The WorldBot 2014

The #Good, the #Bad and the #Ugly

Tommaso Elli ~ tommaso.elli@gmail.com
Carlo De Gaetano ~ carlo.dgtn@gmail.com
Christoph Lutz ~ christoph.lutz@unisg.ch
Myrthe Bil ~ myrthebil@gmail.com
Iris Beerepoot ~ imbeerepoot@gmail.com
Claudio Coletta ~ claudio.coletta@gmail.com
World Cup 2014 breaks Super Bowl record on Twitter
Introduction

➔ **WorldCup 2014**: event with huge Twitter activity

➔ Much of the Twitter stream is structured along **#hashtags**

➔ **Bots** are represented with interesting possibilities of outreach and attention
JUNE 30, 2014, 6:18 P.M.

FIFA has cracked down on that World Cup Twitter bot (and other people tweeting and Vining and GIFing goals) 💪

Last week, we wrote about @ReplayLastGoal, a Twitter account that automatically tweets out a GIF and a video of each goal scored in the World Cup. Great idea! Interesting execution! But also a copyright violation, at least according to FIFA, the governing body that runs the World Cup.

Early Saturday morning, @ReplayLastGoal tweeted a screenshot of an email from Twitter informing Xavier Damman, the account’s creator, that FIFA had filed a request asking Twitter to remove the account for copyright violation.

Hello,

Please be aware that we have received a copyright takedown request seeking the removal of your account. A copy of the notice follows this message.

Pursuant to our policies, Twitter is seeking clarification regarding the request. The notice will be suspended pending our review.

Fuego: Get up-to-the-moment news and see what the future-of-news crowd is talking about and linking to.

Encylo: Our encyclopedia of the future of news. We've got all the most important players in journalism's evolution.

Download the Lab's iPhone app — it's the best way to stay up-to-date on the future of news. It's free and available now in the App Store.
Research Questions

➔ Can we **categorize** the bots surrounding the World Cup?

➔ Can we identify any **behaviour pattern** in the accounts we identify as bots?

➔ (Do different games’ **nationality** attract bots in different ways?)
Methods of Data Collection

→ 1% of Twitter sample in TCAT

→ All teams represented at least once
Methods for Bot Detection

➔ **User account name** and **user description** analysis

➔ **Device/source** analysis

➔ **FF ratio** (Zhu et al., 2012)

➔ **User Co-Hashtag** Analysis

➔ Close Reading
Results
Bot Detection (pt. 1)

➔ User account name and user description detected bots

➔ Sample of ~2500 supposed users
Bot Detection (pt. 1)
Bot Detection (pt. 1)
~ BangBang ~
@94JackyVitry

→ #Team94 → #TeamPSG → Kik : JackyVaan → Snap : jaacky944 → Skype : jacky-v94

Vitry Sur Seine ❤️ 94 ! ❤️

~ BangBang ~ @94JackyVitry · 3 h
Today stats: 521 followers, 208 unfollowers and followed 598 people via Unfollowers.com/?s=dt

Dettagli

Vai al profilo completo
@94JackyVitry are you a bot?

15:10 - 3 lug 2014
@94JackyVitry are you a bot?

- BangBang

15:10 - 3 lug 2014

Rispondi a @94JackyVitry

iosonosempreio @iosonosempreio · 21 h
@94jackyvitry I would say YES YOU ARE!

Dettagli
~ BangBang ~ ti ha ritwittato
21 h: @94jackyvitry I would say YES YOU ARE!

Altri 3 Retweet
Bot Detection (pt. 1)

➔ If you’re out of the mass, you deserve some attention

➔ If you have thousands of friends/followers, but just few tweets, you’re probably a bot

➔ It’s interesting to analyze the friends/followers ratio
Bot Detection (pt. 2)

➔ Users who declare to be a bot

➔ X Axis: Friends/Followers Proportion

➔ Y Axis: Total number of used #tags
We consider declared bots as good bots.
Their metrics put them in the left-bottom corner.
Is that the good bots corner?
The #Good

#CDM2014
@CDM2014_bot
Je suis un bot qui retweet tous les tweets qui parlent de la Coupe du Monde 2014. #CDM2014
#CoupeDuMonde
📍 Brésil

Angel HaxOr
@Hax0rSon
Yes, I’m a freaking bot. Let’s win an Oculus Rift.
📍 I’m a bot
🔗 angelhack.com
Detecting the #Bads?

Users in these clusters have **normal FF ratios** and **don't carry special bot attributes** in their bios or usernames… **still they are bots.**
The #Bad
Detecting the #Ugly?

- The “ugly” works as a residual category, often half “human” half “non-human” (Cyborg)

- They engage followers through specific hashtags (i.e. “potato”) and tend to have more followers than friends

- Their content involves the emotional sphere

- Seems very promising for further qualitative analysis
The #Ugly

RT @orsolawa: Oh c'est chaud patate dès le début !!! #ESPCHI @
Bot Detection (pt. 2)

The diagram shows a scatter plot with the number of tags (#tags) on the y-axis and the Fr/Fo ratio on the x-axis. The data points are color-coded to represent different types of bots:

- Good Bots (turquoise circles)
- Ugly Bots (yellow circles)
- Bad Bots (red circles)

The plot illustrates the distribution and clustering of these bot types based on their characteristics.
Bot Detection (pt. 2)

➔ Maybe yes, we have a good corner.

➔ Let’s run some bigger analysis, within our ~2500 sample!
Conclusion
(and have a happy Comic Sans Day!)

- World cup tweeting is not very botted
- World cup bots are very diverse, it's not all #good and #bad
- Large #ugly category in the middle that's hard to categorize and make sense of
- Boundaries between #ugly and #good are blurry and more difficult to draw than between #good and #bad