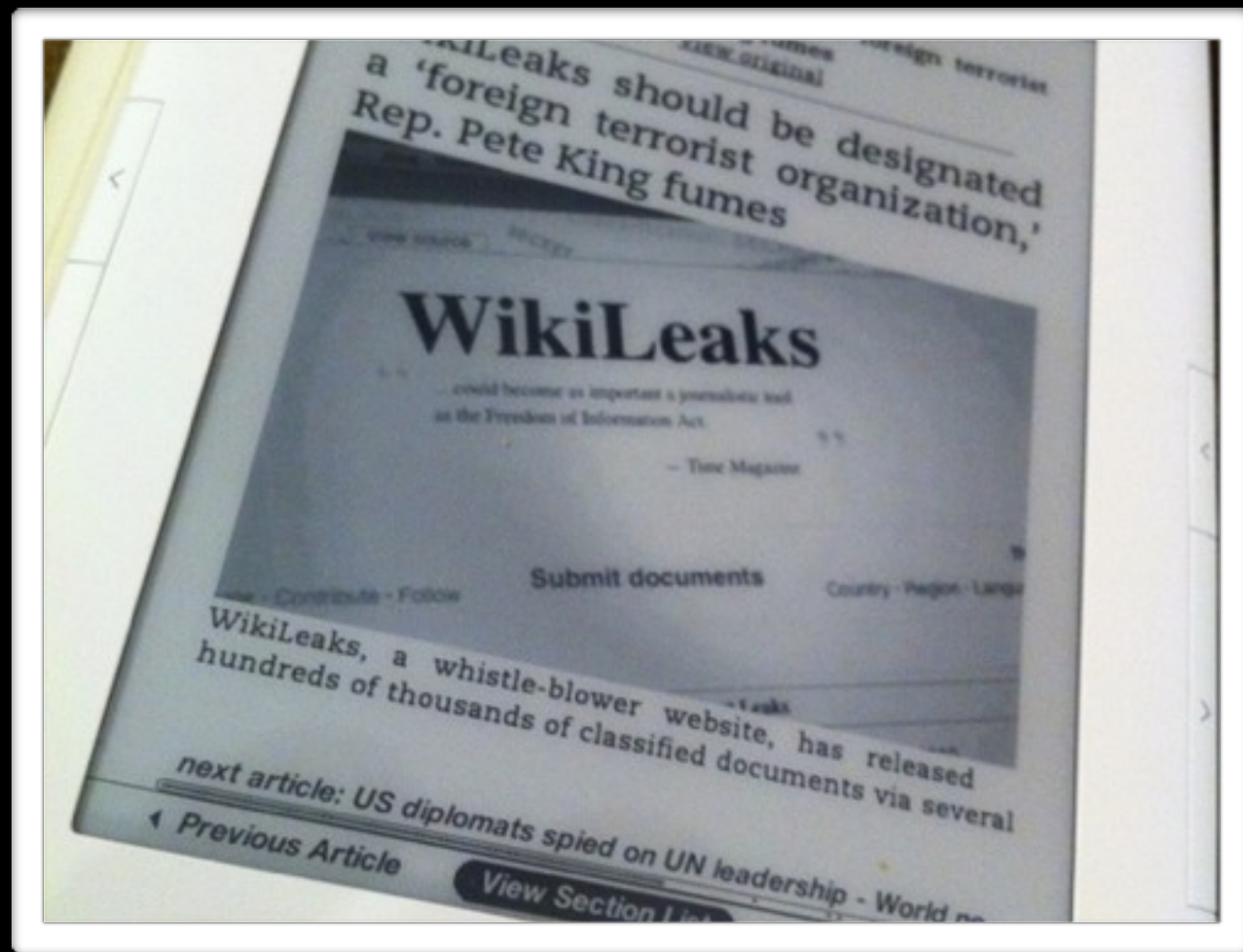


~~Wikileaks~~ Media Research Project

Christine Corbett Moran

@corbett_inc
Christine Corbett
I have not yet developed my position on @wikileaks. I have the cables, the internet, now where to start? where to finish?
4 Dec via Tweetie for Mac
Favorite
Reply
Delete

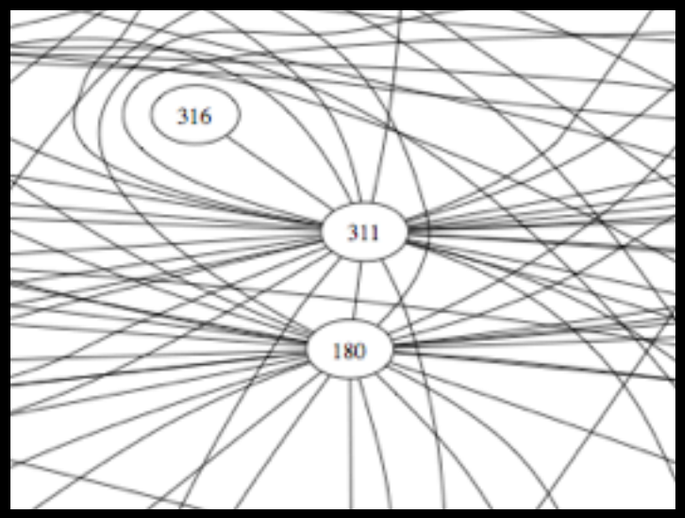
Motivation

@ario
Ario Jafarzadeh

@corbett_inc Start here: http://bit.ly/fnOtWK Finish with balaclavas and molotovs.

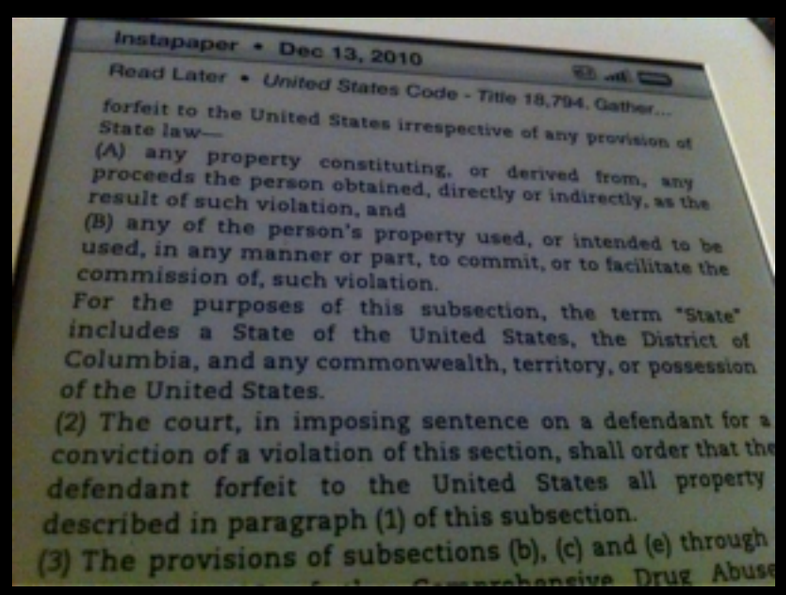
4 Dec via TweetDeck ☆ Favorite ↻ Retweet ↩ Reply

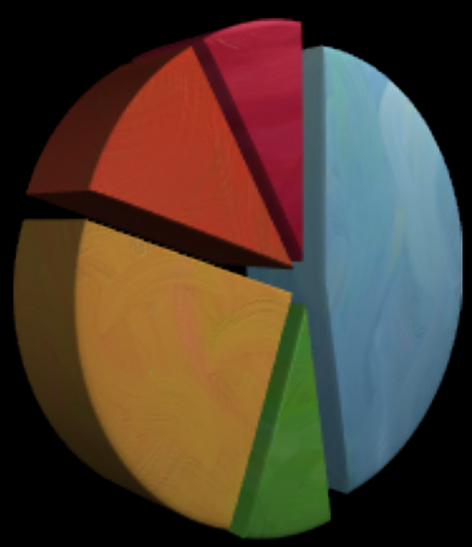
Motivation



folder	cites original sources: key, o=is original, n=no, y=yes, e=extensively, p=partially, l=very limited, s*: cites secondary sources
eading list	y

Methodology

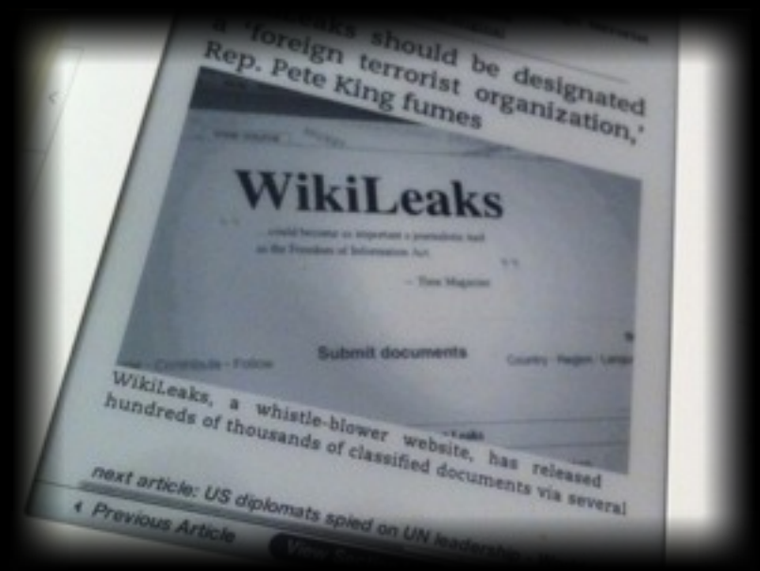


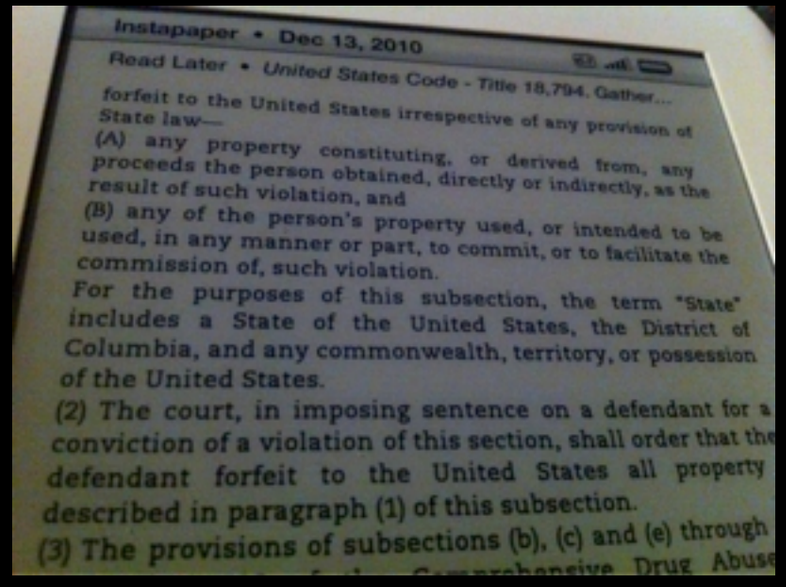


```python
def add_links_to_instapaper(url,username,password=None,wikilinks=False):
    INSTAPAPER_ADD='https://www.instapaper.com/api/add'
    instapaper_params={'username':username}
    if password:
        instapaper_params['password']=password
    # spoofing user agent headers, some sites like Wikipedia require
    opener=urllib2.build_opener()
    opener.addheaders =[('User-agent', 'Mozilla/5.0')]
    original_page=BeautifulSoup(opener.open(url).read())
    all_links={}
    for link in original_page.findAll('a'):
        try:
            link=link['href']
            if not all_links.has_key(link) and valid_link(link,wikilinks):
                instapaper_params['url']=link
                result=urllib.urlopen('%s?%s' % \
                    (INSTAPAPER_ADD,urllib.urlencode(instapaper_params)))
                all_links[link]=True
        except Exception,e:
            continue
```

1. Obtain articles

```
./AddUrlsToInstapaper.py -u
username -p password page_url
```





2. Read articles

Most worthwhile reads

- Julian Assange and the Computer Conspiracy; "To destroy this invisible government" http://cryptome.org/0002/ja-conspiracies.pdf
- WikiLeaks Spokesman Quits http://www.spiegel.de/international/germany/0,1518,719619,00.html
- United States Code: Title 18,794. Gathering or delivering defense information to aid foreign government | LII / Legal Information Institute http://www.law.cornell.edu/uscode/html/uscode18/usc_sec_18_00000794----000-.html
- United States Code: Title 18,793. Gathering, transmitting or losing defense information | LII / Legal Information Institute http://www.law.cornell.edu/uscode/html/uscode18/usc_sec_18_00000793----000-.html
- Changes_in_Guantanamo_Bay_SOP_manual_(2003-2004) http://wikileaks.org/wiki/Changes_in_Guantanamo_Bay_SOP_manual_(2003-2004)
- Daily Press Briefing - December 2, 2010 http://www.state.gov/r/pa/prs/dpb/2010/12/152291.htm

Most worthwhile reads

- On The Media: Transcript of "Leak Proof" (March 13, 2009) http://www.onthemedia.org/transcripts/2009/03/13/04/
- Dept_of_State_Assange_letter.pdf http://media.washingtonpost.com/wp-srv/politics/documents/Dept_of_State_Assange_letter.pdf
- Wikileaks: About page. Current Site or http://web.archive.org/web/20080216000537/http://www.wikileaks.org/wiki/Wikileaks:About
- Wikileaks:Submissions - Wikileaks Current Site or http://web.archive.org/web/20080419013425/http://www.wikileaks.org/wiki/Wikileaks:Submissions
- WikiLeaks and Julian Paul Assange : The New Yorker http://www.newyorker.com/reporting/2010/06/07/100607fa_fact_khatchadourian
- Secret Trafigura report said 'likely cause' of illness was release of toxic gas from dumped waste - Telegraph http://www.telegraph.co.uk/news/uknews/6350262/Secret-Trafigura-report-said-likely-cause-of-illness-was-release-of-toxic-gas-from-dumped-waste.html

Worthwhile Reads

Worthwhile Reads

"To radically shift regime behavior we must think clearly and boldly for if we have learned anything, it is that regimes do not want to be changed. We must think beyond those who have gone before us and discover technological changes that embolden us with ways to act in which our forebears could not."

-Julian Assange and the Computer Conspiracy; "To destroy this invisible government"

Worthwhile Reads

"To radically shift regime behavior
clearly and boldly for if we have lear
that regimes do not want to be change
beyond those who have gone before us
technological changes that embolden us
act in which our forebears could not."

-Julian Assange and the Computer Conspiracy; "To de
government"

"**SPIEGEL:** What is it that you no longer stand behind?
Schmitt: That we promise all of our sources that we will publish their material, for example. Recently, however, we have only focused on the major topics and applied practically all of our resources to them. Take the US Army Afghanistan documents at the end of July, for example. The video of the air strike in Baghdad in 2007, "Collateral Damage," was an extreme feat of strength for us. During the same period of time we also could have published dozens of other documents. And through our rising recognition in the last six months, we have again received a lot of material that urgently needs to be processed and published."

-http://www.spiegel.de/international/germany/0,1518,719619,00.html

Worthwhile Reads

"To radically shift regime beha
clearly and boldly f
that re
beyond
technolo
act in wh

-Julian Assan
government"

"(b) Whoever, in time of war, with intent that the same shall be communicated to the enemy, collects, records, publishes, or communicates, or attempts to elicit any information with respect to the movement, numbers, description, condition, or disposition of any of the Armed Forces, ships, aircraft, or war materials of the United States, or with respect to the plans or conduct, or supposed plans or conduct of any naval or military operations, or with respect to any works or measures undertaken for or connected with, or intended for the fortification or defense of any place, or any other information relating to the public defense, which might be useful to the enemy, shall be punished by death or by imprisonment for any term of years or for life."

-http://www.law.cornell.edu/uscode/html/uscode18/usc_sec_18_00000794----000-.html

-http://www.spiegel.de/international/germany/0,1518,719619,00.html

Worthwhile Reads

Worthwhile Reads

"7. Pervasive spin via language control 'hunger strike' becomes VTF - 'voluntary total fasting'. The word "suicide" is virtually removed from the document. Even an attempted hanging requiring someone to cut down is categorized as "self-harm", despite "self-harm" elsewhere being defined as actions that could not be fatal."

-Changes in Guantanamo Bay SOP manual (2003-2004), Wikileaks

Worthwhile Reads

"7. Pervasive spin

"**QUESTION:** Can you say how many people she's called so far?
MR. CROWLEY: A baker's dozen. (Laughter.)

...

QUESTION: Can you run down the list just quick since you got it right there?

MR. CROWLEY: So we're not going to get to lunch anytime soon. (Laughter.) Okay. President Zardari, President Fernandez de Kirchner,...

...

QUESTION: Does she have any plans to call the vain and vindictive president of Turkmenistan or the godfather-esque president of Azerbaijan this --

MR. CROWLEY: When we have calls we can report out to you, we will, of course, let you know."

-http://www.state.gov/r/pa/prs/dpb/2010/12/152291.htm

"

Worthwhile Reads

"7. Pervasive spin

becom

a

ca

els

fata

-Cha

2004),

"**QUESTION:** Can you say how n

MR. CROWLEY: A baker's doze

...

QUESTION: Can you run down t

MR. CROWLEY: So we're not g

President Zardari, President Fernan

...

QUESTION: Does she have any p

Turkmenistan or the godfather-esque president of Azerbaijan this --

MR. CROWLEY: When we have calls we can report out to you, we will, of course, let you know."

-http://www.state.gov/r/pa/prs/dpb/2010/12/152291.htm

"

"We will not engage in a negotiation regarding the further release or dissemination of illegally obtained U.S. Government classified materials. If you are genuinely interested in seeking to stop the damage from your actions, you should: 1) ensure WikiLeaks ceases publishing any and all such materials; 2) ensure WikiLeaks returns any and all classified U.S. Government material in its possession; and 3) remove and destroy all records of this material from WikiLeaks' databases."

-http://media.washingtonpost.com/wp-srv/politics/documents/Dept_of_State_Assange_letter.pdf

Worthwhile Reads

Worthwhile Reads

Is Wikileaks, as an organization, centralized?

We are regionalizing in an effort to establish a world-wide ethical leaking movement. Regional groups are forming in many countries (see Contact).

Our goal is to build full spectrum of support ranging from business to activists.

While we committed to keep publishing under all circumstances, we will be as open as possible in our policies and practices. The founders have the final say, but this will mainly effect founding documents like the one you are reading now.

-http://web.archive.org/web/20080216000537/http://www.wikileaks.org/wiki/Wikileaks:About

Worthwhile Reads

BOB GARFIELD: But we wondered about the broader implications of a site that publishes sensitive information with next to no editorial discretion. I asked Julian if hypothetically he would publish information sent to his website that could lead to the deaths of innocents, such as, for instance, how to release anthrax into a town's water supply.

JULIAN ASSANGE: Yes, even if there is a possibility that it would lead to loss of life. It's hard to imagine a circumstance where we would get a document and us not publishing it would be helpful. If they were ill motivated, then they could send that in private to terrorist groups, to neo-Nazi organizations, and those organizations could then develop their plans out of the sunlight. And that's the greatest harm.

-http://www.onthemedia.org/transcripts/2009/03/13/04/

Worthwhile Reads

Is Wikileaks

BOB GARFIELD: But we wondered about the broader implications of a site that publishes sensiti
hypothetically he
deaths of innoce
supply.

JULIAN ASSA
It's hard to imag
publishing it wo
private to terrori
develop their pla

-http://www.onthemed

In 2006 Trafigura, a privately-owned, Dutch-registered company that has offices in London, had chartered a tanker, the Probo Koala, to carry a cargo of cheap, dirty petrol known as coker gasoline that it hoped to clean up for a profit.

After using caustic soda and a catalyst to remove sulphur from the fuel, it was left with waste products known as slops that were to be disposed of in Amsterdam. But Trafigura was told that the cost of handling the foul-smelling residue would be far higher than initially thought.

The waste was pumped back onto the ship and in August taken to Abidjan, the biggest city in Ivory Coast, where a local contractor called Compagnie Tommy unloaded it and dumped it at as many as 18 landfill sites across the city at night.

-http://www.telegraph.co.uk/news/uknews/6350262/Secret-Trafigura-report-said-likely-cause-of-illness-was-release-of-toxic-gas-from-dumped-waste.html

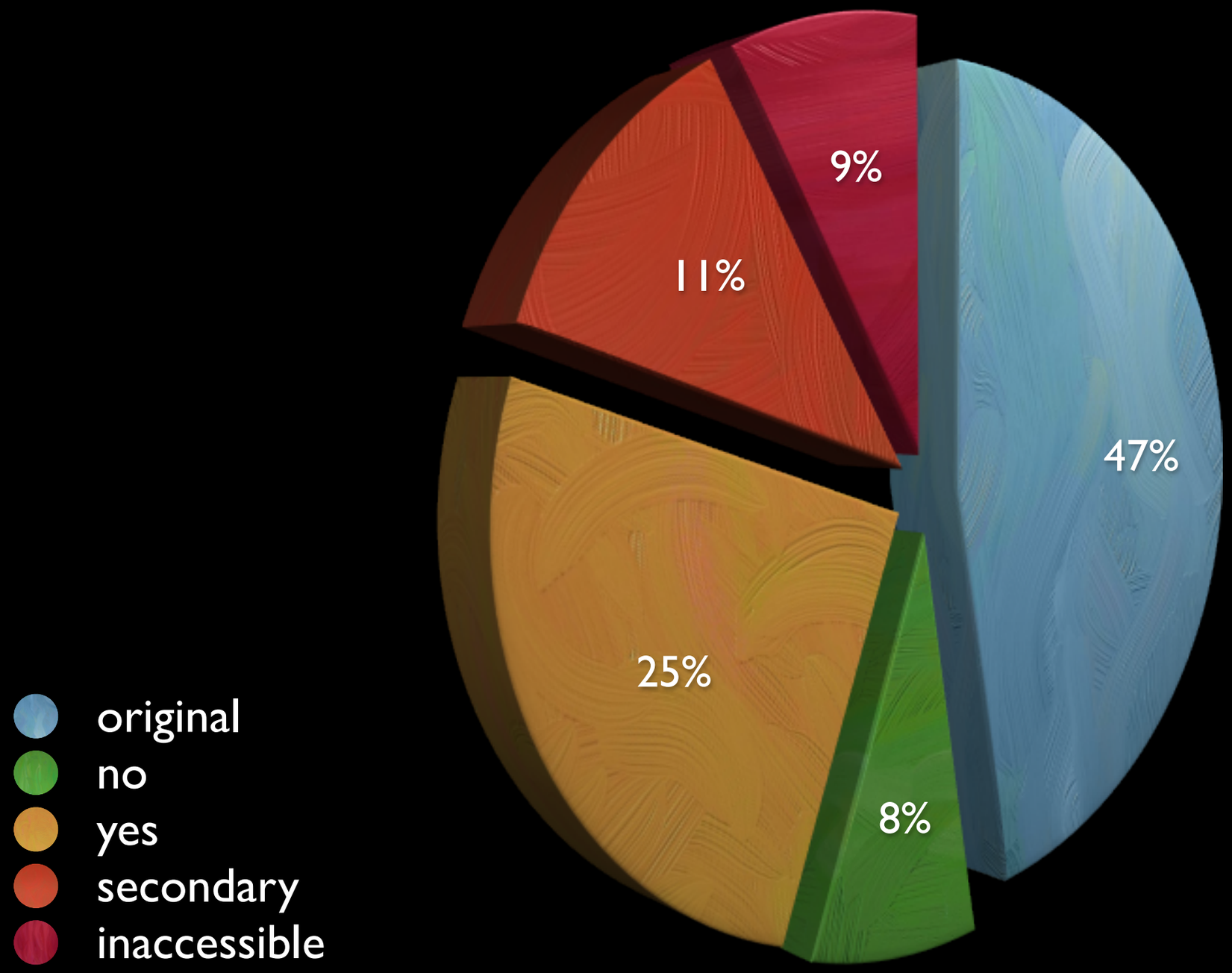
Wikileaks Research Project - Citation classification and assorted notes, original URLS : Sheet 1

link	title	my notes/highlights: forgive typos they were typed quickly on the tiny Kindle keyboard	instapaper folder	cites original sources: key, o=is original, n=no, y=yes, e=extensively, p=partially, l=very limited, s*: cites secondary sources
http://zunguzungu.wordpress.com/2010/11/29/julian-assange-and-the-computer-conspiracy-%E2%80%9Cto-destroy-this-invisible-government%E2%80%9D/	Julian Assange and the Computer Conspiracy; "To destroy this invisible government" « zunguzungu	Read original essay first, think about it, then read the above post. http://cryptome.org/0002/ja-conspiracies.pdf Ja-conspiraces.pdf definitely on the top reading list.	Wikileaks reading list	y
http://news.cnet.com/8301-13578_3-20000469-38.html	8301-13578_3-20000469	CNET article--more valuable for its links to original sources and sections of the espionage act	Wikileaks reading list	y
http://www.firstamendmentcenter.org/analysis.aspx?id=17318	firstamendmentcenter.org analysis		Archive	y
http://www.fas.org/blog/secrecy/2010/06/wikileaks_review.html	Wikileaks Fails "Due Diligence" Review | Secrecy News		Archive	y

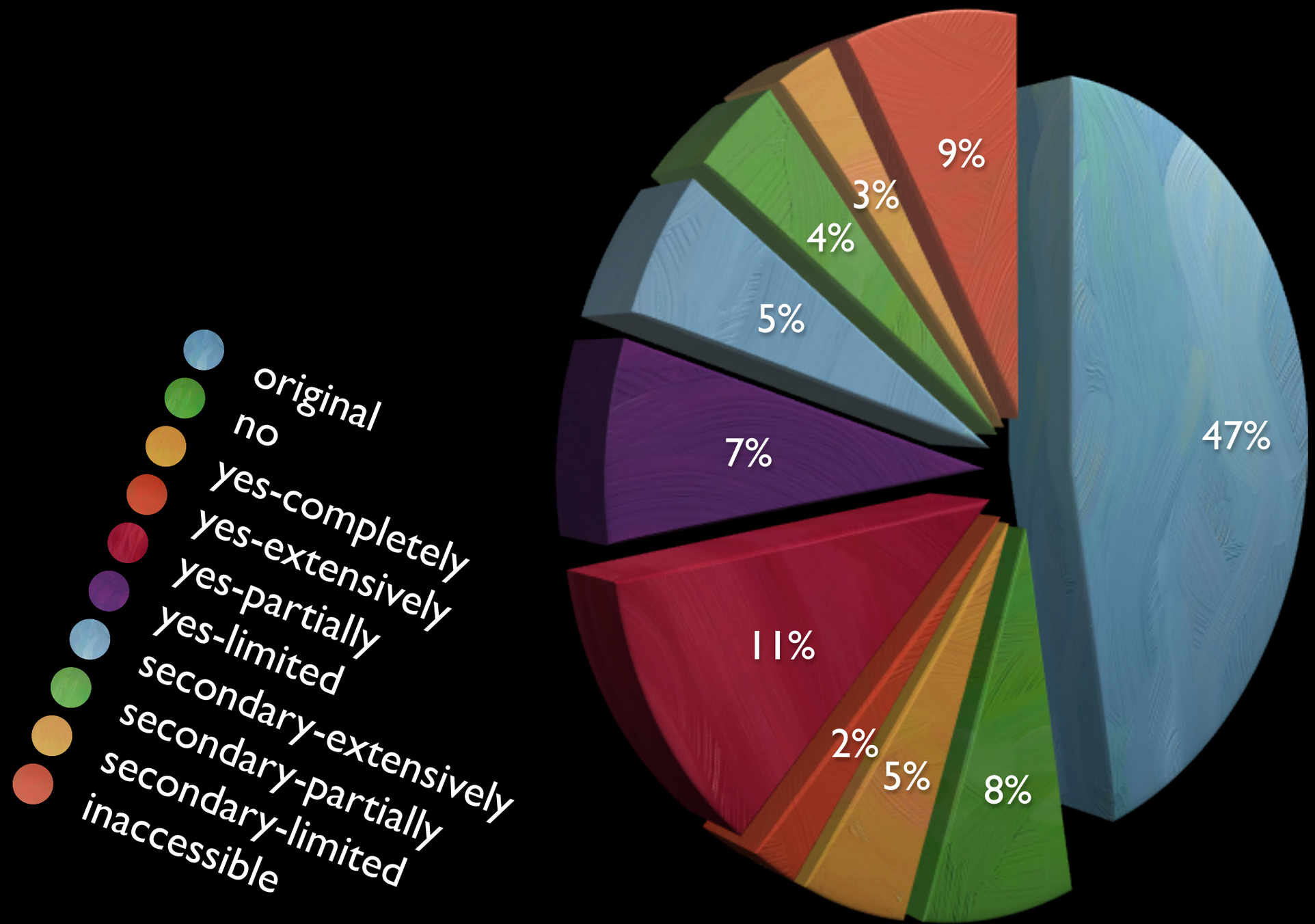
3. Classify articles

http://tinyurl.com/wikipedia-wikileaks-citations

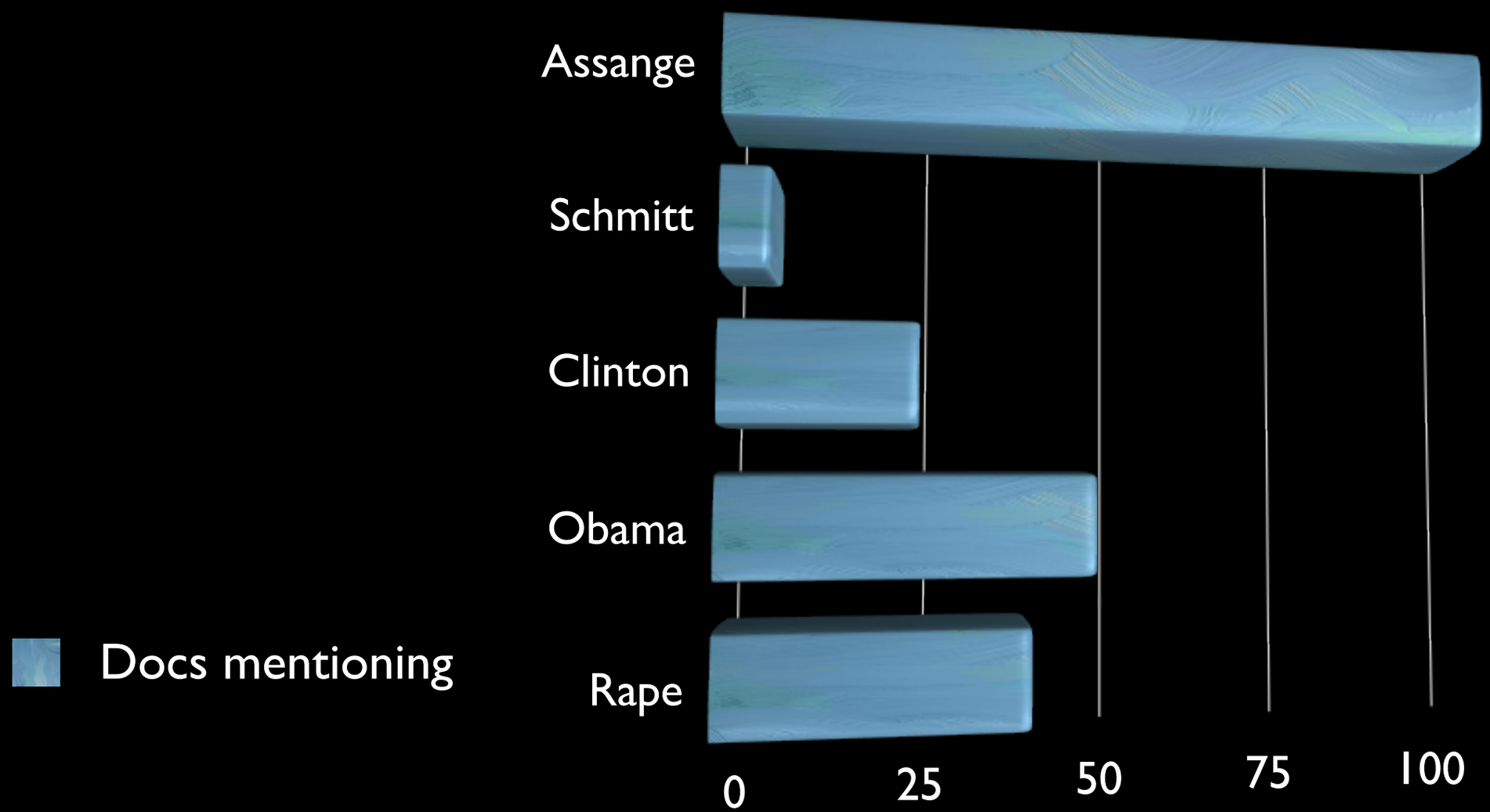
The Numbers



The Numbers


original
no
yes-completely
yes-extensively
yes-partially
yes-limited
secondary-extensively
secondary-partially
secondary-limited
inaccessible
47%
8%
5%
2%
11%
7%
5%
4%
3%
9%

Terms of Interest



Case study: None

Wikileaks Founder Could Be Charged Under Espionage Act, Washington Post, http://www.washingtonpost.com/wp-dyn/content/article/2010/11/29/AR2010112905973.html

Case study: None

Wikileaks Founder Could Be Charged Under Espionage Act, Washington Post, http://www.washingtonpost.com/wp-dyn/content/article/2010/11/29/AR2010112905973.html

> "But the sources, who spoke on the condition of anonymity because the inquiry is rapidly unfolding, said charges could be filed under the act. The U.S. attorney's office in Alexandria - which in 2005 brought Espionage Act charges, now dropped, against two former pro-Israel lobbyists - is involved in the effort, the sources said."

Case study: None

Wikileaks Founder Could Be Charged Under Espionage Act, Washington Post, http://www.washingtonpost.com/wp-dyn/content/article/2010/11/29/AR2010112905973.html

"But the sources, who spoke on the condition of anonymity because the inquiry is rapidly unfolding, said charges could be filed under the act. The U.S. attorney's office in Alexandria - which in 2005 brought Espionage Act charges, now dropped, against two former pro-Israel lobbyists - is involved in the effort, the sources said."

"Former prosecutors cautioned that prosecutions involving leaked classified information are difficult because the Espionage Act is a 1917 statute that preceded Supreme Court cases that expanded First Amendment protections."

Case study: Secondary

Pentagon Papers II? On WikiLeaks and the First Amendment, Law Blog - WSJ, http://blogs.wsj.com/law/2010/07/26/pentagon-papers-ii-on-wikileaks-and-the-first-amendment/

Case study: Secondary

Pentagon Papers II? On WikiLeaks and the First Amendment, Law Blog - WSJ, http://blogs.wsj.com/law/2010/07/26/pentagon-papers-ii-on-wikileaks-and-the-first-amendment/

‘Schauer said the opinion isn’t necessarily “conclusive,” partly because it was issued as a brief “per curiam” decision, and didn’t adopt a categorical rule on when the government might prevent publication of such information. Still, Schauer said that the case, along with two others, Landmark Communications v. Virginia and Bartnicki v. Vopper, “all go in the direction” of dictating that the person at WikiLeaks who got the information would likely elude criminal prosecution or liability unless that person was, said Schauer, “involved in getting the material in the first place.” ’

Case study: Secondary

Pentagon Papers II? On WikiLeaks and the First Amendment, Law Blog - WSJ, http://blogs.wsj.com/law/2010/07/26/pentagon-papers-ii-on-wikileaks-and-the-first-amendment/

‘Schauer said the opinion isn’t necessarily “conclusive,” partly because it was issued as a brief “per curiam” decision, and didn’t adopt a categorical rule on when the government might prevent publication of such information. Still, Schauer said that the case, along with two others, Landmark Communications v. Virginia and Bartnicki v. Vopper, “all go in the d
that the person at WikiLeaks who g
would likely elude criminal prosecu
unless that person was, said Schaue
getting the material in the first plac

The most obvious Supreme Court case to spring to mind is the New York Times v. U.S., the famous “Pentagon Papers” case from 1971. In the case, the Supreme Court ruled that the First Amendment barred the Nixon administration from keeping the New York Times and Washington Post from publishing classified information related to the Vietnam War.

Case study: Primary

Huge Wikileaks release shows US 'ignored Iraq torture', BBC News, http://www.bbc.co.uk/news/world-middle-east-11611319

Case study: Primary

Huge Wikileaks release shows US 'ignored Iraq torture', BBC News, http://www.bbc.co.uk/news/world-middle-east-11611319

"The 391,831 US army Sigacts (Significant Actions) reports published by Wikileaks on Friday describe the apparent torture of Iraqi detainees by the Iraqi authorities, sometimes using electrocution, electric drills and in some cases even executing detainees, says the BBC's Adam Brookes, who has examined some of the files."

Case study: Primary

Huge Wikileaks release shows US 'ignored Iraq torture', BBC News, http://www.bbc.co.uk/news/world-middle-east-11611319

"The 391,831 US army Sigacts (Significant Actions) reports published by Wikileaks on Friday describe the apparent torture of Iraqi detainees by the Iraqi authorities, sometimes using electrocution, electric drills and in some cases even executing detainees, says the BBC's Adam Brookes, who has examined some of the files."

"One document shows the US military was given a video apparently showing Iraqi Army (IA) officers executing a prisoner in the northern town of Talafar."

Citation needed

- WL operating costs: wild speculation by everyone including Assange himself
- Al-Qaeda operative on war logs
- Actual transcript of chat between Lamo and Manning
- ... many many more

Operating Costs

Operating Costs

"So all in all, can you give figures about how much money Wikileaks needs in one year?

Propably[sic] 200 000, that's with everyone paying themselves. But there are people who can't afford to continue being involved fulltime unless they are paid. For that I would say maybe it's 600 000 a year."

-Assange in http://stefanmey.wordpress.com/2010/01/04/leak-o-nomy-the-economy-of-wikileaks/ Interview

Operating Costs

"So all in all, can you give figures ab
money Wikileaks needs in one year?

Propably[sic] 200 000, that's with every
themselves. But there are people who ca
continue being involved fulltime unless th
For that I would say maybe it's 600 000 a

-Assange in http://stefanmey.wordpress.com
2010/01/04/leak-o-nomy-the-economy-of-w
Interview

"Why does it cost $600,000 to put up some documents on a website? Even if Julian Assange did the work himself, (granted it would be slow) it wouldn't cost this much. And I'm sure there would be volunteers who would do it for free in their spare time."
-http://www.smh.com.au/technology/technology-news/wikileaks-shuts-down-unable-to-plug-funding-gap-20100202-n9z4.html

Operating Costs

"So all in all, can you give figures ab
money Wikileaks needs in one year?

Propably[sic] 200 000, that's with every
themselves. But there are people who ca
continue being involved fulltime unless th
For that I would say maybe it's 600 000 a
-Assange in http://stefanmey.wordpress.com
2010/01/04/leak-o-nomy-the-economy-of-w
Interview

"Why does it cost $600,000 to put up some documents on a website? Even if Julian Assange did the work himself, (granted it would be slow) it wouldn't cost this much. And I'm sure there would be volunteers who would do it for free in their spare time."
-http://www.smh.com.au/technology/technology-news/wikileaks-shuts-down-unable-to-plug-funding-gap-20100202-n9z4.html

"Its tremendous success has meant the site has often struggled under the volume of users. It has faced down governments, investment banks and the famously litigious Church of Scientology but paying its operating costs (circa $600,000) has proved its undoing."
-http://www.guardian.co.uk/commentisfree/libertycentral/2010/jan/29/wikileaks-shut-down

Taliban operative

Taliban operative

A Taliban spokesman in Afghanistan using the pseudonym Zabiullah Mujahid said in a telephone interview that the Taliban had formed a nine-member "commission" after the Afghan documents were posted "to find about people who are spying." He said the Taliban had a "wanted" list of 1,800 Afghans and was comparing that with names WikiLeaks provided.

"After the process is completed, our Taliban court will decide about such people," he said.

-http://www.nytimes.com/2010/10/24/world/24assange.html

Taliban operative

A Taliban spokesman in Afghanistan using the pseudonym Zabiullah Mujahid said in a telephone interview that the Taliban had formed a nine-member "commission" after the Afghan documents were posted "to find about people who are spying." He said the Taliban had a "wanted" list of 1,800 Afghans and was comparing that with names WikiLeaks provided.

"After the proces
decide about suc

-http://www.nytimes.c

A Taliban spokesman told Britain's Channel 4 news that the group was sifting through the WikiLeaks documents to get the names of suspected informants and would punish anyone found to have collaborated with the United States and its allies.

-http://watchguardus.com/Wired-Threat-Level/WikiLeaks-Posts-Mysterious-Insurance-File.html

Taliban operative

A Taliban spokesman in Afghanistan using the pseudonym Zabiullah Mujahid said in a telephone interview that the Taliban had formed "commission" after the Afghan docum "to find about people who are spying." Taliban had a "wanted" list of 1,800 A comparing that with names WikiLeaks

"After the proces decide about suc

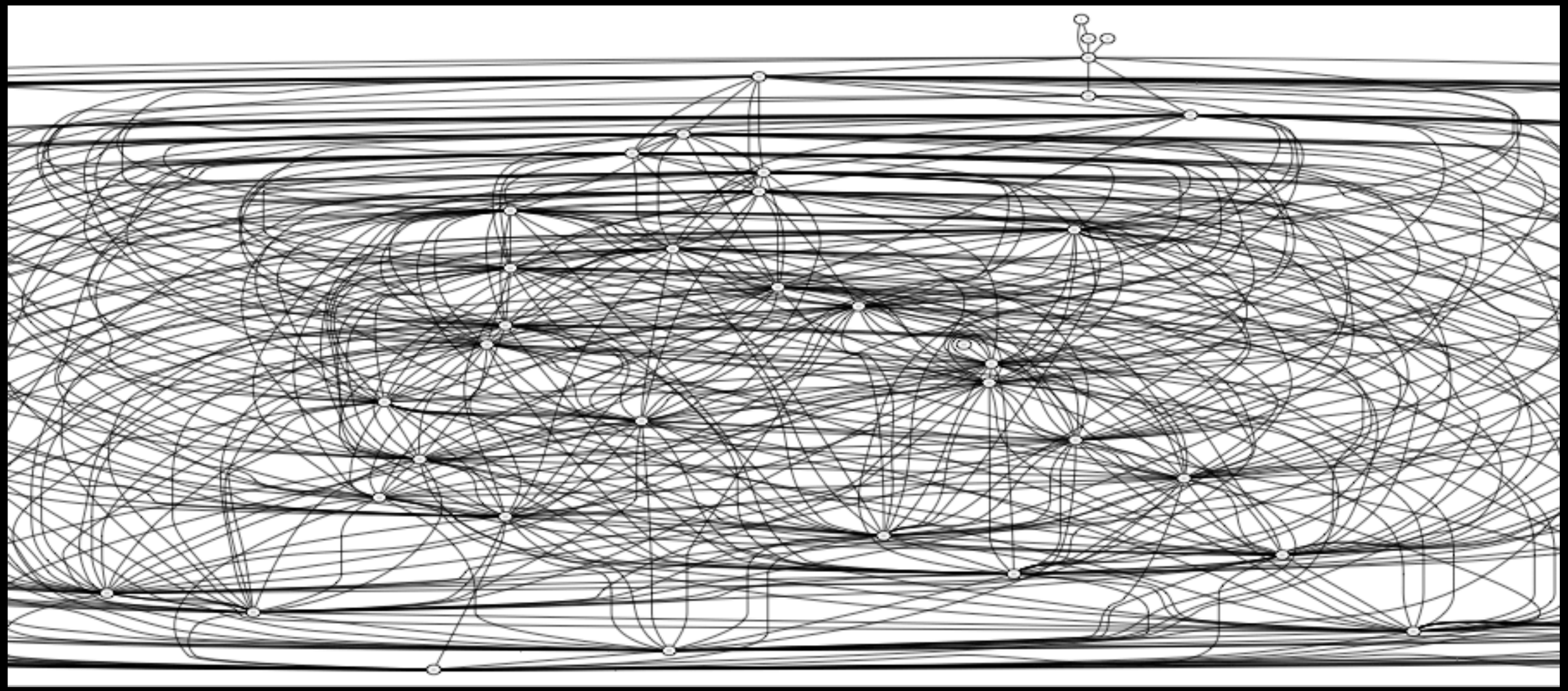
-http://www.nytimes.c

Speaking by telephone from an undisclosed location, Zabihullah Mujahid told **Channel 4 News** that the insurgent group will investigate the named individuals before deciding on their fate.

-http://www.channel4.com/news/articles/uk/taliban+hunt+wikileaks+outed+afghan+informers/3727667.html

A Taliban spokesma the group was sifting through the WikiLeaks documents to get the names of suspected informants and would punish anyone found to have collaborated with the United States and its allies.

-http://watchguardus.com/Wired-Threat-Level/WikiLeaks-Posts-Mysterious-Insurance-File.html

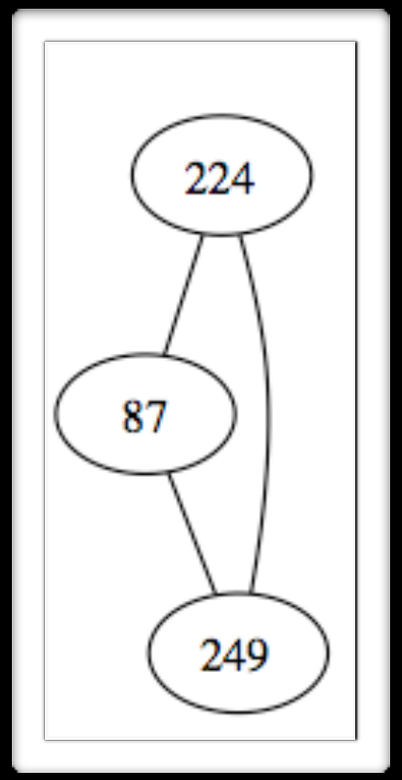


5. Analyze relationships

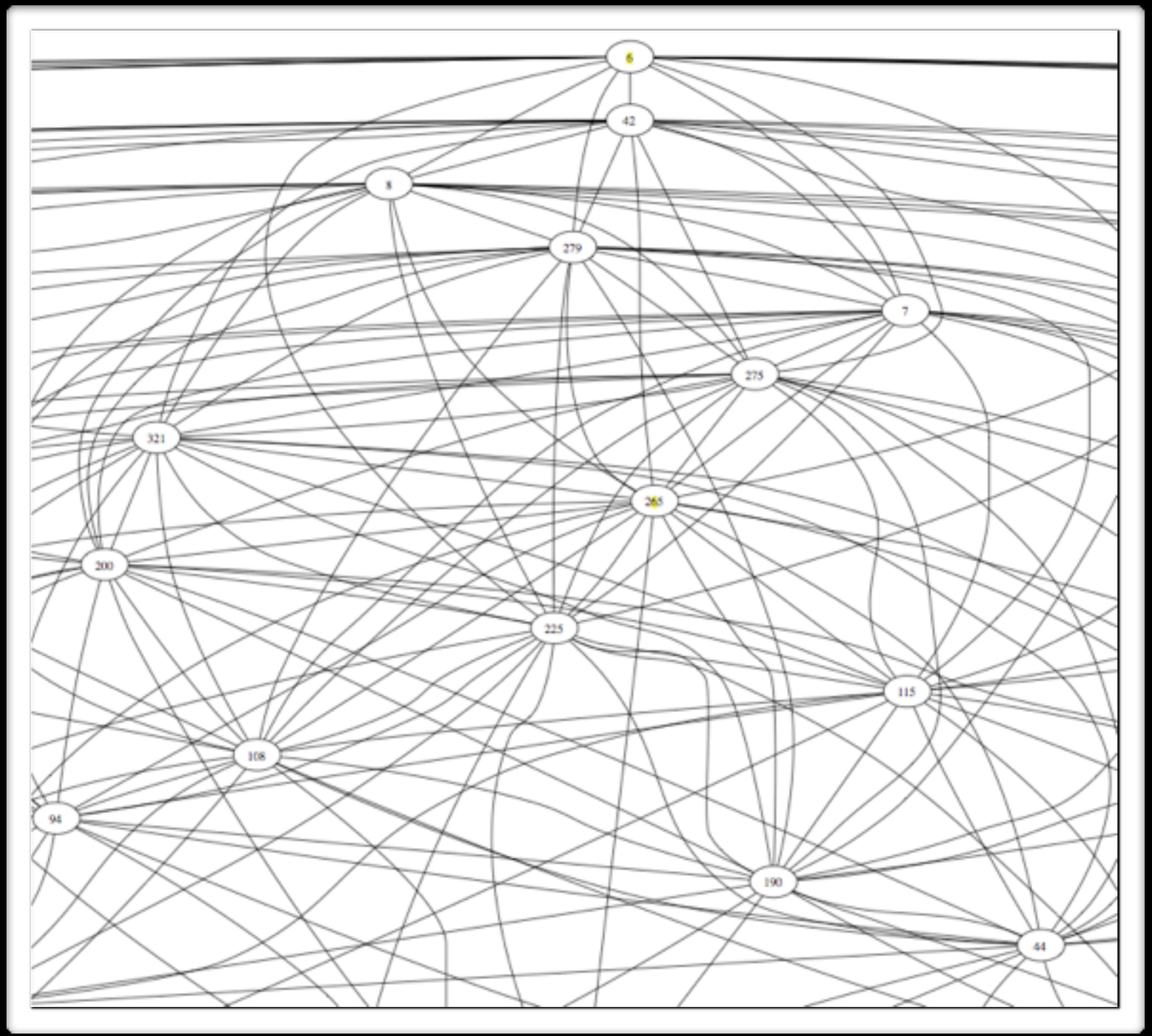
Strategy

- NLTK and training data
- Beautiful soup
- Pygraph
- Graphviz

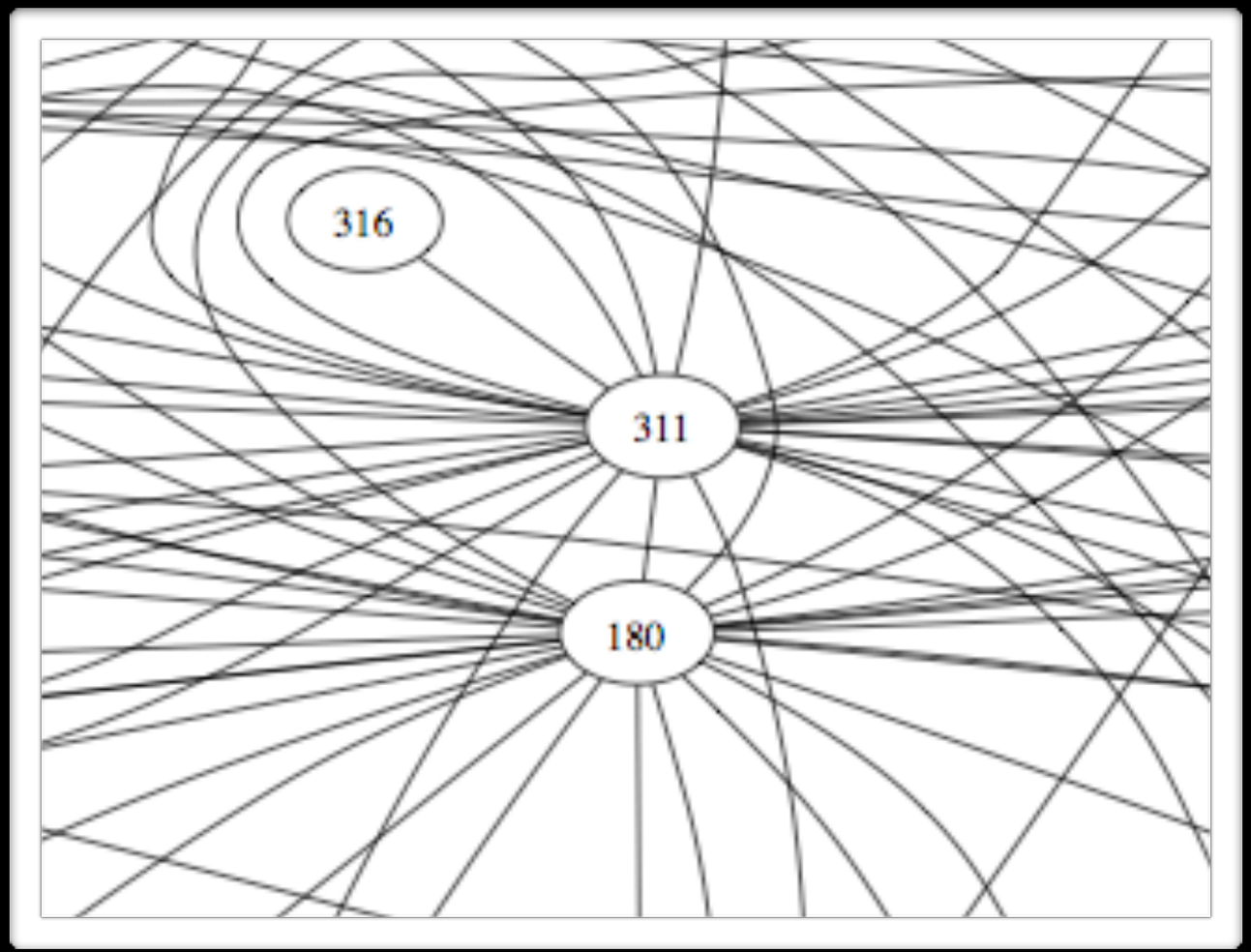
```
# Map sentences to list of docs which contain them
links=defaultdict(list)
# Parser and tokenizer used to split docs into sentences
parser=Parser()
tokenizer = nltk.data.load('tokenizers/punkt/english.pickle')
doc_map=codecs.open('../wl_article_connections.dict','w','utf-8')
# Graph creation, linking any documents which contain the same sentences
gr = graph()
# Mapping sentences to docs which contain them
for i,name in enumerate(os.listdir('.')):
    doc_map.write(u'%d\t%s\n' % (i,name))
    doc=codecs.open(name).read()
    title,doc=parser.parse(doc)
    sentences = tokenizer.tokenize(doc.strip())
    for sent in sentences:
        links[sent]+=[i]
```



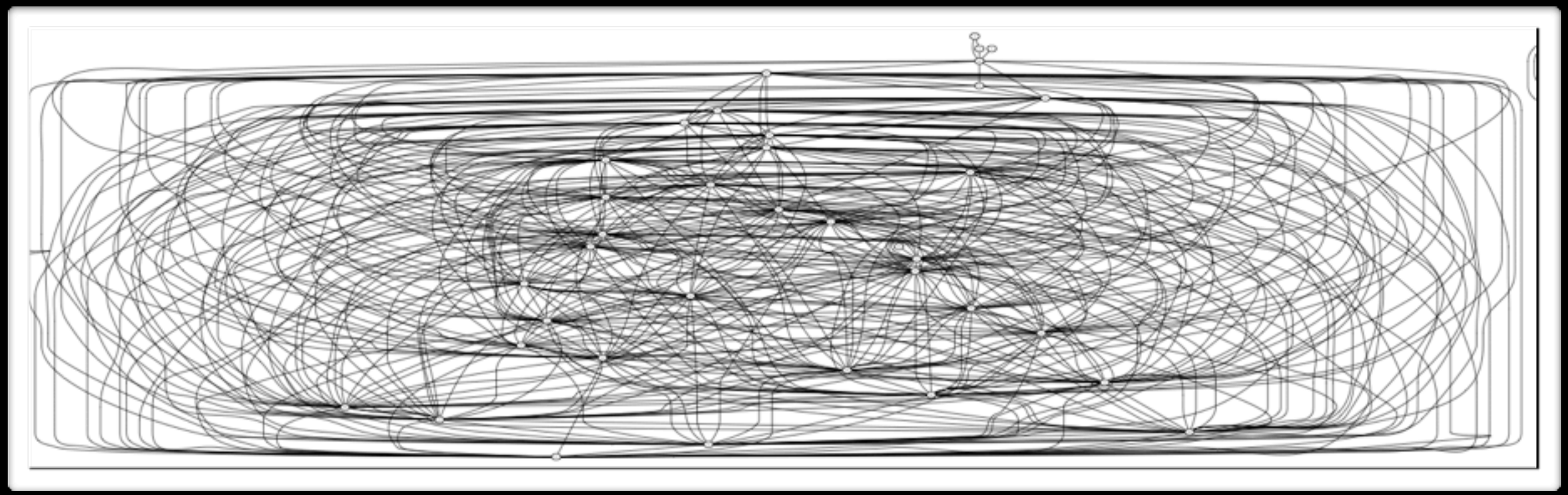
Swedish Articles



New Yorker Profile



Wikileaks About Page



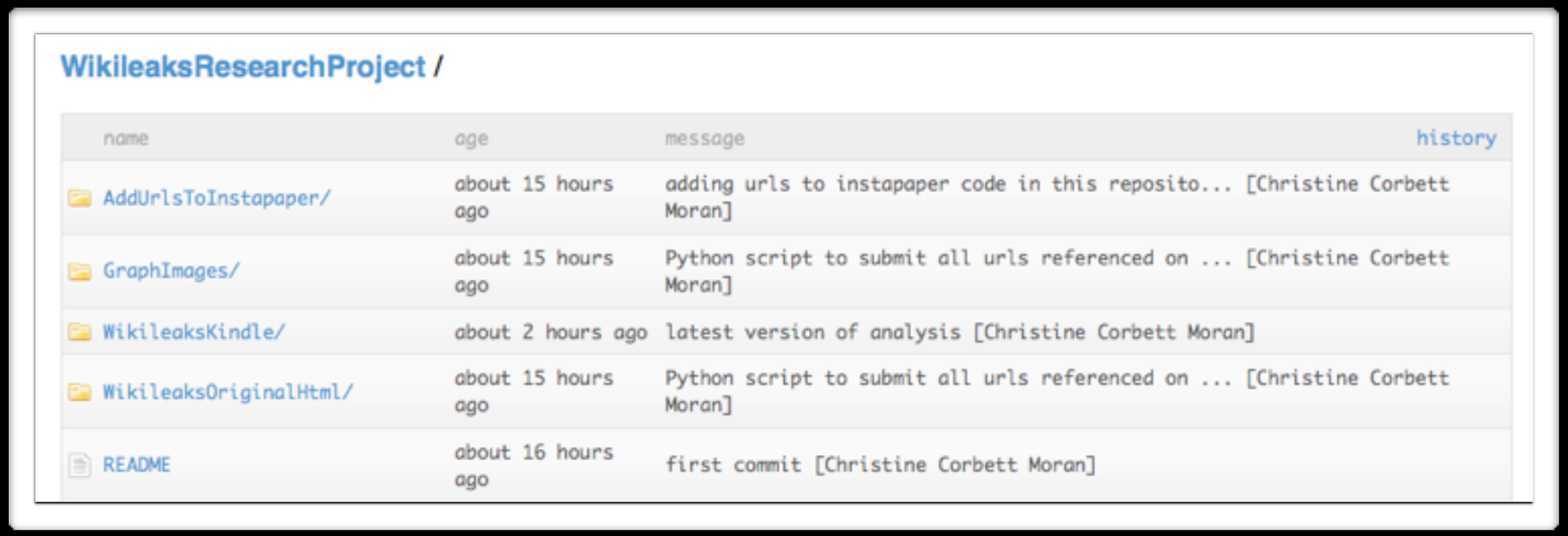
How about that?

Scientific Journalism

"WikiLeaks coined a new type of journalism: scientific journalism. We work with other media outlets to bring people the news, but also to prove it is true. Scientific journalism allows you to read a news story, then to click online to see the original document it is based on. That way you can judge for yourself: Is the story true? Did the journalist report it accurately?"

-Julian Assange in http://blogs.theaustralian.news.com.au/mediadiary/index.php/australianmedia/comments/julian1/

Recreate



WikileaksResearchProject /

name	age	message	history
AddUrlsToInstapaper/	about 15 hours ago	adding urls to instapaper code in this reposito... [Christine Corbett Moran]	
GraphImages/	about 15 hours ago	Python script to submit all urls referenced on ... [Christine Corbett Moran]	
WikileaksKindle/	about 2 hours ago	latest version of analysis [Christine Corbett Moran]	
WikileaksOriginalHtml/	about 15 hours ago	Python script to submit all urls referenced on ... [Christine Corbett Moran]	
README	about 16 hours ago	first commit [Christine Corbett Moran]	

- https://github.com/corbett/WikileaksResearchProject
- contact:
  - **twitter** @corbett_inc
  - **email** corbett@physik.uzh.ch

Extend

- Named entities analysis and other tasks building on NLTK
- Graph-incorporate citation information, strength of connection
- Automatic citation type analysis
- Integrated script
- Level up: point to topic, tells you what to read to educate yourself

Conclusions && Questions && Discussion