Data Observatory: Descriptions

Selena Savić and Yann Patrick Martins

All signals in the database have a textual description. These are short pieces of text describing basic properties, protocol-specific details, replete with technical jargon. Nevertheless, text is a domain that can give us meaningful access to signal qualities through an architectonic relationship. In this first study, we related the self-organizing map of spectral entropy in radio signal audio samples to a statistical model of text using Latent Dirichlet Allocation (LDA) algorithm, popular for topic modelling tasks.[1] We identified nine topics, each taking us to a group of signals that are most relevant to it (Fig 1, left).

I (Selena) began the Study 01 by expressing an interest in the relationship of radio and military. We know many telecommunication technologies, such as the Internet, were initially developed by the military, but is there something new and specific we can learn from this setup? I selected the topic that deals with things military (Fig 2, green highlighted text on the left) and looked at signals that are most representative of it.

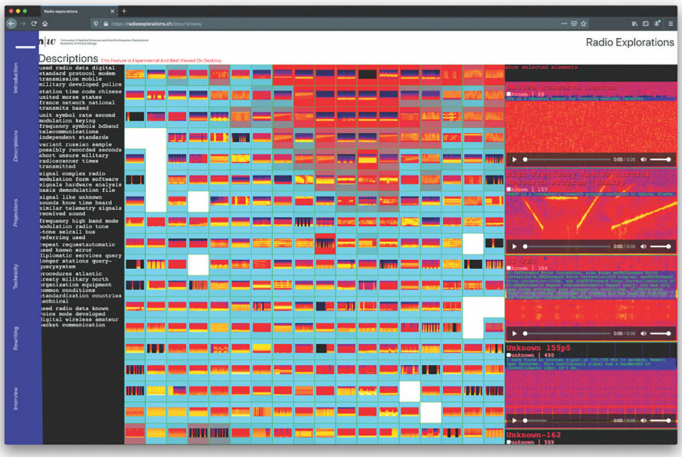
One signal with an interesting spectrogram, High Frequency Active Auroral Research Program (HAARP)[2] drew my attention (Fig 1, right side). It presents radially and diagonally propagating lines, like a drawing. According to the description, it belongs to a research programme

1 Latent Dirichlet Allocation (LDA) is a form of unsupervised learning that views documents as 'bags of words', in which order does not play a role. Topic modelling or topic detection is a machine learning method to discover human-readable topics in text.

2 More on this signal in the database: https://www.sigidwiki.com/wiki/High_Frequency_Active_Auroral_Research_Program_(HAARP), accessed 21.02.2022

studying the properties and behaviour of the Earth's ionosphere. Reading about ionospheric emissions, I learned that climate researchers used data on lightnings to measure the degree of climate change.[3] They found, back in the end of 1990s, a significant correlation between the increase in temperature and in lightning activity in the northern hemisphere of our planet. Within the fingerprint-based SOM, in a cell that is representative of the military topic, HAARP is similar to DUP-FEC-2 and RUM-FEC (Fig 2 and Fig 3, right).

Figure 1: Study 01: Descriptions. Data observatory web-based interface. Topics are listed on the left. Data organized according to fingerprints.



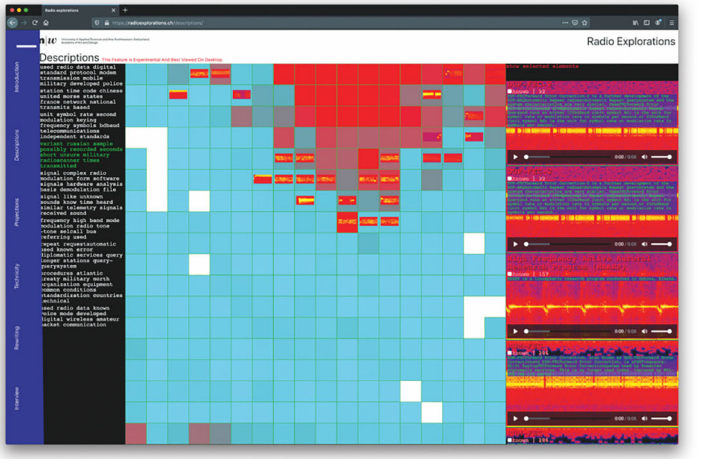
Courtesy of Selena Savić

The 'FEC' in the name of these signals stands for "forward error correction" – an error control method characteristically used in situ-

3 N. Reeve and R. Toumi, "Lightning Activity as an Indicator of Climate Change," *Quarterly Journal of the Royal Meteorological Society* 125, no. 555 (1999): 893–903, https://doi.org/10.1002/qj.49712555507.

ations where retransmissions are impossible. By now, what this cell tells us about military is: *Military communication is tightly connected with diplomacy and intelligence. The connection can be made through techniques of environmental listening (HAARP); the impossibility of retransmission (FEC) is characteristic of many military communication situations*.

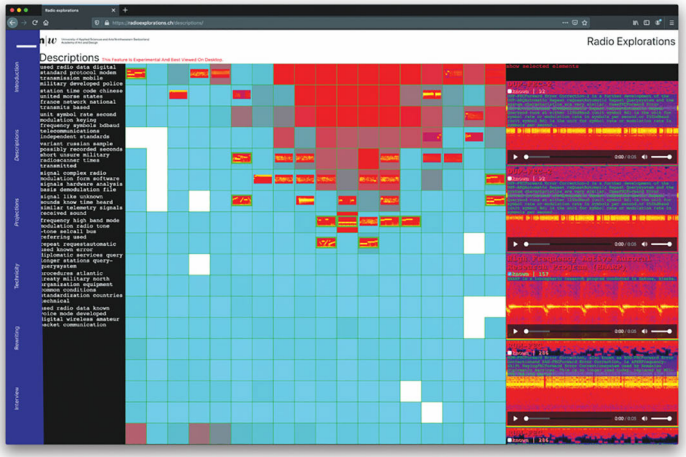
Figure 2: Study 01: Descriptions. Data observatory web-based interface. Selecting a topic (green highlighted text) exposes the most relevant signals.



Courtesy of Selena Savić

The setup of the *data observatory* renders these connections perceptible through an interest in the relation between sonic and textual representation of the signal. The use of a *data observatory* as an instrument to perform the comparison between the recording of a radio signal and its description, starts to unfold stories that would otherwise slip our attention.

Figure 3: Study 01: Descriptions. Data observatory web-based interface. Chunks of one signal, HAARP, are highlighted throughout the map.



Courtesy of Selena Savić