CyberSquirrel 1 

[CS1 Google Spreadsheet](https://docs.google.com/spreadsheets/d/1Z5KhMow9v3L3ST3XfSsi-PewrxqwKYx1uGrrVbzmnjA/edit#gid=622454854)

## Tab 1 - map\_data

The first tab should be pretty self-explanatory there is one line per entry. Each column has header descriptions. Some columns are color-coded (via conditional formatting based on keywords / letters).

### Column A - Date

The date is YYYY-MM-DD because that’s really the only format that sorts properly, unfortunately Google Sheets doesn’t really like that format. A Sheet bug / oddity. It will default to 'automatic' for cells, then try to guess what you mean. You have to force it to plain text to use that format, but you get the value of sorting and doing easy stats. Any errors in sorting or date-related stats messing up, this is the likely culprit.

### Column B - G

B is generally the article title or the contents of the tweet or some other descriptive identifier. That is followed by the type of animal, City, State, country, the URL, and whether not the URL has been saved into the Internet Archive. When adding to the sheet, check archive.org to see if they have a copy. If not, you can instruct it to make one. Due to serious link-rot the last few years, this has become essential.

### Column C - Animal

If you have an unusual animal verify that the animal is listed in the animal section on the ‘Statistics’ tab. If it is not you will need manually add the new animal to the list. You can copy the formula from another animal and just edit the name for ease.

All bird and reptiles are simply listed as ‘Bird’ or ‘Reptile’ in column C. Specific species are then listed in column N under species. For example Bird nests are list as Bird in column C and then Bird nest in column N.

Column H - WayBack Machine

It is important to archive as much as possible into the Wayback machine so that others can attempt to verify our research if necessary. This column has four designators based on Wayback responses.

* Y - Item was successfully archived into the WB
* B - The WB returned a blocked code, usually for a robots.txt
* E - Sometimes the WB returns an Excluded code, usually due to a site specifically requesting the Archive not to archive them.
* N - For whatever reason this item has not been archived and should be revisited.

The WB machine has recently changed its policies around robots.txt and no longer respects them as they once were; However, we still encounter sites blocked by robots.txt but not as much. All B, E, and N codes should be periodically revisited to see if they have changed and can now be archived. Moving forward, also consider archiving via archive.is for redundancy.

Suggest installing a [Chrome plugin 'Save to Wayback'](https://chrome.google.com/webstore/detail/save-to-the-wayback-machi/eebpioaailbjojmdbmlpomfgijnlcemk?hl=en-GB). It is a little buggy but makes sending pages to the archive much much easier than doing it manually.

### Column I - J

After the Wayback Archive designation column is the outage time in minutes and the number of customers impacted. If those fields are blank then the data was not checked, if there is a dash then the data was looked for but was unavailable. Often even when this data is listed in the source it has been approximated. In the past I have erred on the side of accuracy and if a source says ‘power was restored in the afternoon’ I would leave the entry with a dash to indicate it was checked but unavailable. Use your best judgement when entering data based on vague wording.

### Column K

This column is for infrastructure keywords. We currently track about a dozen or so keywords. See the Target section on the statistics tab. A few of the more important keywords IMO are fire, nuclear, and military. You can search for these keywords in the map\_data tab to see some examples. Basically if a keyword is mentioned as being impacted in a source I add it to this column. I am also kind of partial to the ‘Christmas’ keyword which I use anytime a squirrel takes out christmas tree lights. Feel free to add new keywords in the future if you feel they are worthy. If you do, consider backfilling them as well, else that stat will have to come with a caveat like “*tracked since $date*”.

### Column L

We use this column to try to track outages that occur on holidays. Holidays have been problematic because Google sheets doesn’t really have a good calendar integration and they are very U.S. centric. Holidays are based on [the OPM list](https://www.opm.gov/faqs/QA.aspx?fid=e64d74ab-20a3-484c-8682-d2a2b46c22da&pid=c41e6beb-0c14-449d-bde5-355a3a3014cd).

### Column M

This is a general notes column but also where we track fire and deaths. We note when there are casualties by adding e.g. “3 dead” (you can see an example on 2015-11-05).

### Column N

This is used to track sub species of birds or reptiles, this includes bird nests. Other species could be abstracted if we start to see that data (e.g. monkeys).

### Column O

If there is more than one relevant URL that data was pulled from it can go here.

## Tab 2 - Export

I try to export the data monthly to the website. The website map only requires the first few columns so the data is copied from the map\_data tab and pasted here for export to CSV before uploading to the website.

## Tab 3 - Statistics

The picture (currently broken) is a Bald Eagle. Google sheets likes to lose it a couple times a year, we don’t know why.

The total incidents is an example of the buggy nature of adding a new row in map\_data, as the formula for G1 will auto-update. At the time of this writing, the formula is =ROWS(map\_data!52:6804) which should be “2:6804”. Fixed now but this will be the #1 reason for statistics being off.

The world map is very slick, but has been buggy since day 1, primarily in the weighted coloring it uses. It also sometimes will update to use different designations for the country, which I think it has done again since the U.S. is no longer colored.

## Tab 4 - Human Outages

This is kind of a pet project for us both. One of the underlying lessons, humourous or not, is that there are far more outages and deaths by Squirrel than ‘Cyberwar’. This sheet attempts to catalog known and verified issues of cyberwar that lead to outages or death. Be skeptical if a new article comes out claiming such a thing; often times they are debunked or heavily changes months or years later.

This data will ultimately give you any fuel you need to head off arguments that squirrels are not a threat, and let you argue specifically about the number of outages/deaths by each side.

## Tab 5 - Eaton Import

Eaton, a diversified power management company, has a project they call ‘Outage Tracker’. Similar to CS1 they track power outages, unlike CS1 they try to track all outages. We get a copy of their report once a month usually a few days after the first. The data comes as an .xls but all the outages are mixed up, so they need to be sorted and then the columns normalized to our data. They usually have between 10 and 20 animal outages per month. Anywhere from 1 to 10 of these are outages that CS1 missed and does not have listed. Entries are then made for this missed incidents into map\_data.

A green date on the Eaton import tab is an incident that CS1 already had recorded. A red date is an incident that was somehow missed.

## Tab 6 - Research

These are incidents that are likely worth including, but needs a little digging. Either the articles aren’t available or something else is stopping it from simple addition to the main data.

Basically this a dumping ground for articles where we know we can find incidents but haven’t had the time to dig into the research.

## Tab 7 - Strange Shit!

This tab is special to me (Jericho), as it captures all the oddities around animal-related outages. These make for good anecdotes and help people better relate to the project. A lot of these are around damages that did not cause an outage.

## Tab 8 - Anecdotal

A mix of fun stories related to all this, often mentioning statistics about animal-related outages. We also like to capture articles that talk about the defenses power companies use to prevent such damage, especially noting how much they pay (which can be in the millions).

## Tab 9 - Translation

Google has a handy translation feature that lets you add a word, then select/drag the initial cell down to capture all the translations it supports. This was created to help facilitate searching for incidents in countries that we are missing events from.

## Tab 10 - Eaton - Verify Unknowns

The Eaton data has a category for ‘unknown’ cause of an outage. Based on inaccuracies in their animal outage data I know there are probably numerous outages that Eaton lists as Unknown that are in fact animal related. At one point I thought it would be important to save these unknown so that we could go back and see if they were in fact animal related. I have never had time to go through any of these and have actually stopped adding new data to this tab.

## Ideas and Goals

I (Jericho) have spent quite a bit of time trying to get data from countries that don’t appear in the list. I have tried to do this via friends in foreign countries, as well as hiring people via Fiverr. Overall, this has been mostly futile for a variety of reasons. A great example is that one of the words we use commonly translates to something you see in car advertisements in Russia. So Google searches, which weigh heavily based on your geo-located IP, even if you use the TLD-specific Google (e.g. google.ru). Over time, it would be neat to fill in the missing countries.

SR - Yup, long term goal has been to paint the world red, at least one outage in every country.

## Searching

### Google

Google messes with their search algorithm all the time and it is highly localized based on IP address. Also Google news alerts are often either stale or not fresh enough. In addition the alerts sometimes fail to fire for days or weeks at a time. This necessitates occasional manual searching based on date restrictions to find things that Google news alerts missed.

I currently have two Google news alerts. One for simply ‘Squirrel’ and one for

squirrel OR animal OR bird OR snake OR raccoon OR rat OR marten OR beaver OR monkey OR possum OR jellyfish OR osprey OR lizard AND outage OR electrical OR power OR Internet OR fiber

Unfortunately Google has a limit on how long your boolean string can be. Occasionally experiment with different strings and see what results you get.

### Bing

I use the same search strings for Bing news alerts that I do for Google. Bing occasionally finds something that Google missed but usually has way fewer hits in each email. Also Bing news alerts are a lot more problematic than Google in that they usually fail to fire.

### Twitter

The CS1 twitter account has two lists, ‘PC’ and ‘Power Company’. I have forgotten which is which and so now I just add power companies to both lists. I use these lists to do searches via TweetDeck. TweetDeck is the only client I know of that allows you to do searches easily. I have a couple of different tabs with variation of the standard boolean search string.

This string only searches the PC and or Power Company lists

squirrel OR animal OR bird OR snake OR raccoon OR rat OR marten OR beaver OR monkey OR possum OR jellyfish OR osprey OR lizard

and this string searches all of twitter

squirrel AND outage OR fire

I used to have a lot of success with that first string but the last few weeks have been very slow.

Be careful with the second search. If a tweet reporting an outage comes from a power company, local news outlet or government agency I would include it in map\_data. If the tweet comes from a random person I did not feel it had enough authority and so would not include it. Also if a tweet does not give enough information such as date or location I would not include that in map\_data either. I would rather miss an outage than have inaccurate or incomplete data.

## Website

The website is currently hosted at pair.net. It can stay there for the foreseeable future. Any new host would require PHP and SQL to move everything over.

The script for updating the map was written by someone on Fiver in PHP, I really have no idea how it works or how to maintain it. The first time it was written it eventually broke as it exceeded the Google API free lookups. So I had the person rewrite it a few months later. Then it broke when Google started charging for maps access so it was switched to OpenStreet maps and the script was rewritten again. The PHP script triggers off the date of the map\_data.csv file. So when you upload a new map\_data.csv it will automatically have a new creation date which should trigger the PHP script to repopulate the map the first time it is loaded.

Spot Check

Every month after I upload a new map\_data I try to look for one of the new outages on the map from that month to make sure that the data got ported over correctly.

map\_data.csv

Everything runs off of this file. To generate this file copy and paste the first seven columns from the map\_data tab on the google spreadsheet. Paste that into the export tab of the spreadsheet overwriting the existing data. Download the Export tab as a .csv file from Google Sheets. Rename the file new\_map\_data.csv.

Google\_map.php

This is the PHP script that does all the work. There should be no need to touch this script until we switch web hosts. Hopefully when that happens someone will be able to figure out what needs be edited to make it work on a new host. Shouldn’t be to hard.

Index.html

Edit index.html with new stats. Basically just update the date of the last map update and the number of ops for each agent from the stats page. This is pretty basic HTML, nothing fancy here.

## Twitter

See also the Twitter section under Searching.

I don’t tweet every single outage from the twitter account. I try to only tweet recent outages and ones that may seem weird or unusual. I also try to occasionally interact (ie. taunt) power companies and the like. All of this is done as a persona of the Minister of Propaganda of the Cyber Squirrel Army.

When tweeting outages I usually take the format of

Article Title Article URL - City, State

Downtown Bend power restored after three hours, Second animal-related outage of day; 3rd overall <http://www.ktvz.com/news/four-hour-power-outage-hits-busy-eastside-bend-area/665157807> - Bend OR

Once a month, after a map update I will also tweet stats for that month,

Aug ’17 -Total Ops 51, Squirrel 20, Bird 16, Classified 5, Rat 3, Snake 2, Raccoon 2, Cat 1, Bobcat 1. States 23, Countries 8 #cyberwar4ever

And then pin that tweet to the top of the twitter profile. This isn’t necessary but people seem to like it. It also makes the twitter page look like it is being updated on a regular basis.