

**Catching Earworms  
on Twitter:  
Using Big Data to Study  
Involuntary Musical Imagery**



**Dr. Lassi A. Liikkanen**

**@lassial**

Independent scholar,

Design Researcher and Lead Analyst @SC5 Online Ltd. Finland

1.6.2015 Investigating the Music in our Heads symposium at Goldsmiths, London

# In this talk

- Short Case report
- A presentation of an upcoming publication

# CASE STUDY

3

**Goldsmiths**  
**1.6.2015**

@lassial

# Past

- Liikkanen L. & Raaska K. (2013)  
Treatment of anxiety from musical  
obsessions with a cognitive behavior  
therapy tool.  
*BMJ Case Studies.*  
DOI:10.1136/bcr-2013-201064

# **EARWORMS ON TWITTER**

5

**Goldsmiths**  
**1.6.2015**

@lassial

# Forthcoming work

- Liikkanen L, Jakubowski K. & Toivanen J. :  
Catching Earworms on Twitter.  
Accepted to Music Perception

# High level research questions

- Can open data from Twitter be used to inform us on music psychology?
- Can it reveal us something about earworms in particular?

# Big data and musicology

- Huron, D. (2013). On the virtuous and the vexatious in an age of big data. *Music Perception: An Interdisciplinary Journal*, 31(1), 4-9.



# Big data is ...

- **Big volume**
- **Big velocity**
- **Big variety**
  - Stonebaker, 2012

# This is Twitter

- Open social media
- Popular in Western countries
- > 300M MAU
- > 400M Tweets a day



# This is a Tweet

User

140  
character  
max.

Interactions



# The study -Methods

- Gathered all tweets related to earworms from six months
  - Nov 2012 - April 2013
- Create a computational, human-assisted analysis method for analysis
- Validate the computational approach

# Methods - detailed

- Data mining with regular expressions
- Sentiment analysis using bag-of-words approach
- No machine learning

# Preprocessing

1. Reformatting and filtering out unnecessary data fields from the tweets
2. Converting shortened hypertext links into final destinations URLs
3. Filtering for content: filtering out irrelevant data
4. Decoding user location data

# Automated classification

- 1) Random sampling of categorized and uncategorized data
- 2) Manual analysis of the samples
- 3) Rule creation, modification, and removal
- 4) Application of the new rule set

# DATA



# What was accepted?

- English language (set by Twitter)
- Matching keywords: “earworm, ear worm, earworms, stuck tune, haunting melodies”
- Not solely about DJ Earworm

**We captured ~100,00 tweets**

# RESULTS

# Results - overall

- **Twitter does allow extracting insights relevant to music psychology**
- **We did learn new things about earworms**

# Results – Main findings on INMI

- INMI is really ubiquitous
- People share mostly their own INMI experiences
- Experience are described emotionally rich and more negative way than other Twitter topic

# RESULTS IN DETAIL

# Users 56,626

# 173 regions ~ countries

## Appendix 1: List of contributing geographic regions (ISO 3166-1)

Location	No. of contribut.		
United States	21656	Cyprus	39
United Kingdom	17717	Chile	38
Unidentified location	8889	Turkey	38
Canada	4677	Taiwan	36
Australia	3140	Israel	35
India	1517	Greece	34
Ireland	762	Jamaica	34
Philippines	645	Zimbabwe	34
South Africa	639	Jersey	33
New Zealand	605	Peru	33
Mexico	579	Uganda	31
Singapore	564	Czech Republic	29
Indonesia	538	Lebanon	29
Germany	423	Denmark	28
Malaysia	281	Ghana	28
		South Korea	26



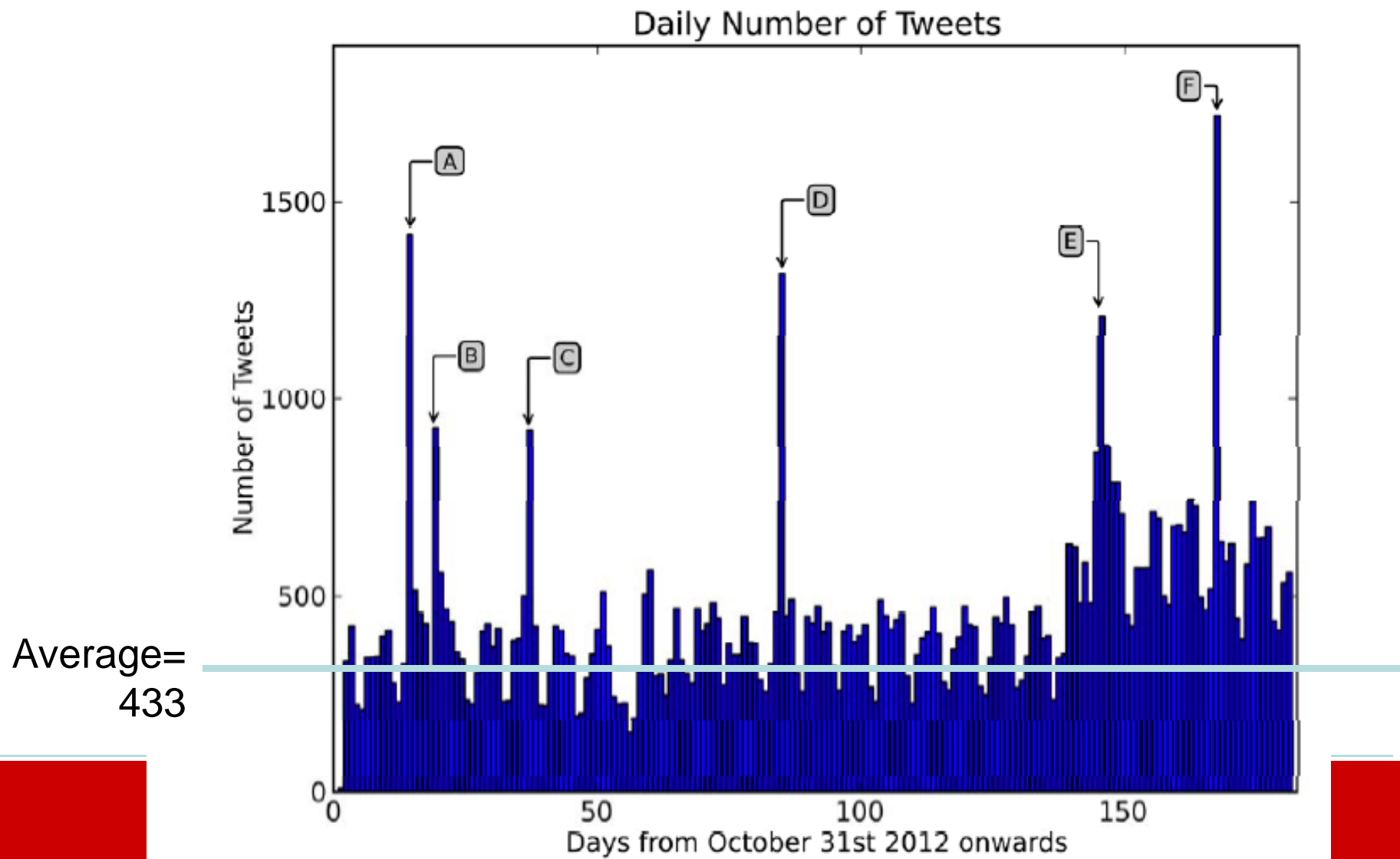
# Final sample

<b>Users</b>	<b>N</b>	<b>%</b>
Number of contributing users	56 626	
Number of locations reported	23 580	
Decoded locations	17 789	75.6%
Number of unique countries	173	

<b>Tweets</b>		
Number of tweets	80 620	
Number of retweets	3 074	
Number of hashtags	39 752	
Number of URL hyperlinks	16 422	
Number of unique hashtags	11 933	30.0%
Number of unique URLs	11 669	71.1%

# When



# Classification

# 51% of the data classified!

Category	Instances	%	User mentions	Included hyperlinks	False positives	False negatives
Unclassified	39 248	48.7%	53.9%	11.6%	-	-
G. My Earworms	22 531	28.0%	44.7%	18.8%	12%	11.5%
F. Music links	9 099	11.3%	36.4%	99.9%	0%	0.0%
D. Earworms information	5 187	6.4%	91.6%	8.8%	2%	3.5%
E. Grateful attribution	4 711	5.8%	5.8%	10.3%	6%	0.0%
H. Viral worms	2 700	3.4%	82.6%	34.7%	2%	0.0%
B. Cure and riddance	2 610	3.2%	51.3%	46.5%	2%	1.5%
C. Earworm nominations	874	1.1%	60.6%	10.1%	20%	2.0%
A. Causal attributions	752	0.9%	1.3%	20.7%	4%	3.5%
Classified M			46.8%	31.2%	6%	3%

# 1. My earworms

28%

# Example

- G. My Earworms
- I have Some Nights stuck in my head but I only know the tune to the chorus so I've been singing it as a weird voodoo African chant instead
  - From **@sodalite** on Sat Mar 30 17:30:29 +0000 2013

## 2. Music links 11.3%

YOUTUBE VIDEO ID	YOUTUBE VIDEO TITLE
IJNR2EpS0jw	Dumb Ways to Die
6q0dsG8fTHY	DJ Earworm Mashup - United State of Pop 2012 (Shine Brighter)
QK8mJJvaes	MACKLEMORE & RYAN LEWIS - THRIFT SHOP FEAT. WANZ (OFFICIAL VIDEO)
NarbqFTrXbE	For the Win - Team Unicorn (Featuring Weird Al Yankovic, Aisha Tyler, Grant Imahara)
8N_tupPBtWQ	Muppet Show - Mahna Mahna...m HD 720p bacco... Original!
5_sfnQDr1-o	Baby Monkey (Going Backwards On A Pig) - Parry Gripp
Z0GFRcFm-aY	R.E.M. - It's The End Of The World
jnlRdu4Wvuk	Lord Sij – Shuup
zvCBSSwgtg4	The Lumineers - Ho Hey (Official Video)
A1VTaAYH3Hc	Badtameez Dil - Full Song - Yeh Jawaani Hai Deewani   Ranbir Kapoor, Deepika Padukone
If5MF4wm1T8	Pop Danthology 2012 - Mashup of 50+ Pop Songs
kfVsFOSbJY0	Friday - Rebecca Black - Official Music Video

# 3. Earworms information

- D. Earworms information: Tweet 8.
- 11.3% of all Tweets
- An "Earworm" describes a song, hook and or chorus that you can't get out of your head.
  - From **@always\_known** on Mon Apr 01 20:32:26 +0000 2013

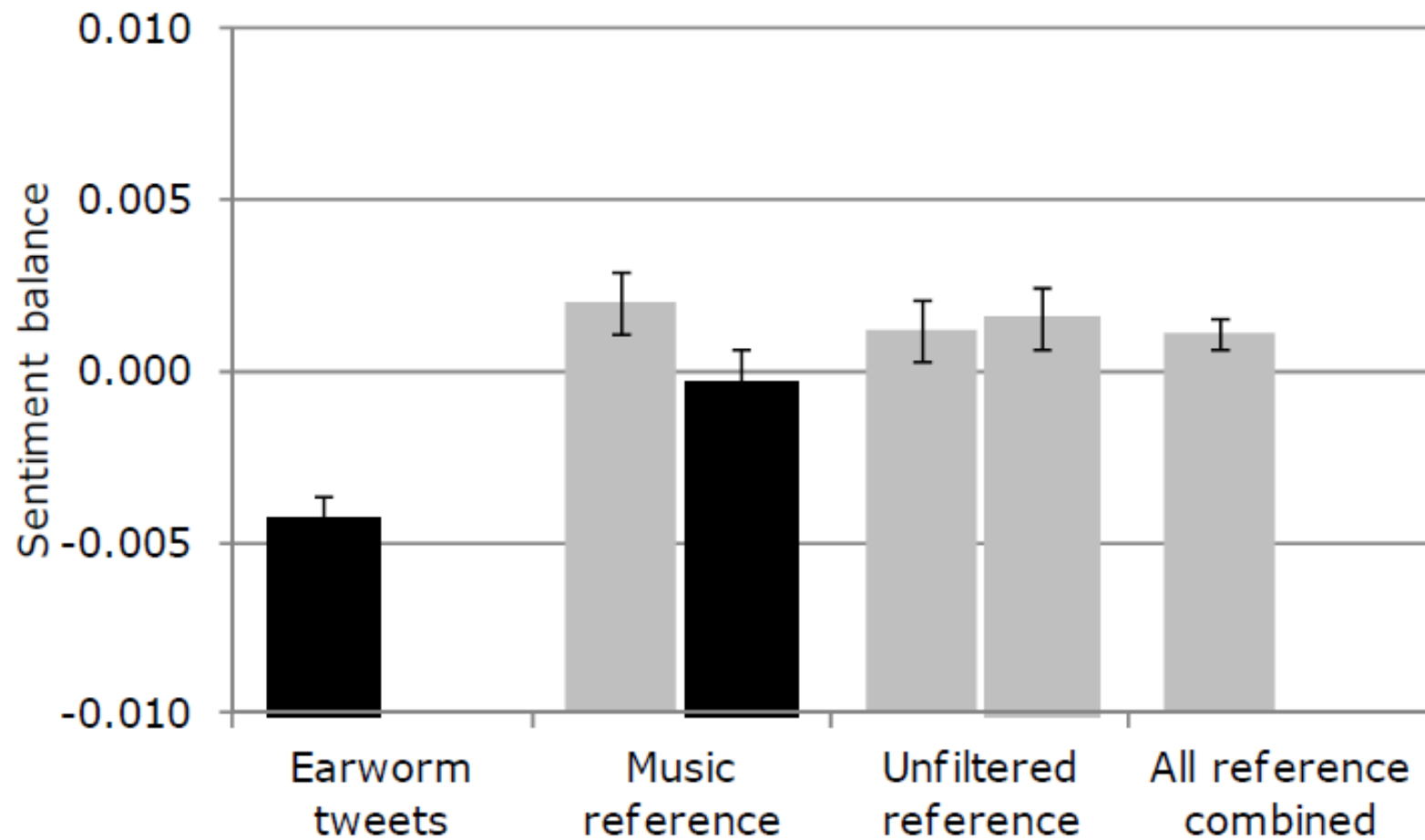


# Sentiment analysis

# On the method

- Comparing earworm-related tweets against non-selected or 'music' related

# Sentiment analysi



# Conclusions and Discussion

# Conclusions

- People share mostly their own INMI experiences
- Experience are described emotionally rich and more negatively phrased than other Twitter discussions

# Conclusions and Discussion

*earworm discussions are somewhat lonesome cries in the micro blogosphere (i.e., Twitter), with infrequent echoes from other users.*

# Conclusions and Discussion

- There is data!
- Is this approach ethical?

From ' Ultimate earworms collection' YouTube playlist

<http://goo.gl/noNpRR>

<https://www.youtube.com/playlist?list=PLFF185AyoHpojr7fZB8d0GwsOlyZAXjuG>

# THANK YOU!

[http://iki.fi/lassial/articles/psychology/World\\_of\\_INMI\\_research](http://iki.fi/lassial/articles/psychology/World_of_INMI_research)