

Polarization, Partisanship and Junk News Consumption over Social Media in the US

[COMPPOP](#) DATA MEMO 2018.1 – Online Supplement / FEBRUARY 6, 2018

Vidya Narayanan
Oxford University
vidya.narayanan@oii.ox.ac.uk
[@vidunarayanan](https://twitter.com/vidunarayanan)

Vlad Barash
Graphika
vlad.barash@graphika.com
[@vlad43210](https://twitter.com/vlad43210)

John Kelly
Graphika
john.kelly@graphika.com
[@apidictionist](https://twitter.com/apidictionist)

Bence Kollanyi
Oxford University
bence.kollanyi@oii.ox.ac.uk
[@bencekollanyi](https://twitter.com/bencekollanyi)

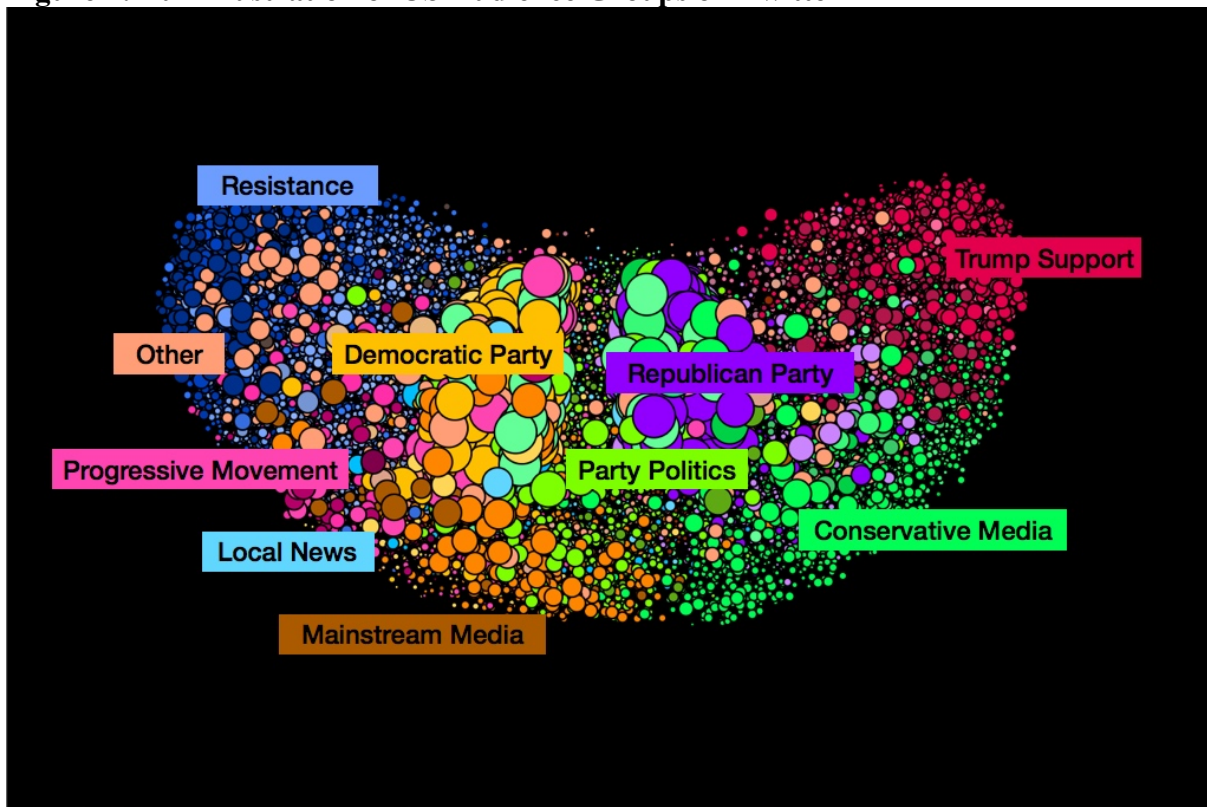
Lisa-Maria Neudert
Oxford University
lisa-maria.neudert@oii.ox.ac.uk
[@lmneudert](https://twitter.com/lmneudert)

Philip N. Howard
Oxford University
philip.howard@oii.ox.ac.uk
[@pnhoward](https://twitter.com/pnhoward)

1. US Audience Groups on Twitter

Each node in this network (Figure 1) represents an account on Twitter. Each node belongs to both a broad group and a smaller segment within that group. A segment is a collection of nodes with a shared pattern of interest while a group is a collection of segments that are politically, culturally, or socially similar. The size of each node is proportional to the number of other nodes that follow it on Twitter. The color of each node is based on its parent segment. Figure 1 is the full color visualization of the US audience groups we found on Twitter.

Figure 1: Full Illustration of US Audience Groups on Twitter



Source: Authors' calculations from data sampled 20/10/17-20/01/18. Note: Groups are determined through network association and our interpretation of the kinds of content these users distribute.

The nodes are placed within the map using a Fruchterman-Reingold visualization algorithm. This works to place nodes on the map according to two principles: first, a “centrifugal force” acts upon each node to push it to the edge of the canvas; second, a “cohesive force” acts upon

every connected pair of nodes to pull them closer together. Table 1 gives a full list of Twitter groups and associated segments considered in this study.

Table 1: Groups and Segments for the Twitter Visualization

Groups	Segments
Conservative Media	Conservative Blogs Conservative Truth GOP Communications Hard Right Journalists Right Wing Media Watch Dogs
Democratic Party	House Democrats State Democrats
Local News	Bay Area News Health News NYC News Texas Journalists
Mainstream Media	Left Wing Political Journalists News Anchors
Other	Brand Digital Marketing EU, Euro Organizations SMM Tech/Finance
Party Politics	Bi-Partisan Research GOP, Democrats Communications House – GOP US Congress US Senate
Progressive Movement	African American Culture and Empowerment Latino Empowerment and Non-profits Race and Grassroots Wage reform and Economic Justice
Republican Party	GOP Organized GOP
Resistance	Anti-Trump Political Analysts Anti-Trump, Human Rights, Humor Impeach Trump Pro Hilary – I’m with Her Resist – Analysts Resist – Trump Resist – Equal Rights Resist – Progressive The Resistance – Not My President Unite Democrats Women’s Issues
Trump Support	Defend Trump/Lifestyle/Humor Patriots for Trump Re-elect Trump Trump Supporters MAGA Trump Train Trump Victory Trump Victory Party

2. US Audience Groups on Facebook

Each node in this network (Figure 2) represents a public page on Facebook. The size of each node corresponds to the number of other nodes that like the page on Facebook. Each node belongs to both a broad group and a smaller segment within that group. A segment is a

collection of nodes with a shared pattern of interest while a group is a collection of segments that are politically, culturally, or socially similar.

Figure 2: Full Illustration of US Audience Groups on Facebook



Source: Authors’ calculations from data sampled 20/10/17-20/01/18. Note: Groups are determined through network association and our interpretation of the kinds of content these users distribute.

Again, a Fruchterman-Reingold visualization algorithm is used to place nodes within the map. Table 2 gives a full list of the Facebook groups and associated segments considered in this study.

Table 2: List of Groups and Segments for the Facebook Visualization

Groups	Segments
Conspiracy	Conspiracy Theories
	Infowars
	International anti-media conspiracy
	Truthers
Democratic Party	Anti-GOP
	Atheism
	Democratic Humor
	Democratic Memes
	Democratic Party
	House Democrats
	Non-profit campaigns

Groups	Segments
Environmental Movement	Animal rights Animal, wilderness conservation Anti-fracking, eco defense Anti-GMO International climate action International Greenpeace Native American, eco resistance
Hard Conservatives	Anti-immigration Anti-liberal Hard Conservatives Media Patriots, Pro-Military, Militia
Libertarians	Anarchist Libertarian Institutes
Military/Guns	Guns, Guns Ownership US Army, Navy, Veterans
Occupy	International Anarchist International Anonymous International Occupy chapters International Occupy movement US Occupy movement
Other Left	ACLU Grassroots movements Immigrant Advocacy Labor reform Organized Labor, Trade Unions US and International LGBT advocacy
Other Non-Political	Food blogs Life coach, inspirational groups Miscellaneous Influencers Music, music clusters, SMM Parenting Spiritualism Un-clustered US and International Art, Art museums
Public Health	Global Health Health organizations Public initiatives US State Departments, US Consulates
Republican Party	GOP Congress GOP Senate
Sustainable Farming	Gardening sustainability Healthy eating bloggers Organic farming Organic food Small farming, self-sufficiency Survivalist Sustainable agriculture and research
Women's Rights	Feminist groups International women's rights Planned Parenthood Reproductive rights

3. Heterophily Index

For every pairing of groups within a network map, a value of heterophily can be calculated. This is a measure of the level of connection between the groups. In order to determine this a ratio is calculated between the actual ties between two groups, compared to the expected ties

between the groups if all the accounts in the map were evenly distributed. The natural log of these ratios is then taken, along with a zero correction to create a balanced index and to ensure that all values are displayed in a positive form.

$$Ratio\ of\ Ratios_r = \frac{\frac{Connections_{pairing}}{\sum_{all\ pairings} Connections}}{\frac{Connections_{pairing}}{\sum_{all\ pairings} Connections}}$$

Expression A: Ratio of Two Ratios

This heterophily index is therefore created through a ratio of two ratios. This ratio reveals whether two nodes have about the proportion of links they should have given its size. This is displayed in Expression A, where a pairing of groups is calculated as having a measure of connections in balance with its share of all the connections.

Half the distribution of possible values from this ratio of ratios ranges from 0 to 1 (corresponding to disproportionately small share of connections in a group given its size) and the other half ranges from 1 to +infinity (a disproportionately large share of connections in a group given its size). However, by taking the natural log of the ratio of ratios the index will become more balanced: from -infinity to 0 becomes less than proportionate share, and from 0 to +infinity becomes more than proportionate share. For example, take a three-group network (A, B and C). If nodes in group A have a total of ten connections, and there are ten nodes in each group, then the expected connections between A and B will be 3.33. If, in reality, the nodes in group A actually have all ten connections to nodes in group B then this connection is stronger than expected. The heterophily score for groups A and B = $10/3.33 = 3.0$. The natural log of this is then taken along with a zero correction across the range of heterophily values.

A greater heterophily index indicates a denser pattern of connections between the two groups. It is important to note however that these scores indicate only first order connections, not second or third order connections.

4. Clustering Algorithm for Determining Groups and Segments

In order to generate segments and groups for each map it is necessary to employ a clustering algorithm. This involves first building a bipartite graph between nodes in the map and the rest of the social medium in question. This bipartite graph provides a structural similarity metric between nodes in the map. This was then used in combination with a hierarchical agglomerative clustering algorithm in order to segment a map into distinct communities. This is a ‘bottom up’ approach whereby each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy. Twitter maps are clustered based on follower-relationships, since mention-relationships have been shown to overemphasize the news cycle and salient external events. Facebook networks are clustered based on page likes.

5. K-core reduction

To identify and map the ‘discussion core’ of the most active, connected, and influential users, we performed a k-core reduction to reduce the total collected set of Twitter users from the initial data collection into a set of well-connected accounts. This produces a maximally

connected subgraph of active nodes with degree of connection at least ' k '. This degree of connection, k , can be thought of as the number of links between each node in the graph. For example, selecting a k value of 0 for the reduction does not remove any nodes from the graph, since each node must have 0 connections or greater. Selecting a k value of 1 would remove all of the nodes that have no connections to other nodes in the graph. Selecting a k value of 2 would remove all nodes with fewer than 2 connections, and so on. A value of k was selected such that the k -core consisted of 12,413 users. This value was found to be a sufficiently large group to represent the major sets of highly active users, but not so large as to make clustering and visualization impractical.

6. Junk News Classification

These sources deliberately publish misleading, deceptive or incorrect information purporting to be real news about politics, economics or culture. This content includes various forms of propaganda and ideologically extreme, hyper-partisan, or conspiratorial news and information. For a source to be labelled as junk news at least three of the following five characteristics must apply:

- **Professionalism:** These outlets do not employ the standards and best practices of professional journalism. They refrain from providing clear information about real authors, editors, publishers and owners. They lack transparency, accountability, and do not publish corrections on debunked information.
- **Style:** These outlets use emotionally driven language with emotive expressions, hyperbole, ad hominem attacks, misleading headlines, excessive capitalization, unsafe generalizations and fallacies, moving images, graphic pictures and mobilizing memes.
- **Credibility:** These outlets rely on false information and conspiracy theories, which they often employ strategically. They report without consulting multiple sources and do not employ fact-checking methods. Their sources are often untrustworthy and their standards of news production lack credibility.
- **Bias:** Reporting in these outlets is highly biased and ideologically skewed, which is otherwise described as hyper-partisan reporting. These outlets frequently present opinion and commentary essays as news.
- **Counterfeit:** These outlets mimic professional news media. They counterfeit fonts, branding and stylistic content strategies. Commentary and junk content is stylistically disguised as news, with references to news agencies, and credible sources, and headlines written in a news tone, with bylines, date, time and location stamps.

Table 3 gives a list of the all junk news sources used for this analysis.

Table 3: List of Junk News Sources

Domain Name	Example URL
100percentfedup.com	http://100percentfedup.com/wow-woman-delivers-knock-out-punch-to-michelle-obama-on-way-to-polls-yes-i-am-black-but-i-am-not-voting-democrati-am-not-on-that-plantation-video/
allenbwest.com	http://www.allenbwest.com/michellejesse/bombshell-new-email-shows-pentagon-tried-to-send-help-in-benghazi-but
americanthinker.com	http://www.americanthinker.com/articles/2016/10/what_kind_of_genius_loses_6_billions_hillary.html
anonews.co	http://www.anonews.co/hillary-clinton-exposed/
barenakedislam.com	http://www.barenakedislam.com/2016/11/03/fbi-sources-believe-clinton-foundation-case-is-likely-moving-toward-an-indictment/
beforeitsnews.com	http://beforeitsnews.com/politics/2016/11/breaking-wikileaks-to-drop-hillarys-33k-deleted-emails-tomorrow-video-2855077.html

Domain Name	Example URL
bipartisanreport.com	http://bipartisanreport.com/2016/09/09/just-in-barbara-bush-verbally-dissects-any-woman-who-votes-for-donald-trump-video/
bizpacreview.com	http://www.bizpacreview.com/2016/11/04/refuse-pawn-huffpo-writer-latino-activist-recants-hillary-support-powerful-op-ed-408408
bredred.com	http://bredred.com/republicans-have-cast-17000-more-votes-than-dems-in-fl/
breitbart.com	http://www.breitbart.com/video/2016/11/06/nate-silver-clinton-one-state-away-losing-electoral-college/
campusreform.org	https://www.campusreform.org/?id=8352
centerforsecuritypolicy.org	https://www.centerforsecuritypolicy.org/civilization-jihad-reader-series/
clintonemail.com	http://clintonemail.com
cnsnews.com	http://cnsnews.com/news/article/susan-jones/americans-not-labor-force-participation-rate-risesdrops
commonsense-conservative.org	http://commonsense-conservative.org/?p=1761
concealncarry.stfi.re	http://concealncarry.stfi.re/forum/forum/general-discussions/2nd-amendent-discussions/778-why-i-carry-a-gun-by-urban-carry-holsters?sf=eyenxnw
conservativedailypost.com	https://conservativedailypost.com/breaking-fbi-confirms-evidence-of-huge-underground-clinton-sex-network/
conservativeoutfitters.com	https://www.conservativeoutfitters.com/blogs/news/92961857-john-kasich-took-202-700-from-george-soros
conservativeread.com	http://conservativeread.com/vote-like-your-guns-depend-on-it-they-do-trump Pence16/#.wcibkuhgady.twitter
conservativereview.com	https://www.conservativereview.com/commentary/2016/11/this-new-ted-cruz-video-is-the-last-obamacare-fact-check-you-will-ever-need
conservativetribune.com	https://conservativetribune.com/wikileaks-reveals-what-bill-hiding/?utm_source=twitter&utm_medium=conserv_tribune
constitution.com	http://constitution.com/iran-reports-u-s-supplying-isis/
crooksandliars.com	http://crooksandliars.com/2016/11/latinos-nevada-and-florida-are-building
dailycaller.com	http://dailycaller.com/2016/10/30/hillary-clinton-knew-she-was-helping-islamists-move-into-power-in-libya/?utm_campaign=atdailycaller&utm_source=twitter&utm_medium=social
dailynewsbin.com	http://www.dailynewsbin.com/?p=26577
dangerandplay.com	http://www.dangerandplay.com/2016/11/02/how-a-mindset-expert-who-hates-politics-got-involved-with-maga3x/
dcclothesline.com	http://www.dcclothesline.com/2016/11/02/confirmed-us-intel-operatives-leaked-clinton-campaign-emails-not-russia/
deepstatenation.com	http://deepstatenation.com/2016/11/oklahoma-republican-calls-for-hillary-clinton-to-be-shot-to-death-screenshots/
dennismichaelylnch.com	http://dennismichaelylnch.com/podestacooking1/
donaldtrumpnews.co	http://donaldtrumpnews.co/news/biggest-star-comes-trump-matthew-mcconaughey-votes-trump/
drudgereport.com	http://www.drudgereport.com/
endingthefed.com	http://endingthefed.com/breaking-bombshell-wikileaks-exposes-democrats-fake-trump-groping-plot.html
eutimes.net	http://www.eutimes.net/2016/10/hillary-clintons-sudden-move-of-1-8-billion-to-qatar-central-bank-stuns-financial-world/
floppingaces.net	http://www.floppingaces.net/2016/10/09/enabler-hillarys-actions-speak-louder-than-trumps-words-guest-post/?platform=hootsuite
freebeacon.com	http://freebeacon.com/politics/mccaskill-fbi-probe-exists-legitimate-question-whether-clinton-broke-law/
frontpagemag.com	http://www.frontpagemag.com/fpm/264540/seven-clinton-policy-priorities-would-devastate-john-perazzo
gotnews.com	http://gotnews.com/breaking-ex-apprentice-summerzervos-paid-500000-gloriaallred-accuse-trump-deal-went-others/
hannity.com	http://www.hannity.com/articles/hanpr-election-493995/watch-undercover-journalist-in-full-burka-15277316/
hotair.com	http://hotair.com/archives/2016/11/10/newt-has-no-time-for-whiny-sniveling-negative-nevertrump-cowards/
hotpagenews.com	http://hotpagenews.com/r/162858

Domain Name	Example URL
infowars.com	http://www.infowars.com/spirit-cooking-clinton-campaign-chairman-invited-to-bizarre-satanic-performance/
inquisitr.com	https://www.inquisitr.com/3680558/new-wikileaks-emails-suggest-bernie-sanders-was-leveraged-into-endorsing-clinton/
joeforamerica.com	http://joeforamerica.com/?p=55358
judicialwatch.org	http://www.judicialwatch.org/press-room/press-releases/judicial-watch-court-hearing-monday-november-7-fbi-clinton-email-records-case/
lawnews.tv	http://lawnews.tv/wikileaks-clinton-camp-worried-uranium-one-deal-being-investigated-hillary-sold-20-of-americas-uranium-to-russia/
lifenews.com	http://www.lifenews.com/2016/10/31/black-pastor-urges-christians-hillarys-deplorables-to-vote-for-donald-trump/
magafeed.com	http://magafeed.com/the_donald-uncovers-dark-connections-between-the-clintons-convicted-child-abductor/
mediaite.com	https://www.mediaite.com/tv/nbc-news-pete-williams-reports-fbi-really-isnt-conducting-clinton-foundation-investigation/
mobile.wnd.com	http://mobile.wnd.com/2016/11/obama-claims-presidents-have-power-to-violate-constitution/#xhmqgbuwzr6y9wd.99
mostdamagingwikileaks.com	http://www.mostdamagingwikileaks.com/
mrctv.org	http://www.mrctv.org/blog/flashback-jesse-jackson-praises-trump-seriousness-and-commitment-diversity-project
nationalreview.com	http://www.nationalreview.com/corner/442059/dont-blame-clinton-trump-2016-wouldve-beaten-obama-2012
naturalnews.com	http://www.naturalnews.com/055879_associated_press_eric_tucker_dishonest_journalism.html
newsbusters.org	http://www.newsbusters.org/blogs/nb/kyle-drennen/2016/11/03/nets-ignore-massive-bombshell-fbi-investigation-clinton-corruption
newsmax.com	http://www.newsmax.com/politics/trump-president-salary-refuse/2015/09/18/id/692155/
nydailynews.com	http://www.nydailynews.com/news/national/calif-woman-accusing-trump-child-rape-break-silence-article-1.2855631
occupydemocrats.com	http://occupydemocrats.com/2016/11/03/fox-news-just-admitted-made-story-hillary-indicted-foundation/
pamelageller.com	http://pamelageller.com/2016/11/must-see-video-undercover-journalist-in-full-burka-is-offered-huma-abedins-ballot-voterfraud.html/
pastebin.com	http://pastebin.com/36q0ykxm
patdollard.com	http://www.patdollard.com/75-of-americans-say-media-biased-for-hillary/
patriotpost.us	https://patriotpost.us/articles/45877
politopinion.com	http://politopinion.com/2016/11/undercover-pv-offered-huma-ballot/
puppetstringnews.com	http://www.puppetstringnews.com/blog/black-voters-for-trump-double-from-2012-election
rasmussenreports.com	http://www.rasmussenreports.com/public_content/politics/elections/election_2016/white_house_watch_nov3
redstate.com	http://www.redstate.com/aglanon/2016/11/07/for-the-sake-of-the-republican-party-dont-vote-for-trump/
redstatewatcher.com	http://redstatewatcher.com/article.asp?id=46969
scooprocket.com	http://scooprocket.com/us/2016/11/01/fbi-dumps-documents-from-bill-clintons-pardon-of-tax-cheat/
shareblue.com	http://shareblue.com/60-million-hillary-voters-will-not-be-silenced/
silenceisconsent.net	http://silenceisconsent.net/explosive-proof-hillary-tied-to-child-trafficker-laura-silby
stateofthenation2012.com	http://stateofthenation2012.com/?p=54773
theamericanfirst.com	http://theamericanfirst.com/video-cnn-interviews-leftist-protester-later-revealed-as-cnn-cameraman/
theamericanmirror.com	http://www.theamericanmirror.com/obama-talks-207-times-campaigning-hillary/
theblacksphere.net	http://theblacksphere.net/2016/11/michelle-obama-deleted-hillary-from-twitter/
theconservativetreehouse.com	https://theconservativetreehouse.com/2016/11/01/epic-rick-santorum-blasts-nevertrump-and-john-kasich-video/
thefederalist.com	http://thefederalist.com/2016/11/04/im-voting-trump-street-fighter-can-take-clinton-machine/#.wb4y8g-asq.twitter
thefederalistpapers.org	http://thefederalistpapers.org/?p=100260

Domain Name	Example URL
thegatewaypundit.com	http://www.thegatewaypundit.com/2016/11/donald-trump-takes-nearly-6-point-lead-crookedhillary-latest-la-times-poll/
theodysseyonline.com	https://www.theodysseyonline.com/calling-millennials-lets-break-party-system
thepoliticalinsider.com	http://www.thepoliticalinsider.com/police-sources-confirm-officials-charges-bill-clinton-pedophile/
therealstrategy.com	https://therealstrategy.com/wikileaks-admits-source-dnc-leaks-murdered-clintons/
therebel.media	https://www.therebel.media/congratulations_america_you_won_your_second_war_of_independence
truepundit.com	http://truepundit.com/breaking-bombshell-nypd-blows-whistle-on-new-hillary-emails-money-laundering-sex-crimes-with-children-child-exploitation-pay-to-play-perjury/
truthfeed.com	http://truthfeed.com/breaking-new-wikileak-bombshell-shows-cnn-asked-the-dnc-what-questions-to-ask-trump/34331/
ukok.page.tl	http://ukok.page.tl/trump-train.htm
usalibertynews.com	http://usalibertynews.com/breaking-la-times-reporter-fired-tweeting-wants-trump-dead/
vaskal.ca	http://www.vaskal.ca/podestafiles
weaselzipper.us	https://www.weaselzipper.us/305882-fbi-agents-are-now-talking-say-comey-stood-in-the-way-of-clinton-email-investigation/
westernjournalism.com	https://www.westernjournalism.com/lawsuit-filed-against-attorney-general-loretta-lynch-over-tarmac-meeting-with-bill-clinton/?utm_source=twitter&utm_medium=mobilefloatingsharingbuttons&utm_content=2016-11-02&utm_campaign=websitesharingbuttons
wnd.com	http://www.wnd.com/2016/10/clintons-black-son-to-make-bombshell-announcement/
youngcons.com	http://www.youngcons.com/clinton-foundation-doctored-memo-to-cover-distributing-watered-down-aids-medicine-in-africa/
yournewswire.com	http://yournewswire.com/nypd-hillary-clinton-child-sex-scandal/