**Telecom Policy, Regulation, & Enforcement | Radio Spectrum**

**Digital Studies and Methods Certificate**

Capstone

J. Stephanie Rose

[JSR67@Pitt.edu](mailto:JSR67@Pitt.edu)

Department of Informatics & Networked Systems

School of Computing & Information

University of Pittsburgh

**ABSTRACT**

When I originally began this course in 2017, I wasn’t sure how different digital methods would be from what I experienced in the typical Computing and Information classroom. During my short time at the University of Pittsburgh, I have been introduced to a multitude of methods and approaches to solve technological problems. These approaches have included everything from algorithmic design and implementation to system design and analysis. However, once, I began this series of courses, I realized that the difference between the Digital Studies and Methods and the typical SCI classroom is a more mindful adoption of methods and approaches that are not – digital for digital sake- but, a careful investigation and understanding of what it is your investigating (which doesn’t necessarily need to be a problem), investigating various tools which may enable you to view your “problem” or area of inquiry from a different perspective, and learning that even “failed paths” or unexpected outcomes are learning opportunities that allow the researcher to reflect more intensively on why and how they are approaching their research.

My area of interest within the Department of Informatics and Networked Systems (DINS) is the automation of spectrum policy. Thus far, my perspective on regulator needs within this area is something that is limited in terms of related literature. One of the few things that scholars within my area can agree on is that the Federal Communications Commission (FCC) can/should be doing a better job. Therefore, I decided to use LIS 3600, LIS 3902, and this capstone as a chance to investigate some of these spectrum issues further.

This paper discusses my journey into a more mindful adoption of digital methods. Put simply, I began with the collection of data, which lead to an investigation and trial & error phase with tools such as Geospatial Information Systems and topic modelling/document analysis, and ends – at least for now – with my adoption of GIS tools that worked best for me, the results of that work, and a final reflection of my journey in its entirety.

**INTRODUCTION**

In the beginning of the digital technology era, policymakers and interested stakeholders discussed at great length the importance of the information highway and its pertinence for the continuance of technologic innovation. Discussions such as these took place in the late 70’s between interested service providers and operators like Time Warner Cable and U.S. West (National Archives II Record Group 417NIIAC-3) – who were interested in bringing a “digital pathway into the home”- just as commonly as they occurred between the Association of Computing Machinery (ACM) while they were positing the information highway and its definitions (National Archives II Record Group 417 NIIAC-6). This information highway – electromagnetic spectrum, also referred to in this paper as radio spectrum or “spectrum” for short – would eventually become the very resource that connects and drives various technologies & services (e.g. cellular devices, Wi-Fi, computers, wearable technologies, home assistants, etc.) in the 21st century.

Within a few years, even more innovative technologies will emerge within the commercial sphere. With such a projected influx of technologies that will be dependent upon electromagnetic spectrum (such as fully autonomous vehicles, fifth generation cellular services, embeddable technologies, and alike), an increased interest in the regulation – and by extension enforcement – of spectrum has come into the forefront regarding present-day discussion and concern of future spectrum management.

My interest in electromagnetic spectrum began four years ago. At first, I was interested in how emerging technological advancements appeared to be a driving force for policy initiatives. Since then, my focus has shifted towards how technology is being regulated and the policies enforced. This is large in part to the research I am conducting with my advisor, Dr. Martin B.H. Weiss and my fellow colleagues where we are investigating wireless automated enforcement measures – from an incumbent perspective and third party regulators such as the Federal Communications Commission. However, for this work, we are more focused on incumbents and the various types of policy implementation (through de facto and/or de jure rights).

Through this work, I began to question how enforcement is being adjudicated currently at the regulatory authority level and if an automated enforcement structure were to be created amongst incumbents, would the Federal Communications Commission be prepared for such an advancement within the policy space. The work that I have been accomplishing for the past year through the courses for Digital Studies and Methods is a culmination of the research I have been working on and hope to use towards my dissertation in the future.

The remainder of this paper is comprised of the following sections; Motivation – within this section, I provide further explanation regarding my selected topic of research. Additionally, within this section I discus what electromagnetic/radio spectrum is, why this topic is important/relevant, and the irony of working on a technology driven research topic which does not traditionally deploy and/or utilize digital methods to analyze problems. Following the motivation section, I discuss the previous iterations of this project. Beginning with the inception of this research in Fall 2017 and ending the section with the work I completed in Spring 2018. This section covers the original project goals I had, the analysis and results at that time, and the future work that I previously hoped to accomplish. Within the points of failure section, I provide a list and explication of digital tools and/or methods that I have attempted throughout this process and why they did not/still not do not work for my captaset. Subsequently, the “points of failure” section is followed by the section entitled “adopted digital methods”. This section discusses which digital methods worked with my data – and me – and provides an explanation as to why these methods were more appropriate for my research. Additionally, I have provided a learning opportunities section which includes workshops and other learning opportunities that have allowed me to think about how to best approach my research utilizing digital methods. Some of these learning opportunities have either disproved or reaffirmed why some of the digital tools and methods were unfit or should be given another try regarding my project and expected outcomes. The next section is about my capstone project and the final phase of my research for the purposes of the digital studies and methods series. Within this section, I provide the final milestone project expectations, progress throughout this semester, and results. The final components of this paper discuss overall limitations – which is separate from the points of failure regarding the digital tools/methods that were not used to accomplish this research, future work – how I plan to expand this research for my dissertation, and lastly, the final reflection and discussion – which recapitulates my research focus and questions, the importance of utilizing digital methods, and the inferences that can be drawn from the results.

**THE DATASET**

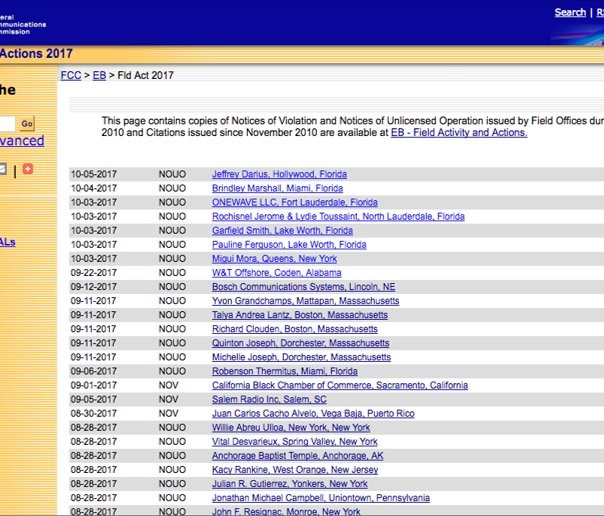


Figure https://transition.fcc.gov/eb/Orders/Welcome.html

At the beginning of this research project, my first course of action/method was collecting data. I came across a [[1]](#footnote-2)database on the Federal Communications Commission’s (FCC’s) transitional website for the Enforcement Bureau (EB). Initially when found it, I was under the assumption that this information was solely focused on violations and enforcement actions adjudicated by the FCC. Later on, after I progressed further in my data collection, I would come to find that this information wasn’t just a list of offenses, but an “archive” of records. This record set began with the inception of the Enforcement Bureau in 1999 and continues to be a repository for present-day EB dealings. The documents collated on this website contained records ranging from policy changes/updates to requests for changing forfeiture amounts and special license considerations – in addition to violations such as spectrum violation and licensees not operating under the terms of their agreement with the FCC.

My primary objective with this dataset was to parse out spectrum specific violations in order to gain a better understanding of how prevalent spectrum interference is with regards to the current technological landscape – which for commercial use is overseen by the FCC. Once I identified my data source and specific subset, I needed to figure out the best way to collect this data in a useful manner. The first constraint that I came across with my dataset, was that each record was encoded as an html document. Typically, with online documents, one can extract the data by scraping, which uses code to parse out specific information needed. This ended up being a cause for concern due to the manner in which it was encoded. Each html file was coded in such a way that it hindered current data collection practices – scraping. The problem with this approach – in terms of my data – was that the information contained in the documents was not organized under any specific tag. Using tags to code within html is a common practice. For example, if I am constructing a document such as this self-reflective essay, I might utilize an <h1> tag to code the title of this paper, but would use the <p> tag to construct each paragraph. Instead, all of the documents’ content (title, case information, etc.) is encoded under the tag <pre> which renders common scraping practices futile-this can be seen in figure 2 on the following page. Therefore, the FCC EB records are not aligned with common coding practices.

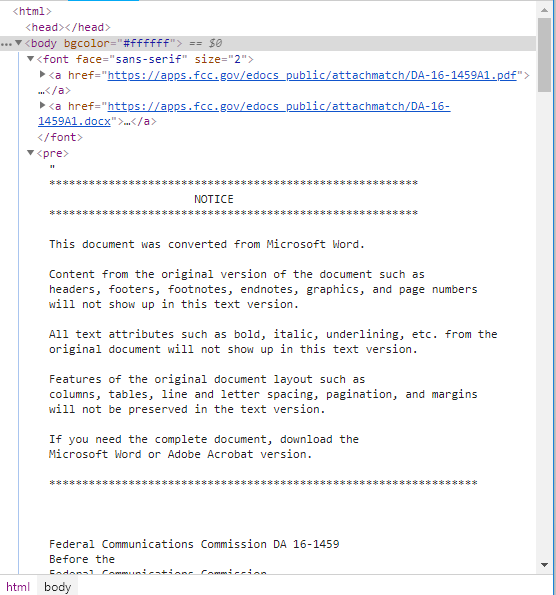


Figure https://transition.fcc.gov/eb/Orders/2016/DA-16-1459A1.html

Upon realizing this, I then decided to go through each record “by hand” in order to construct a spreadsheet with the necessary information to conduct my research. By taking this approach, I realized early on that my project was going to be labor intensive. However, I felt as though it was imperative to do the “leg work” because this is an area - and more importantly, a dataset - that to my knowledge has yet to be investigated within my field.

Additionally, I had the opportunity to conduct some of my research at National Archives II in College Park, MD. My reasoning to do this was twofold. Firstly, I wanted to see if there was a hard copy available of the records that were online – some of the online links were broken or missing information. Secondly, some of the records between the years 2004 and 1999 eluded to records regarding their case prior to 1999. During my visit at the archives, I was hoping to identify these cases in order to have a better understanding of their violation/ offense. Unfortunately, this would be the second data constraint I would endure.

Although the archives were a great experience – one that I would suggest to any young researcher – one set back that I did not consider is that some of the governmental documents are placed on an administrative hold (sometime between 10 and 20 years). Due to this administrative barrier, I was unable to follow up on specific records that pointed to violations prior the 1999 dataset.

Despite this setback, I was fortunate enough to come across spectrum violations from the 1950’s. These cases were adjudicated by the Enforcement Bureau’s predecessor the Field Engineering and Monitoring Bureau. Even though the amount of violations within this archive was limited – no more than 30 or so cases of spectrum interference – the information contained within each casefile provided a type of confirmation of some of the inferences I drew from the online dataset. These inferences included, but were not limited to lack of administrative organization & consistency among records, consistency of enforcement for similar violation types (e.g. purposefully interfering with other frequency bands versus spectrum interference due to equipment malfunction).

Moreover, the archives provided a foundational information regarding the consistency of the business practices of the FCC in relation to spectrum violations and their adjudication practices. Some of the violations recorded from that era were recorded on a form 951. This form provided essential information regarding the case such as date of the complaint, full name and address of the offender, name and address of the complainant, regional office, type of spectrum interference, frequency of spectrum interfered with, etc. However, similarly to the online documentation, this process was also inconsistent. Some of the records did not have a form 951 on file – nor was there mention of one. This lead me to infer, that the record keeping practices haven’t changed much since the 50’s as there is rarely completeness of the violation and in regards to the cases from 1999 and after, the content varies from regional office to regional office. Due to the amount of missing content per violation, some cases needed to be dropped from the overall dataset solely because of insufficient information regarding the incident.

In regards to the online violation records, they did not include the format and/or mention of the form 951. However, some of the information recorded by enforcement bureau agents were similar. Information such as location – and in some cases multiple locations or no location, the responding enforcement bureau regional office, complainant, and frequency of the violation. Yet, this information is not standard across all of the records I read. Some violations included the entire violation history (if it was ongoing for years), previous forfeiture orders and/or penalties, whether the violator was licensed or not, etc. In cases where I could not pin down the specific information that was consistent in most of the records (e.g. entity that violated, location(s) of the violation, complainant, responding regional office, and frequency being interfered) I decided to either change the attribute that I was collecting or dropping the record from my dataset – which caused me extreme consternation.

Changing the attributes of what I was collecting became easier with the more records I reviewed. This was because I was able to “see” a pattern of useful information within each violation that I could then use for analysis. Towards the end of my analysis – the capstone – the attributes I chose to focus on included the name of the person and/or business receiving the violation, the case number – linked the corresponding html document, date of violation – and if not specified the date of the violation notice/enforcement action, city, state, frequency/ explanation of violation – for spectrum implicit cases, penalty – if there was a financial penalty imposed, enforcement type – the publication the violation was filed under (e.g. NOUO, Forfeiture Order, etc.), type of entity (such as a business (BUS), individual (IND), or religious establishment (REG)), enforcement bureau department location, and lastly, whether the violator was licensed/unlicensed.

**PREVIOUS ITERATIONS**

*First Milestone – Fall 2017*

Once I realized the constraints with my data and data collection, I decided that for this iteration of my research, I would review records between the years of 2014 through 2017 – this is prior to my spring visit to the National Archives. My hope was that by reviewing the most current years, the data would provide me with most information pertaining to spectrum violations and enforcement mechanisms that are relevant to present day.

At this juncture of my research, my research questions were seemingly straightforward. My initial questions were:

1. How often does spectrum interference occur?
2. Who are the primary offenders (individuals, businesses, etc.)?
3. What kind of offenses/ interference occurs?

Since this project component, as well as my capstone project overall was conducted in an iterative manner, my preliminary findings – for the first milestone/iteration – that occurred during fall 2017 were derived from the 2017 data (at that time it was all that I had collected by October). Using the preliminary data and results, I decided to attempt Google Fusion Tables for spatial analysis. I selected this tool because it appeared to have features that would allow me to analyze and visualize my data in such a way that I would be able to have outputs that would be more than just “dots on a map”. Through adopting this approach, it appeared as though spectrum interference was not as prevalent among the violations as I had previously assumed. The main offenders during the initial analysis lead me to hypothesize that main offenders were individuals operating unlicensed radio operations from residential areas. Additionally, these offenders appeared to be clustered along the coastal lines. Lastly, the kinds of offenses/type of interference was a varied amongst the different frequency bands – no specific frequency allocation appeared to be the target – shown in figure 3.

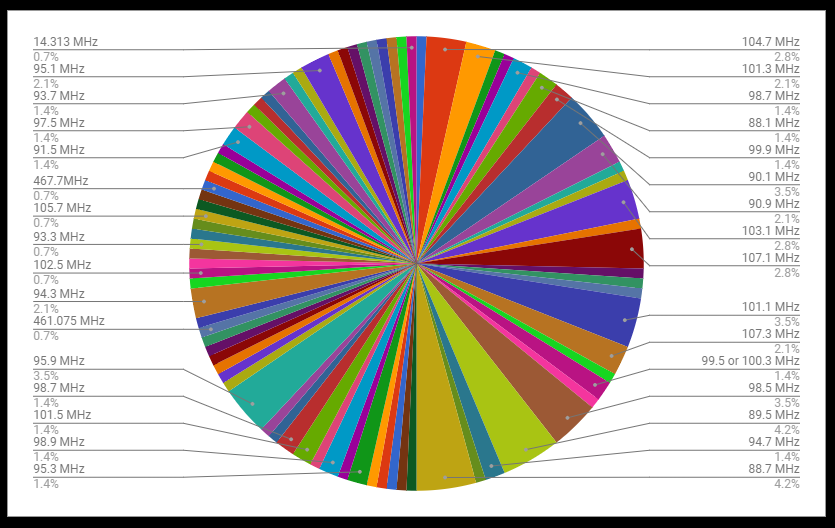


Figure Initial Allocation of Frequency Violations 2017

When I analyzed the data, various frequency bands had been interfered with over time (no band necessarily violated more than any other). Moreover, most of the official notices were published as notices of unlicensed operations (NOUO’s). These NOUO’s typically carried no financial penalty. This was especially interesting because some of these cases were exceedingly harmful (e.g. purposefully causing interference with a public safety band – bands used by police and other first responders, etc.). Based on this information, I began to believe that violations from businesses and/or service providers may be limited to non-existent.



Figure First Attempt at GIS with Google Fusion Tables

My final iteration during this time period had reaffirmed my initial analysis results as there was little change from the 2017 data and that of the additional data I collected from the 2014-2016 dataset. By this point, I had switched tools from Google Fusion Tables to Google Maps. During this shift, my research results remained the same for the most part. Spectrum interference accounted for 25% of all the violations adjudicated by the FCC. The individual actors were still the primary culprits – unlicensed operations (pirate radio), however, there was an uptick in violations that were something other than pirate radio. Subsequently, this changed the results for the kinds of offenses/interferences. The spectrum violations now included excess power usage and continued operations with an expired license.

*Second Milestone – Spring 2018*

Naturally, by the time I needed to continue my research project, I was concerned about not having a large enough dataset to draw substantial conclusions and/or utilize digital methods/ the selected tools in a more meaningful way. Therefore, during this milestone, I focused my efforts on more data collection within the initial iterations. My dataset objective for this semester was to collect records from 2010-2013.

It was during this additional data collection that I began to realize that the corpus of spectrum violations had more of a range than I had previously inferred. Prior to this iteration, I had assumed that spectrum violations would be pretty clear cut. However, as the violations became more nuanced, I realized I needed to not only parse out spectrum violations, but categorize the spectrum cases. It was at this point that I began to the delegate whether a spectrum violation was spectrum explicit or spectrum implicit. I came to this categorization due to the wide array of spectrum significant violations. Some violations were recorded by the FCC as an individual using a device for pirate radio – meaning no license for operation. While other violations may be less specific such as an individual or business selling devices that are unauthorized or have been tampered with in order to operate on frequencies other than authorized by the FCC.

Spectrum explicit cases were reserved for records where the FCC EB identified a clear occurrence of interference and which frequency band the violation took place. Spectrum implicit violations included records where there may have been no frequency band identified, but interference took place, cases where limited or no details were provided yet, the entity could cause interference.

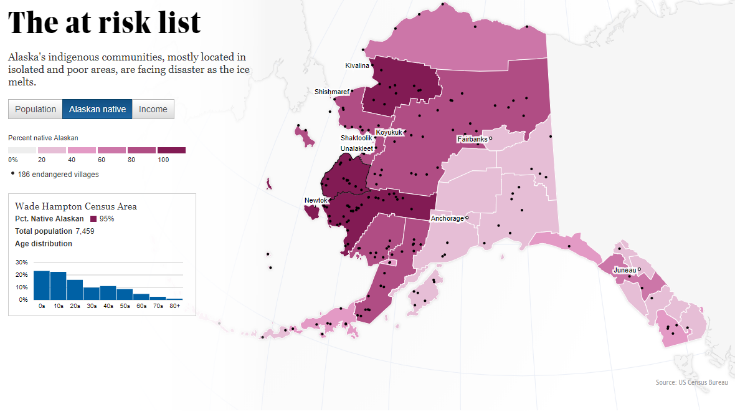
Moreover, I also looked into other digital methods and tools in an attempt to find more sophisticated ways to showcase my findings. During this search, I came across an article published by the Guardian.

Figure https://www.theguardian.com/environment/interactive/2013/may/14/alaska-villages-frontline-global-warming

In this article, they were able to convey their information both spatially and statistically. This display was informative without being cluttered and conveyed the information in an interesting way – using a mixture of a choropleth layer, markers layer and a side bar that included interactive statistics and bar graphs. With my limited knowledge base, their use of GIS was something that I could only hope to aspire to. Therefore, I once again shifted GIS applications. This time, I switched from Google Maps to Tableau in order to have a wider variety of functionality.

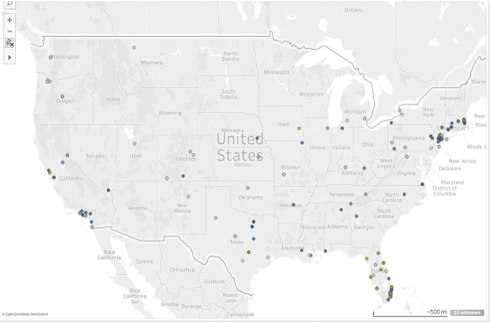


Figure Now Using Tableau

When I use the term functionality, I am referring to what capabilities I have with the tool. With Google Maps, I am able to add my data set and spatially visualize the violations by city, state, etc. and see the map populate based on the occurrence of those records. However, with Tableau, I am able to choose visualizations that are spatial from either a more raster or vector approach. Additionally, I am able to chart, graph, and/or even aggregate the data in the form of a bubble plot. I am not, however, able to embed or showcase this information publicly due to my current student/ free use status with Tableau.

Towards the end of this iteration, my dataset now included records spanning from 2017-2010. I now had roughly 2300 cases and 356 spectrum violations to analyze. My conclusion was that this amount was still not substantial enough to deem spectrum violations as a pervasive violation. During this time, I also ran into user errors with Tableau where I was throwing an error for 165 data points. Despite this hiccup, many of the violations still appeared to be occurring along the coastal areas.

**POINTS OF FAILURE**

When I began this research project, I knew I wanted to analyze and represent my results spatially. Unfortunately, I didn’t exactly have the knowledge base to accomplish this in the same or similar manner as the Guardian article. Furthermore, I knew even less about cartography. With that said, the original tools I reviewed and eventually dismissed included Color Brewer, Scribble Maps, QGIS, and CAST. In addition to spatial analysis, I was also interested in possibly using topic modelling and/or document analysis to further investigate the spectrum implicit cases.

Color Brewer, although interesting, did not allow for the depiction of actual data. In terms of visual representation, this meant that if I decided to implement this tool, I would have been subjected to results that were akin to “dots on a map” which is what I expressly did not want as a final result. Scribble Maps on the other hand, appeared to be promising. Yet, in terms of data collection and learning a new tool from scratch – with no previous experience or knowledge – this application became intimidating and time consuming. QGIS, although free, was also intimidating and would be a learning curve. I feared that I would not have been able to learn it well enough in order to successfully use it for my 2017 iteration. CAST was also labor intensive. Additionally, in order to utilize this tool, I would have needed to download additional components such as GeoDa and PySal – which I was unable to do on my laptop as it was already at its limit with software. Lastly, as discussed in the previous iterations section, I attempted Google Fusion Tables and Google Maps. Both of these applications left me with the “dots on a map” feeling, therefore I ditched these approaches for Tableau where I was able to accomplish more than just spatial visualization.

In regards to the topic modelling and document analysis. I ran into the same error that encountered with the scraping tactic. Due to the encoding of the html documents, applications such as Voyant threw an error when I attempted to analyze the documentation links.

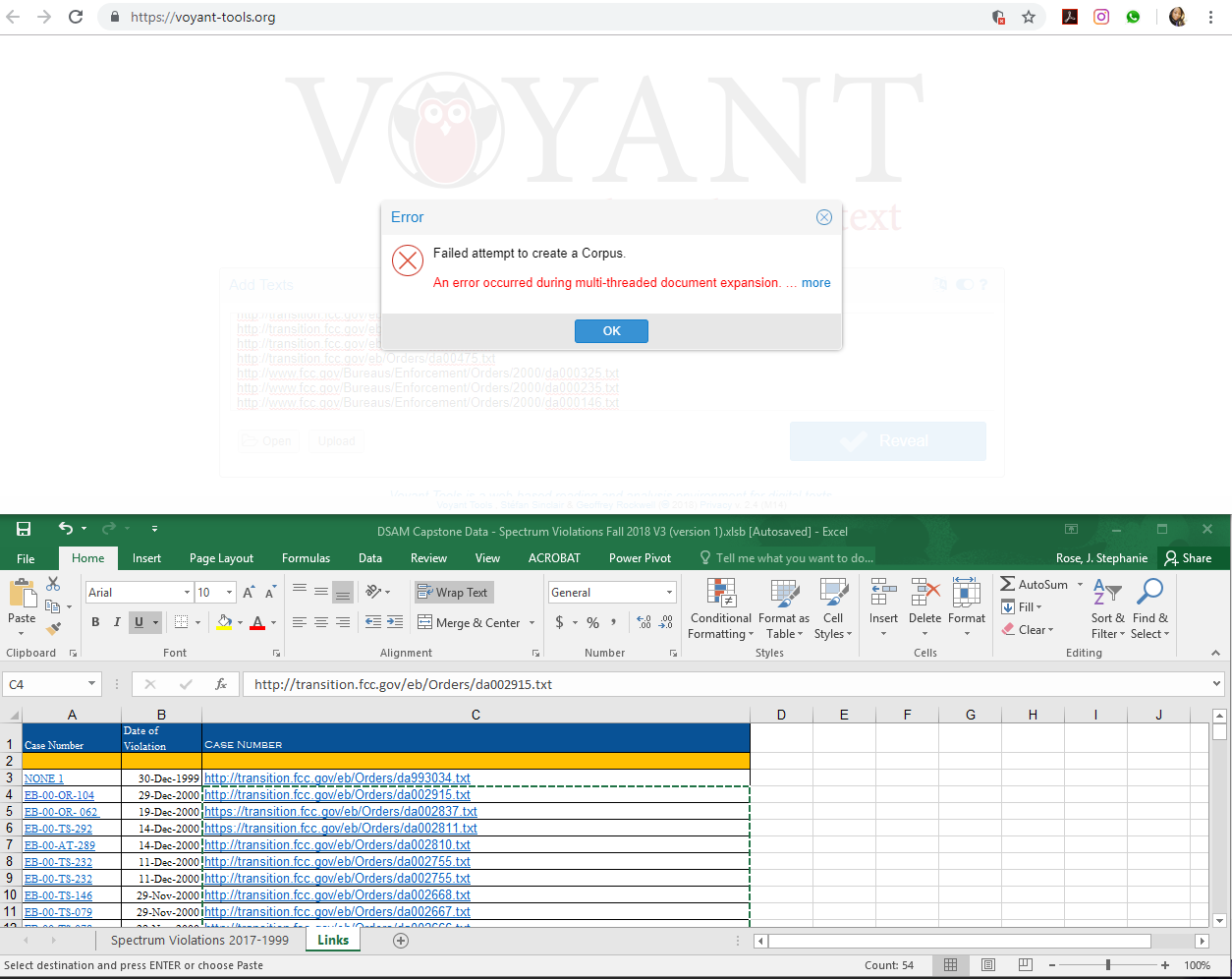


Figure Voyant Error

**ADOPTED DIGITAL METHODS**

When considering time that needed to be allocated to data collection and learning a digital tool, I decided to utilize Tableau to analyze my dataset. Even though Google Fusion Tables and Google Maps were pretty simple to use, they didn’t have the same additional functionality as Tableau. When I first downloaded Tableau and began using it, I was slightly frustrated, however, the more I used it, I realized it was pretty intuitive. The only real issues I ran into with Tableau was that I am unable to embed my work within a website – this would require me to utilize the online server and make my data public (which I do not feel that my research is ready for unlimited public consumption). In order to use a private Tableau server, I would need access to the University’s Tableau server access, which I am unable to do because I am a student and no a primary investigator for a research project.

In terms of document analysis, instead of attempting to analyze each link specifically, I decided to upload the spreadsheet of my dataset. Although I knew it would not yield the same results as the actual FCC records, I was still able to get some interesting results on my dataset.

**LEARNING OPPORTUNITIES**

Due to my severe ineptitude regarding geospatial information systems, I took advantage of opportunities where I could learn more about GIS. These opportunities were great because it provided me the chance to think more spatially and not just in the terms of my own research, but also from the perspectives of others research areas of interest and it provided insight on how other disciplines utilize and deal with implementation of GIS.

Furthermore, it was enlightening to hear what issues others came across and how they were able to resolve their issues – even if it was by using another tool altogether or by changing their approach (and in some cases updating their dataset to fit the parameters of the tool). What I found most helpful, was when some of the workshops required the participants to use paper and writing utensil to scaffold what they wanted to get out of using GIS. This assisted me with thinking with more of a cartography perspective more so than a coding or software driven approach that I began to take with some of the online browser GIS applications. In the below subsections, I provide a brief synopsis of the learning opportunities that I experienced and how each have aided in shaping my thinking in terms of my primary digital method – GIS.

*DATA 101: INTRODUCTION TO MAPPING*

Data 101: Introduction to Mapping was the very first workshop I took in hopes that I would gain knowledge to help me with my digital methods research. Overall, the workshop consisted of about four activities total. The exercises ranged from regional analysis of dogs to mapping a virus/epidemic – I honestly cannot remember the exact problem we were trying to solve- however, I do still have the map (it’s a drawn street view with sections and dots).

During this introduction to mapping, we were given papers, colored pencils and what appeared to be seemingly easy instructions. The one of the papers that was provided to us had a template for a geographic region, the other sheet of paper was a list of dog names. First we were to use stickers and color to annotate what dog names were most popular where. Second, we needed to create a legend indicating the amount of dogs located within each region.

Another activity we did in this workshop was a little more intricate in terms of detail and placement. This exercise required us to view a list of parcels of land, owners, land use, and whether they were Homestead exempt. In this scenario, we needed to map owners on specific parcels of land, however, there were added constraints (e.g. an exempt owner could only purchase a plot of land that near a plot owned by the city – or something to that effect). This activity was more than just counting cute dog names and visualizing according to amount, but required careful study of who the parcel of land belonged to in addition to other attributes that were listed on the information sheet.

Although these tasks may seem easy to some, I realized I am not as familiar with mapping/cartography as I thought I would be. These activities weren’t terribly hard, however, I haven’t had to actively think spatially in years! Aside from the occasional PokemonGo gameplay, I heavily rely on GPS, or some other service/device to “think spatially” on my behalf. So, this workshop ended up being a great way for me to start thinking in a more spatial manner.

SHANNON MATTERN

In this workshop, we broke into groups based on programmatic themes. Thinking in transdisciplinary terms, we attempted to construct a social and structural organization. We discussed topics such as digital inclusion being transformed into digital equity and the subsequent pedagogical challenges.

This workshop enabled me to think in terms of specific methods and skills and posit what different disciplines and fields we can draw information and expertise from in order to gain further knowledge. My key takeaways from this workshop were to critically think about where the intersection of my research lives and what other disciplines I might be able to utilize on order to learn about new methods/approaches and how other might see the “spectrum” problem.

So far, I have investigated my area of interest in terms of information science and its intersection to policy. However, now – and especially after the first workshop- I began to think in more geographical and historical perspectives of my work. Although at the time in my research (spring 2018) I had yet to fully review how History as discipline could assist with further investigating my research, I did begin having more conversations with my Geology department friends about the tools and approaches they used and began to try to understand my data more from that perspective.

LUNCH W/ ANNETTE VEE

Lunch was delicious! Panera is simple, but never disappoints…and neither did my conversation with Annette Vee. Although brief, I got the chance to discuss some of the woes I had with my research. Although the focus of my research for this series of courses has been commercial spectrum regulated by the FCC, for the past couple of years I have been trying to understand the governance of federal spectrum more in order to have a well-rounded picture and research product. Unfortunately, it seems as though many have had issues with pinning down an National Information and Telecommunications Agency (NTIA) employee or even receiving documents through Freedom of Information Act (FOIA) requests.

I left the conversation feeling justified – and less crazy. Some information/ research areas of interest take more time to gather data than others. In some respect, I feel fortunate enough to have been able to come across a dataset directly from the FCC. However, it makes me wonder if I will ever get close to a similar dataset – or even similar issues – for federal spectrum users.

JACK HESSEL

The workshop conducted by Jack Hessel was one of the only workshops that I attended where we dug right into the computational investigation of digital methods. This workshop was pretty intense, but in terms of completely reshaping my perspective on my dataset, it probably gave me the most hope that my data was not a complete dumpster dive and that I might eventually – one days, no time soon – be able to analyze my documents using current digital approaches.

Using the Neural Neighbors data, we navigated machine learning. Throughout this workshop, we ran code without manipulating the source data to see how well it could find a specific image set. Then, we manipulated some images to see if the code would do a better job. I know this is oversimplifying this workshop, but my most interesting takeaway was how the code was able to “learn” on a pixelated level.

By being able to see the code categorize airplanes or cats, by making the images sharper or giving it specific parameters to “look” for, I started to think how this may be able to work in terms of text and the FCC records that I was not able to scrape using code. Essentially, if I turned each document from the FCC into an image, and through manipulating the code…would I eventually be able to have a program that could “read” my documents and provide output for specific terms like MHz and harmful interference? I still have not set aside the time to attempt this method, however, it a future aspiration that I hope to bring to fruition.

WORLD HISTORY SYMPOSIUM (w/ GIS WORKSHOP)

The final event that I attended was the World History Symposium that was held at the start of this current semester. At the symposium, I had the opportunity to experience and engage with scholars utilizing GIS for their research and the specific issues that they had to deal with regarding their datasets and tools, as well as their solutions.

Additionally, during the GIS workshop portion, I was able to see how QGIS – a tool I had previously disregarded – could be used to tell a historical and spatial narrative leveraging story maps. After learning about this plug in for QGIS, it renewed my interest in this tool as future approach to my dataset (at the time of this symposium, I had still not completed my data collection for the entire corpus of violations). However, the end products that were presented piqued my interest in using this application for my dissertation – at which point, I will be investigating all violations that fall under the FCC’s purview.

**MILESTONE III – CAPSTONE | Fall 2018**

The third milestone – my capstone project – is the final iteration of my research project for the digital studies and methods certificate. During this final milestone, I was able to complete the rest of my data collection. My dataset is now a corpus of records ranging from 2017-1999. This means I now have 8040 records pertaining to violations adjudication actions and policies. Out of this dataset, 1250 cases are spectrum violations, which account for 11.8% of the cases. After I completed my data collection, I then cleaned – cleared up any missing information, double checked that what I have on file matches the linked FCC document, and analyzed the data in Google Maps, Tableau and Voyant.

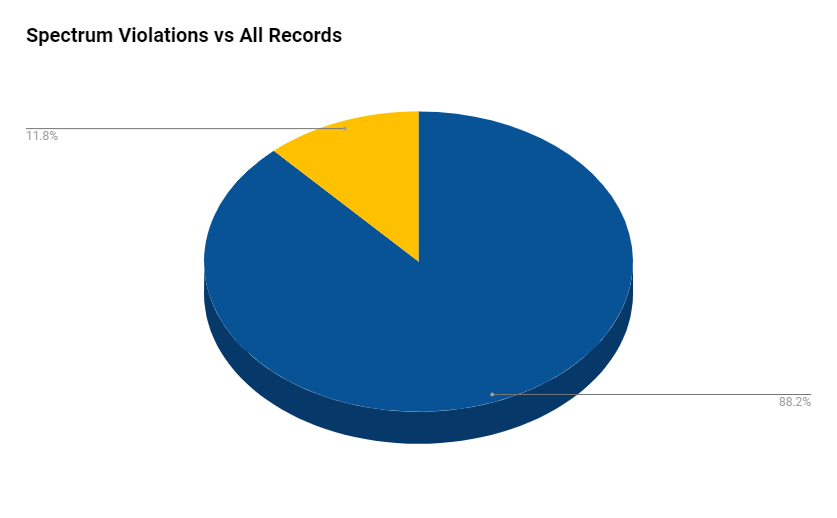


Figure New Data, New Perspective: Spectrum Violation as of Fall 2018

In regards to my research questions, spectrum is still a small and seemingly numerically insignificant violation when compared to the other violations (such as not registering antenna structures, marketing/importing unauthorized devices, and/or defrauding the E-Rate program). Even more so, spectrum explicit violations only make up 4.5% of the overall dataset which causes me to infer that spectrum interference is a low hanging fruit in the grand scheme of violations where the FCC needs to take enforcement action.

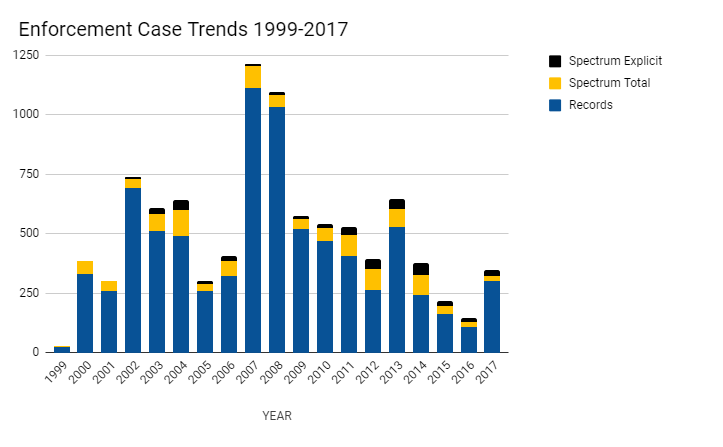


Figure Enforcement Case Trends 1999-2017

Regarding my second research question about who are the main offenders, when considering all spectrum violations there appears to be a shift where individual actors are no longer the primary offenders (34.7%), but business entities appear to be the main culprits (62.4%). This is out of an entity category that included duos, medical entities, municipalities, airports, religious institutions, and educational facilities. This shift can be seen in the figure and chart below.

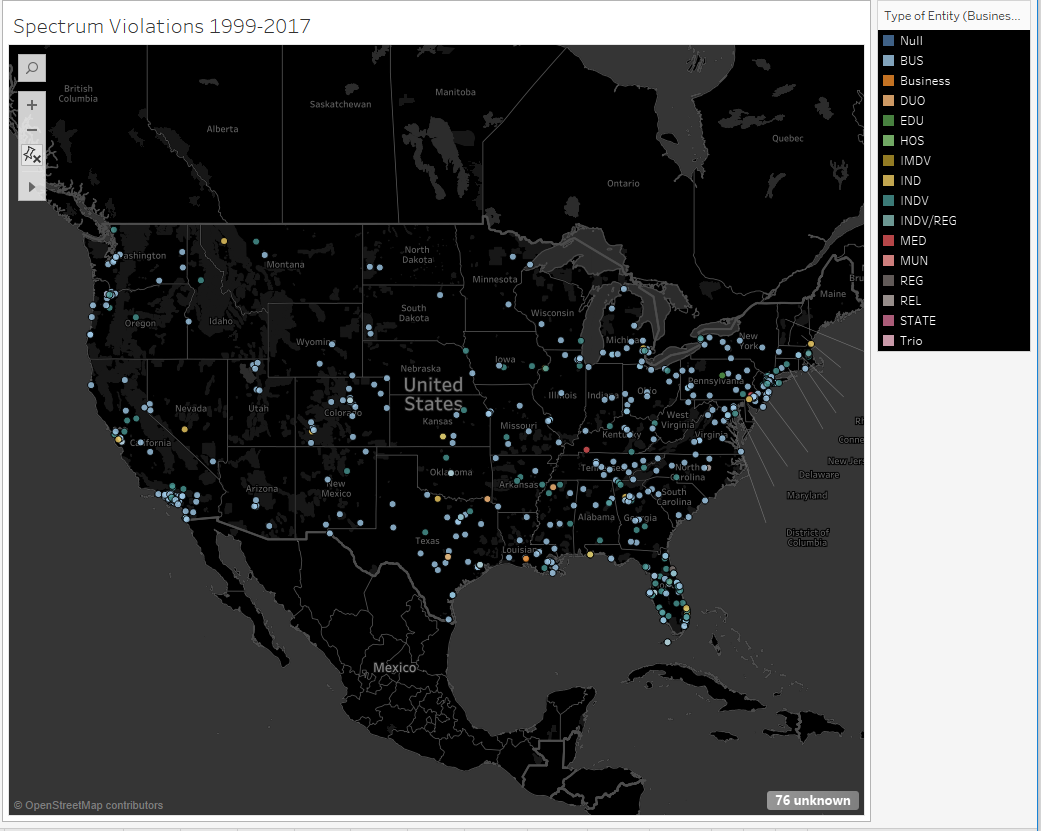


Figure Rise of the Business Violators

Despite the errors – that I am still working on fixing within my data, an increase of BUS points has increased from the first and second milestones of this project. My working hypothesis on this phenomenon is that this increase in business entities as violators is due to their infractions being more nuanced and not necessarily spectrum explicit violations. Through the data, I found that some of the businesses (e.g. hotels) jammed/blocked service in order to promote the use of their own Wi-Fi. Additionally, there are circumstances where businesses may be operating with an expired license, or they are not abiding by their FCC license.

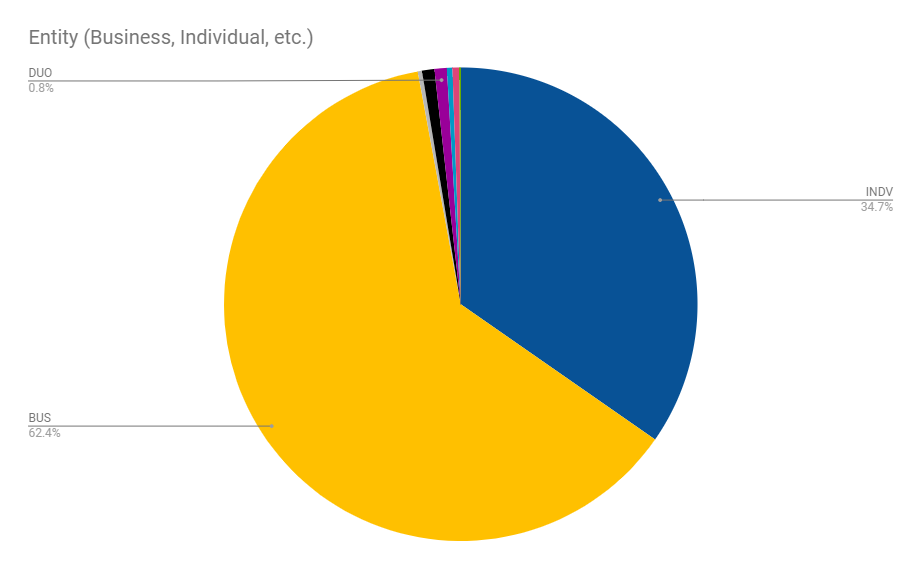


Figure The Data From an Entity Perspective

For my final question – what types of offenses/interference is occurring? – I found that this changed quite drastically from the first and second milestones of this project. Offenses now included violation of operating on a frequency band other than stated within the offender’s license agreement and operation of stations with an expired license (some of these cases were for decades). Pirate radio was still quire pervasive within these types, however, it was interesting to these other violation types arise.

When reassessing the spatial distribution of spectrum violations, it can be observed that these violations are no longer solely clustered along the coastal lines and that the violations have shifted more inland.

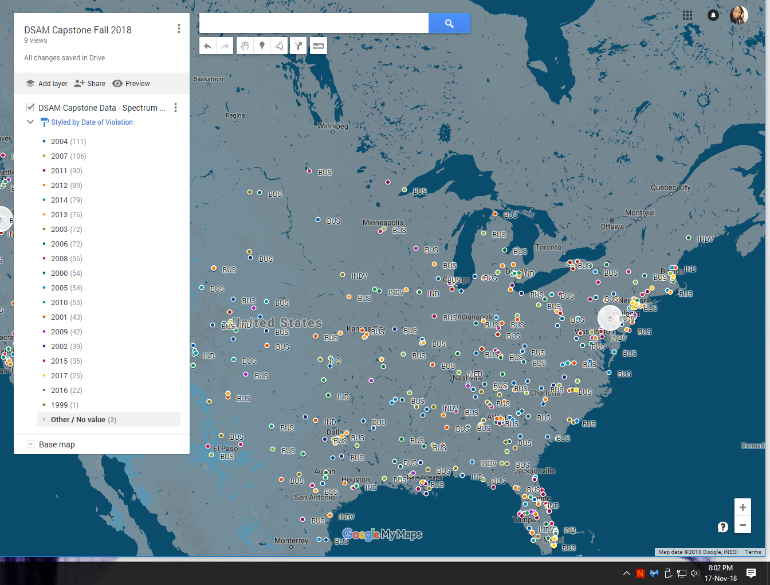


Figure The Data Over the Years - Violations Everywhere

Other observations regarding the spatial analysis include more international actors – which was not observed in milestones one and two. Typically, international violators fall under the spectrum implicit category as they are actively marketing, selling, exporting, and/or providing equipment that has not been authorized by the FCC or intentionally causes spectrum interference such as signal jammers.

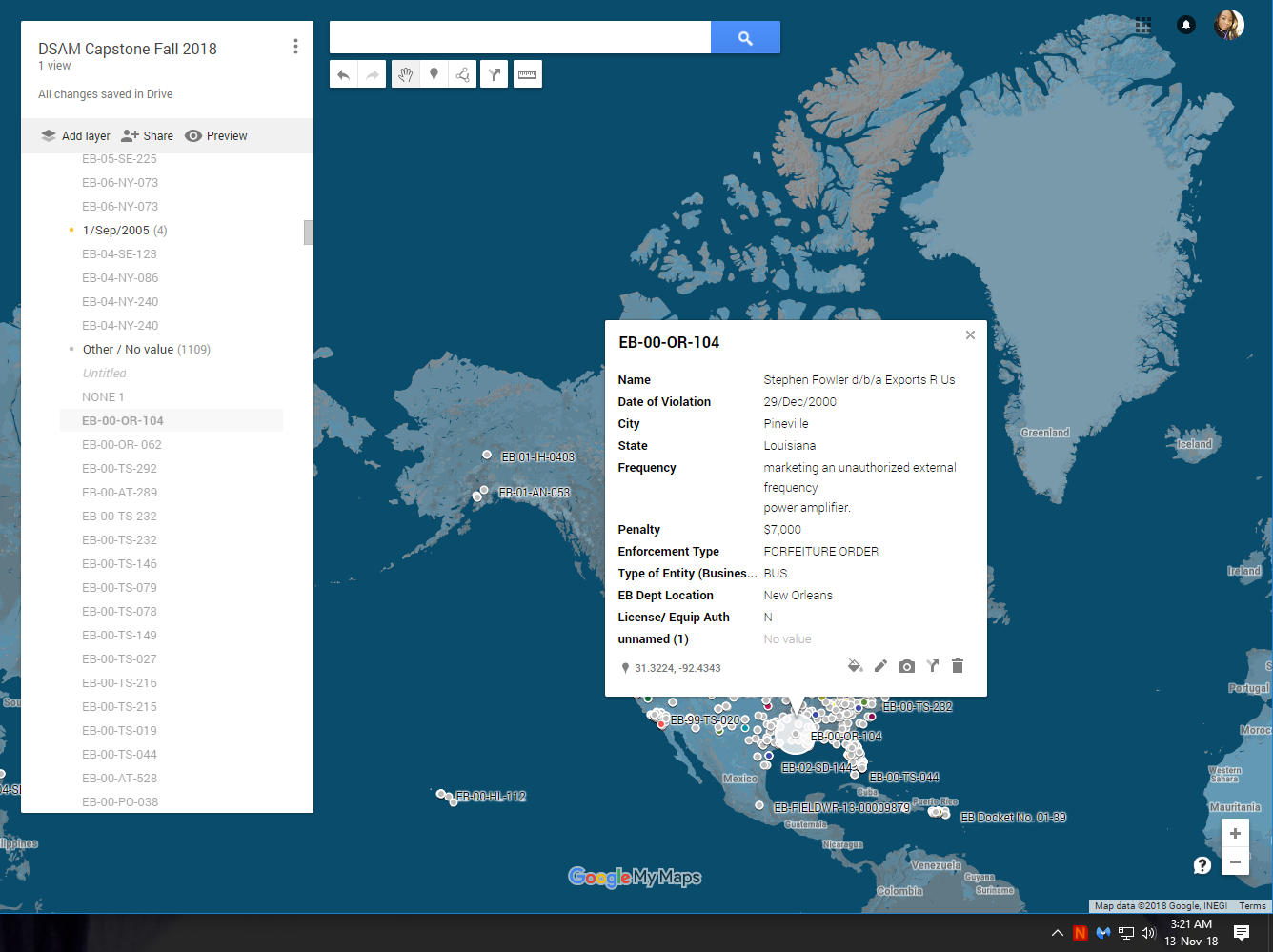


Figure Data Over the Years - Violations Everywhere | A Closer Look

Additional inferences that I drew from the results, include questioning whether spectrum violations are truly not prevalent, or is it possible it appears to not be a pervasive problem due to the lack of FCC regional offices within certain areas. This may provide insight as to why for the first two milestones spectrum violations were so clustered along the coastal areas. There are multiple regional offices located with California and Florida. Additionally, there are few states situated along the North East that are not within proximity to a regional office (e.g. there are offices in the North East are located in New York, Pennsylvania, and Boston).

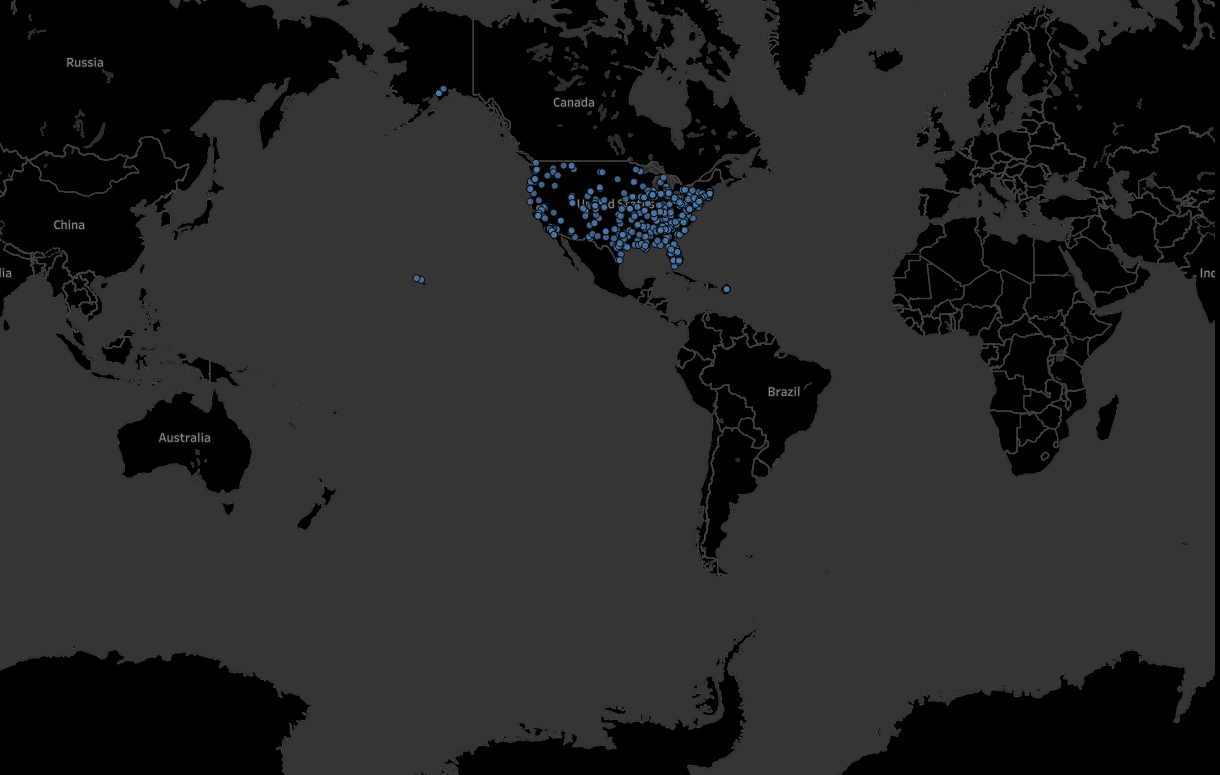


Figure Data Over the Years - Violations Everywhere - A Step Back

Key insights that I had include that spectrum implicit violations are similar to risky driving behaviors (such as drunk or reckless driving). There are quite a few cases where offenders caused interference to public safety frequency bands such as the police department. Furthermore, as the year’s progress (from 1999 to 2017), less financial penalties are imposed on violators and more warnings are issued despite the harmfulness of the spectrum violation. The decrease in financial penalties appear to occur around 2014/2015. As a part of my future work, I plan to map the timeframe of specific policy initiatives such as the Wireless Broadband Plan that occurred under President Obama and other regulatory implementations to see if this change in penalization is occurring due to the policy initiatives being driven by emerging technologies, or if they are occurring due to some other factor.

**LIMITATIONS**

Limiting factors concerning this research project include data disparities, a complete and successful import of the data into a digital tool, user skill level, and informational blackouts. The data disparities observed within this research stem from a lack of consistency within the administrative records extracted from the FCC website. This issue raised several concerns for me, as I wanted to do my due diligence and accomplish a valid, complete, and consistent dataset.

During the collection of data, I often found it hard to pinpoint certain case attributes such as date/time of the actual violation, clarity regarding whether the offenders license status (e.g. licensed, unlicensed, revoked license, or expired license), whether the offenders usage on an unauthorized frequency band caused interference to another entity, the actual location of the violation, and often times there would be a violation record that would include violations spanning decades and frequency bands – however, quite a few of those circumstances were vague or broad generalizations of the offenders ongoing delinquent behavior.

Eventually, I had to make the decision to drop specific cases both from the dataset and also those that threw errors in the GIS tools (similar to the image below). By taking this course of action, I felt a certain uneasiness about the validity of my data.

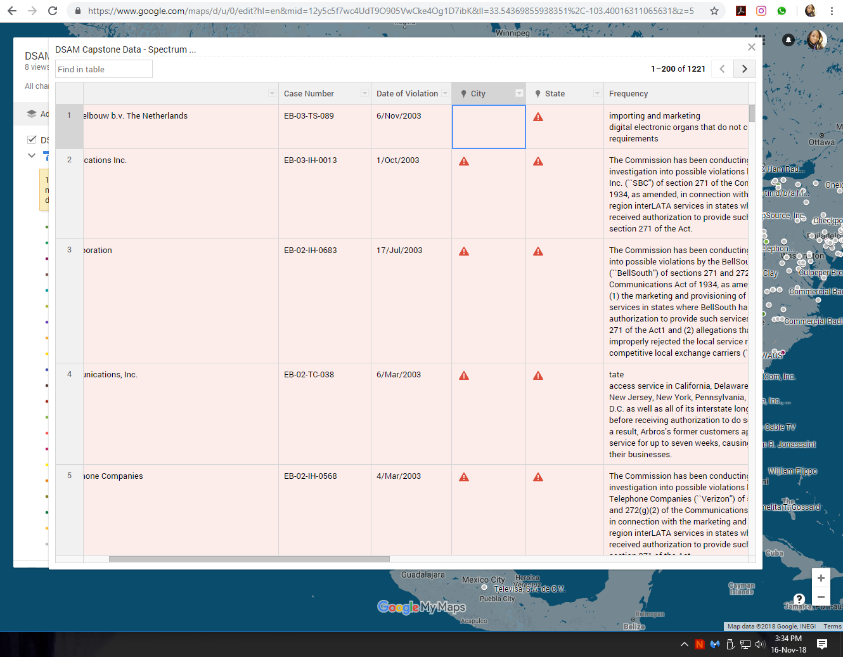


Figure Google Maps Error Fall 2018

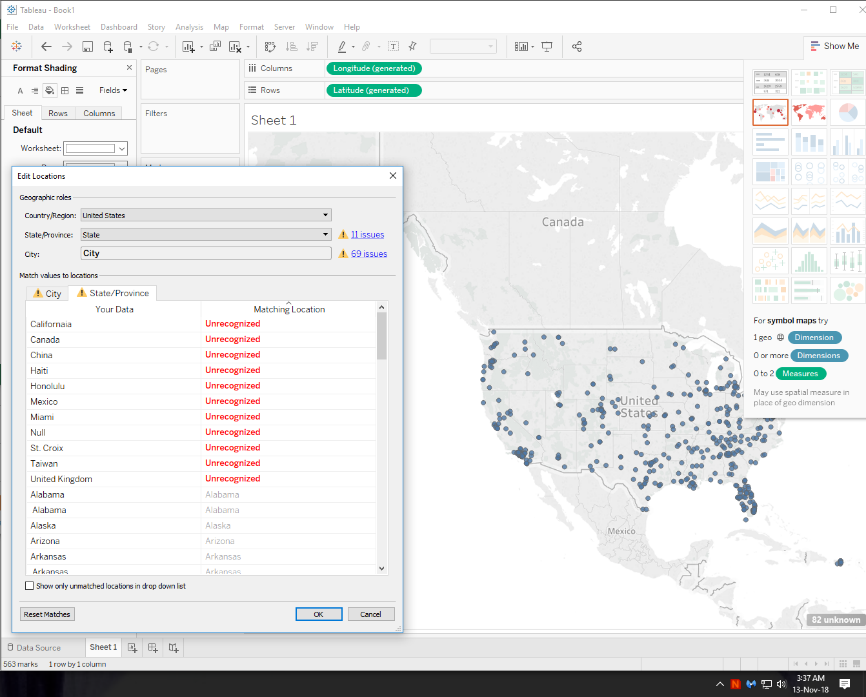


Figure Tableau Error Fall 2018

This uneasiness did not stem from my lack of confidence in my attempt to use my best judgement in order to accomplish an as-complete-as-possible dataset. My issue was in terms of my level of expertise pertaining to the inner workings of the FCC and my lack of expertise regarding their business practices and how they discern types of violations, etc. – I just felt unqualified in making certain determinations regarding governmental records. Moreover, being unable to completely and successfully upload and use my entire dataset added to my unrest. In both digital tools, Google Maps and Tableau, I received errors about certain rows of my data. Attempting to fix and subsequently having to delete certain records made me feel as though I was creating an invalid/ insecure dataset. Furthermore, if my dataset is invalid and/or untrustworthy, how certain should I be of the inferences I am drawing from the results?

I partially feel as though if I had chosen only 2-3 GIS tools from the beginning of my research project in 2017, I would have been able to dedicate more time to getting to know these tools better to where I would fully understand why certain rows of my dataset were being rejected. In the future, I plan on continuing to use Google Maps and Tableau, but also investigating QGIS more. As I keep refining the information in my dataset, I plan to make more time to get more familiar with these tools. Although I have not had as much trouble with Google Maps as I have with Tableau, I realize that with Tableau you get what you pay for. Thus far, I have been familiarizing myself with the free/student trial of this tool, however, when watching the Tableau tutorials and researching this tool online, there are certain features that I am unable to access due to my tier – therefore most of my Tableau visualizations appear to be “dots on a map”, even though I know this tool is capable of much more.

Lastly, the informational black holes were the bane of my existence. At times, it was hard to remain on task with my data collection and focusing on spectrum violations because all of the violations were so enticing. These black hole violations included individuals who had their licenses revoked due to their convictions of child molestation or recipients of the E-rate programming colluding to defraud the federal government and receive additional funds illegally. Because I had to closely read each record in order to assess which cases were spectrum violation, it would sometimes take me an exorbitant amount of time to get through a specific year’s data because all of the violations were important, outrageously idiotic, or severely demented.

**FUTURE WORK**

Despite the information black hole being something I consider a limitation, the amount of close reading I accomplished white collecting my data will be beneficial as I take the next steps of my research within this area. My overall goal is to eventually analyze all of the violations that fall under the FCC’s purview and develop an automated enforcement system to better handle the disparities and inconsistencies between violations and adjudication.

Currently, if an offender has violated the rules/policy of the FCC, the FCC EB either sends agents to the individual and/or the offender receives a letter in the mail regarding their violation or offense. After reviewing roughly 8000 records from the FCC, I am confident that they do not have an automated structure in place as there is no mention of online accessibility in any of the documents.

This approach would be advantageous to the FCC and commercial users as it would provide ease, accessibility, and decrease adjudication response time. I understand that this would be a huge undertaking, however, with the spectrum subset already reviewed, I am confident that I can improve my data cleaning process and increase my GIS skills in time to make this research the crux of my dissertation.

**FINAL REFLECTION & DISCUSSION**

Upon investigating the journey of using my specific area of study within the context of the Digital Studies & Methods curriculum, I have come to the realization that within the field of Computing and Information, we do not typically take time to mindfully reflect and/or assess why we decide to use a specific methodological approach. More often than not, we implement methodologies based on popularity and what others within the field are applying – especially if it is within a similar research area.

Although I feel that investigating a smorgasbord of GIS tools for my capstone ultimately took time from me fully mastering a couple of the tools, I am glad that I had the opportunity to investigate GIS as my digital method. As I stated earlier in the paper, to my – and to my advisor’s knowledge – no one in our field has specifically taken time to research the FCC EB database. Therefore, it is least likely that the application of GIS has truly been used to analyze the pervasiveness of spectrum interference – extending past the assumption that majority of most of all the spectrum interference is caused by individuals and their pirate radio.

By taking a more concerted effort to become more intimate with digital methods and taking the time to see what works best for you and your data, one may discover a new application for their query of research. I feel as though this holds true not just for those within the Computing and Information discipline, but that this is something that may resonate on a trans-, inter-, cross – disciplinary landscape.

Hopefully, through the use of what may appear to be an unconventional/uncommon method within one’s field, we can conduct research on a deeper level. For me specifically, I think it is critical to consider the technologic implications of my work, as well as user perspective and societal impact of regulator decisions. Through the use of GIS, I hope to identify better practices and services on how the FCC can improve their contact and response time with commercial spectrum users (the regular average joes in our society), in addition to efficiently and consistently provide regulation and policy updates to licensees.

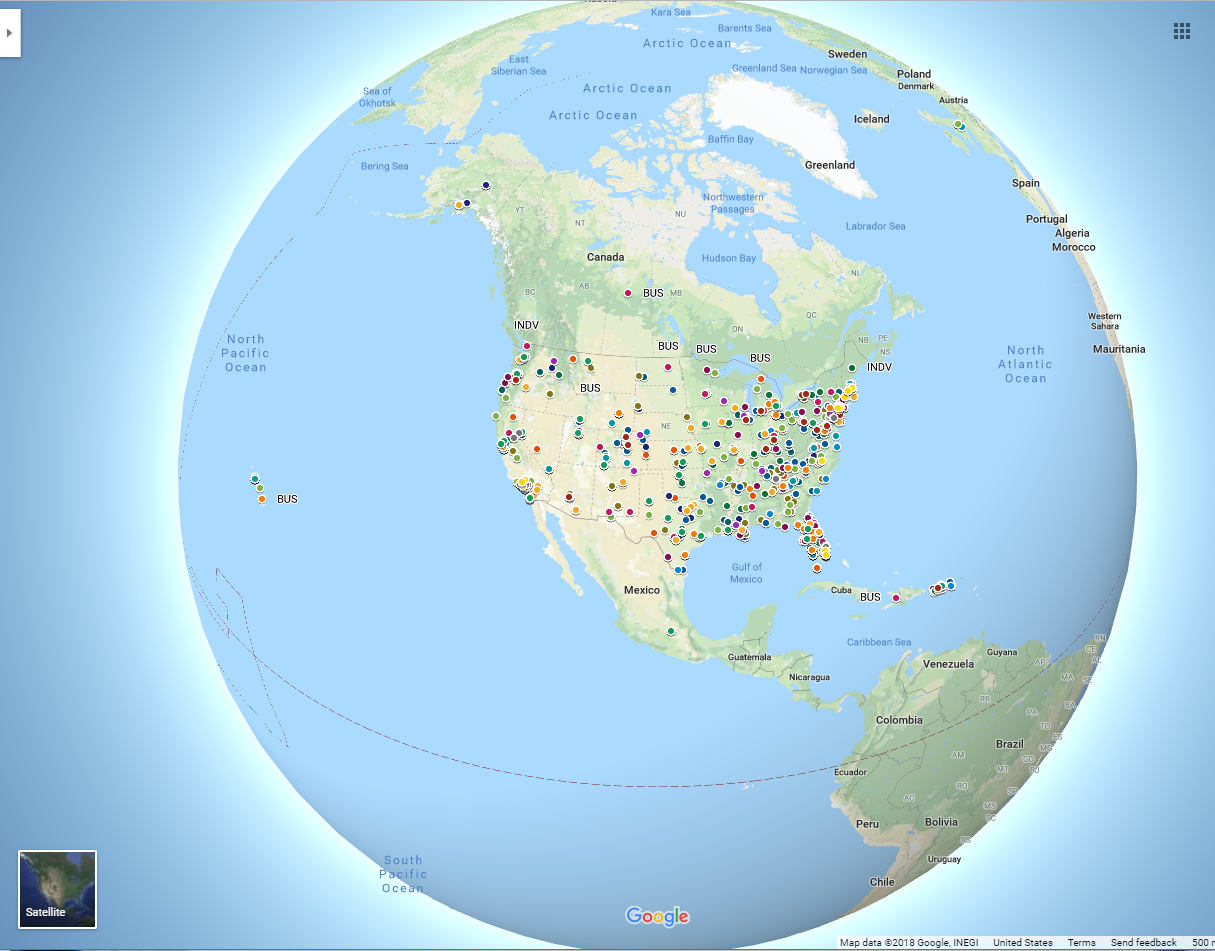
****

Figure Global View 2018...TO BE CONTINUED

**REFERENCES**

Video Recording No. 417.6094432; “N/A” March 26, 1978; Records of National Telecommunications and Information Administration, Record Group 417; National Archives at College Park, College Park, MD.

Video Recording No. 417.6094432; “Computing Machinery (ACM),” March 26, 1978; Records of National Telecommunications and Information Administration, Record Group 417; National Archives at College Park, College Park, MD.

1. Figure 1: Additional information regarding this figure can be found on the online portfolio. When entering the DSAM Capstone online portfolio, click on the project data icon. Once there, you will be able to interface directly to the FCC website. However, although this feature is embedded in the portfolio, clicking on specific links may open up and/or render a new webpage. [↑](#footnote-ref-2)