The Brexit Debate through Social Media: deliberative discussion, or deliberate dysfunction?

Professor Kenneth Benoit  
Head, Department of Methodology, LSE  
(@kenbenoit)

Professor Sara Hobolt  
Chair in European Institutions, European Institute, LSE  
(@sarahobolt)

Dr Jennifer Jackson Preece  
Associate Professor in Nationalism, European Institute, LSE

Dr Jean-Christophe Plantin  
Chair, LSE  
(@JCPlantin)

Hashtag for Twitter users: #LSEworks

@lsepublicevents  
lse.ac.uk/events
The Brexit Debate Through Social Media

Deliberative Discussion, or Deliberate Dysfunction?

Kenneth Benoit
Department of Methodology, LSE
Research Team

LSE Department of Methodology

• EUENGAGE: Dr. Akitaka Matsuo, Research Officer (http://amatsuo.net)

• quanteda (the R package for managing and analyzing text)
  • Dr. Haiyan Wang, Research Officer
  • Kohei Watanabe, Ph.D Candidate (http://koheiw.net)

Imperial College London, Data Science Institute

• Dr. Julio Amador Díaz López, Research Associate
• David Birch, Research Associate
• Prof. Mark Kennedy, Professor
Social media in political communications

• Approximately 400-450 million Tweets per day
• 29.2% of US population are active users
• 20% of UK population (13.1 million in 2016)
• 563 UK MPs have Twitter accounts
87% of UK MPs have Twitter accounts
How Twitter is used for political communication: to send a message

Boris Johnson @BorisJohnson · Jan 13
Global health security will continue to challenge us all. I believe @davidnabarro is the right choice for the next @WHO leader

Donald J. Trump @realDonaldTrump · 15h
Big day planned on NATIONAL SECURITY tomorrow. Among many other things, we will build the wall!
to call attention to an event or report
Henry Dimbleby @HenryDimbleby · Jan 18

Must read. The transcript of @michaelgove's Trump interview is the gift that keeps giving. Worth the subscription to @thetimes on its own.

Henry Dimbleby @HenryDimbleby
Day one is Monday, right? Not Friday or Saturday.

thetimes.co.uk/edition/news/f...
Richard Corbett @RCorbettMEP · 2h
Hammond admits '#Brexit fog' impact on investment in #Britain. No hiding the negative impact of #brexitshambles

Brexit: Philip Hammond warns deal 'will take years' - BBC News
The chancellor warns the process of exiting the EU is likely to last much longer than two years.
bbc.com
Donald J. Trump @realDonaldTrump · 15h

If Chicago doesn't fix the horrible "carnage" going on, 228 shootings in 2017 with 42 killings (up 24% from 2016), I will send in the Feds!
Current project: Brexit-related Tweets

- some 35 million since January
- present analysis: Through end of June (26 million)
- capture based on #hashtags, @usernames, and search terms
<table>
<thead>
<tr>
<th>Hashtags:</th>
<th>Usernames:</th>
</tr>
</thead>
<tbody>
<tr>
<td>#betterdealforbritain</td>
<td>@vote_leave</td>
</tr>
<tr>
<td>#betteroffout</td>
<td>@brexitwatch</td>
</tr>
<tr>
<td>#brexit</td>
<td>@eureferendum</td>
</tr>
<tr>
<td>#euref</td>
<td>@ukandeu</td>
</tr>
<tr>
<td>#eureferendum</td>
<td>@notoeu</td>
</tr>
<tr>
<td>#eusemit</td>
<td>@leavehq</td>
</tr>
<tr>
<td>#getoutnow</td>
<td>@ukineu</td>
</tr>
<tr>
<td>#leaveeu</td>
<td>@leaveeuofficial</td>
</tr>
<tr>
<td>#no2eu</td>
<td>@ukleave_eu</td>
</tr>
<tr>
<td>#notoeu</td>
<td>@strongerin</td>
</tr>
<tr>
<td>#strongerin</td>
<td>@yesforeurope</td>
</tr>
<tr>
<td>#ukineu</td>
<td>@grassroots_out</td>
</tr>
<tr>
<td>#voteteave</td>
<td>@stronger_in</td>
</tr>
<tr>
<td>#wewantout</td>
<td></td>
</tr>
<tr>
<td>#yes2eu</td>
<td></td>
</tr>
<tr>
<td>#yestoeu</td>
<td></td>
</tr>
<tr>
<td>brexit</td>
<td></td>
</tr>
</tbody>
</table>
Frequencies of Tweets

The line plot shows the frequencies of tweets over different months. The x-axis represents the dates from January to July, and the y-axis represents the number of tweets, ranging from $10^4$ to $10^6$. The plot includes two verbs: "share" and "post". The "share" trend is represented by the red line, and the "post" trend is represented by the blue line. The trend for "share" shows a significant increase in July, while the trend for "post" follows a similar pattern but with slightly lower values overall.
Users in data

- 3.6M unique users

- number of tweets:
  - average: 7.2
  - median: 1 (more than 50% had only one tweet)
  - max: 81.1K
### Heavy users

<table>
<thead>
<tr>
<th>username</th>
<th>share</th>
<th>post</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>iVoteLeave</td>
<td>81101</td>
<td>3</td>
<td>81104</td>
</tr>
<tr>
<td>iVoteStay</td>
<td>57615</td>
<td>1</td>
<td>57616</td>
</tr>
<tr>
<td>Col_Connaughton</td>
<td>0</td>
<td>21451</td>
<td>21451</td>
</tr>
<tr>
<td>brexitmarch</td>
<td>0</td>
<td>20873</td>
<td>20873</td>
</tr>
<tr>
<td>MikkiL</td>
<td>18850</td>
<td>1407</td>
<td>20257</td>
</tr>
<tr>
<td>Fight4UK</td>
<td>13785</td>
<td>4777</td>
<td>18562</td>
</tr>
<tr>
<td>JodieActy</td>
<td>13837</td>
<td>1021</td>
<td>14858</td>
</tr>
<tr>
<td>SaraPadmore</td>
<td>14692</td>
<td>20</td>
<td>14712</td>
</tr>
<tr>
<td>Brndstr</td>
<td>0</td>
<td>13449</td>
<td>13449</td>
</tr>
<tr>
<td>marie52d</td>
<td>12884</td>
<td>349</td>
<td>13233</td>
</tr>
<tr>
<td>RoyalNavyNews</td>
<td>2648</td>
<td>10382</td>
<td>13030</td>
</tr>
<tr>
<td>ukleave_eu</td>
<td>9890</td>
<td>2589</td>
<td>12479</td>
</tr>
<tr>
<td>mwengway</td>
<td>10072</td>
<td>2183</td>
<td>12255</td>
</tr>
</tbody>
</table>
Main hashtags

• List of hashtags
  http://rpubs.com/amatsuo/top_hashtags_10000

• c.f List of mentions
  http://rpubs.com/amatsuo/top_mentions_10000
Evolution of hashtags (ranking)

- Top 10 hashtags:
  - #brexit
  - #eu
  - #euref
  - #eureferendum
  - #leaveeu
  - #remain
  - #strongerin
  - #uk
  - #ukip
  - #voteleave

![Graph showing daily ranking of hashtags from January to July]
Evolution of hashtags (ranking)

Daily Ranking (100 = 1st)
Hashtags surging popularity

Daily Ranking (100 = 1st)

hashtag

#bbc  #itv  #nexit  #remain  #takecontrol
#bremain  #maga  #referendum  #sky  #voteremain
Predicting Leave v. Remain

• Method: Naive Bayes classifier, also Support-Vector Machines (SVM)
• Data source: combined tweet corpus at user level
• Creating training data
  • Select “power-users” (more than 100 tweets in the corpus, 15K users)
  • Check the use of pre-determined set of “leave” and “remain” hashtags
  • Calculate the difference in the use of leave and remain hashtags. Construct training data from top and bottom 10% of power users
• Features
  • Use bag of words method (unigrams, remove features used less than 20 users)
Predicting Leave v. Remain (Feature Selection)

• Models are evaluated through
  • Split training data into training and test sets
  • Check the prediction accuracy for test sets

• Three models are tested (prediction accuracy for test set in parentheses)
  • All words (49.4%)
  • Only hashtags (82.8%)
  • Hashtags and mentions (88.2%)
Predicting Leave v. Remain

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remain</td>
<td>9,780,223</td>
<td>36.93%</td>
</tr>
<tr>
<td>Neutral</td>
<td>7,786,297</td>
<td>29.40%</td>
</tr>
<tr>
<td>Leave</td>
<td>8,914,207</td>
<td>33.66%</td>
</tr>
</tbody>
</table>
Patterns of posts, Leave v. Remain

The graph shows the patterns of posts for Leave, Remain, and Neutral sides over time from January to July. The y-axis represents the number of posts, ranging from 1e+04 to 1e+06. The x-axis represents the date from January to July. The lines indicate the trend of posts for each side, with Leave in blue, Remain in red, and Neutral in grey.
MP side predictions by party

MP predictions from Tweets

- Conservative
- Democratic Unionist Party
- Green Party
- Independent
- Labour
- Liberal Democrat
- Plaid Cymru
- Scottish National Party
- Sinn Fein
- Social Democratic and Labour Party
- UK Independence Party
- Ulster Unionist Party
- Sinn Fein

Side: Remained
Neutral
Leave
A few well-known Twitter users

![Graph showing the number of tweets vs. probability of Remain]
Sad liberals spewing hate

#inauguration2017: We're in DC showing you everything going down on this historic day. periscope.tv/w/a1BkOzZNDk5…

Thank God we survived the last 8 yrs

John McCain is a snake. Go away you old goat

Mass impeachment for the traitor democrats
Leave and Remain Features

http://rpubs.com/amatsuo/feature_scores
Leave and Remain Hashtags (Cameron)
Leave and Remain Hashtags (Boris Johnson)
Leave and Remain Hashtags (Gove)
Leave and Remain Hashtags (Farage)
Leave and Remain Hashtags (Corbyn)
Leave and Remain Hashtags (LSE)
Leave v. Remain
Hashtags
SVM classification Europe-wide: Remain
SVM classification Europe-wide: Leave
Hashtag Networks (Top 300, across sides)
Hashtag Networks
(Top 300, across sides)
#leaveeu
Hashtag Networks
(Leave, anti-EU)
Hashtag Networks (Leave, debate)
Hashtag Networks (Remain, economic concerns)
Hashtag Networks
(Reain, parties and political figures)
Visual Analytics tools for Exploring Twitter

• We use the KPMG Data Observatory a 64 screen immersive visualisation environment
• To create an immersive social discussion environment driven by data.
• Interactive queries can be run on the data and displayed in real time to support multi-disciplinary discussion
• 132million pixels allows us to display huge amounts of data for detailed discussion
Retweet Networks

Most mentioned users

Most popular Hashtags

Most popular tweeters

Most original tweets
Community detection

Leave (red)
Remain (purple)

Several communities of leavers
Only one strong remain community

Clear “echo chamber” of leavers
Sentiment Analysis

• Looks up terms from the Linguistic Inquiry and Word Count, a psychological dictionary

• Contains categories about:
  • positive and negative emotion
  • politics
  • power
  • quantitative language
  • tentative language
  • sadness
  • future v. past orientation
example: “reward” language

```
data_dictionary_liwc[['reward']]
```

<p>| | | | | | | | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>access*</td>
<td>accrue*</td>
<td>accumul*</td>
<td>achievable</td>
<td>achieve*</td>
<td>achievei*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>acquir*</td>
<td>add</td>
<td>added</td>
<td>adding</td>
<td>adds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>advantag*</td>
<td>adventur*</td>
<td>amass*</td>
<td>approach</td>
<td>approached</td>
<td>approaches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>approaching</td>
<td>award*</td>
<td>benefit</td>
<td>benefits</td>
<td>best</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>bets</td>
<td>better</td>
<td>betting</td>
<td>bold</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>confident</td>
<td>confidently</td>
<td>crave</td>
<td>craving</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>dares</td>
<td>daring</td>
<td>desire*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>earn</td>
<td>earned</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>excite</td>
<td>excited</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>fulfill*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>good</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>jackpot*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>obtaining</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>perfect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>positive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>promot*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>steal*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>successfully</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>109</td>
<td>took</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>wagers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Positive v. Negative Emotion Language

- Remain
- Neutral
- Leave
Power Language

Remain

Neutral

Leave
The Brexit Debate through Social Media: deliberative discussion, or deliberate dysfunction?

Professor Kenneth Benoit
Head, Department of Methodology, LSE
(@kenbenoit)

Professor Sara Hobolt
Chair in European Institutions, European Institute, LSE
(@sarahobolt)

Dr Jennifer Jackson Preece
Associate Professor in Nationalism, European Institute, LSE

Dr Jean-Christophe Plantin
Chair, LSE
(@JCPlantin)

Hashtag for Twitter users: #LSEworks
@lsepublicevents  lse.ac.uk/events