Analyzing Public Comments to the FCC on Net Neutrality

Jacqueline Antwi-Danso and Tarini Konchady

December 14, 2017 — The Federal Communications Commission (FCC) rolls back Title II oversight of Internet Service Providers (ISPs)

TITLE II--COMMON CARRIERS

PART I--COMMON CARRIER REGULATION

SEC. 201. [47 U.S.C. 201] SERVICE AND CHARGES.

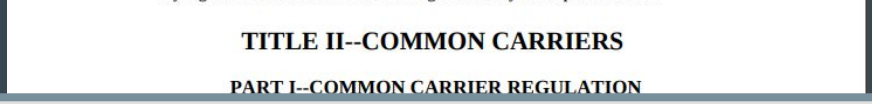
(a) It shall be the duty of every common carrier engaged in interstate or foreign communication by wire or radio to furnish such communication service

upon reasonable request therefor; and, in accordance with the orders of the Commission, in cases where the Commission, after opportunity for hearing, finds such action necessary or desirable in the public interest, to establish physical connections with other carriers, to establish through routes and charges applicable thereto and the divisions of such charges, and to establish and provide facilities and regulations for operating such through routes.

(b) All charges, practices, classifications, and regulations for and in connection with such communication service, shall be just and reasonable, and any such charge, practice, classification, or regulation that is unjust or unreasonable is hereby declared to be unlawful: *Provided*, That communications by wire or radio subject to this Act may be classified into day, night, repeated, unrepeated, letter, commercial, press, Government and such other classes as the Commission may decide to be just and reasonable, and different charges may be made for the different classes of comunications: *Provided further*, That nothing in this Act or in any other provision of law shall be construed to prevent a common carrier subject to this Act from entering into or operating under any contract with any common carrier not subject to this Act, for the exchange of their services, if the Commission is of the opinion that such contract is not contrary to the public interest: *Provided further*, That nothing in this Act or in any other provision of law shall prevent a common carrier subject to this Act from furnishing reports of positions of ships at sea to newspapers of general circulation, either at a nominal charge or without charge, provided the name of such common carrier is displayed along with such ship position reports. The Commissioner may prescribe such rules and regulations as may be necessary in the public interest to carry out the provisions of this Act.

SEC. 202. [47 U.S.C. 202] DISCRIMINATION AND PREFERENCES.

(a) It shall be unlawful for any common carrier to make any unjust or unreasonable discrimination in charges, practices, classifications, regulations, facilities, or services for or in connection with like communication service, directly or indirectly, by any means or device, or to make or give any undue or



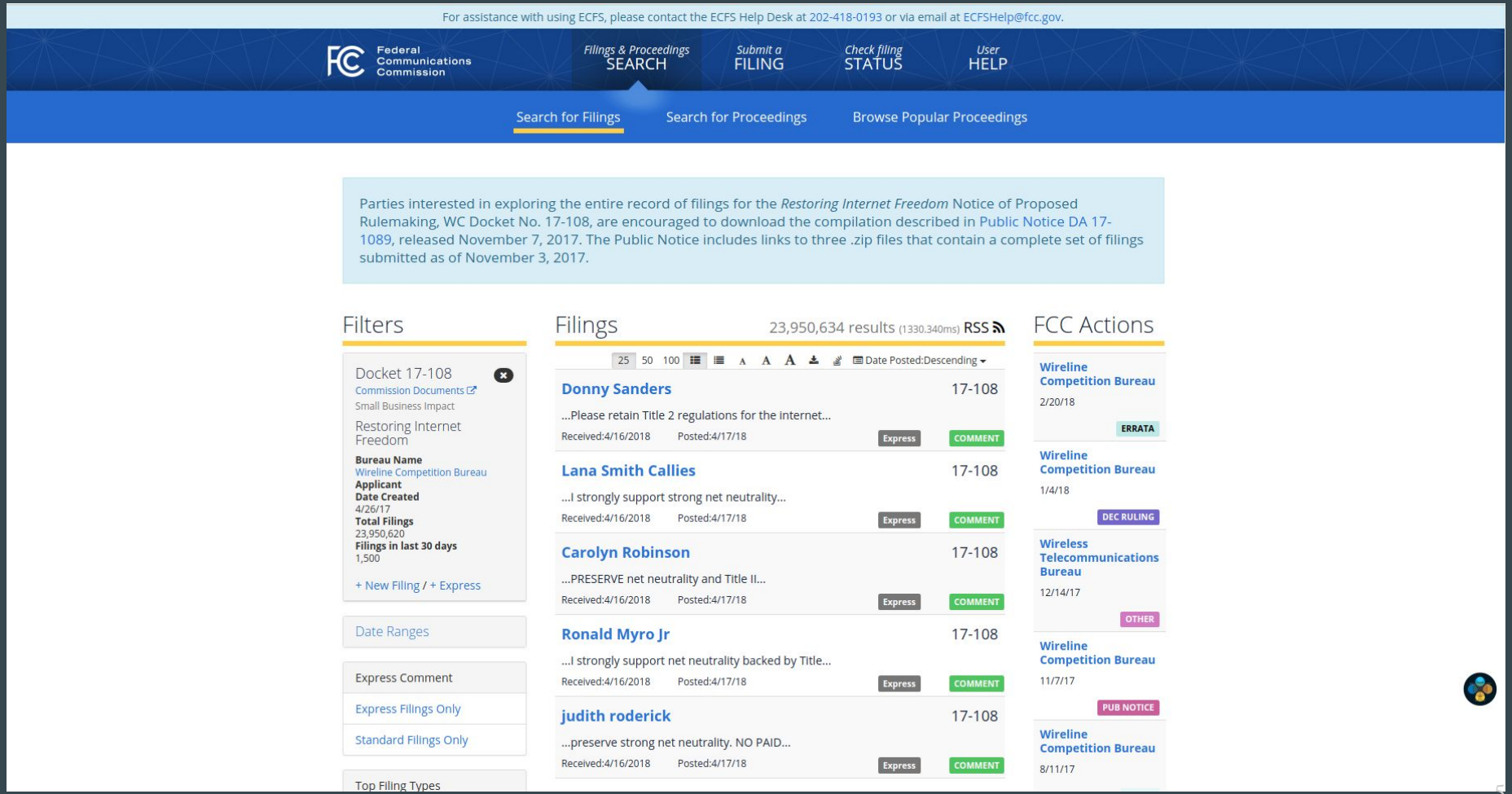
TITLE II--COMMON CARRIERS

PART I--COMMON CARRIER REGULATION

DON'T GIVE CERTAIN ENTITIES PREFERENCE WHEN PROVIDING SERVICES

SEC. 202. [47 U.S.C. 202] DISCRIMINATION AND PREFERENCES.

(a) It shall be unlawful for any common carrier to make any unjust or unreasonable discrimination in charges, practices, classifications, regulations, facilities, or services for or in connection with like communication service, directly or indirectly, by any means or device, or to make or give any undue or


For assistance with using ECFS, please contact the ECFS Help Desk at 202-418-0193 or via email at ECFSHelp@fcc.gov.
Federal Communications Commission
Filings & Proceedings SEARCH
Submit a FILING
Check filing STATUS
User HELP
Search for Filings
Search for Proceedings
Browse Popular Proceedings
Parties interested in exploring the entire record of filings for the Restoring Internet Freedom Notice of Proposed Rulemaking, WC Docket No. 17-108, are encouraged to download the compilation described in Public Notice DA 17-1089, released November 7, 2017. The Public Notice includes links to three .zip files that contain a complete set of filings submitted as of November 3, 2017.
Filters
Docket 17-108
Commission Documents
Small Business Impact
Restoring Internet Freedom
Bureau Name
Wireline Competition Bureau
Applicant
Date Created
4/26/17
Total Filings
23,950,620
Filings in last 30 days
1,500
+ New Filing / + Express
Date Ranges
Express Comment
Express Filings Only
Standard Filings Only
Top Filing Types
Filings
23,950,634 results (1330.340ms) RSS
25 50 100
Date Posted:Descending
Donny Sanders 17-108
...Please retain Title 2 regulations for the internet...
Received:4/16/2018 Posted:4/17/18 Express COMMENT
Lana Smith Callies 17-108
...I strongly support strong net neutrality...
Received:4/16/2018 Posted:4/17/18 Express COMMENT
Carolyn Robinson 17-108
...PRESERVE net neutrality and Title II...
Received:4/16/2018 Posted:4/17/18 Express COMMENT
Ronald Myro Jr 17-108
...I strongly support net neutrality backed by Title...
Received:4/16/2018 Posted:4/17/18 Express COMMENT
judith roderick 17-108
...preserve strong net neutrality. NO PAID...
Received:4/16/2018 Posted:4/17/18 Express COMMENT
FCC Actions
Wireline Competition Bureau
2/20/18
ERRATA
Wireline Competition Bureau
1/4/18
DEC RULING
Wireless Telecommunications Bureau
12/14/17
OTHER
Wireline Competition Bureau
11/7/17
PUB NOTICE
Wireline Competition Bureau
8/11/17

For assistance with using ECFS, please contact the ECFS Help Desk at 202-418-0193 or via email at ECFSHelp@fcc.gov.

Federal Communications Commission

Filings & Proceedings SEARCH | Submit a FILING | Check filing STATUS | User HELP

Search for Filings | Search for Proceedings | Browse Popular Proceedings

Parties interested in exploring the entire record of filings for the *Restoring Internet Freedom* Notice of Proposed Rulemaking, WC Docket No. 17-108, are encouraged to download the compilation described in Public Notice DA 17-1089, released November 7, 2017. The Public Notice includes links to three .zip files that contain a complete set of filings submitted as of November 3, 2017.

Filters

Docket 17-108
Commission Documents
Small Business Impact
Restoring Internet Freedom
Bureau Name
Wireline Competition Bureau
Applicant
Date Created
Total Filings
23,950,620
Filings in last 30 days
1,500
+ New Filing / + Express

Date Ranges

Express Comment
Express Filings Only
Standard Filings Only

Top Filing Types

4/26/17
Total Filings
23,950,620
Filings in last 3

Received:4/16/2018 Posted:4/17/18 Express COMMENT

Ronald Myro Jr 17-108
...I strongly support net neutrality backed by Title...
Received:4/16/2018 Posted:4/17/18 Express COMMENT

judith roderick 17-108
...preserve strong net neutrality. NO PAID...
Received:4/16/2018 Posted:4/17/18 Express COMMENT

FCC Actions

Wireline Competition Bureau
2/20/18
ERRATA

Wireline Competition Bureau
1/4/18
DEC RULING

Wireless Telecommunications Bureau
12/14/17
OTHER

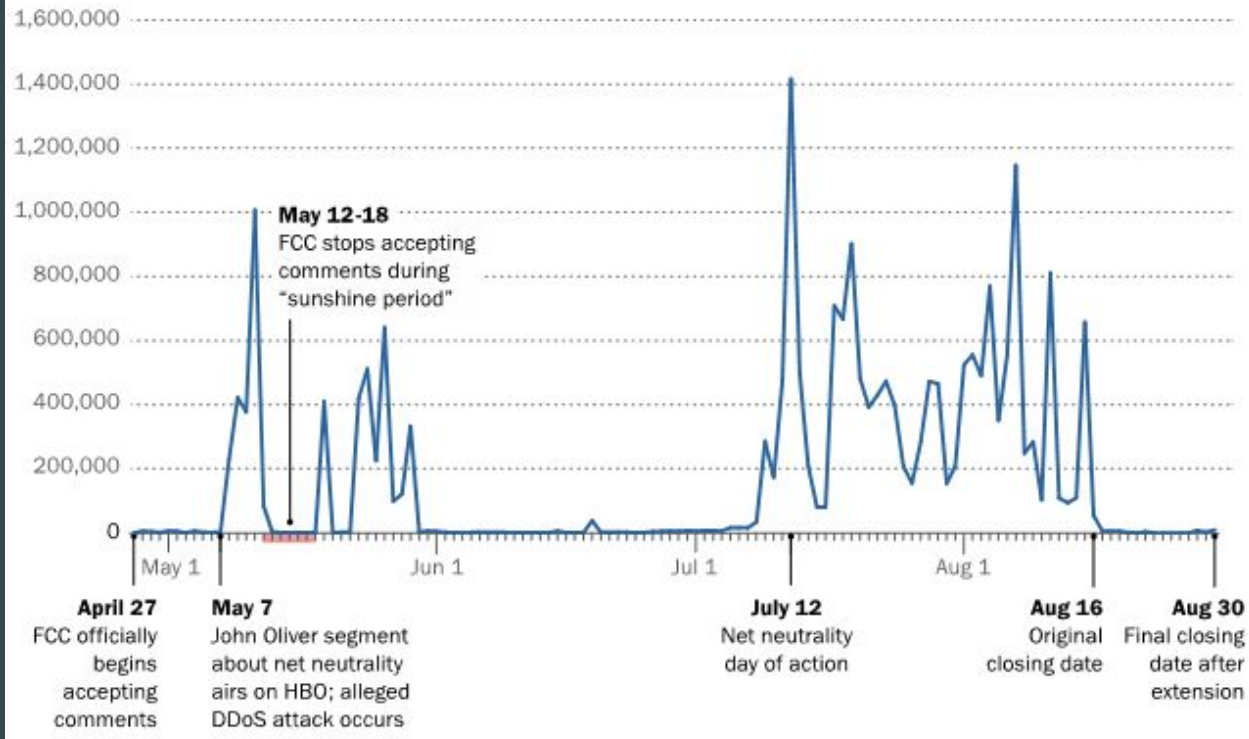
Wireline Competition Bureau
11/7/17
PUB NOTICE

Wireline Competition Bureau
8/11/17

Net Neutrality II: Last Week Tonight with John Oliver (HBO)
NET
NEUTRALITY
1:49 / 19:32

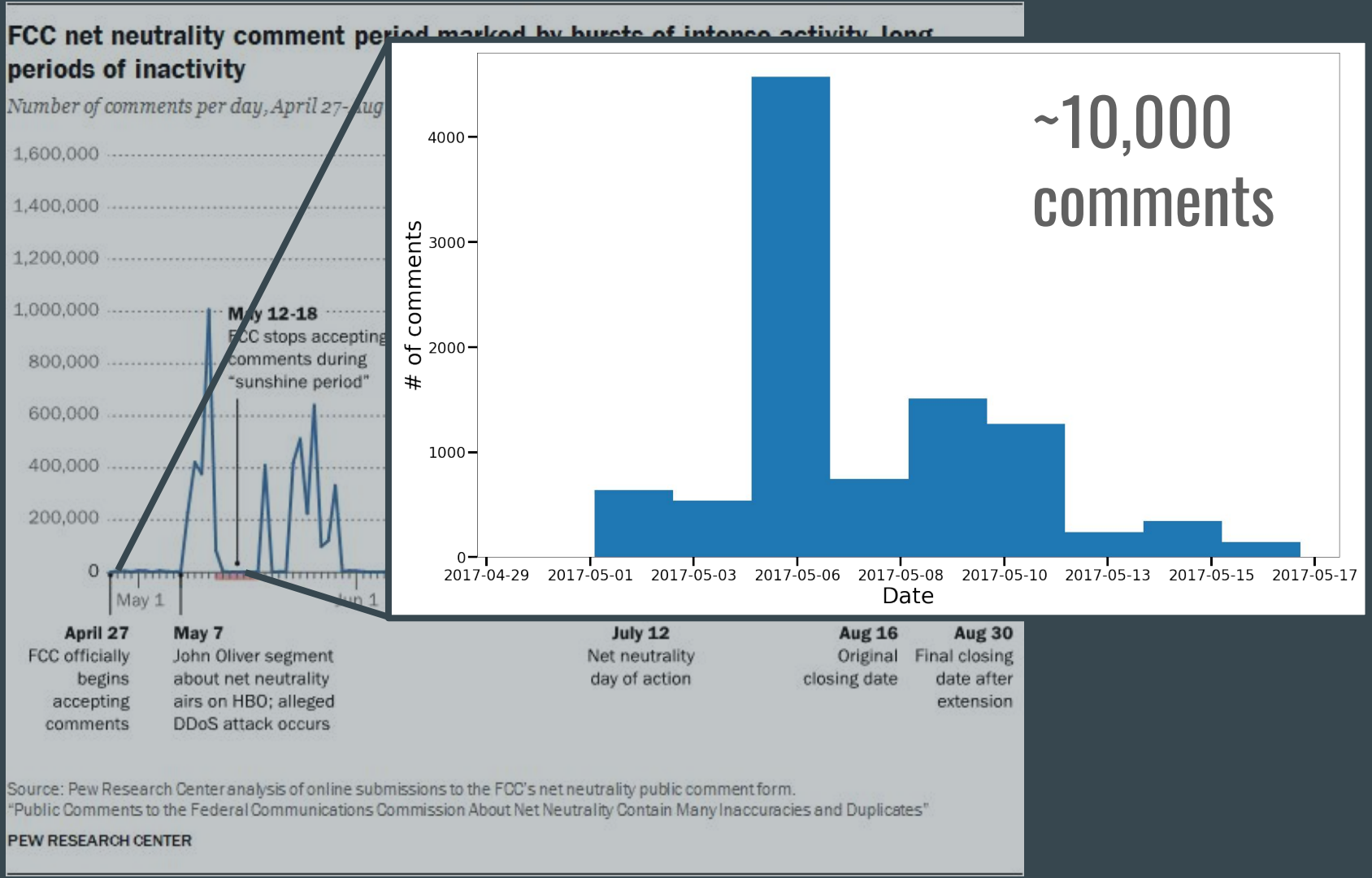
FCC net neutrality comment period marked by bursts of intense activity, long periods of inactivity

Number of comments per day, April 27-Aug. 30, 2017



Source: Pew Research Center analysis of online submissions to the FCC's net neutrality public comment form.
"Public Comments to the Federal Communications Commission About Net Neutrality Contain Many Inaccuracies and Duplicates"

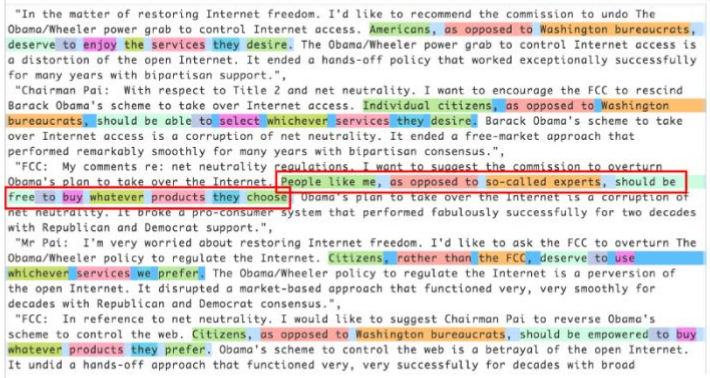
PEW RESEARCH CENTER


FCC net neutrality comment period marked by bursts of intense activity, long periods of inactivity
Number of comments per day, April 27-Aug
1,600,000
1,400,000
1,200,000
1,000,000
800,000
600,000
400,000
200,000
0
May 1
May 12-18
FCC stops accepting comments during "sunshine period"
April 27
FCC officially begins accepting comments
May 7
John Oliver segment about net neutrality airs on HBO; alleged DDoS attack occurs
July 12
Net neutrality day of action
Aug 16
Original closing date
Aug 30
Final closing date after extension
~10,000 comments
# of comments
4000
3000
2000
1000
0
2017-04-29
2017-05-01
2017-05-03
2017-05-06
2017-05-08
2017-05-10
2017-05-13
2017-05-15
2017-05-17
Date
Source: Pew Research Center analysis of online submissions to the FCC's net neutrality public comment form.
"Public Comments to the Federal Communications Commission About Net Neutrality Contain Many Inaccuracies and Duplicates"
PEW RESEARCH CENTER

Jeff Kao Follow
Data Scientist. Author of WaPo cited Net Neutrality study https://tinyurl.com/fcc-nn. Alum at Metis SF, Columbia Law and UWaterloo
Nov 23, 2017 · 10 min read

More than a Million Pro-Repeal Net Neutrality Comments were Likely Faked

I used natural language processing techniques to analyze net neutrality comments submitted to the FCC from April-October 2017, and the results were disturbing.



"In the matter of restoring Internet freedom. I'd like to recommend the commission to undo The Obama/Wheeler power grab to control Internet access. Americans, as opposed to Washington bureaucrats, deserve to enjoy the services they desire. The Obama/Wheeler power grab to control Internet access is a distortion of the open Internet. It ended a hands-off policy that worked exceptionally successfully for many years with bipartisan support.",
"Chairman Pai: With respect to Title 2 and net neutrality. I want to encourage the FCC to rescind Barack Obama's scheme to take over Internet access. Individual citizens, as opposed to Washington bureaucrats, should be able to select whichever services they desire. Barack Obama's scheme to take over Internet access is a corruption of net neutrality. It ended a free-market approach that performed remarkably smoothly for many years with bipartisan consensus.",
"FCC: My comments re: net neutrality regulations. I want to suggest the commission to overturn Obama's plan to take over the Internet. People like me, as opposed to so-called experts, should be free to buy whatever products they choose. Obama's plan to take over the Internet is a corruption of net neutrality. It broke a pro-consumer system that performed fabulously successfully for two decades with Republican and Democrat support.",
"Mr Pai: I'm very worried about restoring Internet freedom. I'd like to ask the FCC to overturn The Obama/Wheeler policy to regulate the Internet. Citizens, rather than the FCC, deserve to use whichever services we prefer. The Obama/Wheeler policy to regulate the Internet is a perversion of the open Internet. It disrupted a market-based approach that functioned very, very smoothly for decades with Republican and Democrat consensus.",
"FCC: In reference to net neutrality. I would like to suggest Chairman Pai to reverse Obama's scheme to control the web. Citizens, as opposed to Washington bureaucrats, should be empowered to buy whatever products they prefer. Obama's scheme to control the web is a betrayal of the open Internet. It undid a hands-off approach that functioned very, very successfully for decades with broad

Spot the fake comment. Surprise – they're all fake.

Never miss a story from **Hacker Noon**, when you sign up for Medium. Learn more
GET UPDATES

REPORT

NOVEMBER 29, 2017

Public Comments to the Federal Communications Commission About Net Neutrality Contain Many Inaccuracies and Duplicates

Fully 57% of comments used temporary or duplicate email addresses, and seven popular comments accounted for 38% of all submissions

BY PAUL HITLIN, KENNETH OLMSTEAD AND SKYE TOOR

Correction: This report initially noted that 450,000 comments were submitted to the FCC during its previous open comment period on net neutrality. That data point was based only on the initial comment period, spanning Feb. 9-July 18, 2014. The FCC subsequently reopened the comment period through Sept. 15, 2014, and the report now reflects the total number of comments received during the entirety of the 2014 public comment period. In addition, a reference to John Oliver in a sentence referring to the most popular pro-net-neutrality comment has been removed. Pew Research Center has issued a statement regarding concerns raised about this analysis.

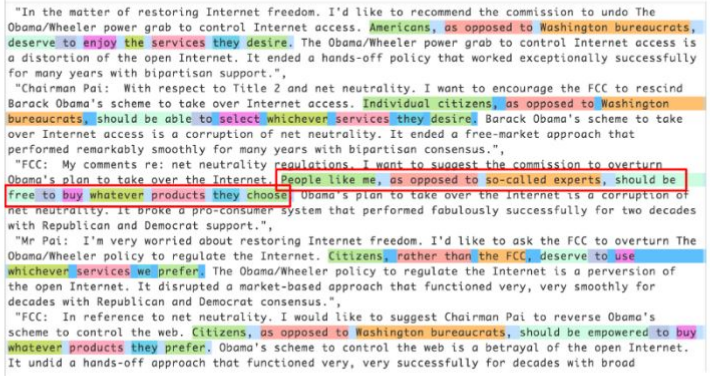
For the second time in less than four years, the U.S. Federal Communications Commission (FCC) is considering regulations regarding net neutrality – the principle that internet service providers must treat all data the same, regardless of the origin or purpose of that data. Opponents of net neutrality regulations argue that ISPs should have the right to prioritize traffic and charge for their services as they wish. Meanwhile, supporters of net neutrality suggest that so-called fast lanes are anti-competitive and would prevent start-ups and smaller companies from competing with more well-

REPORT MATERIALS
Complete Re
TABLE OF CONTENTS
Overview
Many submissions c false or misleading p information
Many submissions h organized efforts to the commentary per
The comment period marked by bursts of activity and long stre few submissions
Acknowledgments
Methodology
RELATED
PEW RESEARCH CENTER 2017
Internet/Broad Fact Sheet

Jeff Kao Follow
Data Scientist. Author of WaPo cited Net Neutrality study https://tinyurl.com/fcc-nn. Alum at Metis SF, Columbia Law and UWaterloo
Nov 23, 2017 · 10 min read

More than a Million Pro-Repeal Net Neutrality Comments were Likely Faked

I used natural language processing techniques to analyze net neutrality comments submitted to the FCC from April-October 2017, and the results were disturbing.



"In the matter of restoring Internet freedom. I'd like to recommend the commission to undo The Obama/Wheeler power grab to control Internet access. Americans, as opposed to Washington bureaucrats, deserve to enjoy the services they desire. The Obama/Wheeler power grab to control Internet access is a distortion of the open Internet. It ended a hands-off policy that worked exceptionally successfully for many years with bipartisan support.",
"Chairman Pai: With respect to Title 2 and net neutrality. I want to encourage the FCC to rescind Barack Obama's scheme to take over Internet access. Individual citizens, as opposed to Washington bureaucrats, should be able to select whichever services they desire. Barack Obama's scheme to take over Internet access is a corruption of net neutrality. It ended a free-market approach that performed remarkably smoothly for many years with bipartisan consensus.",
"FCC: My comments re: net neutrality regulations. I want to suggest the commission to overturn Obama's plan to take over the Internet. People like me, as opposed to so-called experts, should be free to buy whatever products they choose. Obama's plan to take over the Internet is a corruption of net neutrality. It broke a pro-consumer system that performed fabulously successfully for two decades with Republican and Democrat support.",
"Mr Pai: I'm very worried about restoring Internet freedom. I'd like to ask the FCC to overturn The Obama/Wheeler policy to regulate the Internet. Citizens, rather than the FCC, deserve to use whichever services we prefer. The Obama/Wheeler policy to regulate the Internet is a perversion of the open Internet. It disrupted a market-based approach that functioned very, very smoothly for decades with Republican and Democrat consensus.",
"FCC: In reference to net neutrality. I would like to suggest Chairman Pai to reverse Obama's scheme to control the web. Citizens, as opposed to Washington bureaucrats, should be empowered to buy whatever products they prefer. Obama's scheme to control the web is a betrayal of the open Internet. It undid a hands-off approach that functioned very, very successfully for decades with broad

Spot the fake comment. Surprise – they're all fake.

Never miss a story from **Hacker Noon**, when you sign up for Medium. Learn more
GET UPDATES

o the Federal
ommission About
ain Many
uplicates

nporary or duplicate email
omments accounted for 38%

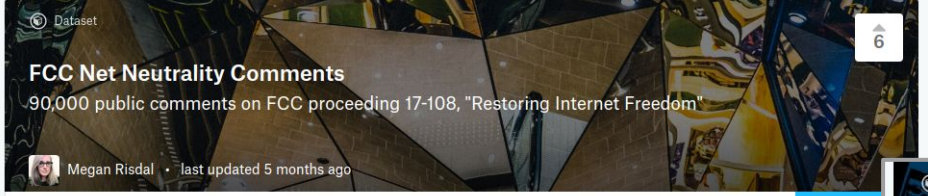
0 comments were submitted to the FCC during its
hat data point was based only on the initial
he FCC subsequently reopened the comment
reflects the total number of comments received
iod. In addition, a reference to John Oliver in a
trality comment has been removed. Pew
concerns raised about this analysis.

S. Federal Communications Commission (FCC) is considering regulations regarding net neutrality – the principle that internet service providers must treat all data the same, regardless of the origin or purpose of that data. Opponents of net neutrality regulations argue that ISPs should have the right to prioritize traffic and charge for their services as they wish. Meanwhile, supporters of net neutrality suggest that so-called fast lanes are anti-competitive and would prevent start-ups and smaller companies from competing with more well-

REPORT MATERIALS
Complete Re
TABLE OF CONTENTS
Overview
Many submissions c
false or misleading
information
Many submissions h
organized efforts to
the commentary per
The comment period
marked by bursts of
activity and long stre
few submissions
Acknowledgments
Methodology
RELATED
PEW RESEARCH CENTER
2017
Internet/Broad
Fact Sheet

Our Objective

To quantify “spamminess” of comments based on comment text and submission identifiers



Overview Data Kernels Discussion Activity Download (8 MB) New Kern

1 Files (8.07 MB)

fcc_comments.csv

Download All

fcc_comments.csv
79.03 MB • Updated 5 months ago

Download

About this file
Help us describe this file

Preview (first 100 rows) Column Metadata Column Metrics

confirmation_number	Help us describe this column...	Numeric
id_submission	Help us describe this column...	Numeric
express_comment	Help us describe this column...	Numeric
date_submission	Help us describe this column...	DateTime
date_disseminated	Help us describe this column...	DateTime
date_received	Help us describe this column...	DateTime
text_data	Help us describe this column...	String
city	Help us describe this column...	String

Overview Data Kernels Discussion Activity Download (194 MB) New Kernel

1 Files (193.76 MB)

proc_17_108_uniq...

Download All

proc_17_108_uniques_clustered_full.csv
2.61 GB • Updated 5 months ago

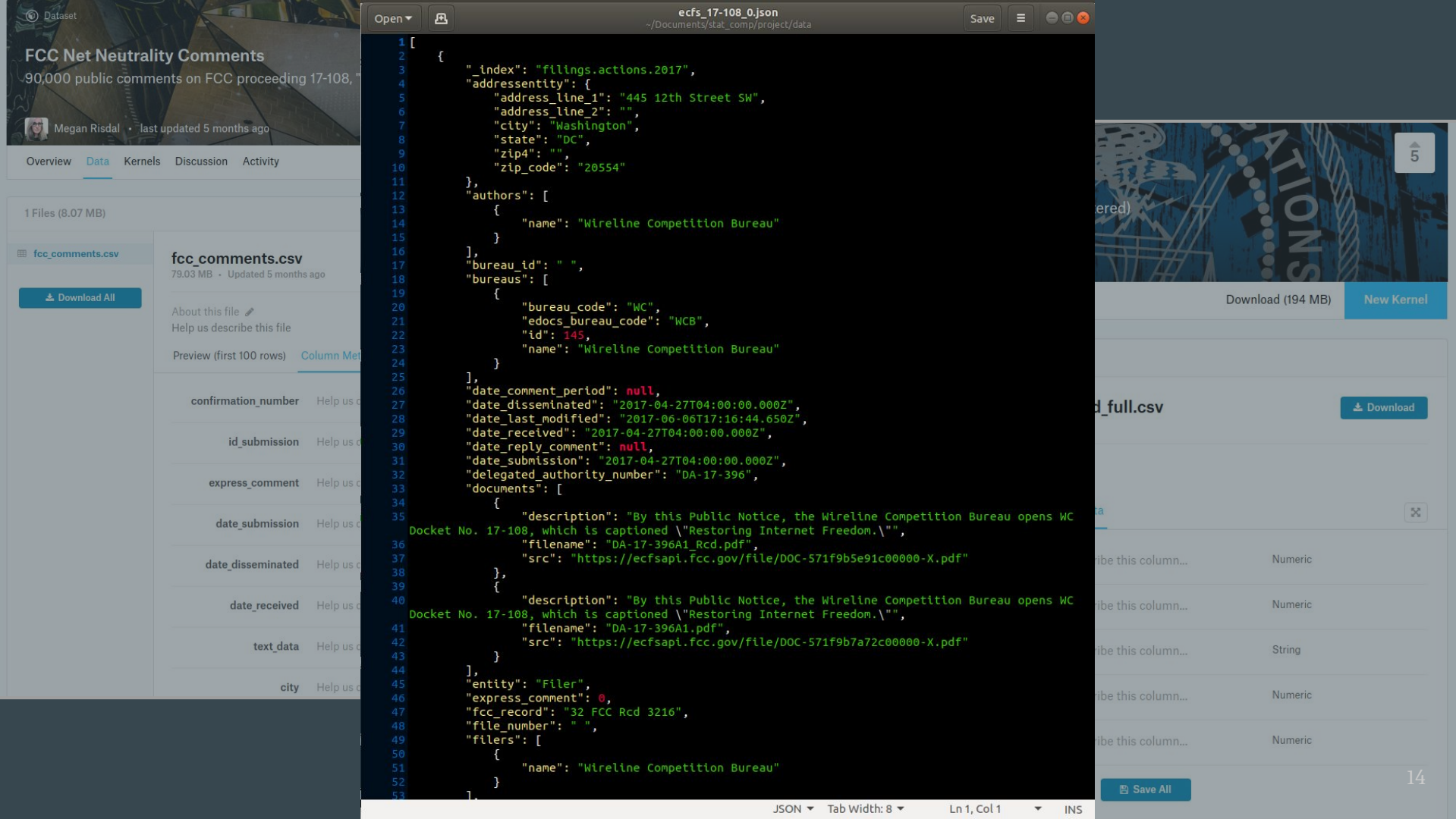
Download

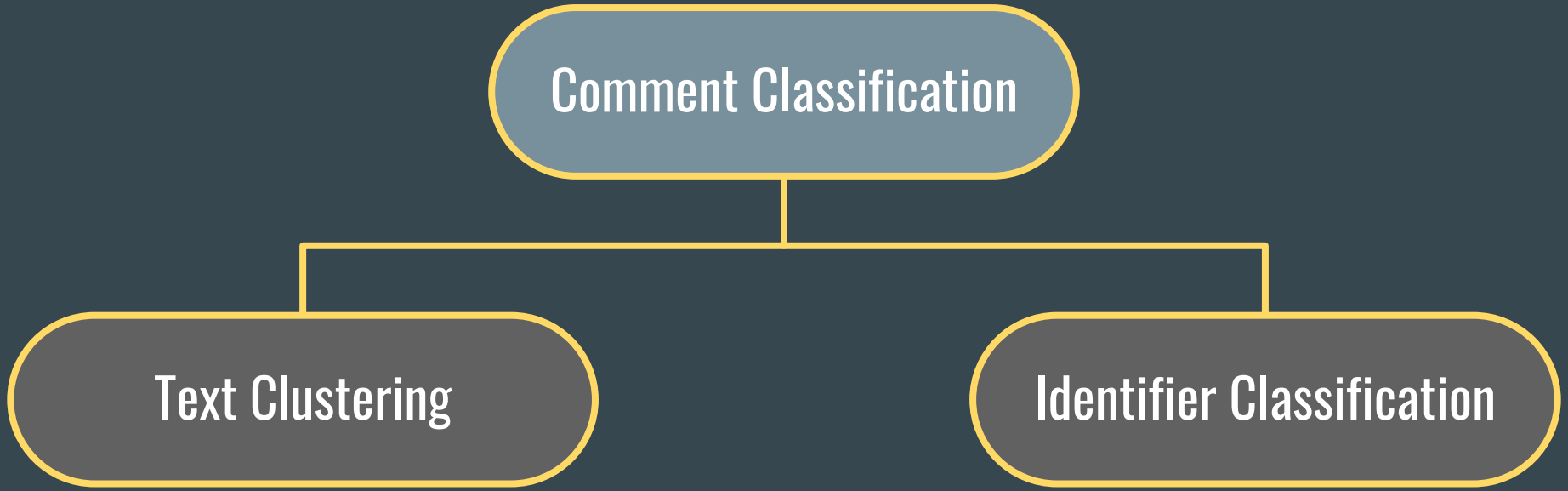
About this file
Help us describe this file

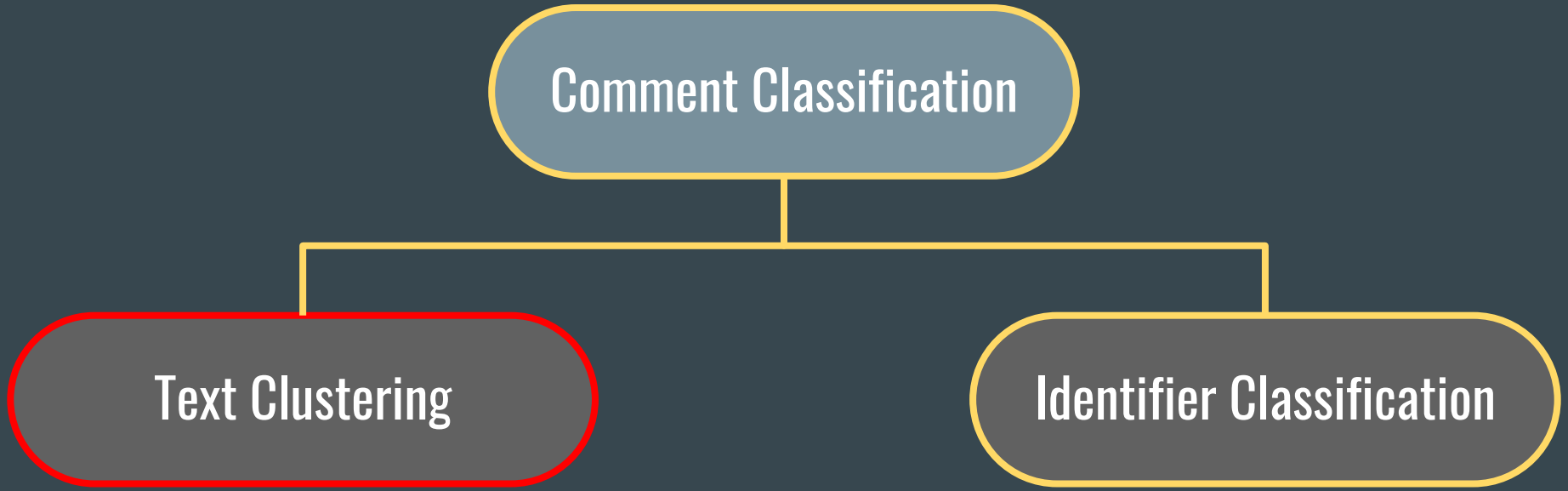
Preview (first 100 rows) Column Metadata

level_0	Help us describe this column...	Numeric
level_1	Help us describe this column...	Numeric
text_data	Help us describe this column...	String
docid	Help us describe this column...	Numeric
dupe_count	Help us describe this column...	Numeric

Save All


Dataset
FCC Net Neutrality Comments
90,000 public comments on FCC proceeding 17-108, "
Megan Risdal • last updated 5 months ago
Overview Data Kernels Discussion Activity
1 Files (8.07 MB)
fcc_comments.csv
Download All
fcc_comments.csv
79.03 MB • Updated 5 months ago
About this file
Help us describe this file
Preview (first 100 rows) Column Met
confirmation_number
id_submission
express_comment
date_submission
date_disseminated
date_received
text_data
city
Open
ecfs_17-108_0.json
~/Documents/stat_comp/project/data
Save
[
{
"_index": "filings.actions.2017",
"addressentity": {
"address_line_1": "445 12th Street SW",
"address_line_2": "",
"city": "Washington",
"state": "DC",
"zip4": "",
"zip_code": "20554"
},
"authors": [
{
"name": "Wireline Competition Bureau"
}
],
"bureau_id": " ",
"bureaus": [
{
"bureau_code": "WC",
"edocs_bureau_code": "WCB",
"id": 145,
"name": "Wireline Competition Bureau"
}
],
"date_comment_period": null,
"date_disseminated": "2017-04-27T04:00:00.000Z",
"date_last_modified": "2017-06-06T17:16:44.650Z",
"date_received": "2017-04-27T04:00:00.000Z",
"date_reply_comment": null,
"date_submission": "2017-04-27T04:00:00.000Z",
"delegated_authority_number": "DA-17-396",
"documents": [
{
"description": "By this Public Notice, the Wireline Competition Bureau opens WC Docket No. 17-108, which is captioned \"Restoring Internet Freedom.\"",
"filename": "DA-17-396A1_Rcd.pdf",
"src": "https://ecfsapi.fcc.gov/file/DOC-571f9b5e91c00000-X.pdf"
},
{
"description": "By this Public Notice, the Wireline Competition Bureau opens WC Docket No. 17-108, which is captioned \"Restoring Internet Freedom.\"",
"filename": "DA-17-396A1.pdf",
"src": "https://ecfsapi.fcc.gov/file/DOC-571f9b7a72c00000-X.pdf"
}
],
"entity": "Filer",
"express_comment": 0,
"fcc_record": "32 FCC Rcd 3216",
"file_number": " ",
"filers": [
{
"name": "Wireline Competition Bureau"
}
],
JSON Tab Width: 8 Ln 1, Col 1 INS
Download (194 MB)
New Kernel
d_full.csv
Download
Numeric
Numeric
String
Numeric
Numeric
Save All
14


Comment Classification
Text Clustering
Identifier Classification


Comment Classification
Text Clustering
Identifier Classification

Text Clustering

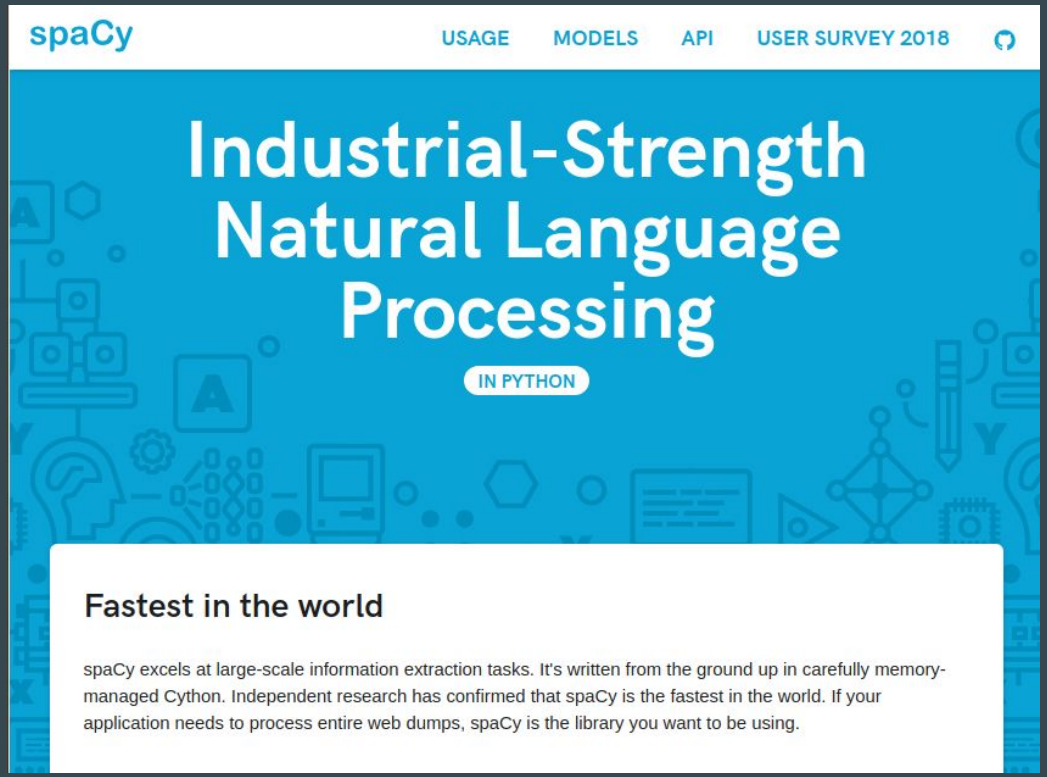
- Vectorize comment text
- Cluster based on text vectors
- Identify signature strings in clusters to determine their stance

Text Clustering

- **Vectorize comment text**
- Cluster based on text vectors
- Identify signature strings in clusters to determine their stance

spaCy

- Tool for Natural Language Processing (NLP)
- Written in cython—a C extension for Python to improve performance
- Amongst other things, can convert text to vectors



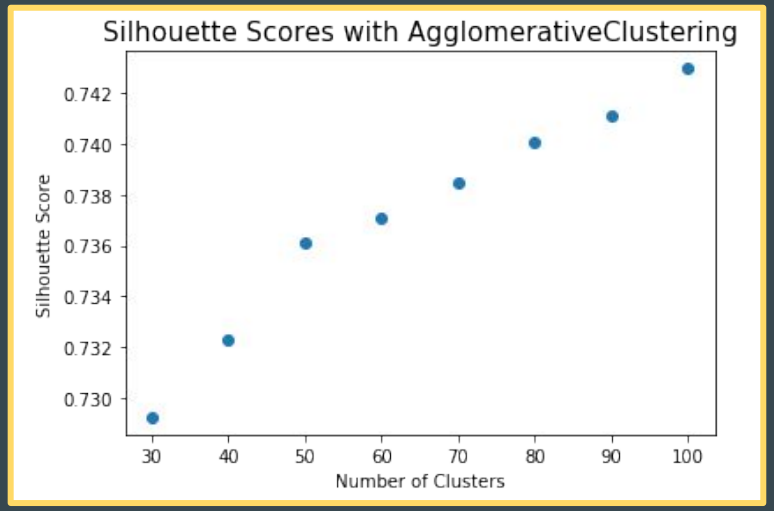
https://www.analyticsvidhya.com/blog/2017/04/natural-language-processing-made-easy-using-spacy

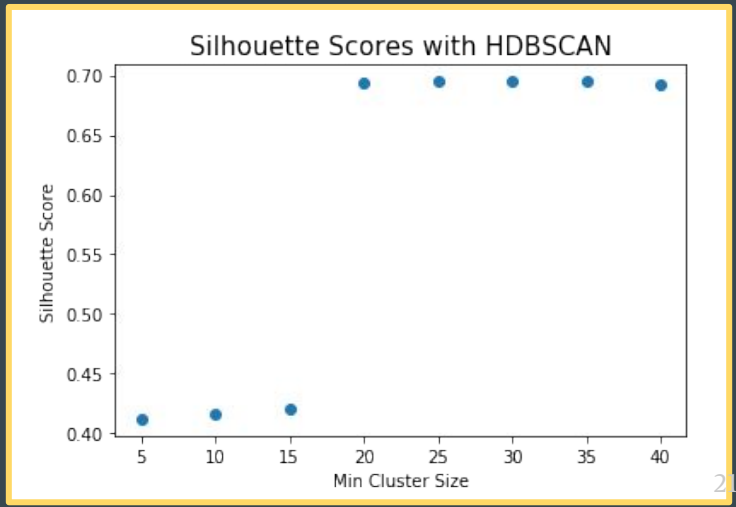
Text Clustering

- Vectorize comment text
- **Cluster based on text vectors**
- Identify signature strings in clusters to determine their stance

Clustering

- Compared clusterers Kao used:
  - AgglomerativeClustering
  - HDBSCAN
- Used silhouette scores to evaluate clusterers (scores can range from -1 to 1; higher scores indicate better clustering)
- Chose AgglomerativeClustering with 50 clusters

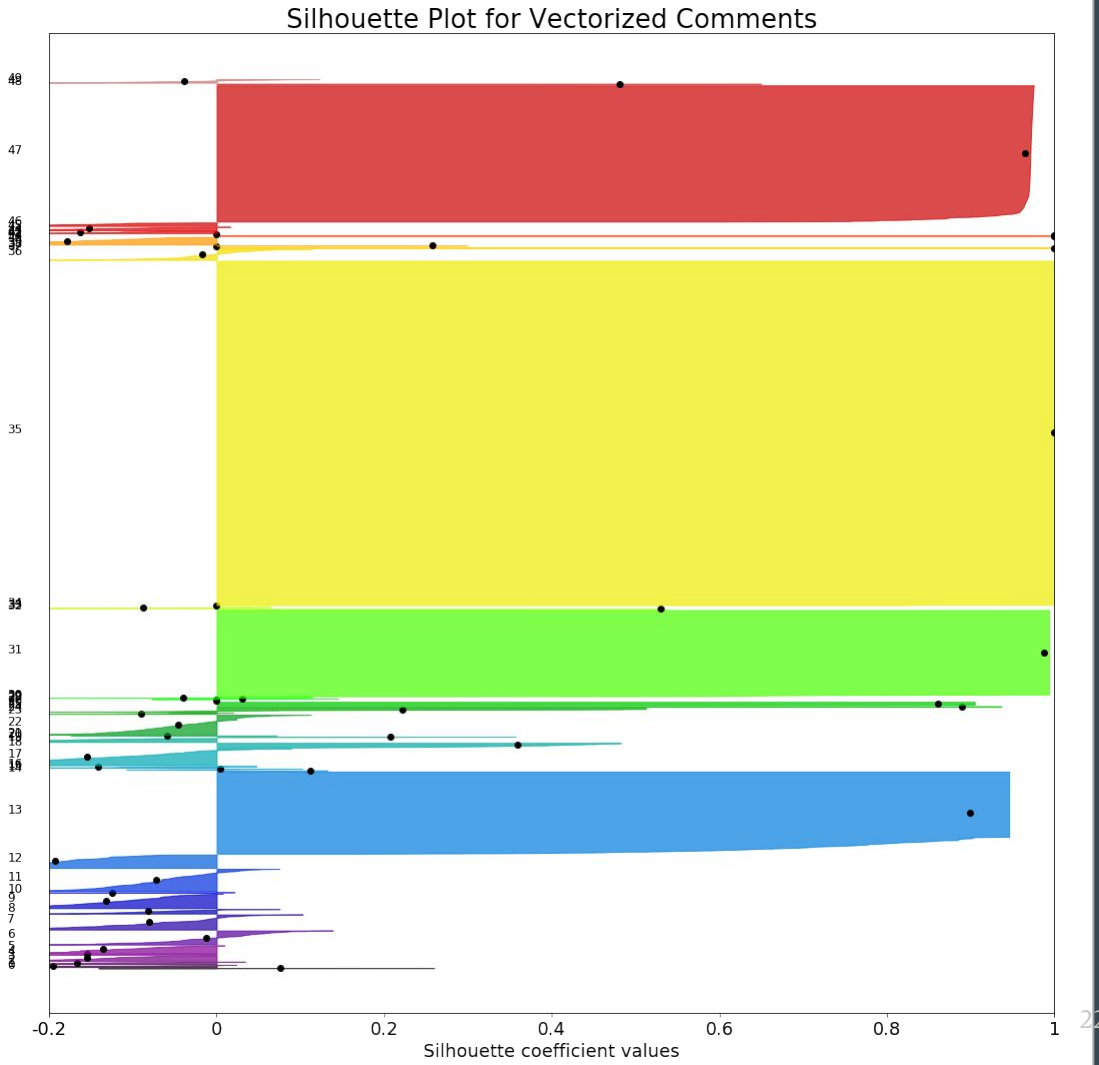




Clustering

From scikit-learn: “The silhouette plot displays a measure of how close each point in one cluster is to points in the neighboring clusters and thus provides a way to assess parameters like number of clusters visually.”

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$



Text Clustering

- Vectorize comment text
- Cluster based on text vectors
- **Examine comments in clusters to determine their stance**

Clustering

Cluster 0: 4 comments

	Unnamed: 0	cluster_id	confirmation_number	text_data
452	452	0	201704282250600103	NO NO NO on this Bill...
5221	5221	0	20170428296512159	No NOT end Net neutrality!
5294	5294	0	20170428008619682	No NOT end Net neutrality!
8781	8781	0	20170429216947370	Leave my intranetz alone!\r\nChristopher

Cluster 35: 4062 comments

	Unnamed: 0	cluster_id	confirmation_number	text_data
1290	1290	35	20170428073789141	I was outraged by the Obama/Wheeler FCC's deci...
1291	1291	35	20170428047230153	I was outraged by the Obama/Wheeler FCC's deci...
1292	1292	35	20170428551902680	I was outraged by the Obama/Wheeler FCC's deci...
1293	1293	35	20170428487725518	I was outraged by the Obama/Wheeler FCC's deci...
1294	1294	35	201704280288329190	I was outraged by the Obama/Wheeler FCC's deci...

(very much anti-net neutrality)

Clustering

Cluster 47: 1607 comments

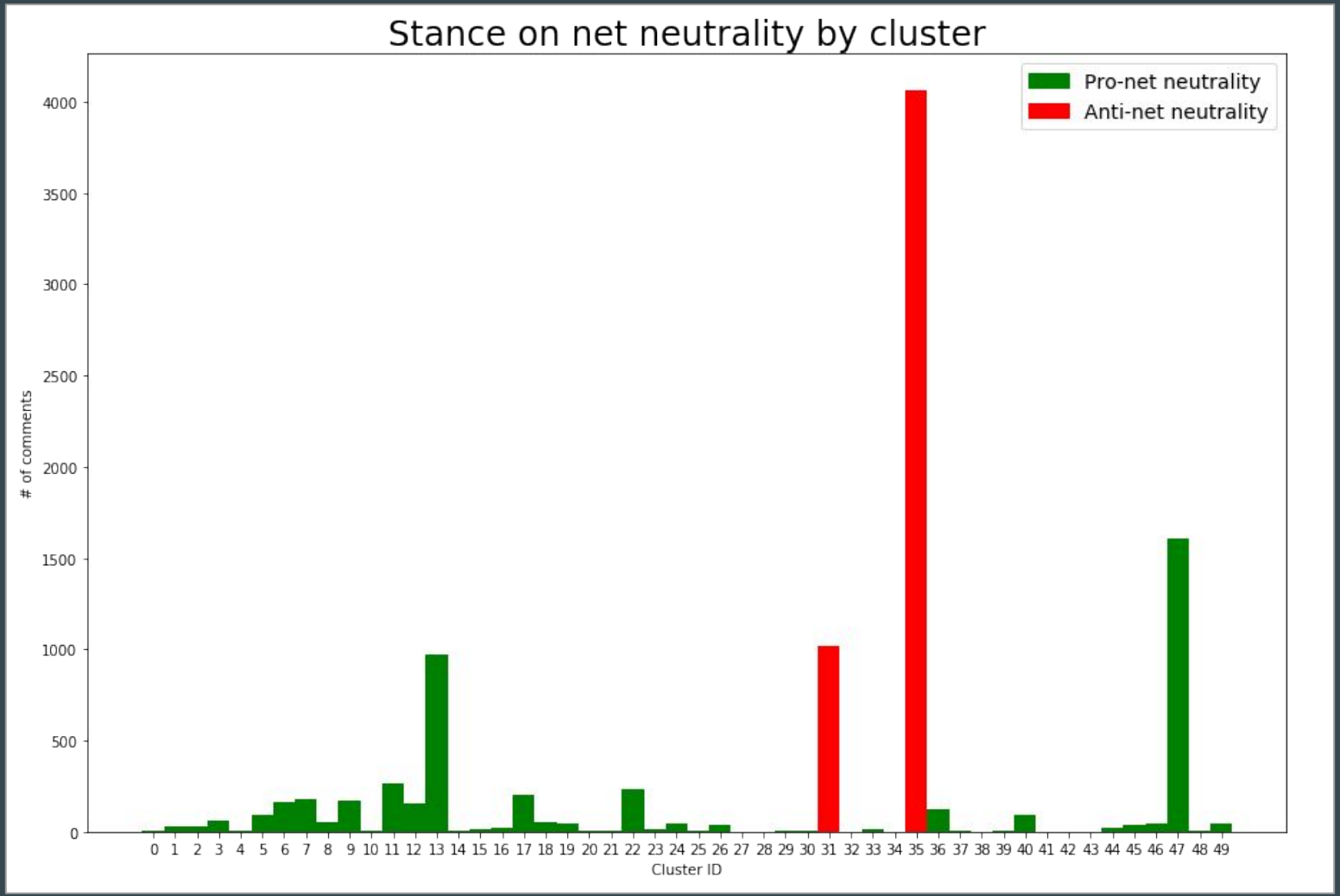
	Unnamed: 0	cluster_id	confirmation_number	text_data
3430	3430	47	20170428790410777	The FCC Open Internet Rules (net neutrality ru...
5358	5358	47	201704282450902566	The FCC Open Internet Rules (net neutrality ru...
5411	5411	47	2017042823303654	The FCC Open Internet Rules (net neutrality ru...
5412	5412	47	2017042867624899	The FCC Open Internet Rules (net neutrality ru...

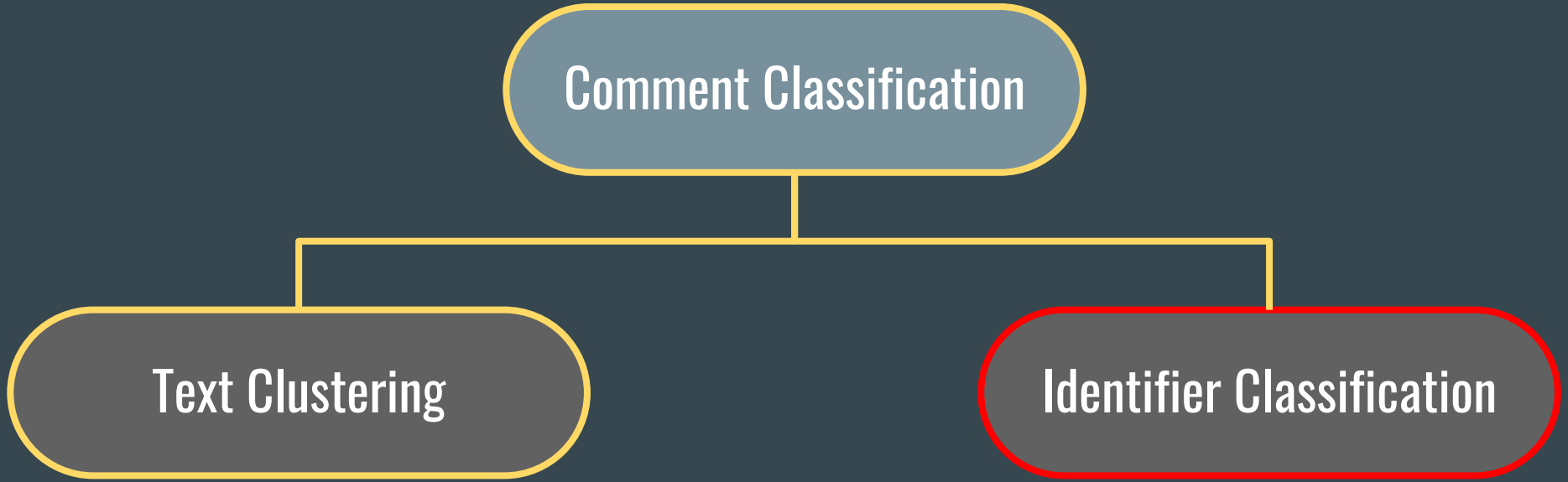
Cluster 31: 1016 comments

	Unnamed: 0	cluster_id	confirmation_number	text_data
16	16	31	20170427291878313	Obama's Title II order has diminished broadban...
17	17	31	20170427069381932	Obama's Title II order has diminished broadban...
18	18	31	20170427209520206	Obama's Title II order has diminished broadban...
19	19	31	2017042744785141	Obama's Title II order has diminished broadban...
20	20	31	20170427141621159	Obama's Title II order has diminished broadban...

Cluster 13: 969 comments

	Unnamed: 0	cluster_id	confirmation_number	text_data
5914	5914	13	201704292338019826	It's my understanding that the FCC Chairman in...
5918	5918	13	201704292706108323	It's my understanding that the FCC Chairman in...
5919	5919	13	20170429076675752	It's my understanding that the FCC Chairman in...
5922	5922	13	201704291732305323	It's my understanding that the FCC Chairman in...
5931	5931	13	20170429188152819	It's my understanding that the FCC Chairman in...


Stance on net neutrality by cluster
Pro-net neutrality
Anti-net neutrality
# of comments
Cluster ID
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49
0 500 1000 1500 2000 2500 3000 3500 4000



- Text vectorization
- Agglomerative clustering
- Sentiment analysis

Identifier Classification : Motivation

Comments were from a wide variety of sources:

Type	Most Likely Filer(s)
Unique	Individuals
Form Letters	Individuals, organizations, advocacy groups, bots?

Identifier Classification : Motivation

Type	Most Likely Filer(s)
Form Letters	Individuals, organizations, advocacy groups, bots?

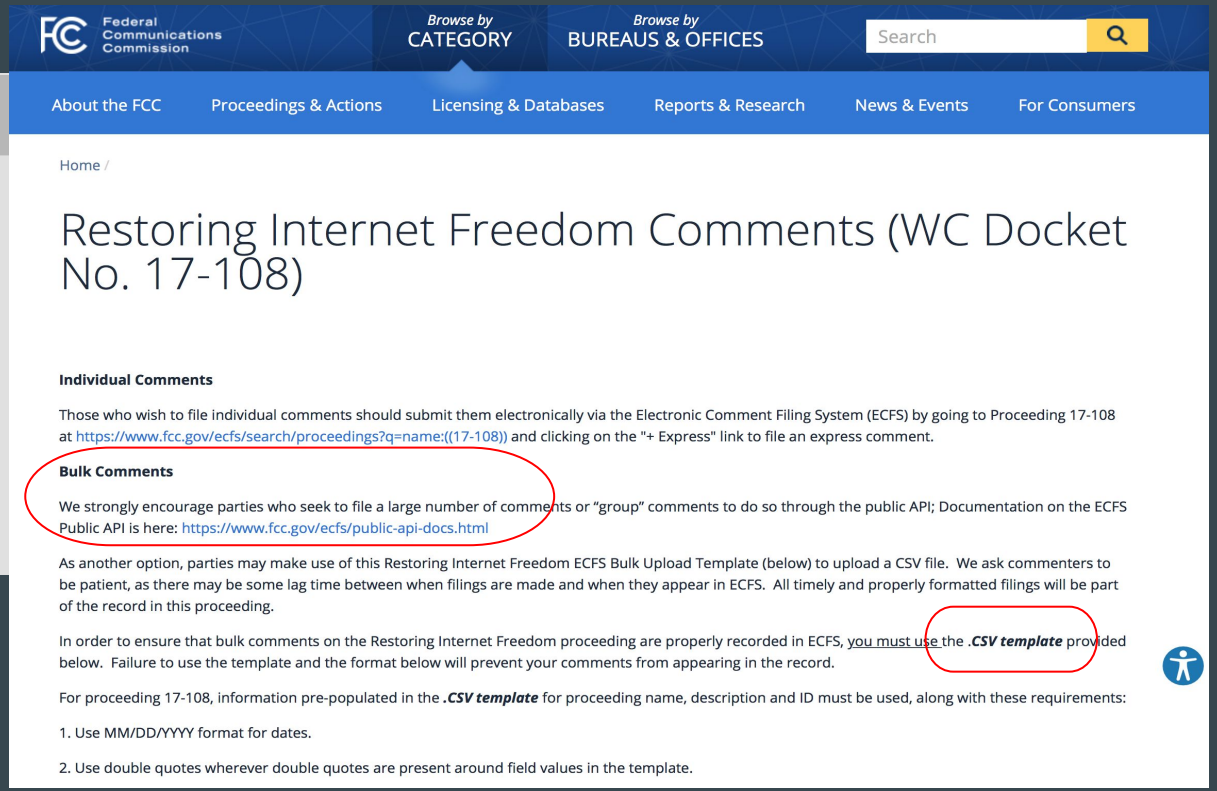
Comment count : ~ 70,000 (out of a set of 100,000)

"The unprecedented regulatory power the Obama Administration imposed on the internet is smothering innovation, damaging the American economy and obstructing job creation..."

Identifier Classification : Motivation

Comments were from a wide variety of sources:

Type
Unique
Form Letters
API submissions
CSV submissions



Identifier Classification : Motivation

Comments were from a wide variety of sources:

Type	Most Likely Filer(s
Unique	Individuals
Form Letters	Individuals, organizations, advocacy groups, bots?
API submissions	Organizations, advocacy groups, bots
CSV submissions	Organizations and groups, certain kinds of workers

Identifier Classification

- Gathering identifying information
- Generating “spamminess” metric
- Analyzing trends based on metric and clustering

Identifier Classification

- **Gathering identifying information**
- Generating “spamminess” metric and model
- Analyzing trends based on metric and clustering

Identifier Classification : Gathering Identifying Information

Brief Comment	Before leaving office, the Obama Administration rammed through a massive scheme that gave the federal government broad regulatory control over the internet. That misguided policy decision is threatening innovation and hurting broadband investment in one of the largest and most important sectors of the U.S. economy. I support the Federal Communications Commissions decision to roll back Title II and allow for free market principles to guide our digital economy.

Identifier Classification : Gathering Identifying Information

Filing Detail

ID	19108101515230907	**Proceeding**	17-108
Name of Filer	beyonce rihanna		
Type of Filing	COMMENT	**Filing Status**	DISSEMINATED
Viewing Status	Unrestricted		
Date Received	Aug 4, 2017	**Date Posted**	Aug 10, 2017
Address	12 n barton trl	**City**	Batavia
State	IL	**ZIP**	60510

Unrealistic name

Address doesn't exist

Brief Comment

Before leaving office, the Obama Administration rammed through a massive scheme that gave the federal government broad regulatory control over the internet. That misguided policy decision is threatening innovation and hurting broadband investment in one of the largest and most important sectors of the U.S. economy. I support the Federal Communications Commissions decision to roll back Title II and allow for free market principles to guide our digital economy.

Identifier Classification : Gathering Identifying Information

Filing Detail

ID	19108101515230907	**Proceeding**	17-108
Name of Filer	beyonce rihanna		
Type of Filing	COMMENT	**Filing Statu**	
Viewing Status	Unrestricted		
Date Received	Aug 4, 2017	**Date Poste**	
Address	12 n barton trl	**Cit**	
State	IL	**ZIP**	60510

Brief Comment Before leaving office, the Obama Administration rammed through a massive scheme that gave the federal government broad regulatory control over the internet. That misguided policy decision is threatening innovation and hurting broadband investment in one of the largest and most important sectors of the U.S. economy. I support the Federal Communications Commissions decision to roll back Title II and allow for free market principles to guide our digital economy.

Identifier Classification : Gathering Identifying Information

Filing Detail

ID	10509156931656	**Proceeding**	17-108
Names of Filers	The Flash,Batman,Wonder Woman,Aquaman,Cyborg,Superman		
Type of Filing	COMMENT	**Filing Status**	DISSEMINATED
Viewing Status	Unrestricted		
Date Received	May 9, 2017	**Date Posted**	May 9, 2017
Address	Central City	**City**	Central City
State	PA	**ZIP**	17331
Brief Comment	Thats just a dumb thing to do		

Identifier Classification : Gathering Identifying Information

- Is the name realistic?
- Are the address, state, and city realistic?

Tool : spaCy NER

Identifier Classification : Gathering Identifying Information

- Is the name realistic?
- Are the address, state, and city realistic?
- Email confirmation?
- Do the name and email address match up?

Name: DONALD DEFOUW

Email: deerhunterd42@hotmail.com

Identifier Classification : Gathering Identifying Information

- Is the comment from a form letter?
  - 25 form letters
- Does the person's information appear in any of the following:
  - Lists of celebrities and popular people
  - Impersonated people
  - Data breaches

Identifier Classification : Gathering Identifying Information

Sources:

- Celebrity list : IMDB

Identifier Classification : Gathering Identifying Information

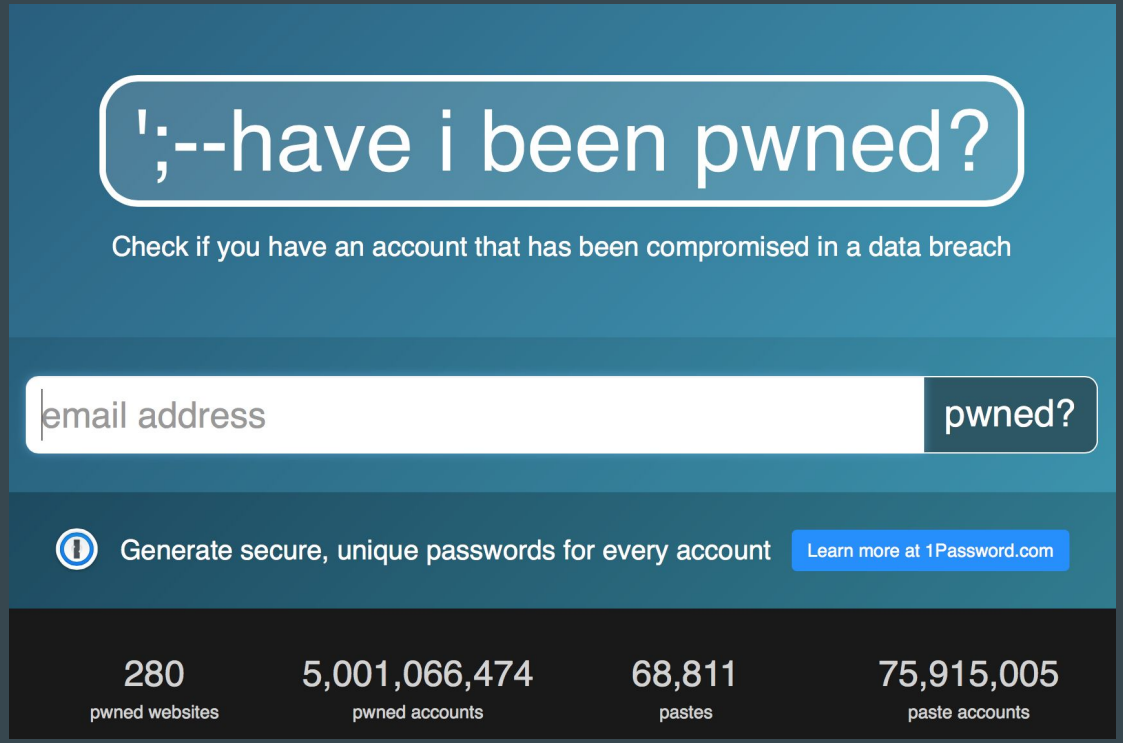
Sources:

- Celebrity list : IMDB
- Impersonated people : Fight for the Future

Identifier Classification : Gathering Identifying Information

Sources:

- Celebrities
- Impersonated
- Breached accounts



Identifier Classification

- Gathering identifying information
- **Generating “spamminess” metric**
- Analyzing trends based on metric and clustering

Identifier Classification : Generating a “Spamminess” Metric

- Realistic name
- Email and name match up
- Email confirmation

- Breached account
- Celebrity
- Makes impersonated list

Identifier Classification : Generating a “Spamminess” Metric

- Natural model : Bernoulli Trial

$$X_i \sim Bernoulli(p)$$

- Several trials ~ Binomial (n, p)
  - Parameters :
    - n (number of trials)
    - p (success probability in each trial)

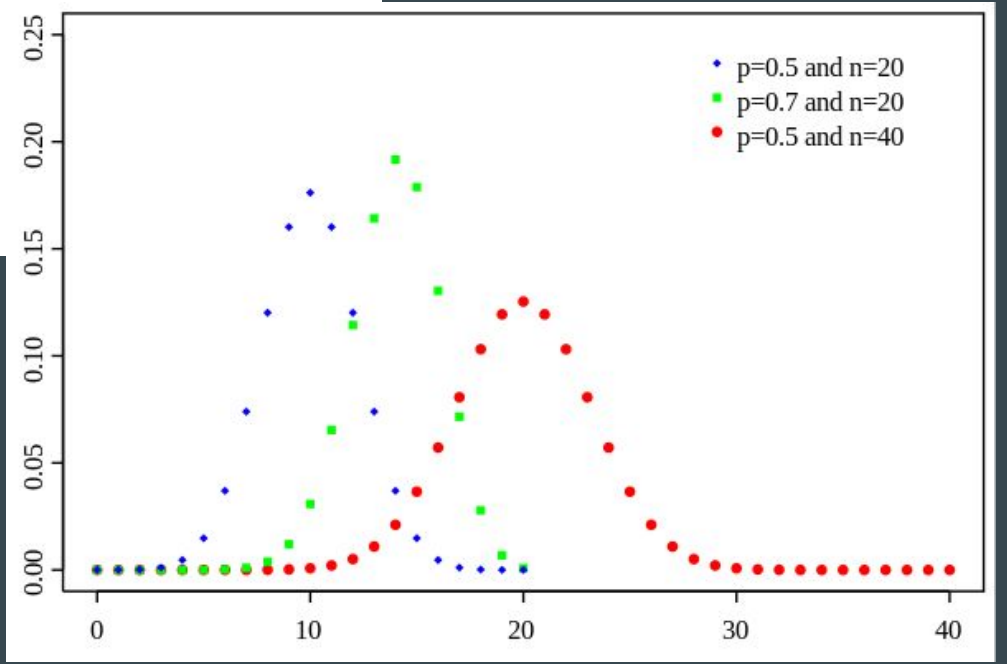
Identifier Classification : Generating a “Spamminess” Metric

$$Pr(k; n, p) = \Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

for $k = 0, 1, 2, \ldots, n$, where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Source : Wikipedia



Identifier Classification : Generating a “Spamminess” Metric

- To make p intuitive:
  - Positive feature : 1 if yes, 0 if no
  - Negative feature : 0 if yes, 1 if no

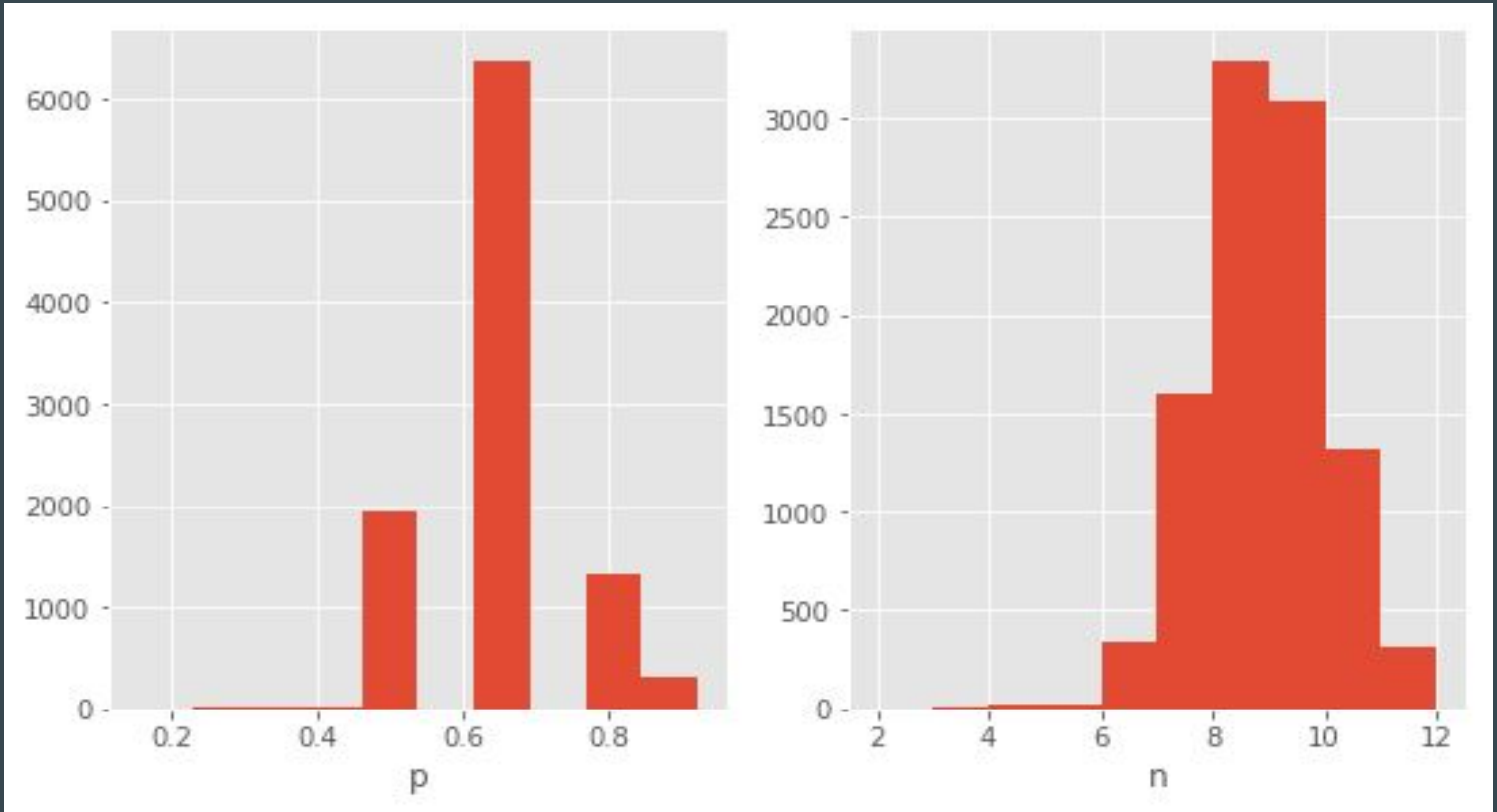
- Example:
  - Realistic name : 1
  - Fake account : 0
  - p = ½

First Name : Jacqueline
Last Name : Antwi-Danso
Breached : yes

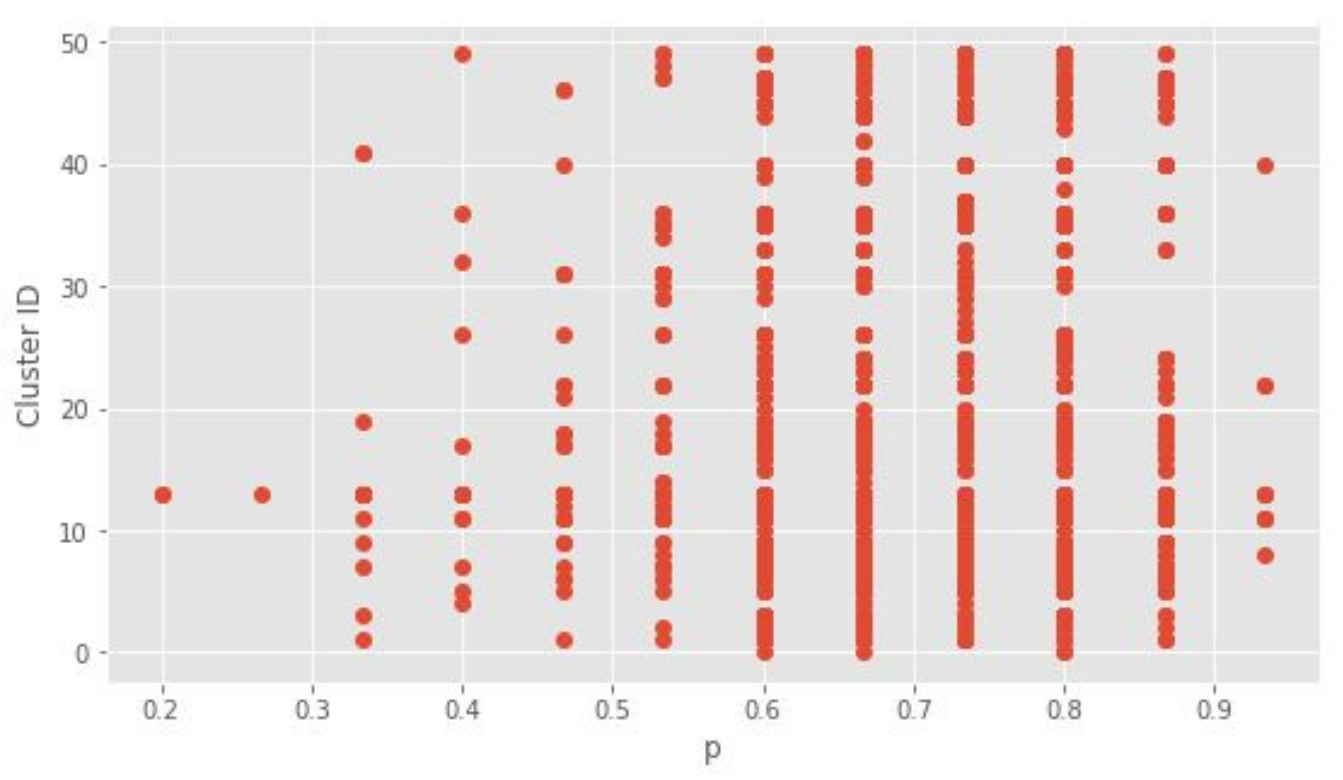
Identifier Classification

- Gathering identifying information
- Generating “spamminess” metric and model
- **Analyzing trends based on metric and clustering**

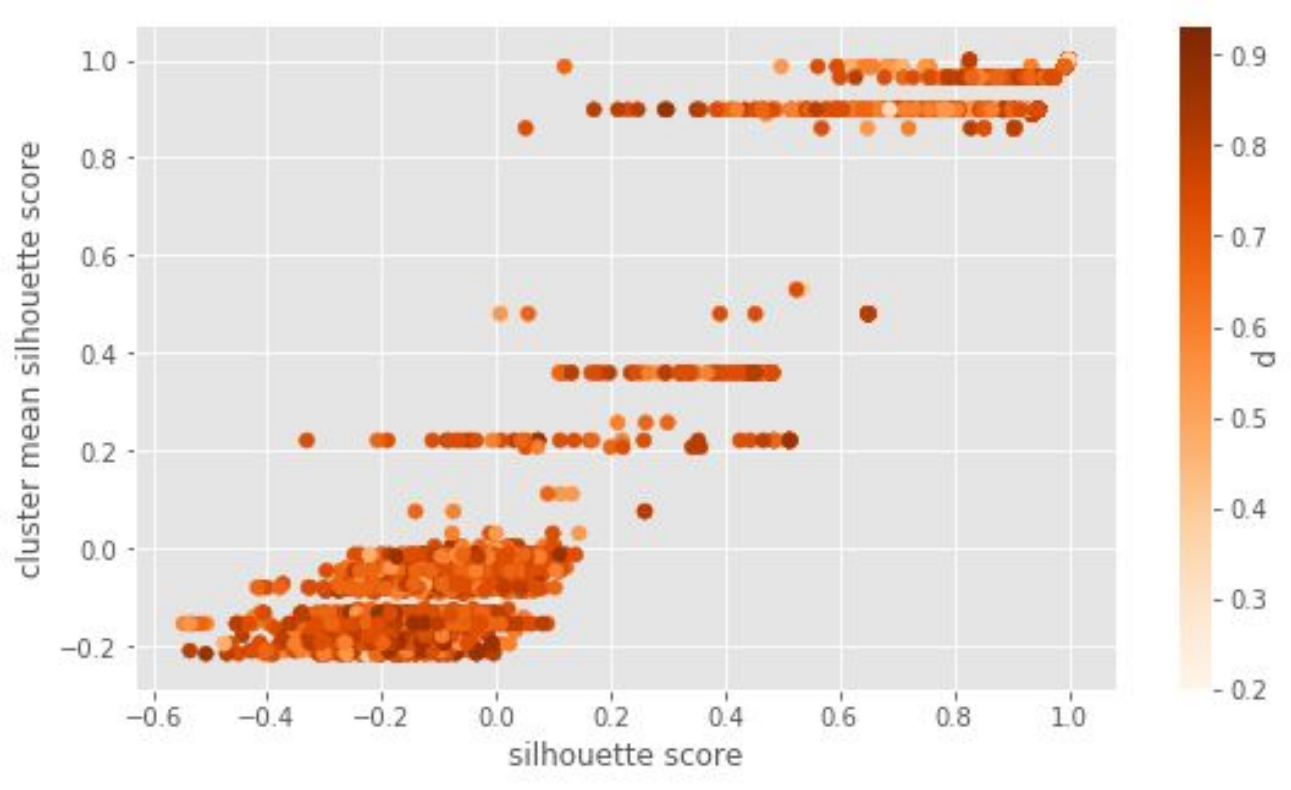
Identifier Classification : Analyzing trends



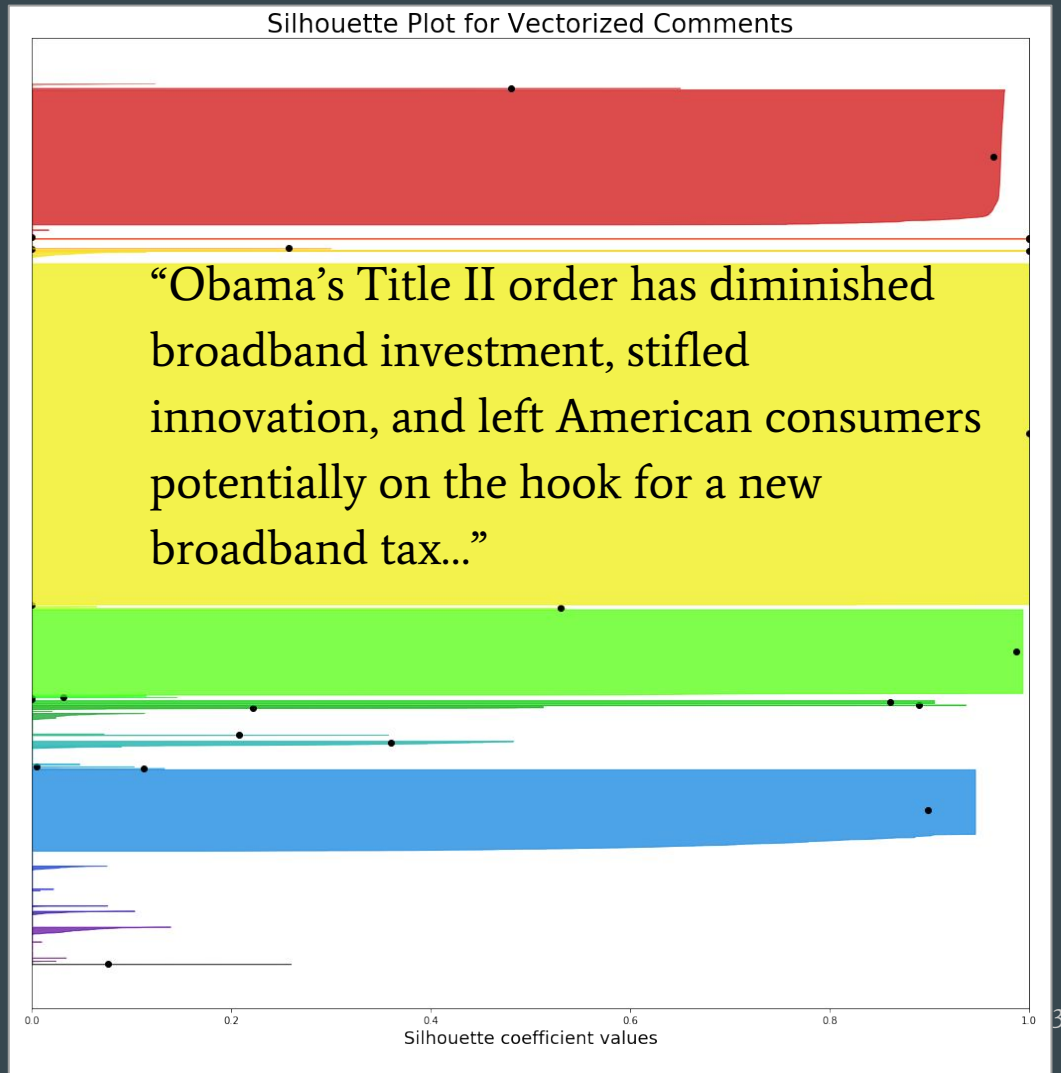
Identifier Classification : Analyzing trends



Identifier Classification : Analyzing trends



Identifier Classification : Analyzing Trends



Identifier Classification : Analyzing Trends

```
Name: DONALD DEFOUW

Email: deerhunterd42@hotmail.com
Address: 2353 West Lenoir Dr.  Lenoir NC 28645
Breached: 0
p: 0.533333333333
silhouette score: 0.993535362778
Email confirmation:  nan
Comment: Obama's Title II order has diminished broadband
```

```
Name: Terrence  Pugh
Email: tmjpugh@gmail.com
Address: 911 E Roosevelt st  Weslaco TX 78599
Breached: 0
p: 0.666666666667
silhouette score: 0.030075112596
Email confirmation:  True
Comment: I am small business owner in IT commerce space. I depend on
ovide pro Bono education to k-12 learner's interested in STEM. Allow
```

Limitations of Analysis and Improvements

Limitation

- Binomial model may be too simplistic

Possible Remedy

- Give some features more weight than others
- Employ Bayesian inference

Limitations of Model and Improvements

Limitation

- Too simplistic
- **Inaccurate with little information**

Possible Remedy

- Give some features more weight than others
- Employ Bayesian inference
- **Obtain more identifying data (e.g. census records)**

Limitations of Model and Improvements

Limitation

- Too simplistic
- Inaccurate with little information
- NLP and sentiment analysis accurate 85% of the time

Possible Remedy

- Give some features more weight than others
- Employ Bayesian inference
- Obtain more identifying data (e.g. census records)
- **Improve spaCy by training on more data.**

Limitations of Model and Improvements

Limitation

- Too simplistic
- Inaccurate with little information
- **NLP and sentiment analysis accurate 85% of the time**
- **No metric for cluster assignments**

Possible Remedy

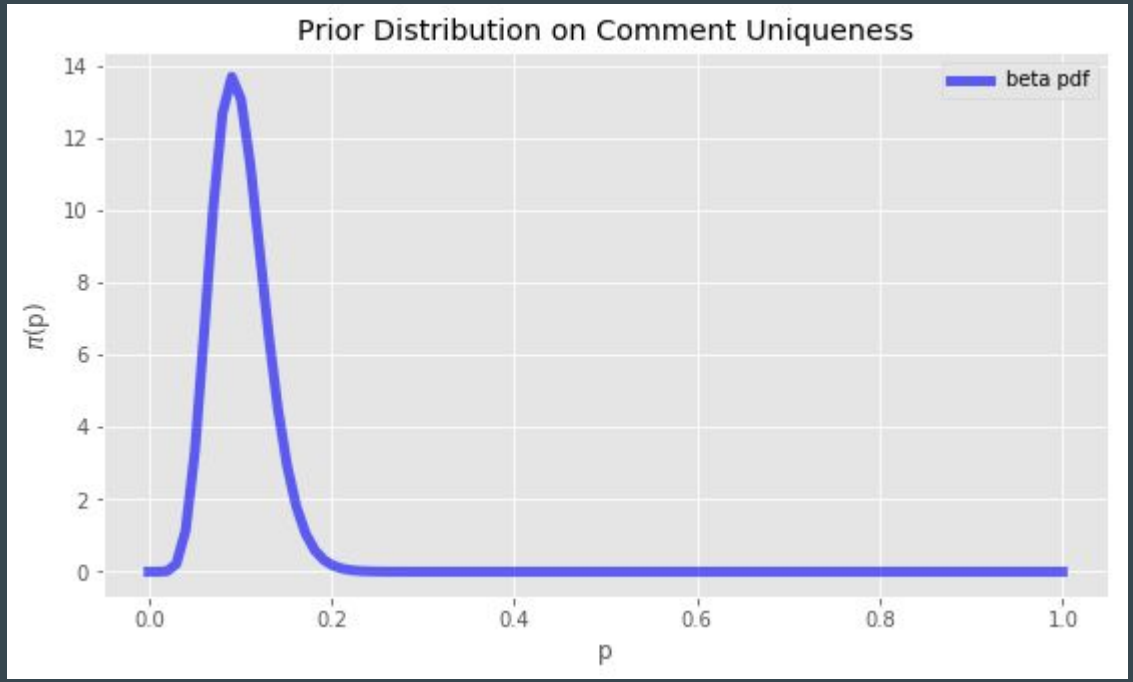
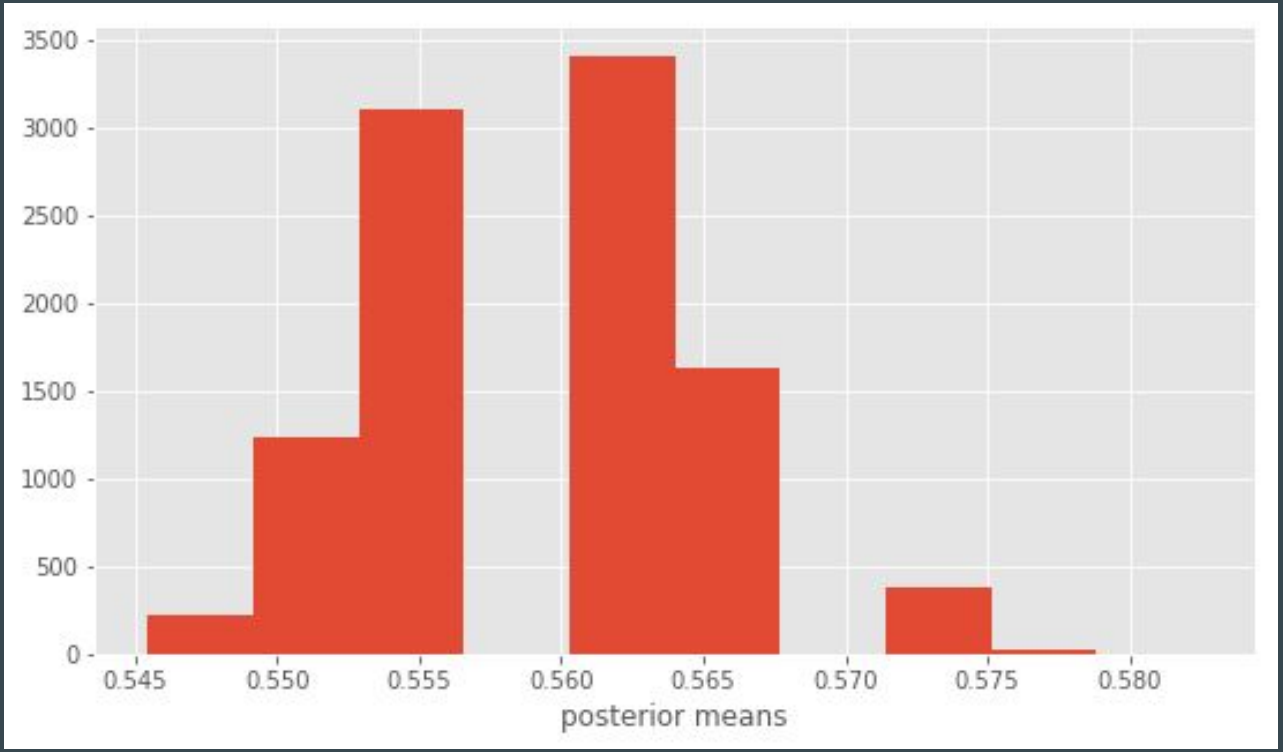
- Give some features more weight than others
- Employ Bayesian inference
- Obtain more identifying data (e.g. census records)
- Improve spaCy by training on more data.
- **Use metric for model selection e.g. AIC**

Analyzing Public Comments to the FCC on Net Neutrality

...

Jacqueline Antwi-Danso and Tarini Konchady

Bayesian Analysis



Bayesian Analysis