Exploring citizen-government interactions through analysis of Twitter Data
Big-Data and Data Analytics:

• The Internet of Things (IoT), Big Data, Social Media, Data Analytics or in a broader term “the digitalization” is the “Steam Engine” of our times.

• Digitalization is a comprehensive, social game changer and in fact, a paradigm shift whose inflection point is already past (Boorsma, 2017).

Big-Data and Data Analytics:

- As with many things, digitalization entered the domains of urban studies too.
- Big data and their potential use in urban settings are reflected in the concept of “Smart City”
Smart City: 
While a definitive framework is yet to be established – all smart city promoters agree that there must be a meaningful partnership among the Citizens, Governments, Businesses and Academia in order to be successful in defining and implementing Smart Cities.
Smart City:
IEEE Smart Cities website suggests 6 components of a Smart City:

- smart economy
- smart mobility
- smart environment
- smart people
- smart living
- smart governance
This discussion is focused on the last component:

**smart governance**

With an aim to measure the engagement and/or relationship between a local government entity and its citizens.
Exploring citizen-government interactions through analysis of twitter Data

Howard County, Maryland was selected as the area of interest for this exploration.

Howard County government is active in social media and post messages about government events and news regularly.

The diverse citizens with above average education and income were thought to be responsive and concerned about their governments' activities.

Therefore, Howard County seemed to be a good candidate for this study.
Exploring citizen-government interactions through analysis of twitter Data

Attempted to examine if tweeter messages used by Howard County Govt. and/or tweeter interactions between the governments and citizens can be used to track the level of involvement of citizens with their government (and vice versa) about County operations.

Used Tweeter data that were freely accessible

All the analysis were done in ‘R’ – an open source software
Data Collection and cleanup:

Using the function `lookup_coords` in the library 'rtweet', the bounding box coordinates of Howard county was collected. The coordinates would be used to filter tweets to find county specific tweets only.

```r
$place
[1] "Howard County, MD"
$box
   sw.lng sw.lat ne.lng ne.lat
-77.18711 39.10313 -76.69732 39.36916
$point
   lat lng
39.28735 -76.96431
attr("class")
[1] "coords" "list"
```

Most frequently used twitter accounts by County government were collected from the Howard County website (https://www.howardcountymd.gov/)

[1] "HOCOGov"   "HOCOGOVEexec"   "HCPDNews"   "HCFERS"
[1] "HC_JonWeinstein" "HOCOBOEMaryland"
[7] "JenTerrasa"
Data Collection and cleanup:
Government twitter accounts were then used to find the associated twitter users and their followers (i.e. the citizens who have interests in government tweets).

The total number \(143,478\) of Howard County followers is impressive compared to the County population \(313,414\).
Exploring citizen-government interactions through analysis of twitter Data

Data Collection and cleanup:
A closer look:

A closer look at the data shows that the Police Department (HCPDNews) is an outlier with 108,072 followers. So on the surface it might seem citizens pay close attention to their government, apparently they are concerned about a specific agency that deals with **crime, safety and traffic control**.
Data Collection and cleanup:

Going deeper:

Functions were created to collect and evaluate citizens' tweets relevant to government

- Recent tweets sent out by each of those government accounts were collected
- All the hashtags used by government accounts and major/obvious hashtags were identified.
- search_tweets function of rtweet library was used to collect the citizen tweets.
- In order to select the tweets that were possibly generated as responds/reactions to government tweets, the most recent common hashtags used by the Howard County government were used.
- To control the citizen locations, the bounding box coordinates of the County were used as query parameter.
- The citizens tweets were separated from government tweets by comparing the users_id of the tweets.
Exploring citizen-government interactions through analysis of twitter Data

Data Collection and cleanup:
Citizen tweet locations:
Exploring citizen-government interactions through analysis of twitter data

Government tweets at a glance:

A good number of government tweets were liked and retweeted. Out of 2065 original tweets, 1239 were retweeted 3794 times, and 1463 were favorited 4690 times. The statistics here suggest a good response to government tweets.
Exploring citizen-government interactions through analysis of twitter Data

Government tweets at a glance:

Frequency of government tweets

[Graph showing consistent frequency of government tweets across months]
Exploring citizen-government interactions through analysis of twitter Data

collecting and evaluating citizen tweets:

![Frequency of citizen tweets]

Inconsistent
Exploring citizen-government interactions through analysis of twitter Data

Tweet Text mining:
- Two clean Corpus (for citizen tweets and for government tweets) were created
- Texts were analyzed to see if similar terms are common in them

![Graphs showing the 30 most frequently used terms in tweets sent out by the government and citizens.](image)

30 most frequently used terms in tweets sent out by the government and citizens. No significant match are seen between these two sets of words (terms)
Exploring citizen-government interactions through analysis of twitter Data

Tweet Text mining:
- Term-document Matrix for citizen and government tweets were created
- Dendrograms were drawn for both

Government dendrogram shows some association of the words. There were no association of words or no distinct clusters in the citizens tweets suggesting no focused discussion on certain topics but many scattered interests.
Exploring citizen-government interactions through analysis of twitter Data

Tweet Text mining: Evaluation through Wordclouds

Govt. tweets suggest some relevant words concerning government operations, such as project, DPW (public works), meeting, permit, construction, improvement, public, repair, maintenance etc. Citizen tweets seems very diverse and nothing really stands out.

The common words found in the cloud such as amp, join, holiday, tonight, work etc. are very general and does not seem to suggest any interaction between the government and the citizens.
Exploring citizen-government interactions through analysis of twitter Data

Tweet Text mining:  Topic model comparison

Both government and citizens tweet texts were grouped under five topics each, and the 10 most frequent terms related to each topics were plotted to examine if there were any similarities between the topics and terms that would suggest any interaction:

All the topics and related terms in government and citizen tweets do not show any similarities
Tweet Text mining:

Compare document-term-matrix of government and citizen tweets

Very low similarities, which could be due to very general words as seen in the commonality cloud above
Summary:

All the analysis above point to the fact that both the County government and the citizens need to take initiatives for effective communications.

The number of tweets sent out by the government and the number of totals followers they have are encouraging, which suggest both the willingness and environment are there to use social media such as tweeters for better communication between Howard County government and its citizens.

The huge number of followers of Police Department means people, in general, are naturally drawn to stories or news that have quick and explicit impact on them such as a crime event or accidents.

Therefore departments like Planning and zoning etc. that have significant influence on citizens' future livelihoods but are not immediately felt should be more proactive to connect to the citizens.
Summary:
As is inherent with a paradigm shift, new questions and challenges were raised concerning the use, advantage and impact of big data on urban systems and urban living. There are lot of concerns and questions – such as privacy, security, legal issues, data ownership etc. to name a few.

The mindsets of both the citizens and the governments need to be changed, which requires:

communication, open dialogues and openness

Let’s TALK ......

BLOG:  https://khanmehdi.wixsite.com/mdsmartcity

Smart Cities Meet-up Group, Maryland
Exploring citizen-government interactions through analysis of twitter Data

Thank you !!!