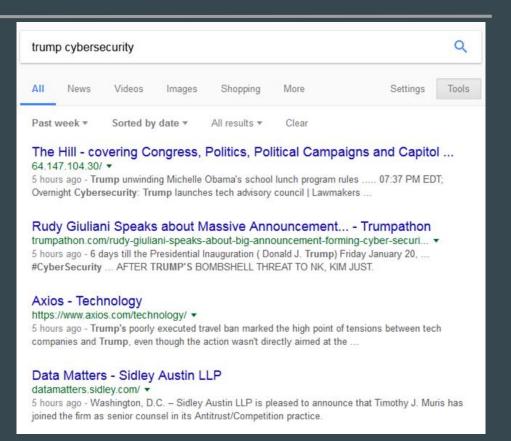
# CySpyder

•••

Chase Brown & Joe Dodson

### What's the problem?

 No straightforward way to search multiple sites for cybersecurity articles



#### Our Solution - CySpyder

 An easy-to-use interface that streamlines searching for cybersecurity articles

#### Article Title

ISC2 Issues White House Cybersecurity Executive Order Recommendations

The Implications Behind Proposed Internet Privacy Rules

InterContinental Hotel Chain Breach Expands

Why Brand Trumps Tech in C-Level Conversations

Microsoft: Foreign Surveillance Requests Under FISA Shot Up in 2016

ShadowBrokers Expose NSA Access to SWIFT Service Bureaus

Fake News at Work in Spam Kingpin's Arrest?

One of World's Most Wanted and Prolific Alleged Spammers Arrested

ShadowBrokers Dump More Equation Group Hacks, Auction File Password

China-Based Threat Actor APT10 Ramps Up Cyber Espionage Activity

FCC Privacy Rule Repeal Will Have Widespread Security Implications

Trump Signs Repeal of ISP Privacy Rules

Trump Extends Obama's EO for Sanctioning Hackers

Industry Braces for Repeal of ISP Privacy Rules

Post-FCC Privacy Rules, Should You VPN?

Should Trump Tackle Air-Gapped Critical Infrastructure?

Privacy Advocates Vow to Fight Rollback of Broadband Privacy Rules

Trust, Cloud & the Quest for a Glass Wall around Security

# Key Features

### Minimalistic Search Design



### Easy to Navigate Search Results

Article Search		- 0	×
File Help			
Key Article Subjects	Date	Article Title	Î
Click on an article to see more info	04/27/17	Attack Method Highlights Weaknesses in Microsoft CFG	
ener on an arrest to see more into	04/14/17	Exploit Kit Activity Quiets, But Is Far From Silent	
	04/12/17	Critical Security Updates from Adobe, Microsoft	
	04/11/17	Microsoft Patches Three Vulnerabilities Under Attack	
	03/23/17	LastPass Fixes Serious Security Flaw in Chrome, Firefox Extension	าร
	03/17/17	VM Escape Earns Hackers \$105K at Pwn2Own	
	03/16/17	Hackers Take Down Reader, Safari, Edge, Ubuntu Linux at Pwn2O	w
	03/14/17	Adobe Fixes Six Code Execution Bugs in Flash	
	03/14/17	Patch Tuesday Returns; Microsoft Quiet on Postponement	
<u>'</u>	03/14/17	Adobe, Microsoft Push Critical Security Fixes	
Most Mentioned Phrases in Results	03/08/17	In a Cybersecurity Vendor War, the End User Loses	
≻(53): Microsoft	02/27/17	Google Discloses Another 'High Severity' Microsoft Bug	
>(39): Windows	01/11/17	Adobe, Microsoft Push Critical Security Fixes	
	01/10/17	Microsoft Issues Record Low Number of Patch Tuesday Bulletins	
➤(31): Adobe	01/10/17	Two New Edge Exploits Integrated into Sundown Exploit Kit	
5 (27), Fl-1	01/05/17	Experts Warn of Novel PDF-Based Phishing Scam	
≻(27): Flash	12/15/16	Microsoft, Google to Block Flash by Default in Edge, Chrome	
≻(25): Microsoft Edge	12/14/16	Flash Player Bug An Eavesdropper's Delight	
	12/13/16	Microsoft Patches Publicly Disclosed IE, Edge Vulnerabilities	
	11/22/16	Microsoft Cutting Off SHA-1 Support in February for Edge, IE 11	
	11/15/16	More Than Half Of Android Devices Run Outdated Browsers	
New Search Edit Search Save Search	11/15/16	\/Mware Patches \/M Escane \/ulnerahility	
	<		> ٧

### Data Analysis

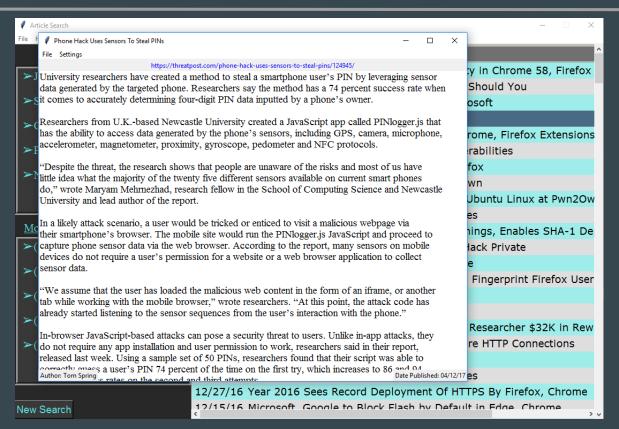
#### **Key Article Subjects**

- **≻**Skype
- **≻**Microsoft
- **>**SKYPE
- **≻**Facebook
- ➤ Information Security

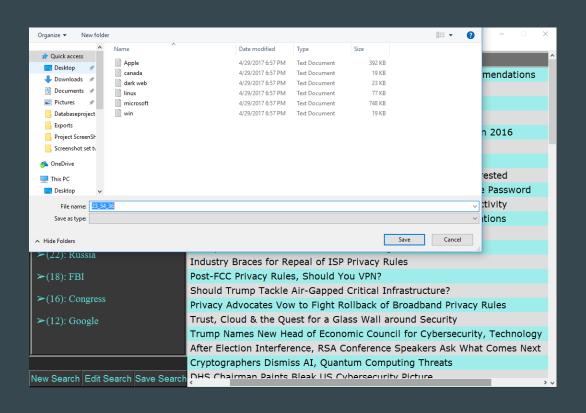
#### Most Mentioned Phrases in Results

- >(53): Microsoft
- **>**(39): Windows
- **>**(31): Adobe
- >(27): Flash
- >(25): Microsoft Edge

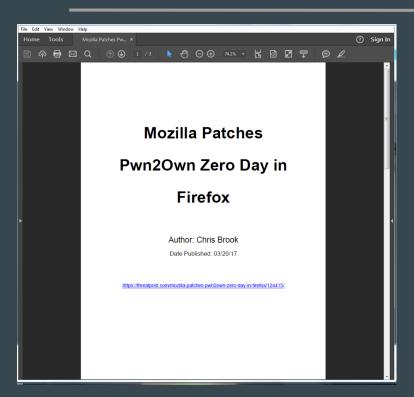
#### Quick Viewing of Articles

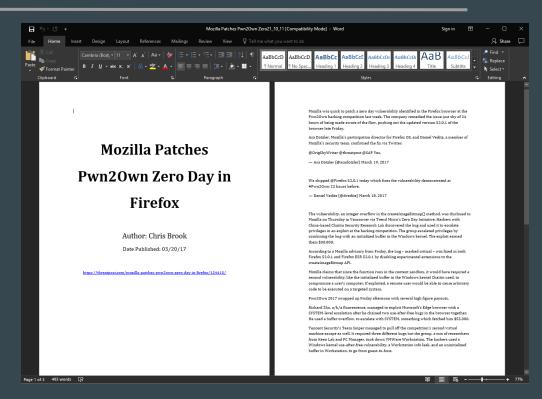


#### Ability to Save/Open Previous Searches



### PDF & Docx Export Options





## Behind the Scenes

#### Data Retrieval

- Web Scraping
  - Automated way to download and extract data from a website
- Crawlers/Spiders
  - o Bots built to navigate a website
  - Each one is custom built
- CySypder's crawlers deployed separately on the Cloud
  - Runs automatically on a daily schedule
  - Keeps data up-to-date and readily available at all times
  - Easier program maintenance

#### Data Storage

- Uses MongoDB
- Data Model
  - Spiders use a specific model to enforce consistent data storage
- Duplicate Prevention
  - Unique ID created by hashing the title and URL of each article
- Database Cloud-hosted by mLab

#### Data Access & Searching

- RESTful API used to request data
  - Created with Node.js and Swagger
  - Hosted on Apigee
- Data Access
  - GET requests to database with user query parameters
- Database Query
  - Use of regular expressions for searching key components of articles

### Data Analysis

- SpaCy
  - Python library to do part-of-speech tagging
- Article Subjects
  - Found by looking at POS tags
  - Tries to group phrases based on semantic similarity
- Subjectivity Score
  - Classifies each sentence as objective or subjective
  - Calculates average of all sentences

### The All-Important GUI

- Created with Python
  - Uses Tkinter GUI framework
- Uses multithreading for I/O and analysis