Measuring Twitter Chat Participant Engagement: #LiveFitNOLA Example

Appendix 1: How to visualize and measure engagement level for hashtag-based Twitter conversations

Kristina M. Rabarison, DrPH, MS
Merriah A. Croston, MPH
Naomi K. Englar, BA
Connie L. Bish, PhD, MPH
Shelbi M. Flynn, BA
Carolyn C. Johnson, PhD, FAAHB
**Analytical Process**

- Data source: Full 75-min transcript (744 tweets)
- Mapped participant engagement
  - 135 Twitter users (66 participants + 69 mentioned usernames)
  - 474 mentions (engagements between Twitter users)

**Diagram:**

- Symplur
- Open Refine
- Gephi
SYMPLUR
Register a Hashtag and Collect Twitter chat Transcript
How to submit a hashtag with Symplur:

Submit your hashtag The Healthcare Hashtag Project

Authenticate
Step 1

Select Hashtag Type
Step 2

Submit Information
Step 3

Get Confirmation
Step 4

You need to Sign in with Twitter before you can submit

Your hashtag should:

- be healthcare related
- be of value to the healthcare community
- be of a certain usage level by multiple people
- be unique enough that its intended conversation isn’t drowned out by other uses of the same hashtag
How to collect hashtag-based Twitter two-way communication transcript from Symplur:

- Search for the hashtag of interest in Symplur (e.g. #LiveFitNOLA)

Why the Healthcare Hashtag Project?

Discover Where The Healthcare Conversations Are Taking Place
Discover Who To Follow Within Your Specialty Or Disease
Discover What Healthcare Topics Are Trending In Real-Time

The Community by the Numbers

1,095,886,540 Tweets
15,360 Topics
8,008 Hashtags
3,086 Contributors
### #livefitnola Hashtags

<table>
<thead>
<tr>
<th>Hashtag</th>
<th>Type</th>
<th>Topics and Description</th>
</tr>
</thead>
</table>
| #LiveFitNOLA | Healthcare Tweet Chats | chronic disease, community health, exercise, new orleans, nutrition, Obesity, physical activity  
#LiveFitNOLA Twitter Chat is hosted by Fit NOLA and Tulane PRC on the 1st Thursday of every month 12-1pm CST. We'll talk about health & wellness topics relevant to New Orleans with a new guest host each month that can add their lens of insight to the selected theme.... |
Open hashtag page

#LiveFitNOLA Tweet Chat

What is #LiveFitNOLA?

Healthcare Topics
- chronic disease
- community health
- exercise
- new orleans
- nutrition
- Obesity
- physical activity

Related Hashtags
- #COS15
- #NHConference
- #icn2013
- #YWM2014
- #MOmortality
- #DiabetesInterview
- #childobesitychat
- #PowerofToday

- Last chat: Thursday 3rd September 10:00 AM PDT
- Next chat: Thursday 1st October 10:00 AM PDT

#LiveFitNOLA is a healthcare tweet chat hashtag submitted by @FitNOLA

Scroll down to the bottom of the page to set date and time

#LiveFitNOLA Twitter Transcript and Analytics

Start: 03/05/2015 10:00 AM
End: 03/05/2015 11:15 AM

Get Transcript

Time Zone: All times are Pacific Time/San Francisco -0700 GMT. Convert
Click on “Get Transcript” to load the transcript within the specifically defined timeline (Example transcript page: www.tinyurl.com/LiveFitNOLAMarch52015)

Two options to collect data and time specific transcript:
- Copy and paste the hashtag transcript from Symplur into an Excel file, saved as a .csv file.
- Download R from www.r-project.org/. Then, use and edit the R code provided in Appendix B.
Open Refine
Prepare dataset for network visualization and analysis
How to format transcript data with OpenRefine:

- Download and install OpenRefine ([www.openrefine.org/download.html](http://www.openrefine.org/download.html))

**Download OpenRefine**

You will find on this page a list of OpenRefine distributions and extensions available for download. Are we missing something? Want to fix a typo? You can submit changes (pull request) from here.

**Official Distribution**

Read the installation instructions

**OpenRefine 2.6**

This is the first beta release of OpenRefine 2.6 on Aug 27, 2013. A change log is provided on the release page.

- **Windows kit**, Download, unzip, and double-click on `google-refine.exe`. If you’re having issues with the above, try double-clicking on `refine.bat` instead.
- **Mac kit**, Download, open, drag icon into the Applications folder and double click on it.
- **Linux kit**, Download, extract, then type `./refine` to start.
Launch OpenRefine and create a new project

Choose files from your computer and click on “Create Project” button
- Remove the first column, named “Column” and contains numbers

---

**744 rows**

<table>
<thead>
<tr>
<th>Show as:</th>
<th>rows records</th>
<th>Show: 5 10 25 50 rows</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Column</th>
<th>source</th>
<th>target</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td>Facet</td>
<td>Welcome to the 1st #LiveFitNOLA chat! Introduce yourself &amp; let us know who's here! Guest host @Healthfitmag is w/ us.</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td>Text filter</td>
<td>RT @TulanePRC: Prizes for most engaged #LiveFitNOLA participants include gear &amp; gifts from @FITByYou. @Healthfitmag also has gift from @MasÅ™</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td>Edit cells</td>
<td>Thanks for having us as the 1st guest host for #LiveFitNOLA monthly chat! Look forward to hearing how everyone stays healthy and fit!</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td>Edit column</td>
<td>RT @TulanePRC: Welcome to the 1st #LiveFitNOLA chat! Introduce yourself &amp; let us know who's here! Guest host @Healthfitmag is w/ us.</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td>Transpose</td>
<td>Split into several columns...</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td>Source</td>
<td>Add column based on this column...</td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td>View</td>
<td>Add column by fetching URLs...</td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td>View</td>
<td>Add columns from Freebase...</td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td></td>
<td>View</td>
<td>Rename this column</td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td></td>
<td>Reconcile</td>
<td>Remove this column</td>
<td></td>
</tr>
</tbody>
</table>

Move column to beginning
Move column to end
Move column left
Move column right

---

Measuring Twitter Chat Participant Engagement: #LiveFitNOLA Example, Rabarison KM et al.
- Transform all the contents the “source” column to lowercase

- Repeat the step above for the “target” column
Transform all the contents the “source” column to be preceded by “@” symbol.
Transform all the contents the “target” column

- Copy: `filter(value.split(/[^a-z0-9-_@#]/),i,i.startsWith("@")).join(``,")``
- And paste in the Expression field below, to extract Twitter usernames mentioned in each tweet.
Split the contents of the “target” column, to create set of source and target pairs
Fill down the contents of the “source” column to finalize the relationship between source and target @usernames.
### Measuring Twitter Chat Participant Engagement: #LiveFitNOLA Example, Rabarison KM et al.

<table>
<thead>
<tr>
<th>Show as:</th>
<th>rows</th>
<th>records</th>
<th>Show: 5 10 25 50 rows</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>source</td>
<td></td>
<td>target</td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>Facet</td>
<td>healthfitmag</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Text filter</td>
<td>planeprc</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>Edit cells</td>
<td>Transform...</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Edit column</td>
<td>Common transforms</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>Transpose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>Sort...</td>
<td>Fill down</td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>View</td>
<td>Blank down</td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Reconcile</td>
<td>Split multi-valued cells...</td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td></td>
<td>Join multi-valued cells...</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster and edit...</td>
<td></td>
</tr>
</tbody>
</table>
- Dataset for network visualization and analysis is now complete

- Export the new dataset with the complete “source” and “target” columns as a .csv file
How to visualize and collect a Hashtag-based two-way communication network with Gephi:

- Download Gephi (www.gephi.github.io/)

**The Open Graph Viz Platform**

Gephi is an interactive visualization and exploration platform for all kinds of networks and complex systems, dynamic and hierarchical graphs.

Runs on Windows, Linux and Mac OS X. Gephi is open-source and free.

Learn More on Gephi Platform »

Download FREE Gephi 0.8.2-beta

Release Notes | System Requirements

- Features
- Quick start
- Screenshots
- Videos
Open Gephi and create a new project

Select the “Data Laboratory” then “Data Table”
- Import Spreadsheet, the .csv file you saved at the end of OpenRefine step
  - Select “Comma” from the “Separator” drop down menu
  - Select “Edge table” from the “As table:” drop down menu
- Click the “Next” button in the import window
- And make sure you check the “Create missing nodes” option.
Go back to the “Overview” tab and note how many “Nodes” and “Edges” are in the network. (Ex: Nodes: 135, Edges: 474).

- Nodes represent the network members, which include the Twitter chat participants (here, 66 participants) and any other Twitter usernames they mentioned during the chat (here, 69 additional usernames mentioned).

- Edges represent the relationships or connections between the nodes. In this case, edges are the connections formed when a #LiveFitNOLA chat participant mentioned another participants or another Twitter username.
Under the “Statistics” tab, run some simple metrics. For the purpose of this exercise, the two important metrics are: “Average Degree” and “Modularity”

- Average degree is the average total number of mentions in the Twitter chat network, regardless of direction.
- Modularity is a community detection algorithm, which identifies the number of communities created in the Twitter chat network based on the number of engagements around particular nodes.
Under the “Partition” and “Nodes” tabs, click on the green refresh arrows.
- Choose “Modularity Class” as the partition parameter
Under the “Ranking” and Nodes tabs, choose “Degree” as the rank parameter.
Set the node sizes by clicking on the red diamond button, under the “Ranking” tab

- e.g. min size = 10, max size = 100
Choose a layout to visualize the Twitter chat network map.
- The layout choice is based on preference.
- For the #LiveFitNOLA Twitter Chat, we first used the “Yifan Hu” layout, then “Noverlap” to remove the overlapping nodes in the network map.
Back to the “Data Laboratory” and “Data Table” tabs, export Nodes table and save.
Network statistics:

- Import the Nodes table file you saved from the last step into Excel
- Run summary statistics for Degree, In Degree, Out Degree
- Identify your usernames of interest (e.g. @TulanePRC and @FitNOLA) and note the number of In Degree and Out Degree they had.
  - Number of in degree = number of incoming communications
  - Number of out degree = number of outgoing communications
#LiveFitNOLA Network
Analysis results
Total number of nodes = 135
Total number of edges = 474
Average network degree = 7 (13.4)
Degree range = 1 - 101

Number of communities = 5
- 33.33% centered around @TulanePRC
- 31.85% centered around @FitNOLA
- 13.33% centered around @HealthFitMag
- 14.81% centered around @CDC Chronic
- 6.67% centered around an unidentified user

Measuring Twitter Chat Participant Engagement: #LiveFitNOLA Example, Rabarison KM et al.
Definition of engagement:

- **Engagement** in two-way communications on Twitter is defined as the number of incoming and outgoing mentions between users.
  - A mention can be a direct mention, a retweet, or a reply
  - In other words the incoming and outgoing interactions between two Twitter users within a network bound by a hashtag and a specific timeframe.

- **Overall engagement** for a Twitter user of interest (ex: @TulanePRC) is the total number of two-way communications that user was involved in, regardless of direction.

- **Outgoing mention** is the number of times a Twitter user of interest mentions another user in a tweet.
  - For example: @TulanePRC’s outgoing communication is the **total number of times @TulanePRC retweeted or replied to another Twitter user**.

- **Incoming mention** is the number times a Twitter user of interest was mentioned in another user’s tweet.
  - For example: @TulanePRC’s incoming communication is the **total number of times @TulanePRC other Twitter users retweeted or replied to @TulanePRC**.
Engagement ratio:

- **Engagement ratio** compares a network member’s total number of outgoing mentions to their total number of incoming mentions.

\[
@\text{username} \text{ Engagement ratio} = \frac{\text{@username out degree}}{\text{@username in degree}}
\]

- @username out degree: The number of outgoing mentions for a network member.
- @username in degree: The number of incoming mentions for a network member.
Return on engagement (ROE)

- **ROE** measures the engagement gain or loss generated relative to the amount of engagement invested.
  - In other words, the engagement gain or loss calculated as incoming mentions related to amount of outgoing mentions invested.

\[
@\text{username} \text{ ROE (\%)} = \frac{\text{@username's total incoming mentions} - \text{@username's total outgoing mentions}}{\text{@username's total outgoing mentions}} \times 100
\]
#LiveFitNOLA engagement levels:

<table>
<thead>
<tr>
<th></th>
<th>Outgoing</th>
<th>Incoming</th>
<th>Out:In</th>
<th>ROE</th>
</tr>
</thead>
<tbody>
<tr>
<td>@TualnePRC</td>
<td>32</td>
<td>45</td>
<td>1 to 1</td>
<td>40%</td>
</tr>
<tr>
<td>@FitNOLA</td>
<td>63</td>
<td>38</td>
<td>2 to 1</td>
<td>-40%</td>
</tr>
<tr>
<td>@HealthFitMag</td>
<td>40</td>
<td>33</td>
<td>1 to 1</td>
<td>-18%</td>
</tr>
<tr>
<td>@CDCChronic</td>
<td>8</td>
<td>19</td>
<td>1 to 2</td>
<td>138%</td>
</tr>
</tbody>
</table>
Thank You!

For questions or comments:
Kristina M. Rabarison, DrPH, MS
KRabarison@cdc.gov
For more information, contact CDC
1-800-CDC-INFO (232-4636)

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.