# SIGID Wiki Archival and Knowledge Practices

*Interview with Carl Colena*

*Carl Colena is a software engineer, digital signal processing hobbyist and the current administrator of the SIGID wiki website and radio signal digital archive. The interview was conducted by Selena Savić and Yann Patrick Martins in May 2020.*

**Selena Savić and Yann Patrick Martins:** The Signal Identification Guide (SIDIG) wiki is an organized database of information on radio signals. It contains data on signals' characteristics such as frequency and bandwidth, modulation type, as well as short descriptions, audio samples and waterfall plots. How was this project started and who supported it?

**Carl Colena:** A community of hardware hackers were able to turn the inexpensive digital television dongle (DVB-T) into a software defined radio (SDR) hardware. The discovery of these affordable TV dongles that can be used like a spectrum analyser, has significantly dropped the barrier of entry for amateur radio enthusiasts and other technically literate individuals, to explore radio signal reception and hacking.

Carl Laufer, the owner of the RTL-SDR blog,[1] started the Signal Identification Guide (SIGID) wiki[2] project in 2014. The goal was to create a blog for all interested people who were using RTL-SDR to look around the radio signal spectrum and understand what they are receiving.

1 RTL-SDR (RTL2832U) and software defined radio news and projects https://www.rtl-sdr.com/ (accessed 22.06.2022).

2 Signal Identification Guide wiki https://www.sigidwiki.com (accessed 22.06.2022).

There was previously no centralized database describing these signals. The SIGID wiki website started as a collection of the information about radio signals that was held among a community of radio enthusiasts who wanted to explore radio space and understand what they see and hear.

I was among the people who purchased an RTL-SDR enabled dongle around that time. I was interested in analysing Wi-Fi signals because I was having issues with network reception on campus where I was studying. I was planning to use the TV-dongle to search for a place on the campus to work from.

With my equipment, I could actually 'see' the radio spectrum, and I wanted to know more about all these different signals. I soon became fascinated with all these radio signals that exist around us, which we are not aware of. I had similar questions to many users of the RTL-SDR blog when I first came across SIGID wiki. I saw this as an opportunity to both find and contribute information from other sources. I added signals from my own research; I added signals which someone else would identify on another website, I added additional signal samples. Most of the signals that you see on the website today are pages that I have written. Eventually, around a year later, I became the website administrator.

I then started to work on reforming the website structure. I split the long list of all signals, which contained some 200 entries, and introduced categories.

## Known and unknown signals

**SS:** There are two major categories of radio signals on SIGID wiki: known and unknown (Figure 1). The first ones are those that serve identification. I could record a signal and turn to the wiki to compare it with those in the database to find one that matches. Or, if I do not find a match, I could upload my signal's data to the wiki as "unknown" and have it hopefully identified by someone else. How does this work in practice, how do you narrow down the search and how can you be sure you found a match?

**CC:** If someone recorded a signal and cannot find a match on the SIGID wiki, there is a number of different routes to take. Most people would create a page in the unidentified section, and upload whatever traits that they have recorded to help identify it. We provide a form to do this (Figure 2), and the more information you provide the easier it will be for other users to identify a new signal.

**SS:** This suggests that signals in the database are most probably "identifiable", just not yet identified. How do those in the unknown category pass to the known one? Conversely, how can one challenge a signal as incorrectly identified?

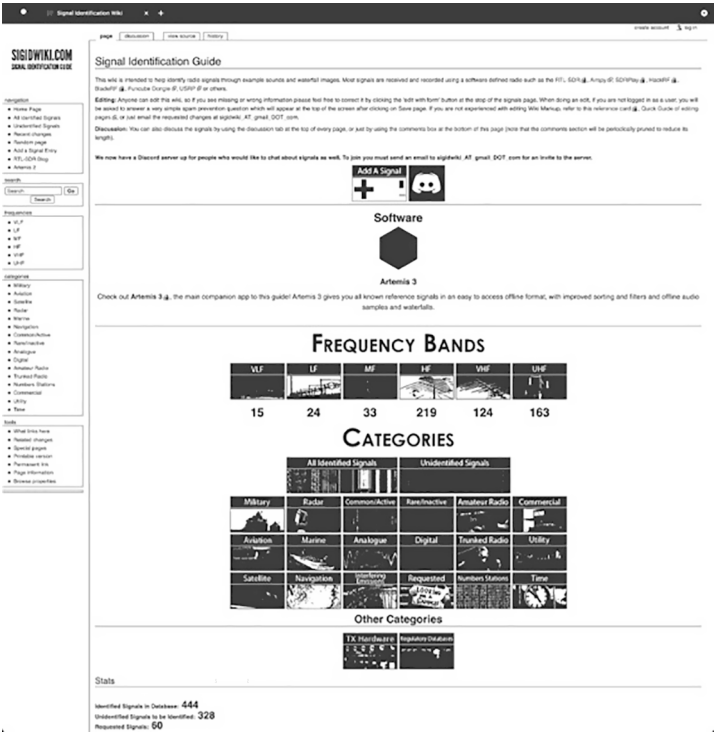
**CC:** Identification is a community effort. Someone from the user community, including myself, could look through unknown signals and recognize a signal they know. We often find out that a signal in the unidentified section corresponds to a newly identified signal on another website – it simply had not been identified at the time when it was added to SIGID wiki. We then add it to the collection of samples that showcase what this signal looks like in the environment. Most signals do not have a static representation, especially if they are transmitting data: they may have different modes and phases. Oftentimes audio samples on the SIDIG wiki are snippets of these transmissions, so there may be a mode or behaviour of the signal that is not captured within this short timeframe of the sample. Unidentified signal samples might therefore be a snippet of a known signals that was undocumented.

Contesting already identified signals happens as well in the user community. A user recently flagged a signal, stating which type of analysis they used and showing how these do not match up. This works on an evidence-based approach. If somebody would say that they have stronger evidence, demonstrating that a signal is not what it is identified as, that should be taken seriously to correct its identification.

**SS:** Is the aim, or one of the aims, of the SIGID wiki to resolve unknown signals' identity?

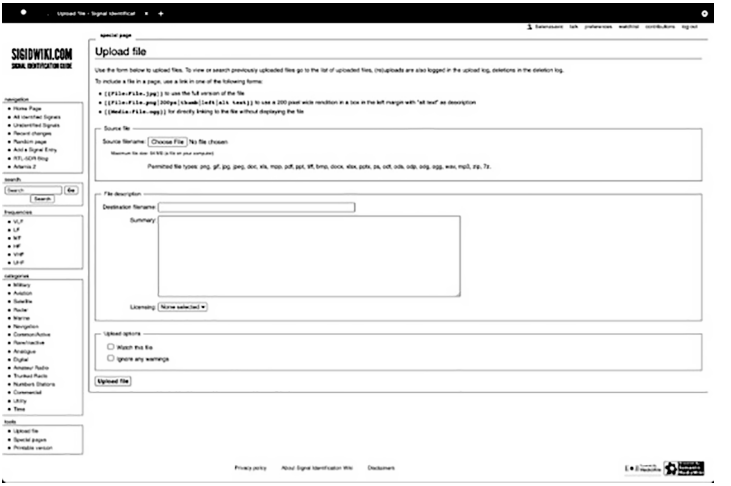
**CC:** Yes, that is certainly one of the aims. For the most part of the work that I did, the aim was also to serve as an archive for all of the signals that have been used for transmission, but also what could be natural emissions, like lightning spherics, for example. The SIGID wiki acts both as a holistic guide, as well as a kind of a 'museum' of different radio signals that have existed. We can observe in the archive how we have transitioned from lower band signals that are not able to carry as much information, to complex signals that can broadcast television.

*Figure 1: SIGID wiki website.*




Screenshot from 01.04.2022, made by Selena Savić

*Figure 2: SIGID wiki signal upload form.*




Screenshot from 01.04.2022, made by Selena Savić

## Signal knowledge

**SS:** How do conventions on archival techniques of tagging, categorization and metadata emerge in the SIGID wiki?

**CC:** The categories emerged out of a need for organizing signals we collected, trying to group them into meaningful sections. I first categorized them by application. This is reflected in the broader community of listeners, with groups that are specifically dedicated to, for example, satellite reception, or to public safety trunked radio channels. Categories are thus based on listener communities that emerged around certain types of transmissions. The question remains whether the users of a signal should define that signal. Categories are a conversation point within the community of users of the SIGID wiki. There are different proposals to update the categories to better align with objective metrics.

In terms of consistency, some of the categories are hard to disjoin. One such category is the digital. Radio signals are inherently not digital: they are continuous waves propagating in space. However, a large group of signals are digital in terms of the information that they transfer, such as cellular communication. Conversely, certain signals are used for digital purposes, but are analogue in nature. Telemetry devices can broadcast signals that are basically a tone, a sine wave whose frequency shifts based on changing pressure in a pipe. These are typically used in industrial systems, where the tone transmits the status of equipment: discreet analogue tones that could be purely analogue. Another example of an analogue signal is radar. Observations such as distance to buildings, or airplanes in the sky are not digital by any means. Signals are inherently not digital physically (in their transmission), but they can be used to communicate digital information.

**SS:** There is a lot of information on signals that is not quantitative, such as descriptions, application, historical use. Who writes these descriptions and other information, who checks if it is correct? Which resources are commonly used?

**CC:** I wrote the vast majority of signal pages on the SIGID wiki website. The descriptions are typically about answering high-level questions such as 'what is the signal used for?', 'who uses it?' and so on. It is similar to a Wikipedia article about fish, for example. I typically give an account of a signal's history, its' characteristics, any unique identifying trait. This knowledge is based on years of writing about signals. I include all additional information I might have, such as links to video samples on YouTube or the sources where I got information from, additional images or samples. I try to do as much fact checking as I can on my own.

The information comes from different places, typically second-hand sources like historical accounts. Some information, especially about commercial and military signals, tends to be quite opaque. It is hard to verify. Some information can be validated with first-hand sources, from a government or a technical white paper, but they do not always exist. There are certain companies that have heaps of information on radio

signals and publish their databases of signals. I often used Wavecom[3] as a source of relatively reliable information. There are companies that specialize in signal intelligence (SIGINT), a field commonly used by governments and military to gather actionable data from emitted radio signals, for surveillance or intelligence purposes. Another source that I go to, at least in the United States, is the Federal Communications Commission (FCC). All telecommunication devices that are sold in the United States have to have an FCC ID which corresponds to a page with information about the manufacturer, the use of the equipment, frequencies and bands it is allowed to transmit on. Each government has their own regulatory agency, and you can use it to get information about anything that is wireless.

**SS:** Which properties do you look at in a radio signal?

**CC:** For each kind of signal there are different ways of finding out features. Sometimes I dive into the actual signal features, the structures. One of the features of communication-based signals is symbol rate. For example, Morse code is made of pauses and dots. With digital data, it is the timing between symbols that makes sense between two frequencies.

When analysing a signal, I look for its general characteristics, where it may be recorded, the type of modulation that it uses. I narrow it down further into specific features in terms of the spacing of tones, the speed at which symbols are being sent. Any unique characteristic that stands out is useful to both characterize a signal and potentially identify it.

**SS:** Which methods and tools do you use in signal analysis?

**CC:** Radio waves, an electromagnetic phenomenon, cannot be directly experienced by humans. We have to use extended methods of sensing to observe this domain. The first thing I always do is a visual and acoustic

3 See Wavecom online decoder website: http://www.wavecom.ch/content/ext/DecoderOnlineHelp/default.htm#!worddocuments/acars.htm (accessed 21.06.2022).

inspection. I can look at the waterfall plot live, as I am listening to the signal, establishing a connection between what I hear and what I see. Even though you are experiencing the signal through these proxy forms, you understand something about radio waves. I do not know if the other senses would make sense for that.

I mostly use free software tools for visual inspection, like SDR#,[4] or HDSDR[5] to replay the signal and look at its visual characteristics. I play with the Fast Fourier Transform (FFT) bin size to take a look at how this signal looks with inversed temporal resolution. The spectrogram, a short-time Fourier transform (STFT), can be a tool to derive different 'perspectives', on signals. Depending on the perspective you use, a signal will appear to have something distinguishing, or not. It is completely based on what you use to look at it. If you were to look at the signal itself, in raw I/Q recording of energy, the question would be what you could really understand. That is the reason we have tools, like the Fourier transform, to translate and understand the frequency components.

I could explain this through two primary forms of analogue radio communication: amplitude modulation (AM) and frequency modulation (FM). Depending on whether you use amplitude or frequency as a method for information transfer, you will look at the changes in the carrier wave differently. If we were to take a frequency modulated signal, and amplitude demodulate it, we will see in the spectrogram a simple straight line, no information at all. Thus, the reception mode, the perspective for which you receive the signal, enables you to observe whether and how it encodes and transmits information.

**SS:** On the Wiki, within first level categorization by application (e.g. military or aviation), there is second-level distinction between active and inactive signals. Are those still broadcast by some infrastructure that remains active but no longer serves any purpose? Could we still hear the inactive signals or is this archive material?

---

4 See AIRSPY website, https://airspy.com/ (accessed 21.06.2022).

5 See High Definition Software Defined Radio (HDSDR) website, http://www.hdsdr.de/ (accessed 21.06.2022).

**CC:** Inactive signals are archival material. The recordings in the SIGID database were made years ago, in the 1990s or earlier. It is hard to say whether inactive signals will ever be transmitted again, given that some signals require specialized equipment that may not be around anymore. Whether a signal is considered active or not will depend on the community's input about the signal use by its designated user; and whether it can be received in the wild. There is an established community of listeners, both shortwave and more general amateur radio listeners, who monitor the amateur radio bands. Part of their interest is to keep unauthorized or problematic people off the amateur radio bands, since this space is kind of self-policed. They also listen to the whole band, and they note observations on transmissions they receive.

With regards to inactive signals, there is no hard threshold. A lot of the early signals on this page were used by diplomatic services in the '80s and '90s, before the internet became the principal mode of telecommunication. Diplomatic services used these dedicated radio channels because they needed a resilient and secure way of communicating back to the home country. I would not expect these to be used nowadays, given that you can send an encrypted email instead.

**SS:** Can you tell us more on the use of radio signals outside of transmitting information, such as for example for motion detection?

**CC:** A good example of this are ionospheric sounders. Similar to the way bats use the echo in a cave to orient themselves, radio waves can transmit signals to get the sounding of the ionosphere that surrounds the earth. This is typically done to understand the magnetic field of the earth at different points and time of the day, and in response to different stimuli. These systems use a big transmitter, broadcasting up into the sky. Receivers located at different locations pick up the reflections from these signals. They are not listening for the signal itself, but to see what the reflection from the signal looks like compared to the signal that was sent. This difference is then used to measure and characterize the atmosphere. Radio waves are used here purely as energy.

Amateur radio operators do something similar, which is called QSO.[6] They use low energy transmissions like a sport, trying to send a signal at the lowest energy possible to the furthest possible distance. They send out a short message identifying the operator, over and over again. It is not really useful in terms of information, but when someone receives it, they will note the distance and time period of the broadcast, giving a sense of local radio environment.

Research in the context of motion detection is typically interested in non-intentional emissions. These are emissions from things like a computer monitor, or the electrical wiring in a house, emissions that exist as a consequence of other processes (e.g. supplying energy to devices). This is typically used in the security field, both for information and physical security. Each key on a keyboard produces a distinct system interrupt in the computer circuitry, and it can be used to reconstruct the typing. This has been experimentally proven possible.

## Natural radio

**SS:** Could it be the case that some of the signals in the 'unknown' SIGID wiki category come from natural sources such as ionospheric emissions?

**CC:** Certain traits are unmistakeable signs that a signal is not a natural transmission. A signal with a strong wave patterns or complex structure is probably not natural, but it could be. From what I have seen, naturally occurring emissions are typically short. This depends on what one is listening to. For example, a meteor going through the Earth's magnetic field creates an electromagnetic wave at a given frequency with a drift in the frequency component.[7]

6 Contact (amateur radio) Wikipedia page: https://en.wikipedia.org/wiki/Contact_(amateur_radio) (accessed 21.06.2022).

7 See entry on Ionized Meteor Trails on SIGID wiki: https://www.sigidwiki.com/wiki/Ionized_Meteor_Trails (accessed 21.06.2022).

There might be some unidentified signals that are of natural origin. A ubiquitous radio signal that appears in high frequency range is called the 'whistler'. Similar to the meteor trails, electromagnetic discharges in the ionosphere generate whistling sounds that look like drifts of energy in terms of their frequency composition, sliding in and then disappearing. Those would certainly be natural phenomena.

There was another source that I found pretty interesting while looking at very low frequency emissions, singles of Hertz to sub-Hertz. A colleague of mine who worked at the South Pole on installing an extremely low frequency antenna, the purpose of which is to continuously record this frequency band. These antennas pick up a lot of natural phenomena like lightning, or earthquakes. Those are electromagnetic emissions coming from the Earth. Such natural emissions are typically longer than the ones at the higher frequencies, because they are actually slower. Higher frequencies tend to come from the outside of the Earth: pulsars, meteors, outer space phenomena, while very low frequency natural transmissions come from the Earth itself.