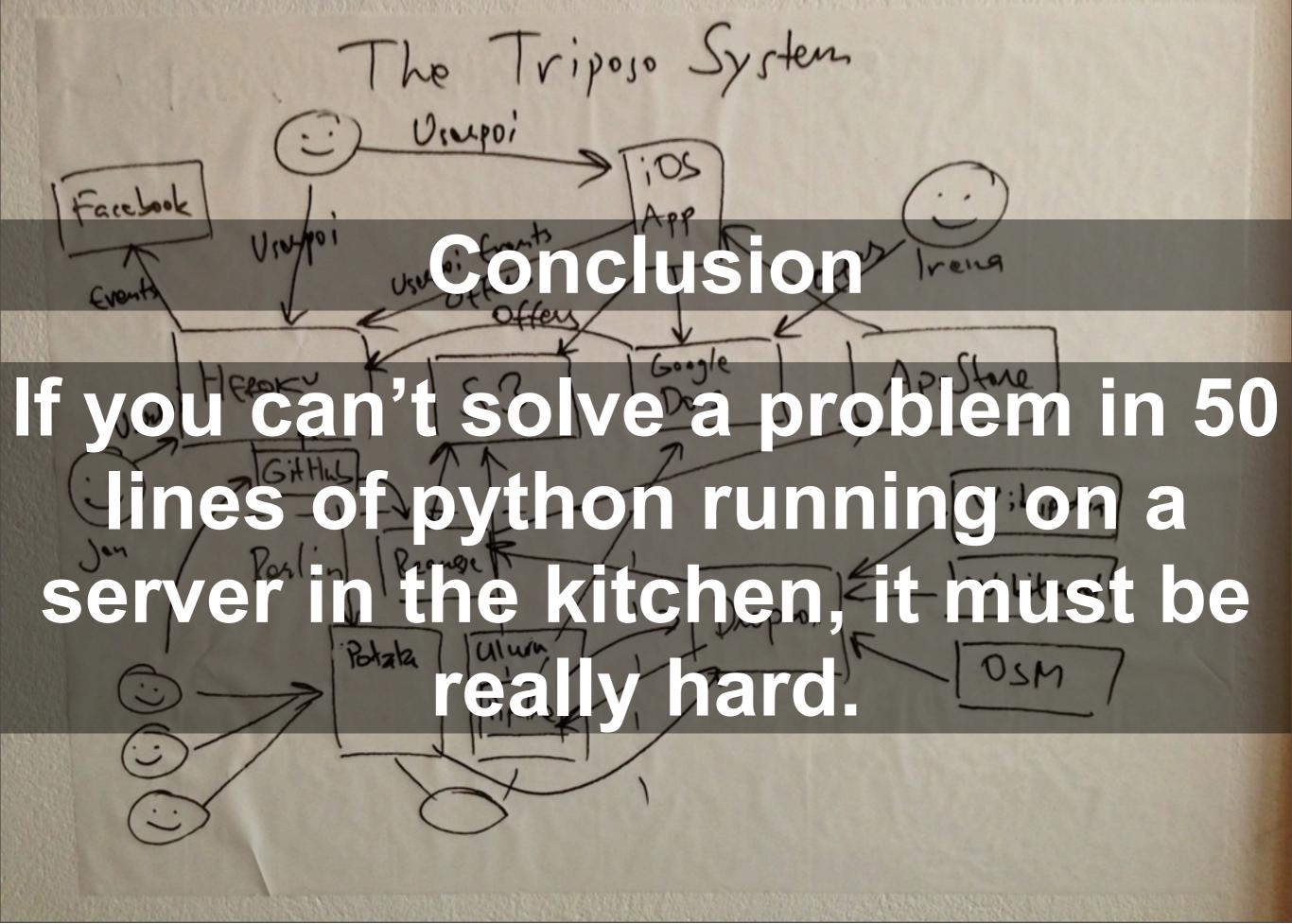
#### Personalization

#### Personalization











#### Thanks!

Douwe Osinga (@dosinga)