

Analyzing the Croatian Tweetosphere

Mislav Cimperšak, Damir Boras, PhD

Faculty of Humanities and Social
Sciences, University Zagreb



Summary

- Croatian tweetosphere
- Research
 - Twitter API
 - Collecting data
 - Falsely identified
 - Data deviations
- Website
- Further work

Croatian Tweetosphere

- Only opened accounts
- Assumption ~ 20 000 users
 - Based on Slovenian user count (~9400)

Research

- News
 - Formal vs. User created
- “Live infographics”

Twitter API

- API limitations
 - *Anonimously*: 150 requests per hour
 - *Authenticated*: 350 requests per hour
- Python_twitter

Collecting data

- Collected since April 2011
- Searching global public timeline for specific Croatian words
 - Trying to exclude similar languages
- Users collected hourly
- Search for new users run daily

Falsely identified

- Purging collected users
 - 1739 identified
 - 792 Croatian

Data deviations?

	Users	T / week	Avg / user
HR	792	52514	66.31
global	200 million	1 billion	5

	Less then 100 followers
HR	34%
global	97%

	Less then 100 friends
HR	37%
global	81%

Data deviations explained

- 41% of global users have not tweeted since they created account

	Users	T / week	Avg / user
HR	792	52514	66.31
Global	118 million*	1 billion	12.2

- The more you tweet...

Website

- <http://faust.ffzg.hr/tvitamin>

Further work

- Increasing user set
 - Identifying based on location
 - Some more obtrusive methods limited with API
 - Manual?
- Distinction between corporate and private users

Resources

- Twitter Statistics (March 2011.)
 - <http://blog.kissmetrics.com/twitter-statistics/>
- Twitter Statistics On Its 5th Anniversary (March 15, 2011)
 - <http://www.onlinemarketing-trends.com/2011/03/twitter-statistics-on-its-5th.html>
- Twitter Stats- Up and Away! (May 24, 2011)
 - <http://imadnaffa.blogspot.com/2011/05/twitter-stats-up-and-away.html>
- Twitter API
 - <http://apiwiki.twitter.com/>
- Python-twitter
 - <http://code.google.com/p/python-twitter/>