Cover Letter

Dear Torben Elgaard Jensen,

Our names are Alexander Luis Manuel Siegfried, Alfred Lund Felumb and Peter Zinck Munksgaard and we have taken the opportunity to submit our article: "Exploring Electronic Music: Developing a Data Practise *Thing"*, to you as our supervisor and in hopes of getting it published in the DASTS database of excellent student papers. We do so as the article is conceptualized as a modest countermove to issues regarding algorithmic decision making in *recommender systems* online. In practical terms it is an experiment of combining Digital Methods and Participatory Design to develop an exploration tool for electronic music. From a theoretical perspective it is a take on how to operationalize the descriptive language of Actor-Network Theory in a developmental process. In the sense that we have attempted to take on an active role in *doing* socio-technical analysis as a strategy of adaptation and reflection within an entrepreneurial practice. This is relevant for our fellow students, as it offers an alternative to describing the work of others, in which students often find themselves in the precarious situation of being dependent on the willingness of external collaborators. Apart from this, it might also be beneficial for other Techno-Anthropology students to be met with an honest account of the challenges involved with doing innovation in the context of a semester project. Especially for those interested in doing Participatory Design or Digital Methods themselves.

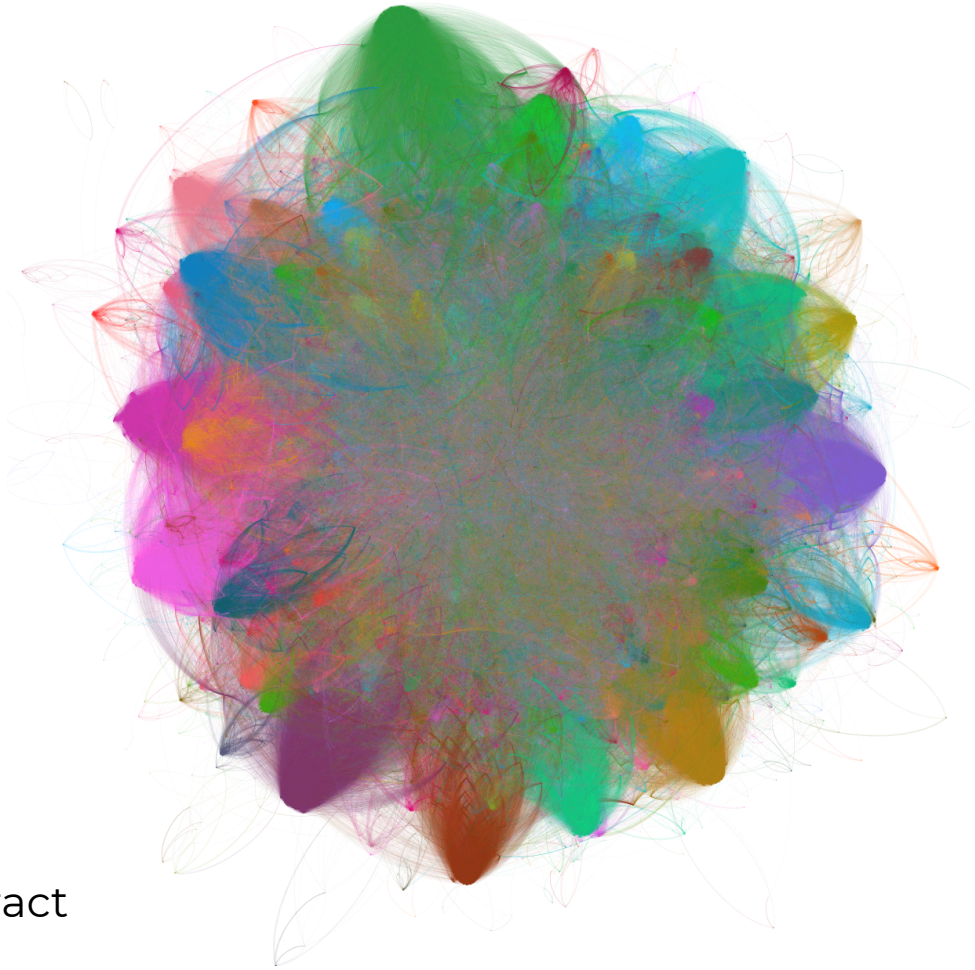
The article is original and has not been published before, nor is it under consideration for publication at another journal. Finally, we have no conflicts of interest to declare.

Contact: asiegf17@student.aau.dk, afelum17@student.aau.dk, and pmunks17@student.aau.dk

Exploring Electronic Music: Developing a Data Practice *Thing*

by Alexander Luis Manuel Siegfried, Alfred Lund Felumb and Peter Zinck Munksgaard

Word count: 7072



Abstract

This article is a story of an innovation process wherein, we as designers, present an honest account of all the obstacles and sometimes failures that also goes into the design of new things. We present the resistances we have faced in infrastructuring an environment that accommodates a continuous design of an exploration tool for data practices concerning Electronic Music. As such, this article is both an auto-ethnographic account of our own design process - and an evaluation of our experiment with combining Participatory Design and Digital Methods with the sensibilities of Actor-Network Theory.

Special thanks to Torben Elgaard Jensen for supervising this project, and to Poppy Jones for providing technical assistance throughout the project.

Introduction

In this article we've set the somewhat lofty goal of initiating development of a novel way to explore the veritable sea that is electronic music. The basic idea is to use visual network analysis to map music releases and their associated artists and labels.

Using the metaphor of 'sea' for electronic music is a deliberate choice, simply because of the sheer amount that exists and is continually being produced. This vastness is the core conception of the field and we imagine that it brings about some of the issues that we seek to address. Namely that it would be reasonable to assume that browsing all of the music manually (not to mention, in a meaningful way) is impossible and that this in turn urges the necessity of algorithmic assistance. Currently, the main options for getting such 'algorithmic assistance' is by engaging with the software giants like Spotify and Google / Youtube. In such engagements, users can at best expect to face suggestions, and at worst be the unwitting target of an automated stream of algorithmically selected content. Such types of algorithmic intervention, brings into question themes about 'loss of agency and transparency' as a result of efforts to maximize user retention. A symptom of which is the glooming presence of *recommender systems* (Helberger et al., 2018). This is not to say that we are of the opinion that all users *have* to be enthusiastically engaged in content exploration, but that we would like to frame our work as a modest countermove to algorithmic decision making. This way, we seek to aid people who are already engaged in exploring the 'sea' of electronic music.

But why electronic music you might ask? Firstly, our basic inspiration for this project stems from a general interest in electronic music, as one of our group members is an avid enthusiast. Secondly, we find it to be an interesting case in its variety and abundance of music genres and styles (Heuguet, 2020). Hence, in the case of genres and styles, literary sources discuss what can be characterized as 'classification disputes' as seen in Wiltsher (2016): Is it by investigating common musical and thematic features? Is a given piece of music or artist defined by its affiliation to one (or more) genre(s) or is it the other way around? Or is it a matter of investigating the cultural and historical context surrounding the music at the point of release? While these questions are certainly interesting and worth investigating, it would be a disservice to our academic colleges in musicology for us to pursue. Rather, to accommodate our professional strengths, we seek to circumvent genres entirely - on a conceptual level at least. The idea is to focus on the relational structure of all

involved actors (music, artist and labels) without any presuppositions of how they *should* be arranged.

> After having conferred with a student in musicology at University of Copenhagen (KU), our critiques of genres were met with modest confirmation. Additionally, it was cemented further, as the student pointed out, that the music industry has many factors going into genre distinctions - e.g. whether or not music is on billboards in the United States has significant influence on if it is considered Pop or not (C.E. Landberg, personal communication, April 26, 2021).

The inspiration for having this focus is heavily based on an old sociologist's dream of the *Single Level Analysis*, which entails making no distinction between micro and macro layers of organization (Munk, 2019a; Latour et. al., 2012). Here we explicitly bet on the notion that it's possible to utilize digital means to do so - or at least approximate it. That being said, 'going digital' has caveats of its own. Namely, that any sort of work with digital tools or platforms implies the study of these in and of themselves as well (Munk, 2019b). In other words: describing how data is generated on the web, implies the description of how it is shaped by the platform it is created upon (Weltevrede & Borra, 2016; Birkbak & Munk, 2017). In doing so, we delve into the Digital Methods of our field, that are usually ascribed to Controversy Mapping (Venturini, 2010) or for showcasing Multiplicity (Jensen et al., Accepted/In press). From what we have described so far, the shorthand framing of doing so, is to translate our Actor-Network Theory (ANT) sensibilities into an object that is usable and tangible to people not familiar (Law & Lin, 2020). This sentiment is akin to the Socratic conversation Latour had with a doctoral student, in which Latour explains how focus has to be diverted from reproducing the clichés that explanatory models tend to produce (2004). Moreover, that the order of business in ANT, is to 'just describe' - a sentiment echoed in Controversy Mapping as well (Venturini, 2010). In that sense, we argue that genres and styles can be seen as the cliché in the case of electronic music. However, this does not mean that we posit that genres are insignificant, but that our experiment is to use a different descriptor. Furthermore, it is important to stress that our argument is not that investigating digital relations alone is better than or can replace the ethnographic fieldwork that ANT usually entails. Rather, we consider it part of a scientific process in which we are adding another *layer of mediation* for describing and exploring electronic music (Latour, 2014). With this, we try to stay in line with the idea that scientific development exists at the end of a very long series of transformations - between which no step

is an exact replica of the previous (Latour et al., 2012; Latour, 2014, p. 348; Law 2003). This puts further emphasis on the need to document the transformations in this project - especially considering the technical nature of Digital Methods.

An interesting implication of circumventing genres, is that it can be seen as the removal of said culturally and historically rich framing from the description of music. From a Visual Network Analysis perspective, what we would end up with are clouds of connected words which we have no basis for understanding (Jacomy, 2020). Under normal circumstances this can be said to be the aim of a Digital Methods research project - to trace and analyse the clustering of network data. Yet, in our case, we seek to make available said type of mediation for others. In a sense this means that the design of the thing we're making is dependent on outside expertise. Hence from a procedural perspective, the overarching inspiration for achieving this is Participatory Design[1] as described by Björgvinsson, Ehn and Hillgren (2012). We choose this as an ANT-inspired design approach that seeks to include users, and facilitate continuing design *after* design. In theoretical terms, this means thinking of the facilitation of a design process as *infrastructure* around the construction of *socio-material assemblies* - a *Thing* with a capital *T* (Björgvinsson et al., 2012). In a sense, a process of *heterogeneous engineering* (Law & Lin, 2020; Law, 1987).

From all these considerations our problem formulation is as follows:

How can we use Digital Methods and Participatory Design to develop an exploration tool for electronic music?

Heterogeneous Engineering

So what is it we are doing? Even mustering all our naïve enthusiasm about Participatory Design, it's neither *quite invention* nor *innovation* in the strictest sense. The process has quite simply been too materially heterogeneous, and at the same time, the idea of a happy customer holding an end product also seems a long way off. So what *can* we call it? As this in many ways is the first few steps of potentially hundreds, it seems too early to attribute any definitive labels. But, following Akrich, Callon and Latour through the harsh light of retrospective, our journey has been one of both *interessement* and resistance, of betrayals and alliances, of numerous adaptations and streams of accusations - in this sense; a process of *innovation* (2002a, 2002b).

[1] Another direct inspiration for our approach, comes from the work of TANTlab on developing a framework for doing Participatory Data Design (Jensen, et al. Accepted/In press).

This is why the following story won't be a tale of heroic entrepreneurs or genius designers. It will be a story of a confrontation with our own hybris; of the challenges we encountered and the not-quite-rational-but-very-urgent choices we had to make to overcome them. Ultimately, how these trials and tribulations have affected where we stand now. As such, an "experimental innovation process" might be the most fitting label we *are* capable of producing right now. The "experimental" part pertaining both to our ambitions of novelty and our entry into Digital Methods as "skilled beginners". The goal thus seems basic in principle: "just" get the data, "just" do the visualizations and "just" do Participatory Design. Deliberate oversimplifications aside, as anyone who has done Digital Methods, design or any kind of research for that matter can probably attest: things are rarely this straightforward. So let's start from the beginning.

Dumpster Diving on Discogs

The first step was thus: "just get some data". Of course, first we had to find a source. The requirements for this were also quite simple: a) the data had to be more or less readily available (i.e. no obtuse barriers for data collection); b) the data had to be relatively plentiful - no sense in looking at relational data if we can only get a handful; and c) ideally, the data would be accompanied by as much metadata as possible - to make room for later uses. With this in mind, Dicsogs seemed like an ideal candidate. Discogs being at once both an online record store and a user-generated discography (hence the naming) could thus likely accommodate all of our data needs. A quick look at an arbitrary Discogs entry confirms this (Figure 1). It contains plenty of useful information about the release. Ranging from: artist and title (1), music styles and date (2) or even community metrics (3). Discogs even had a friendly

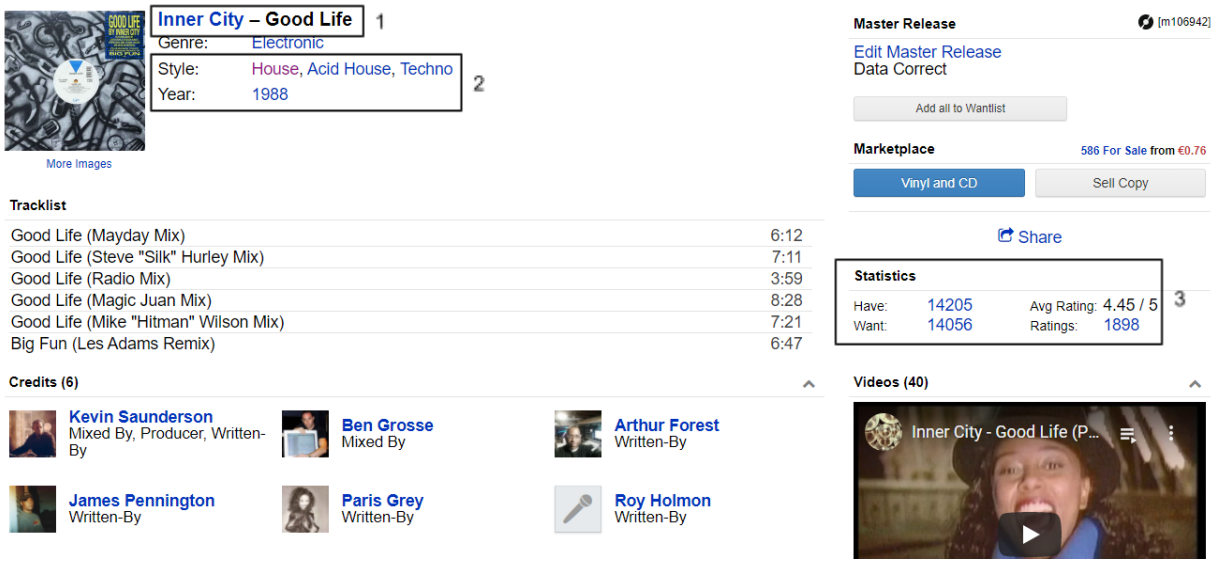


Figure 1. A closer look at the Discogs page for "Good Life" by Inner City. 1) Artist and release title, 2) music styles and data, 3) community metrics. Screenshot from: https://www.discogs.com/Inner-City-Good-Life/master/106942

looking API[2] - which within Digital Methods usually translates to an easy way of getting access to structured data from a site. True to form however, these can often be a mixed blessing; allowing data collection without the need for a scraper[3], but at the same time making said data collection rest squarely on the premises of the site in question (Perriam et al., 2020).

Hence, we quickly stood before the first crossroad. See, Discogs, besides having an API, also dump all their data on a monthly basis. We mean “dump” in a very literal sense, as the files were basically a HTML print of the entire website. The very ‘raw’ nature of the format, however, also meant that we would have access to practically *all* the data we could ever want - more than five millions worth of electronic releases; complete with metadata about artists, labels, music styles and so on! This had an admittedly alluring effect on us - the more data, the better, we thought. In practice though, we found that text editors like Excel and Notepad++ couldn't even open the files. Even with purpose built software like Sublime Text, 60 gigabytes worth of XML file is quite a lot to handle for most computers. Software and hardware limitations aside, what made matters worse, was that the files had to be filtered down to only the parts relating to electronic music. To do so, we would have to manually find patterns in the data and write code to extract it. The way things went, even if we managed to open the files, the format made them very hard for humans to read (Figure 2).

Figure 2. An excerpt from the first data dump on Discogs, concerning “Stockholm” by the Persuader.

After a lengthy process of searching through various XML parsers, looking through code repositories, and trying to identify the correct “carrots” (“<>”) of the file - we began looking towards the alternative. In contrast, the Discogs API seemed practically inviting. As a mirror of the site's own search function, the API would (on paper) allow us to get *only* the data we wanted, and in a nice clean JSON file as well! (see figure 3) In reality, the API wasn’t nearly as easy a solution as we could have hoped for.

[2]Application Programming Interface (API), in short an interface with the software application (here Discogs) that defines the types of call or requests that can be made from the application, the format of these and so on (Munk, 2019).

[3] (Web) Scraper is a tool that automatically grabs specific information from a webpage (Munk, 2019).

```
{
        "country": "Germany",
        "year": 2021,
        "format": [
          "Vinyl",
          "12\""
        ],
        "label": [
          "Misfit Melodies",
          "Pets Publishing",
          "Minds On Fire",
          "SMV Schacht Musikverlage GmbH & Co. KG",
          "WAS-Word And Sound Medien GmbH",
          "R.A.N.D. Muzik",
          "Misfit Melodies",
          "Misfit Melodies"
        ],
        "type": "master",
        "genre": [
          "Electronic"
        ],
        "style": [
          "House",
          "Nu-Disco"
        ],
        "id": 1990639,
        "barcode": 4251804123471,
        "user_data": {
          "in_wantlist": false,
          "in_collection": false
        },
        "master_id": 1990639,
        "master_url": "https://api.discogs.com/masters/1990639",
        "uri":
"/Catz-N-Dogz-Gerd-Janson-Modern-Romance/master/1990639",
        "catno": "MFM08",
        "title": "Catz 'N Dogz & Gerd Janson - Modern Romance",
        "thumb":
"https://img.discogs.com/NRy3CqnjUjwWzGqxqELV58_Oypw=/fit-in/150x
150/filters:strip_icc():format(jpeg):mode_rgb():quality(40)/disco
gs-images/R-17065878-1613130780-3055.jpeg.jpg",
        "cover_image":
"https://img.discogs.com/t1Z_JtPKqy_zAJGXg7fVtugOXcc=/fit-in/600x
573/filters:strip_icc():format(jpeg):mode_rgb():quality(90)/disco
gs-images/R-17065878-1613130780-3055.jpeg.jpg",
        "resource_url":
"https://api.discogs.com/masters/1990639",
        "community": {
          "want": 54,
          "have": 78
        }
```


Figure 3. An excerpt from the JSON file gathered through Postman, concerning "Modern Romance" by Catz 'N Dogz & Gerd Janson

Because Discogs API itself mirrors the search function of the site, it also has the same performance limitations. That is, the search function and the API both use *pagination*. Pagination, in short, means that the site returns search results divided into a number of pages - here at the maximum of 100 items or releases per page. While inconvenient for anyone with data science goals, it hardly constitutes a disaster. What does though, is the caveat: "up to a maximum of 100 pages per search".

> To put it differently, any API query we made could only return a maximum of 10.000 items. In the words of the Discogs staff member "rodneyfool" at their development forum "We've found 10.000 results to be a sufficient amount for most use-cases [...]" (n.d.).

The problem should be quite clear by now: as researchers, *not* users, going from five million data points to 10.000, is quite disheartening.

Hence the dilemma. On one hand, the XML dumps represented a greater technical challenge to get working - but had the benefit of not restricting our data collection. On the other hand, the API approach appeared much more manageable. While we perhaps couldn't get *all* the data we wanted, the data we could get, was just a few clicks away. In the end, after much internal deliberation in the group - and decidedly without all the information needed to make an informed decision - we chose the API approach.

Persuading the Discogs API

The second difficult (and quite urgent) decision was thus: how to limit our data inquiry?

> In the end, we settled on 12″ vinyl releases. We chose this delimitation by referring to the potential collaborators in the Copenhagen area. In doing so we pondered: who could

> have a use for such a “Thing” we were making. We found an abundance of DJ’s, producers, record labels, record stores and radio shows that could have been viable candidates. A common thread among all these is that they had a vested interest in niche music that was printed on vinyl, which among other things, are the remains of a time when digital audio formats were outperformed by the analogue (Cornell, 2016).

To limit the search even further, we also opted for only using so-called “Master Releases”[4], meaning that we also avoided any problems with duplicates. Down to 202.000 entries. Sadly, still a lot more than the maximum 10.000. After trying many different ways of delimiting the search, we found that going through the releases year by year would yield us a maximum of 96 pages per year - and would thus allow us to get all 202.000 releases without cutting further corners.

“Easy enough” you might think, “now we just have to make some 50 API queries, and then we’ll have our data”? Not quite. To avoid rate limitations we had to obtain authorization from Discogs, a process which is meant for businesses wanting to automate recordkeeping of their music stock - that is to say; *not* research. One newly minted Discogs application later and a whole mess of technical troubles with obtaining authentication tokens; “...and we were in” (as hackers say in the movies).

Surely, then we can make our API requests and be on with our business? But sadly, once more, pagination got in the way. Not only could we *only* get up to a hundred pages worth of releases, we could also *only* get them *one page at a time*. Enter the first real *social* alliance of the project: Poppy the Programmer (who also happens to be the girlfriend of one of the authors - but that’s besides the point). With a bit of JavaScript, a self-iterable variable was made, so that each subsequent query would get a new page each time. The program we used for the authorization and query making process (Postman) even allowed for *multiple runs* of a query in one go! Time consuming, but doable.

“Finally” you might think - having successfully negotiated an alliance between the Discogs API and Postman - “we can now surely get the data!” Again, not quite. See, Postman is an excellent tool for *doing* the API requests themselves, but as mentioned, is not exactly made for social data science purposes. So while we could easily do hundreds of requests with a single click, the software is only meant to *test* API’s, not extract data from them. Enter Poppy

[4] “Master Releases” is Discogs way of handling multiple versions of the same release. Say a release was handled by a different record company when distributed to different countries or that there exists both a 7″ and a 12″ version of the same release. In such cases Discogs users will list the different versions separately, but also create a Master Release page, listing all the different versions of the same release.

the Programmer again. This time, instead of doing a variable for the page count, we realised it was possible to write data *directly* to a variable. Thus bypass the restrictions on data extraction. The only thing left was the somewhat laborious process of going through all the API requests year by year and manually combining the data to a single JSON master file.

Right, "now that we have the data", we thought, "we just have to put it into Gephi[5] and visualize it!". Easier said than done. In fact, we met even more resistance just getting the data *into* a network format, than with harvesting it in the first place.

All roads lead to Gephi

Our primary idea for the project was to make a relational map of electronic music. In its basic form, this would entail doing a network graph; plotting artists (A), labels (L), releases (R), and their links or *edges* (see figure 4). On a technical level, this was a matter of translating our data into something Gephi could understand. On paper, making network visualisations in Gephi is not too complicated. You give it two CSV files; a list of nodes and a list of edges. Et voilà, you have a Network Visualization!

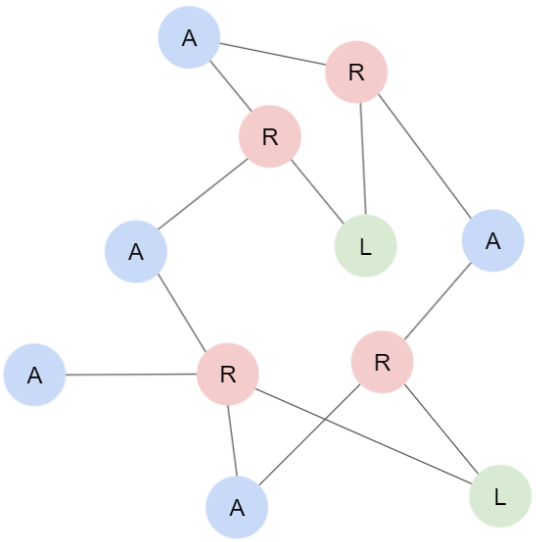


Figure 4. A schematic example of how our network is structured.

In principle, converting our JSON data to Gephi CSV shouldn't be too hard either - especially with a little Python know-how and some help from Poppy the Programmer. But this is where the choice to work with the user-generated data came back to bite us. The data being made by users might not seem problematic at first glance (it might even be beneficial from the right perspective): But where the XML file had a "carrot" for everything, our API JSON only had one line for both the artists *and* the title of the release (see figure 5).

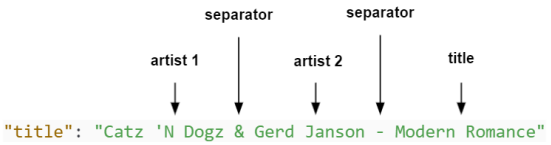


Figure 5. Example of how artists and release titles are separated in the original Discogs API data file

[5] A popular and flexible network graphing tool (Jacomy et al., 2014).

“No problem” we thought, “we'll just split the artist and title at the EM-dash; and, in cases of multiple artists, by any “and” or “&”. “Easy!”. Into Gephi our newly minted CSV’s go.

“Betrayal!”

“Reading Error!”. “This is not a real CSV file”, Gephi accuses us, “I cannot possibly read this!”. Back to the proverbial drawing board again. It turns out that any arrandt quotation marks or commas break the file format. A quick snippet of code later the problem is fixed. Back goes the files into Gephi.

“Betrayal!”

“Encoding Error!”. “This file isn’t in the utf-8 encoding format you told me it was” cries Gephi, “I cannot possibly read this!”.

Back to Python again. Because, to our great inconvenience, although all of our files, through all the steps up until now had been encoded with “utf-8”; if you don’t *explicitly* tell Python that, it stops being the case.Back into Gephi again. This time it’s a success! We get our network. But something seems off. Our three most prominent nodes (by a magnitude of hundreds) are “Various [artists]”, “Untitled [release]” and “Not on Label”, respectively (see figure 6).

Later, when talking to a record store owner, we find out that, in a historical context at least, electronic music differs from other genres in an interesting way: In contrast to Pop or Rock music for

Figure 6. A viewing of our first network, where node size is determined by degree. The coloration of nodes is as follows: Orange indicates artists, purple releases and green labels

> example, where it is often the norm to know who the lead singer is, in DJ culture, it was often the opposite. It turns out musicians generally wanted to 'let the music speak for itself' rather than it becoming popular because of the momentum fame might bring. Whether because they wanted to keep the records they played a secret so that other DJ's couldn't copy their style; or just because they wanted to change the style of music they played. This in turn means that the average electronic musician would have at least a handful of aliases[6]. (Kamper, M., personal communication, April 30, 2021).

The network seemed practically unreadable, so we had to intervene - empiricist ambitions be damned (Marres, 2015). By removing all of the different naming artifacts, we finally got a 'readable' network. But hold on, something is amiss again. This time, all of the labels on the artist nodes looked weird. Names were arbitrarily split and the most prominent "artists" in our network this time, were single letters [figure 7]. Back to Python once again. Turns out, separating artists in the data was more complicated than we thought. As the Discogs page for a given release is user-generated, it also means that *anyone* writing said page, can use *any* number of ways to split artists apart. Not just "&" or "and". Through a process that can best be described as "one part programming and three parts trial and error", we constructed a list of regular expressions that would (hopefully) allow us to separate artists from one another. It goes as follows: either lower- or uppercase, with or without punctuation marks, and with a variable amount of whitespace: "&", "and", "featuring", "feat", "ft", "presents", "aka", "with", "w/", "vs", "x", "+", ":", "/" and finally ",".

Network Projections and Memory Troubles

Now that our data was clean - no more encoding errors, no more format breaking commas, no more naming artifacts - one might be tempted to say that Gephi was finally "recruited" into our *Thing*. But as you already might have begun to expect, nothing in this process ever went according to the plan (if there ever was one). In Visual Network Analysis the goal is often (on the technical side) to identify "clusters" in your data. These are, structurally[7] and/or topologically distinctive areas, where the density of nodes is

[6] For a humorous but not at all atypical example, see the German musician Uwe Schmidt, or as his Latin Dance alter ego is known: "Señor Coconut".

[7] Here "structurally" refers to using Gephi's "modularity class" algorithm (Blondel et al., 2008; Lambiotte et al., 2009) to look at how how nodes group together based on edges in common, as opposed to looking at the layout created by an energy repulsion algorithm (Jacomy, 2020)

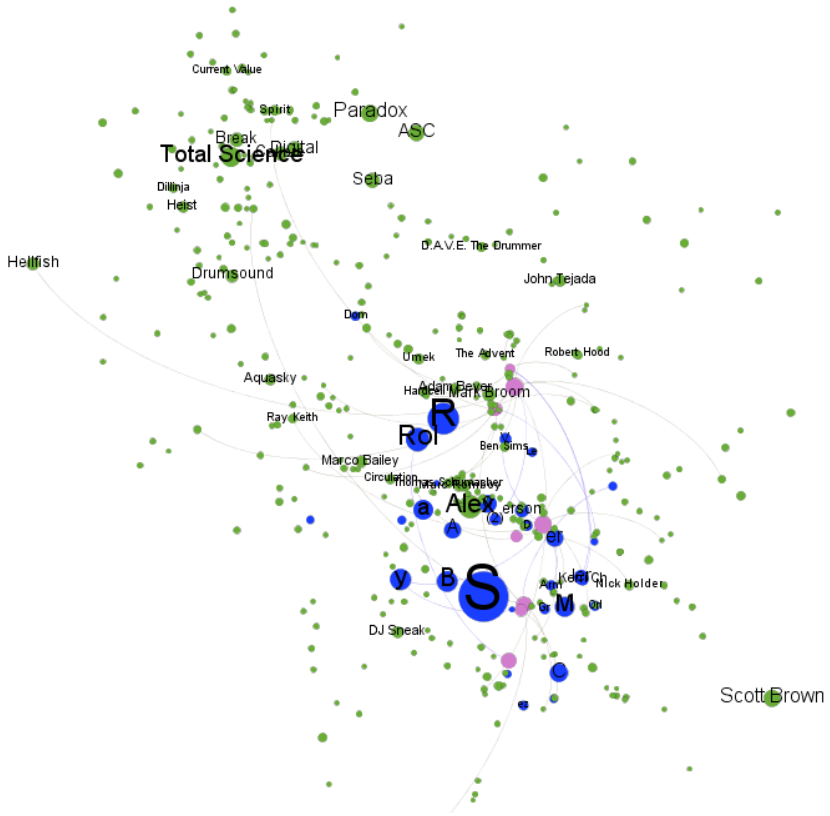


Figure 7. Excerpt from our network, only showing the 200 nodes with the highest degree. Node size is set after degree, color by type. The blue colored nodes show example of single letter "artist" or artist with broken-up names.

higher than other places (Jacomy, 2020). On the qualitative side, the goal is then to "explore" what makes these clusters appear, and what this can tell us about the research subject. In many ways, this influenced both what we wanted out of our Gephi networks, and what we wanted to (literally) *see* in the network layout. As such, we mainly expected our network to be a "topography of electronic music", where we, the proverbial cartographers, could "map" the various areas. And hopefully: to find clusters representing alternative categorizations as compared to genres. But this wasn't the case. Despite our best efforts, the closest thing to cluster separation we achieved was a network more resembling an "eyeball" than anything else (see figure 8). No nice, clear, colour coordinated clusters, just a big messy hairball (Jacomy, 2020).

So what should we do, we thought. Just give up? Or maybe retrace our steps all the way back to the data dumps and hope for better results next time? In the interest of time and our own sanity, we went for another approach. See, technically speaking, our network can be described as "bi-partite" (or tri-partite to be precise) (Totet, 2020). This means that there are multiple types of nodes, and that nodes of the same type cannot link to each other.

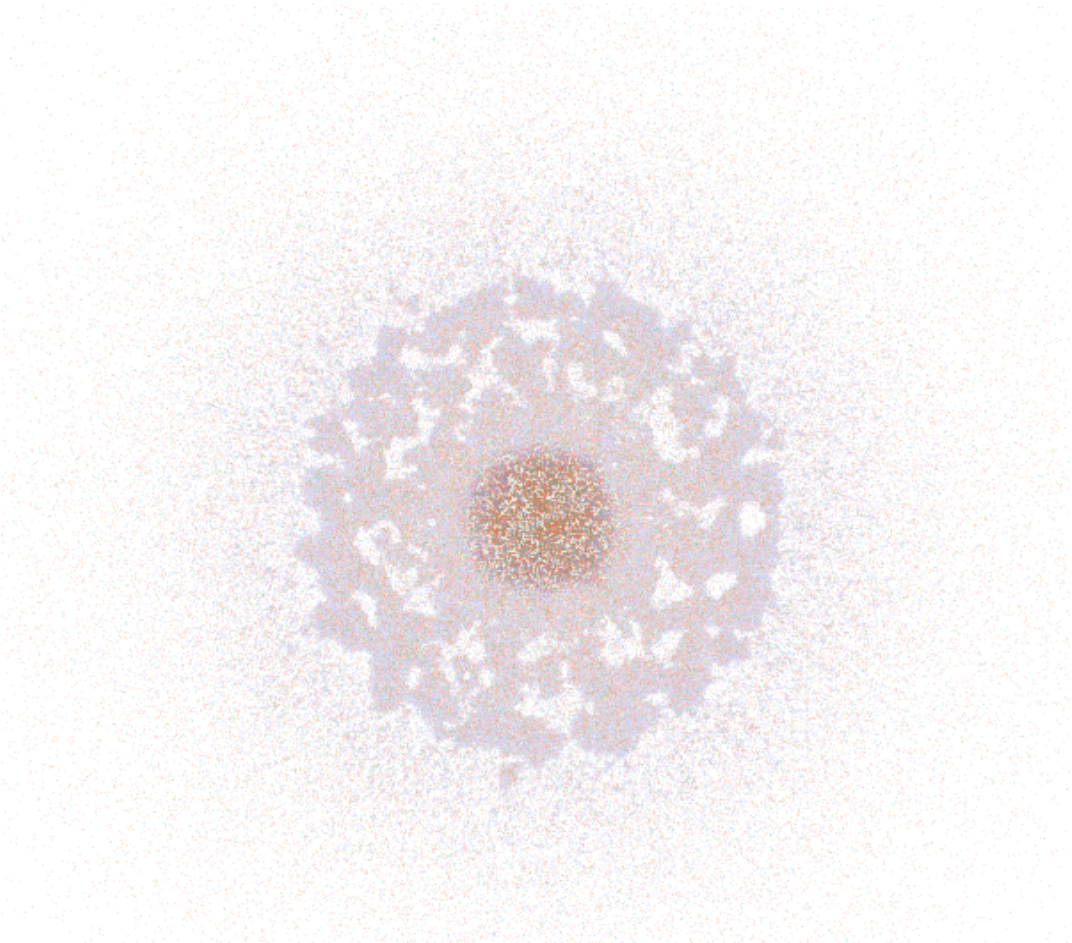


Figure 8. Rendition of our first "successful" network visualization. Node color is set by type, size by degree. Specialized with ForceAtlas2, with increased edge weight to promote clustering. Furthermore, the network is cleaned of "faulty" nodes, naming artifacts and filtered by giant component. Original network size: 273243 nodes and 406546 edges. Cleaned: 269174 nodes and 387546 edges

To be concrete, the release nodes are only connected to artist- or label nodes, not other release nodes. And vice versa (see figure 9).

The important point here is that this property of the network (its "bi-partiteness") would allow us to project one part of the network on to another (Totet, 2020). In practice this means that one type of node in the network, the labels for example, can be "projected" onto another type of node in the network - in this case releases (see figure 10). This would then mean that any two or more releases having the same label, would then gain a direct edge between them, while the corresponding label node would be removed. We could then repeat the same process with artists and releases, and produce a network of *only* release nodes - where all edges represent connections to common labels or artists by proxy. What's even more intriguing, was that Gephi even has a plugin (Multidmodenetworkstranformations) that can do these projections for you.

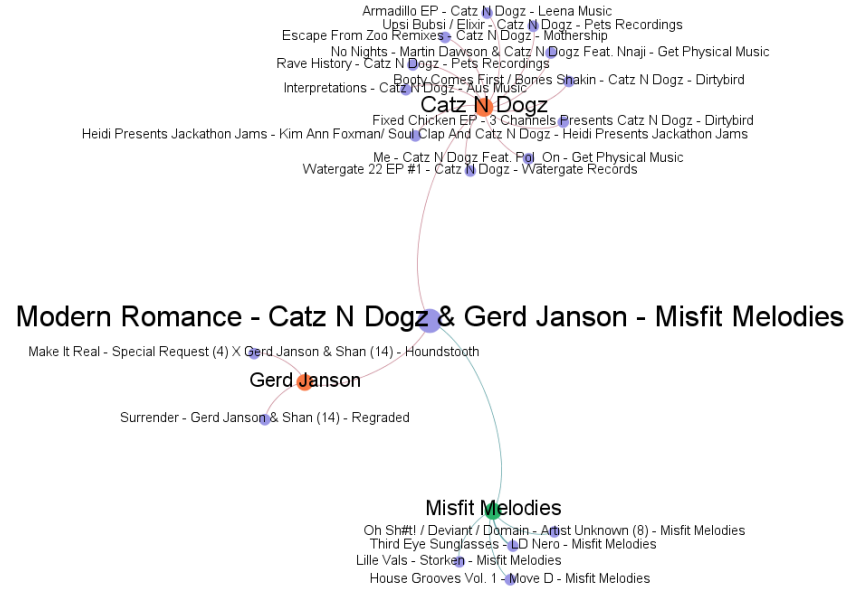


Figure 9. Excerpt of the network surrounding the release "Modern Romance" by Catz 'N Dogz & Gerd Janson. The network shows the primary release, the artists and label connected to it - and the the releases immediately connected to said artists and label.

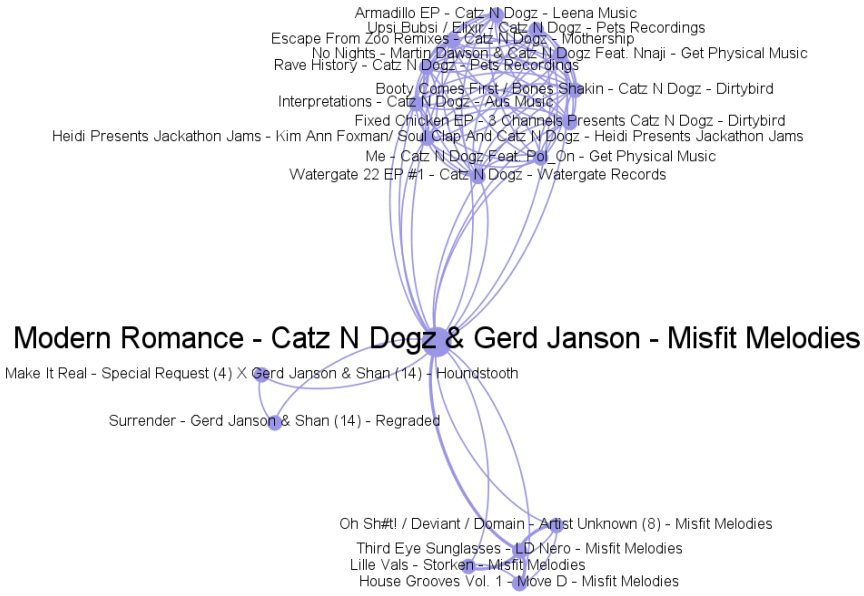


Figure 10. Excerpt of the network surrounding the release "Modern Romance" by Catz 'N Dogz & Gerd Janson. This network has been "projected" with the "multimodenetworktransformation" plugin for Gephi. This means that the label and artist have been removed, and edges between all releases connected to it have been made instead.

“Betrayal!”

“This network is too big” says Gephi, “I cannot possibly process this many things at once!”

Even with 32 gigabytes of RAM at our disposal, and all the tricks we could muster to make the files more j- there was no persuading Gephi to project our networks for us. See Gephi is built on the programming language Java, and Java is (so we learned) not very good at memory management. So despite our best efforts there was seemingly nothing to do within the frame of Gephi itself. So once more, we went back to Python. And with the help of the Python library NetworkX, and a small script: we made the files necessary to construct what would *hopefully be the final* network. And as it turns out, this time it actually worked (see figure 11)! Despite problems with merging the various files, and with finding the right layout settings: we finally had cluster separation, both topologically and structurally![8] We even had all the metadata we wanted - including hyperlinks to all the Discogs pages of the releases. Still, the network was far from perfect. All the metadata made the network cumbersome to deal with, and due to the network projection, we had gone from 7.000 edges to more than 4.000.000. So even if we had separable clusters, in practice, the process of exploring these would have been undoable within a reasonable time frame. To make matters worse, while most of the individually visible clusters or modularity classes we looked at could be meaningfully described as converging around specific labels or music styles — A very large part of the network was simply too topologically “flat” - too intermingled - to make sense of. Even though this is all very interesting from a Digital Methods perspective; the nature of the network made it practically unusable for our design intentions.

[8]We went from some 80,000 modularity classes, to 100 and from a complete “hairball” to a partial harbaill with visible clusters around the “corona” of the network.

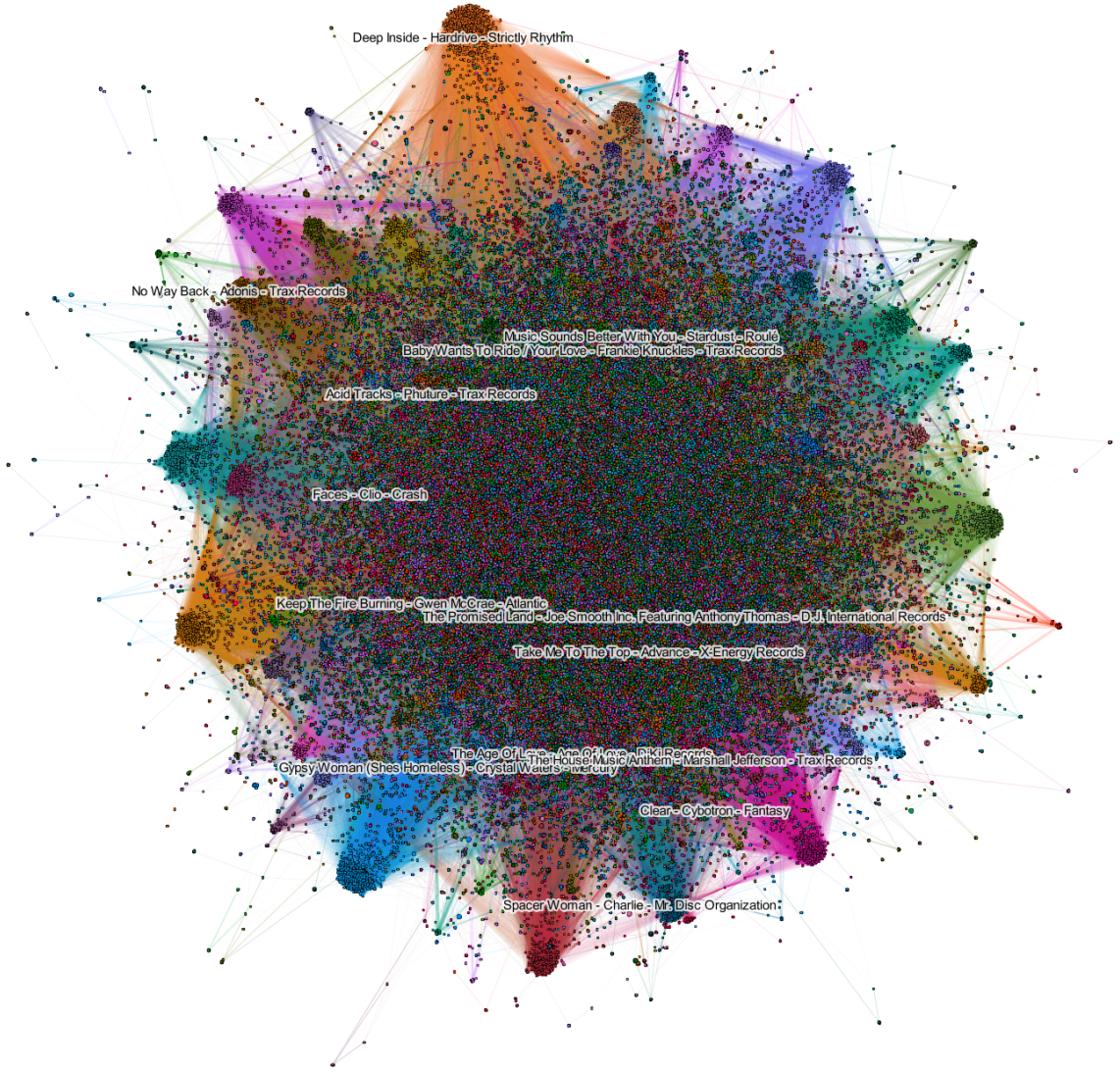


Figure 11. This was the final network we made before the workshop. Its edges are made from the the projected networks of "Artist-Releases" and "Labels-Releases. Size is set by our new "Popularity Metric" (difference between "want" and "have" community metrics). Layout is ForceAtlas2 with increased edge weight and node overlap to promote clustering. Color is set with the "Modularity Class" algorithm. The visible labels of some nodes, are the 14 biggest releases (in terms of the "Popularity Metrics") in the order of "Title" - "Artists" - "Labels".

Designing with Visual Networks

"What now", you might think, "what do we do with a too-big-to-manage visualization, and no real solution in sight?" Well, remember how this section of the paper is called "Heterogenous Engineering" not "Gephi Engineering". While we had been hard at work on the data side of things, we had also been trying to find potential collaborators. Although not everyone was as easily convinced by the idea of a genreless music map.

> Speaking of which, the first recordstore to respond to our pleas, was not necessarily sceptical of our approach - but

described our project as “antithetical” to their own business. Where our project rested on the idea of removing genres from the picture, this store used the various tall tales and historical anecdotes about genres as the main sales pitch for second-hand records. That is to say, the customers don't need to be able to navigate the music themselves, if the one curating it for them is already an expert in the field. While obviously true, it very much ran counter to our idea.

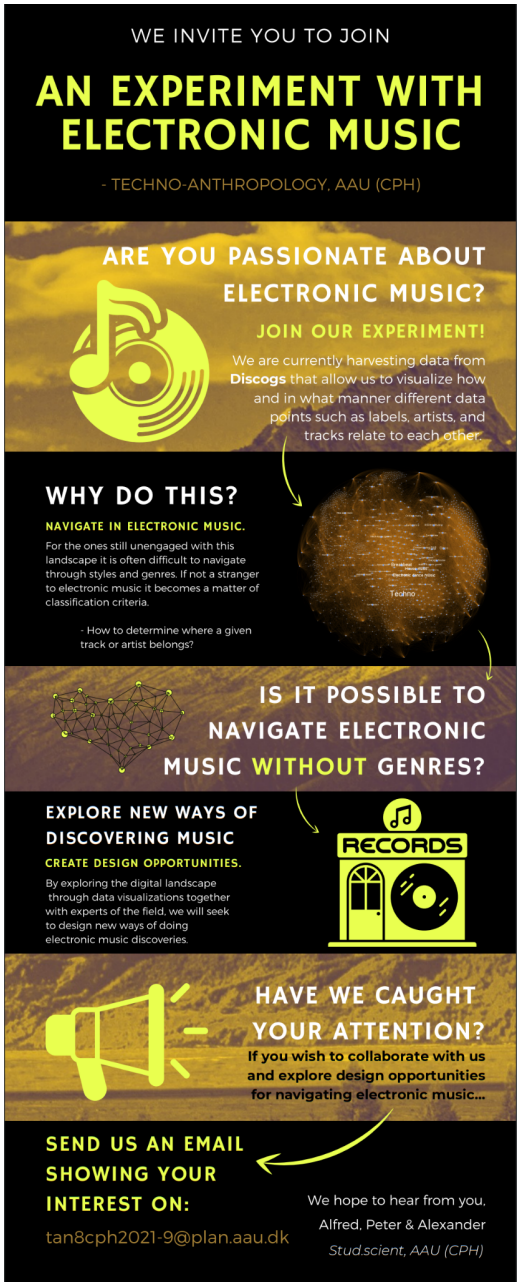


Figure 12. Poster we made as part of the effort to find people or records stores to collaborate with.

But luckily we did find a partner in crime: Proton Records. Proton Records is both a booking agency and a record store, specialising in curating and promoting adventurous and alternative electronic music - music that might otherwise go below peoples radar. As such, the idea of a “non-deterministic” exploration tool seemed a great fit with Proton's other goals. A big part of Proton's claim to fame is not necessarily having the biggest inventory or the cheapest prices, but rather having a unique and interesting music selection. As Proton puts it:

> “Proton is a Copenhagen-based group of curators promoting electronic and adventurous music. Proton aims to further expand the independent music scene in Copenhagen, nationally and internationally by local and international non-profit initiatives.” (Proton, n.d.).

Part of Proton’s business model is also figuring out which old records might be popular reprints. So when we first met with

Anders, the co-owner of Proton - besides being met with unusual amounts of friendliness, enthusiasm and music knowledge - we also found both the first potential use case *and* major reconfiguration of the initial design concept. See, as trance is currently having a bit of renaissance (according to Anders), reconfiguring our maps to find trance music viable for reprint, seemed ideal. This was an easy adaptation in principle: we could identify "popular"[9] trance releases, then look at its immediate neighbors and qualitatively determine their interest. This approach, however, is quite a departure from looking at clusters in the entire network. Which got us thinking: what if the lack of readability of our network wasn't a disadvantage? What if we could actually use its relative "unfinishedness" as a design opportunity?

Following Participatory Data Design, we could probably call our construction a "flexible visualization" (Jensen, et al. Accepted/In press), but words like "prototype" (Björgvinsson et al. 2012), "low-fidelity mockup" (Jensen, et al. Accepted/In press) or even "boundary object" also apply (Star & Griesemer, 1989; Jensen, et al. Accepted/In press). Regardless of terminology, our path seemed laid out. Instead of doing the usual approach with Visual Network Analysis of printing out giant posters of the networks and annotating them by hand; we would do something different. Something which seemed very clever at the time - but not quite so much in retrospect. We would invite Anders directly into "the machine room" of Gephi, and together find out which of the various options, filters, and metadata we had available, would be conducive for a potential design. We would, in the best participatory fashion we could facilitate, *de-construct* our *Thing*, collaboratively, and explore any design possibilities that might arise from this process. And then - along the lines of Latours idea of the *anti-program* (1990) - figure out which elements would be necessary to *re-construct* our *Thing without* our presence. Of course this all seems very rosy and it would be downright uncharacteristic to the story we are trying to tell, if everything went according to plan. So of course it didn't.

A machine room workshop

The first very practical problem was how to transport our decidedly *not* practical networks, from the harddrive of a stationary computer, to a record store on Nørrebro. Already, one might begin to dread not printing out the networks instead. Luckily, we were able to borrow a computer competent enough to handle the network files. Or so we thought. When we got to the store, excited to show off all our hard work - we were betrayed by Gephi.

[9] Based on how many people on Discogs "want", "have" or both, a specific release

Again. Although the borrowed computer could manage the files on a data level, albeit sluggishly, the rendition of the networks themselves didn't prove so collaborative. They would freeze and force a restart at the most inopportune moments. Maybe the networks had worked fine without too much tampering, but our constant 'fiddling' with settings and filters proved too much. When we finally got Gephi under control - more due to patience than technical guile - an hour and a half had passed, and no one had the stamina left for the laissez-faire explorative mindset we had hoped for. But, it wasn't all disaster - a hard learned lesson about the usability of network maps aside - the parts of our work we did manage to show, worked well as part elicitation device and part proof of concept. As such, we did find two relatively qualified design outcomes: Building on the idea of tracing hype and relations with trance music; the central design would be to make a web based application that would allow the exploration of an "ego network[10]" of a specifically selected release. This network could then be further filtered as necessary, by specifying which sub-styles, timespans or popularity metrics to include or exclude. An important functionality, following the necessity of usability, would be to showcase the various metadata of a release in a meaningful way - most importantly, a direct link to the Discogs page of a release — That the experience of going from exploring the network to listening to it, would be only a few clicks away. The final functionality discussed was the idea to import an inventory list directly from Discogs, to either use as the seed for the relational networks, or as the object of the exploration itself.

Apart from the design, two possible use cases were also discussed. In the first use case, this would mean a record store manager inputting a currently popular release, and the application then creating a network of the release and its neighbors, out to a variable depth (see figure 13 for an ego network we explored in the workshop). The important point being, that this type of network could act as an alternative to clicking your way through Discogs manually. Ideally you would then be able to explore releases that would not necessarily appear through a regular search. The other use case would be that of a music enthusiast who wants to find new interesting music. In this case, the input would be a favourite track or a release they want to find similar music to. Then, instead of relying on the use patterns of other people, the network would provide a relational look into music that associates with the given input release.

[10] "Ego networks" are networks constructed from only a part of the entire network that connects to only a single specific node, the "Ego". In this case, depth refers to how many "steps" away from the initial node you take in terms of constructing the ego network.

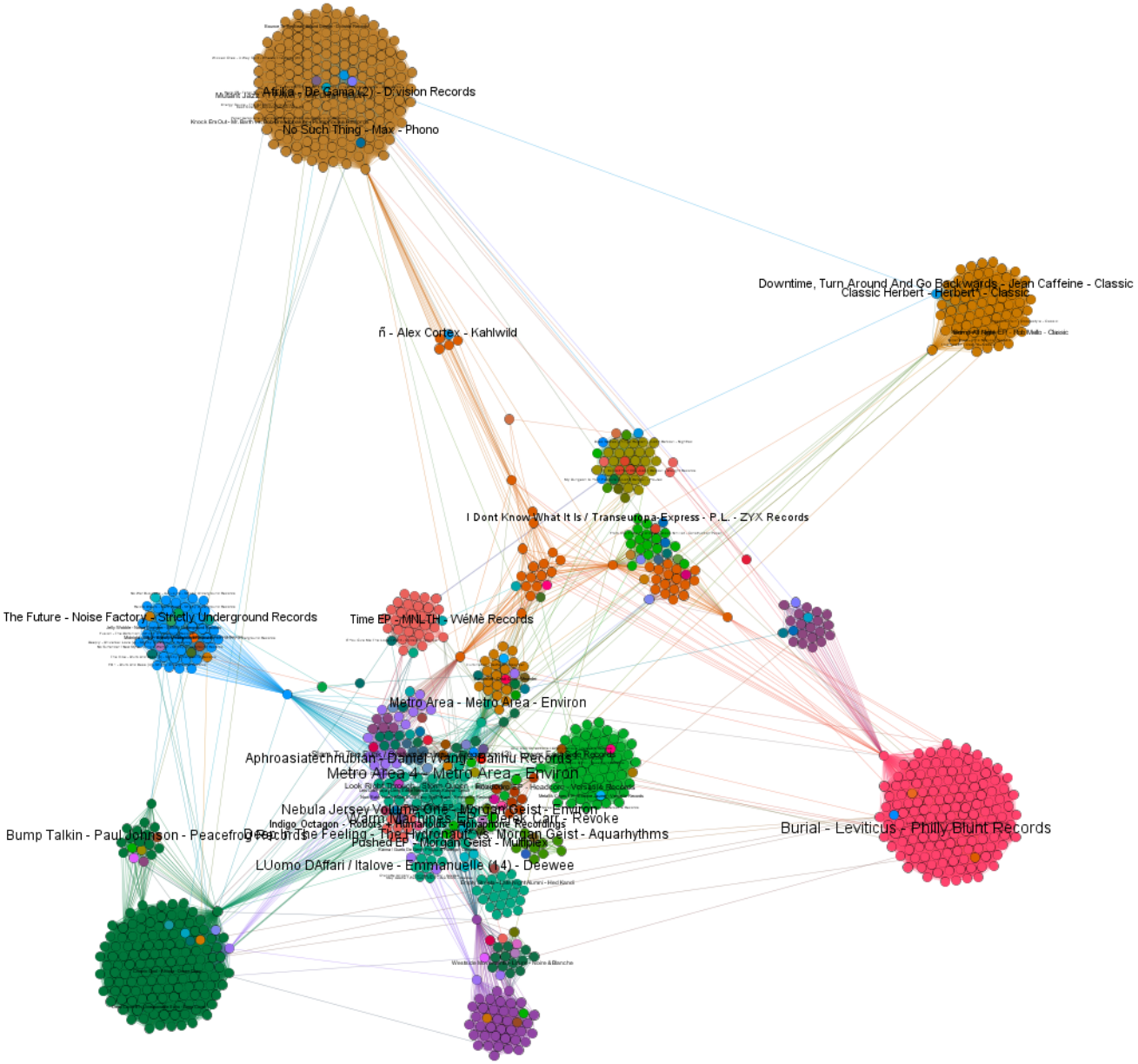


Figure 13. One of the ego networks (surrounding "Metro Area" by Metro Area) we explored as part of the workshop. Color is inherited from the "big" networks modularity classes, but the network itself has been spatialized again. The "named" nodes show the most "popular" releases in this network. The labels of these nodes are ordered as follows: "Title" - "Artists" - "Labels".

Discussion

In the beginning of this article, we promised to account for how our initial design concept has changed throughout the innovation process. Both for the sake of transparency and as an attempt to confront our own hybris. Through this process, both the *Thing* and the idea it was born from, has gone through a number of adaptations, transformations, and translations. But “all translations betray” and all manipulations means giving up one thing in the pursuit of another (Law, 2003; Latour, 2004). So what are the “unaccountables” of our process, the “loose ends” of our network?

From the first step of data collection, sacrifices had to be made. Either due to the specifications we had to make, the user-generated nature of the data, or that bits and pieces were simply lost along the way towards the network visualizations. The end result is the same. Multiple times, when either Anders or we attempted to find certain releases in the network - releases that rightly *should* have been there - we could not find them. Even though they appeared in the original dataset or on Discogs; they were nonexistent in the visualization itself. Similarly, in the process of making the projected networks, *something* went wrong. Again, whether it was due to the data, lapses of judgement on our part, or simply the aggregate of technical steps where something can go amiss - we don’t know. But the outcome is that some parts of the network *make sense* - either directly through sharing common denominators or, qualitatively speaking, for experts such as Anders. Other times, some connections in the network seemed placed *almost* at random. And to make matters worse, as of now, we don’t even know where to put the blame.

But, all that being said, the design concept and the prototype, is more than the sum of its faults. Not all of its particularities can be explained away with reference to technical resistance. Most of the major changes come from design insights. The first major adaptation was the departure from the idea of a printable map with annotated clusters. In practice, we found the sea of electronic music to be too vast, too interconnected, for this approach to be fruitful. Instead, as an alternate course of action, we adapted to using ego networks centered on a single release, by which much more manageable network visualizations were made - both technically and analytically speaking. And more importantly, these ego networks are *much more* user-friendly, than dealing with the giant hairballs that make up the full networks.

Secondly, the idea of making maps that *completely* eschew genres or styles, also went more or less out the window. Even when trying our best, genres can’t really

be removed from the picture - and neither is it necessarily desirable to do so. Just as artists and labels are largely interchangeable in terms of how the networks are ordered. According to Anders this is partly due to artists using aliases (oftentimes several), and to artists sticking to the same labels. And vice versa. Hence, our goal was changed once again. It became a matter not of removing style and genre descriptors, but creating networks that didn't *rely* on them. Networks where potential users could *themselves* choose how to delimit their search - be it either by genres and styles, or by other metrics such as country, year or popularity.

Arguably the most important lesson, however, was the first hand experience of the difference between a researcher and a user exploring a network visualization. While we can make as technically impressive or ambitious visualizations as we want, none of it matters if they are not *accessible* (Vikkelsø, 2007) to future users. To that end, we think the most important question we can currently keep in mind while continuing the work on a prototype, is the same question we asked Anders during the workshop: "what would happen if we or Gephi weren't there?" In other words, what is the *anti-program* to the *program* of ours or Gephi's presence (Latour, 1990)? And more importantly, how can we *delegate* this *program* to the design we are working on (Latour, 1992)?

So where does this leave us now? For one, we can try to make the *Thing* we have made - *in all its multiplicity* - accessible to others (de Laet & Mol, 2000). If not to open the *blackbox*, then at least to make it *open source*. That is, if we want to hold ourselves to the standards of Participatory Design - or to the idea of making a countermove to algorithmic decision making. In this sense, the least we can do is make our prototype flexible enough to openly invite *re-designs* and *re-appropriations*, while still being stable enough to fulfill its multiple use cases (Star & Griesemer, 1989; Björgvinsson et al., 2012). To which end, we have made a github repository , where all the technical details of the project will be made accessible for anyone else interested.

Conclusion

In this project we have set out to develop an exploration tool for electronic music, by combining Participatory Design and Digitals Methods - all through the sensibilities of Actor-Network Theory. As such, the original idea was to do network visualizations that circumvented genres for their layout, but instead used the relations between artists, labels and releases as the basis. In this sense, the point was to translate our ANT-sensibilities into a material form, to allow future users to explore the music without relying on prior categorizations.

In the process of developing this idea, it has, however, changed quite a bit. Although the purely technical resistance we met with repurposing user-generated data for research has left its mark - it was the participation of one record store owner, Anders, that made the biggest impact on the design. Technical limitations aside, work on a prototype can continue with three strong design prescriptions: Firstly, the objective of using network visualizations should *not* be to analyze and annotate clusters in the *entirety of the* dataset, but rather to explore the *ego networks* of *specifically* chosen releases. Secondly, it is not necessarily meaningful to strive towards removing genres from the networks entirely; rather, the goal should be to make visible and flexible, the various ways the ego networks can be delimited. Finally, all of these technical and conceptual considerations are meaningless, if the networks themselves are not *accessible* to other people than researchers. As such, the most important task in making said prototype, is to *delegate* the role that we - and Gephi - play in making visually and contextually manageable networks, to an external tool. A further obligation, towards both the theory, design and people involved in it, is also to make both our technical struggles and the results, available to other interested parties. If not to be picked up after the short lifespan of a semester project, then at least to make transparent the steps we went through to get to where we are now.

References

Akrich, M., Callon, M., Latour, B., & Monaghan, A. (2002a). The key to success in innovation part I: The art of interessement. International Journal of Innovation Management, 6(2), 187–206. https://doi.org/10.1142/S1363919602000 55

Akrich, M., Callon, M., Latour, B., & Monaghan, A. (2002b). The key to success in innovation part II: The art of choosing good spokespersons. International Journal of Innovation Management, 6(2), 207–225. https://doi.org/10.1142/S1363919602000 562

Birkbak, A., & Munk, A., K. (2017). Digitale metoder. Hans Reitzels Forlag.

Björgvinsson, E., Ehn, P., & Hillgren, P. (2012). Design Things and Design Thinking: Contemporary Participatory Design Challenges. *Design Issues, 28*(3), 101-116. Retrieved April 7, 2021, from http://www.jstor.org/stable/23273842

Blondel, V., D., Guillaume, J. -L., Lambiotte, R., Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* (10), P1000

Cornell, K. (2016, April 19). *6 Reasons Why vinyl is popular again*. Turncore. https://www.tunecore.com/blog/2016/04/why-vinyl-is-popular-again.html

de Laet, M. & Mol, A. (2000). The Zimbabwe Bush Pump: Mechanics of a Fluid Technology. *Social Studies of Science*, *30*(2), 225–263. https://doi.org/10.1177/030631200030002002

Helberger, N., Karppinen, K., & D'Acunto, L. (2018). Exposure diversity as a design principle for recommender systems. *Information, Communication & Society,* 21(2), 191–207. https://doi.org/10.1080/1369118X.2016.1271900

Heuguet, G. (2020). Electronic music history reloaded: Ishkur's online "Guide to electronic music 3.0". *Sound Studies* (6:2, pp. 275-279). https://doi.org/10.1080/20551940.2020.1794332

Jacomy, M. (2020, September 26). Hidden structures in hairballs, and how to see them. Reticular. https://reticular.hypotheses.org/1809

Jacomy, M., Venturini, T., Heymann, S., Bastian, M. (2014). ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLOS ONE* 9(6): e98679. https://doi.org/10.1371/journal.pone.0098679

Jensen, T. E., Birkbak, A., Madsen, A. K., & Munk, A. K. (Accepted/In press). Participatory Data Design: Acting in a digital world. In G. Downey, & T. Zuiderent-Jerak (Eds.), Making and Doing STS MIT Press

Lambiotte, R., Delvenne, J.-C., Barahona, M. (2009). Laplacian Dynamics and Multiscale Modular Structure. *Networks.*

Latour, B. (1990). Technology is Society Made Durable. *The Sociological Review (Keele)*, *38*(1_suppl), 103–131. https://doi.org/10.1111/j.1467-954X.1990.tb03350.x

Latour, B. (1992). Where are the missing masses? The sociology of a few mundane artifacts. In W. E. Bijker (ed.), *Shaping Technology/building society: Studies in sociotechnical change* (pp. 225–258), MIT Press

Latour, B. (2004). On using ANT for studying information systems: a (somewhat) Socratic dialogue. In Avgerou,C., Ciborra, C. & Land,F.F. (Eds.), *The Social Study of Information and Communication Study* (pp. 62-76). Oxford University Press.

Latour, B. (2014). The More Manipulations, the Better. In Coopmans C., Vertesi J., Lynch M., & Woolgar S. (Eds.), *Representation in Scientific Practice Revisited* (pp. 347-350). MIT Press. https://doi.org/10.7551/mitpress/9780262525381.001.0001

Latour, B., Jensen, P., Venturini, T., Grauwin, S., and Boullier, D. (2012). 'The whole is always smaller than its parts' – a digital test of Gabriel Tardes' monads. The British Journal of Sociology 2012 Volume 63 Issue 4, 590-615.

Law, J. (1987). Technology and Heterogeneous Engineering: The Case of Portuguese Expansion. Science & Technology studies, 28(5), 169–202. https://doi.org/10.22394/0869-5377-2018-5-169-199

Law, J. (2003). Traduction/Trahison: Notes on ANT. Centre for Science Studies. Lancaster University, Lancaster. http://www.comp.lancs.ac.uk/sociology/papers/Law-Traduction-Trahison.pdf

Law, J., & Lin, W. (2020). Care-ful Research: Sensibilities from STS. Retrieved April 7. 2021 from http://heterogeneities.net/publications/LawLin2020CarefulResearchSensibilitiesFromSTS.pdf

Munk, A. K. (2019a). Four styles of quali-quantitative analysis: Making sense of the new nordic food movement on the web. Nordicom Review, 40(s1), 159-176. https://doi.org/10.2478/nor-2019-0020

Munk, A., K. (2019b). Mapping controversies with digital methods: scrapers, crawlers, APIs. https://medium.com/@EthnographicMachines/mapping-controversies-with-digital-methods-scrapers-crawlers-apis-17e0c96c340a

Marres, N. (2015). Why map issues? On controversy analysis as a digital method. Science, Technology, & Human Values, 40(5), 655–686. https://doi.org/10.1177/0162243915574602

Perriam, J., Birkbak, A., & Freeman, A. (2020). Digital methods in a post-API environment. *International Journal of Social Research Methodology*, 23(3), 277-290. https://doi.org/10.1080/13645579.2019.1682840

Proton (n.d.), About. http://www.pr-t-n.org/about/

Rodneyfool (n.d.). 404! Oh no! - access limit of 10000 releases on explore discogs …. https://www.discogs.com/forum/thread/747186#7438700

Star, S.L., & Griesemer, J.R. (1989). Institutional Ecology, `Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. Social Studies of Science, 19(3), 387-420. https://doi.org/10.1177/030631289019003001

Totet, M. (2020, November 15). K-partite and Bipartite Graph (Multimode Networks Transformations). Github. https://seinecle.github.io/gephi-tutorials/generated-html/k-partite-bi-partite-graph-en.html

Venturini, T. (2010). Diving in magma: how to explore controversies with actor-network theory. *Public Understanding of Science (Bristol, England)*, *19*(3), 258–273. https://doi.org/10.1177/0963662509102694Vikkelsø, S. (2007). Description as Intervention: Engagement and Resistance in Actor-Network Analyses. *Science as Culture,* 16(3), 297–309. https://doi.org/10.1080/09505430701568701

Wiltsher, N. (2016). The Aesthetics of Electronic Dance Music, Part I: History, Genre, Scenes, Identity, Blackness. *Philosophy Compass*, 11(8), 415– 425. https://doi.org/10.1111/phc3.12333

Weltevrede, E., & Borra, E. (2016). Platform affordances and data practices: The value of dispute on Wikipedia. Big Data & Society. https://doi.org/10.1177/2053951716653418

WORKSHEETS:
Exploring Electronic Music: Developing a Data Practice Thing

Alexander Siegfried, Alfred Felumb & Peter Munksgaard

TAN8, Group 9, 2021

11164 words

INDEX

Preface

The compilation of worksheets is to accompany the admission of our article. It should not necessarily be read chronologically, but instead it is to function as a reference work elaborating on the foreshadowing and literary foundation as well as the general positioning and analytical takeaways.

Problem Area

Side project: Classification disputes about electronic music on Wikipedia

The following is an excerpt from the project we did about “Genre Controversies” on Wikipedia, as part of the Mapping Controversies course we attended. While the subject matter differs from the actual semester project (Controversy Mapping, not design) - it still serves as both our first venture into the field of electronic music, and as the basis for our conceptions of some of the field's problems. See appendix A for further elaboration and the Controversy Atlas we made as part of said project.

> Our inspiration for this [side] project is partly based on our own interest in electronic music, as well as an academic fascination with genre *classification disputes*. As Latour states, controversy mapping is in many ways the ANT scholars' wet dream (APA). Electronic music being in many ways digitally native also helps with the snugness of this fit. However, on a more practical level, if we view the act of determining the genre for a piece of music or album, as a sort of ‘emic ordering process’, a process of associating and dissociating your taste or style from other from other genres and musicians and their association - then Genre Controversies is in many ways an ideal ANT case. As such this [side] project is a matter of figuring out how we can use wikipedia to study genre controversies within electronic music (in media specific ways!), not make sweeping statements about the nature of genres or controversy.
>
> So what do we mean by ‘Genre Controversies’ and how have we gone about investigating them?

By Genre Controversy we mean situations where two or more parties disagree about what constitutes a genre - or more precisely, how to demarcate between genres and their hierarchical order. At first glance, this might seem like an insurmountable task. Luckily, Wikipedia is very helpful in this regard. Wikipedia not only has pages dedicated to all manner of genres and music styles, it also has several ways of categorizing and ordering them. The first step was thus to find a proper starting point for studying our Genre Controversies. At that, the article "List of electronic music genres" proved to be the ideal candidate (*not* the category "Electronic music genres") as it both serves as a comprehensive list of genre articles on wikipedia, but also as the starting point of our investigation. The next step was to figure out how to study these genre pages. Here the talk pages of the genre articles proved to be the important part. As each talking page (rightly so, according wikipedia's own guidelines) is used as a way to avoid 'edit wars', and instead politely discuss disagreements, with the goal of finding a middle ground. In many ways then, talk pages seem an ideal place to study wikipedia controversies.

The first talk page to investigate was obviously from the "List" article itself - which indeed proved to be a variable treasure trove of discussions. Interestingly, these (more or less) all seemed to spark from a single editor claiming to be an authority on electronic music - which thus resulted in a number of discussions about genre notability (what genres qualify to be included on the list). To no surprise this reflects Wikipedia's nature, as communities of editors, no lone wolves writing entire articles by themselves. The notability discussions, however, also highlight another aspect of genre controversies: the lack of stable authorities. Rarely can one single reference be counted on to 'prove' a genre's existence or qualities. Instead users rely on everything from personal experience, arguments about music culture or geographical origins, to the genre's BPM range and the use of specific synthesizers, instruments or samples. Interestingly, the 'closure mechanism' (APA) that seemed to repeat itself in the case of notability discussions, was the *media specific* argument of whether a genre had its own wikipedia article or not. As a result, all listed genres now *have* to have another genre article hyperlinked ("bluelinks") - and redlinks *have* to have links to genre articles in other languages.

The next step was as such 'just' a matter of putting two and two together. We knew genre pages were the place to look and that the "List of electronic music genres" provided a (relatively) exhaustive list of genre pages. We thus mapped *all* the Wikipedia pages of *all* the genre articles, in the form of two networks based on the text contents' computed LDA Topics and related words or articles. In conjunction, the two networks provided a relatively comprehensive map of genre controversiality within electronic music on Wikipedia. Firstly, by comparing topics and their related words, the map showed a pattern suggesting that some genres (primarily House and Electronic Rock) might be more controversial than others. Secondly, by looking for words indicating discussions (e.g. "genre", "fact", or even "discussion") and talk page activity (i.e. timestamp indicators such as "utc"); what topics they linked most closely to - and in turn, what articles related to that topic. As such we got an indication of what genre articles might be more controversial than others. While outside the scope of this [side] project to investigate them all, a quick glance at a 'controversial' candidate ("Deep house"), seemed to confirm the suspicion. The talkpage contained arguments about (for example) whether Deep house is an overcategorization of other subgenres (e.g. "Detroit techno" or "Chicago house"); what the proper BPM range is; if another music platform (www.beatport.com) caused this overcategorization; and so forth. No of which seemed likely to be resolved in a finite sense anytime soon.

If anything, this seems to confirm an initial concern we had: that genre controversies tend to be perpetual. That the scarcity and instability of closure mechanisms, the personal nature of music taste and the ever evolving nature of electronic music seem to be fertile grounds for controversy. While this might be disheartening in the sense of conflict resolution, it would also seem to open an ideal space for design - luckily something we have the rest of our semester project to resolve.

Review of: Wiltsher, N. (2016). The Aesthetics of Electronic Dance Music, Part I: History, Genre, Scenes, Identity, Blackness. *Philosophy Compass*, 11(8), 415–425. https://doi.org/10.1111/phc3.12333

Authorship
Nick Wiltsher is associate professor at Uppsala University in Sweden. As a philosopher by education, Wilsther specializes in Aesthetics and philosophy of art and philosophy of mind. Within this field, Wilsther tackles questions of race and gender, phenomenology.

Overview
This article is part one of an exploration of the facets of electronic dance music. It seeks to address the problem of how to differentiate real or authentic dance music from mainstream or commercial clubbing. Specifically it discusses whether it is a matter of history and genre-definition; a matter of defining scenes or subcultures; or a matter of blackness. In doing so, main genres of dance music are characterized and are compared with philosophers' accounts of genre, sociologists' ways of thinking about cultures and links it to critiques of whitewashing of electronic dance music. This way, the article can be seen as a philosopher's experiment to weigh in on the debate of blackness within electronic dance music. In this review, focus will primarily be on the discussion about how genres can be defined in various ways, as our interests do not include how identity can be tied into music nor identity politics.

Authenticity and Definition
At the onset of the article, Wilsther makes a brief problematization of defining genre distinctions. Doing so, it's suggested that authenticity is used as the main descriptive move. Firstly, because it has the benefit of both engaging inquiry of how dance music can be characterized. And also that it brings into question how authenticity is discussed in literature.

History and Genre
Wiltsher describes how electronic dance music has a long and quite thickly described history. Specifically how there in a canonical view exists four genres archetypes (house, techno, hardcore and garage), that are all defined by the specific timeperiods, places and musical features. Furthermore, each archetypical genre proliferates finically distinguished sub-genres. In light of

this, Wiltsher narrates a scholarly debate about the ‘hardcore continuum’, which posits that it’s simply true that new genres evolve from existing ones. In this sense, authenticity can be used effectively by tracing lineages of genres. It’s posited that ‘genetic’ relations indicate authenticity by connection to subcultures. Finally, he introduces the Aestheticians approach in which genres can be described as the “..sufficient accumulation of standard features is enough to classify a work as belonging to a certain genre.”(p. 417). To this, Wiltsher argues the conundrum that: genres define works, and works define genres. In this light, Wiltsher posits that it might be fruitful for Aestheticians to also investigate how genres constitute works in their features.

Takeaways

The main takeaway from this article is that it’s complicated business to discuss genre distinctions. This can be attributed to how well developed the history of the field is or the scholarly discussion about genre demarcation. It might be very interesting to bring a Visual Network Analysis perspective on this issue, yet it’s not the order of business in this project. We seek to develop a tool for exploration rather than definition. Therefore, we will not pursue the particularities of genre distinctions beyond that they exist in the databases we’re working with.

Theoretical Sensibilities

In the sections below, we discuss the participatory character of our design approach and our general positioning from within the field of STS-studies. The literature is 'merely tools' to investigate the state of affairs and convey or close in on an "uniquely adequate account of a given situation" - as Latour puts it (2004). Literature reviews of selected articles can be found at the end of this document.

Participation and Visual Network Analysis?

Participatory Design (PD) is in its essence all about arrangíng situations or devising methods that allow stakeholders and "issue experts'' to take part in the decision-making and design early on; to unfold their problems or concerns towards the subject; and to generate insights and build *Things* together. These activities involve an enactment, a mediation, and an adaptation. Basically, the goal is to involve both the users "existing skills" and "tacit knowledge" in the design of new artifacts. To do this we are inspired by the notion of infrastructuring (Björgvinsson et al., 2012) and strive towards supporting an environment that enables a process of 'design-after-design. In our view, a process of infrastructuring is complemented by a process of *interessement*. We want to form alliances and mobilize interests from relevant stakeholders. Enrolling such socio-material alliances is a matter of trial and error entailing many open negotiations and adaptations. It is a process filled with accusations and resistance that stabilizes the design through a continual re-shaping of what *it could* and *could not* become (Akrich et al., 2002a; Akrich et al., 2002a; Akrich et al., 2002b). We want issue experts as allies in our data practice, which makes the literature on Participatory Data Design (PDD) an inspiration as to what activities that can be performed in data practices with participants. Even though we *do not* have prosperous circumstances of doing data sprints because of some technical and physical circumstances, we can still build our approach in reflection to the literary work. It is very akin to the process of PD, but with the different focus also of generating insights, while developing useful tools or visualizations based on the collectively selected data (Jensen et al., Accepted/In press). The use of "low-fidelity mock-ups" seems very similar to the use of "design Things", which in both cases are described as 'Boundary Objects'. Following their example we can attempt to 'open the blackbox' by incorporating a large degree of flexibility in the co-creation of multiple

perspectives (Jensen et al., Accepted/In press). All this while ensuring an understanding of the technical choices (being) made and by explaining the formatting (Munk, 2019). Especially when applying “Single Level Analysis”, we should also be attentive towards our account of the interconnectivity embedded in the data set. It is apparent to think critically and reflectively about the appropriation of digital entities by qualitatively “situating” quantitative patterns (Munk, 2019). The concept of “monads” also directs our attention towards the interconnectivity, once again. The “ways” elements are connected and could be associated differently in terms of the aggregates and the starting point or anchor. We are therefore encouraged to circulate “differently conceived wholes” as partial totalities that make trajectories emerge - as associated “successive attributes” (Latour et al., 2012). Because data visualizations can be persuasive; we need to guard ourselves (and participants) against overinterpretation, while questioning what type of “materialization” we want to perform (Jensen et al., Accepted/In press). After all, our design process has the ambition of being grounded in actual *work practices* or *use cases*, so the design has to make sense in *their* world.

ANT-sensibilities

From Latour we learn that it makes no sense to detect and determine content and pertaining contexts. ‘Things’ are hybrid characters that can be investigated through retrieving the delegated script and by performing ‘de-scriptions’. Do not, he posits, confuse a description of a network with the act of de-scripting (Latour, 2004). The latter is concerned with what translations and transformations that are present. When de-scripting (as Madeleine Akrich would put it) or deconstructing (by Bruno Latour) one should look for associations and substitutions of the ‘action program’ - attuning to “what” also constitutes the artifact and “what” diminishes or destabilizes the ‘Thing’. When following the flow of action, complexity furthermore arises when attending to the anti-programs (Latour, 1990; 1992). Objects are as he puts it “a bit more complicated, folded, multiple, complex, and entangled” than what the ‘objectivist’ would have them be. The “object” makes no sense in itself, but becomes meaningful from the gathering surrounding it (Latour, 2004).

It is in many ways, as Law and Lin (2020) describe it, a matter of being sensible about the “otherness”, uncertainty and iterative nature of the design process. Our Thing is not forming a coherent ‘whole’, which makes it a

meaningful act to embrace resistance and alternative ways the course of action could unfold. We are urged to be care-ful when we are to present our insights, as translations of empirically situated concerns. The findings most likely contain “multiple realities” and a lot of “non-coherence”, so a sensitivity is posed to this “otherness” (Law & Lin, 2020).

Our approach could, with the words of especially Michel Callon and John Law, be called “heterogeneous engineering” as a dependency on both non-human and human actors interacting in “juxtapositions”, that are transformed into a network of our “system building” (Law, 1987). We have established alliances with various issue experts and important actors from the field - ranging from a record store owner, the online record store and discography Discogs, our own computers, and note the least, the network visualization tool Gephi. As an approach with much resemblance to low-fidelity prototyping it is important for the design object to be easily and quickly manipulated in order to explore its various design possibilities together - as such, our visualizations have to be *flexible* in their construction, but still express our initial idea meaningfully. It bears many similarities to the framework of Ehn, Hillgren & Björgvinsson (2012) with their “design-after-design”, as a chain of potential boundary objects allowing participants to set up their own “infrastructuring” process that supports future designs that are applicable in use-situations. Both conceptualizations are concerned with heterogeneous matters of concern interacting through a process of building relationships. Both the design and the use of the artifact is a product of the network of alliances surrounding it.

Technical Protocol (Programming and Gephi)

Overview

1. Get data
    a. Query Discogs API via Postman, for all electronic music 12" vinyls, year by year.
2. Format and filter data
    a. Index and filter data to contain necessary metadata (title, label, styles, year, country, want/have community metrics and uri) with Alteryx
    b. Create node list file with metadata, and two cleaned adjacency list files (for artists and labels respectively) via custom python script
3. Create network visualizations
    a. Project adjacency list files via NetworkX and *bipartite projection*
    b. Combine with node list file to create final network file
    c. Spatialize with ForceAtlas2, isolate Giant Connected Component and colour nodes by modularity class.

Data Collection via Postman

The goal of using Postman, was to both make the process of obtaining authentication from Discogs and the making of API the queries, easier. Through cross referencing with Discogs' API's documentation and the site itselfs search function, we found that we had to refine our search queries to work around Discogs performance limitation and their pagination system. In our case we found the necessary specification to be: 12" Master releases on vinyl. Furthermore, to not get more than 100 pages per Search API call, we had to do each year's query separately.

1) Create an updating variable for the page number; to allow the query to call a different page of the pagination each time run.

```
pm.environment.set("page_number", 1 +
parseInt(pm.environment.get("page_number")));
```

2) Write the output of each query to the variable "r" in a JSON format; to allow collection of the data afterwards.

```
let responses = pm.collectionVariables.has('r') ?
JSON.parse(pm.collectionVariables.get('r')) : [];

console.log(responses);
```

```
responses.push(pm.response.json());

pm.collectionVariables.set('r', JSON.stringify(responses));
```

	VARIABLE	INITIAL VALUE	CURRENT VALUE	...	Persist All	Reset All
☑	r		[{"pagination":{"page":1,"pages":84,"per_page":100,"items":8388,"urls":{"last"...			

3) Run the “Search” API query through the postman interface with the following settings (variables marked in “{{}}” and defined in the “collection variables” settings):

GET {{url}}/database/search?per_page=100&type=master&page={{page_number}}&genre=electronic&year={{year}} **Send**

Params ● Authorization Headers (9) Body Pre-request Script ● Tests ● Settings ● Cookies

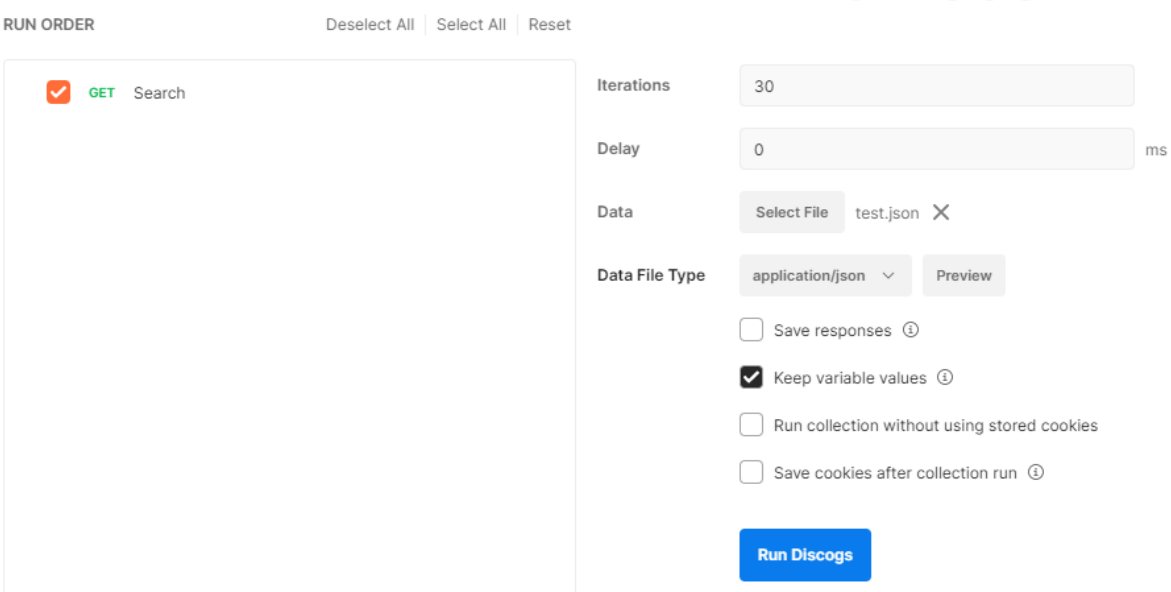
Query Params

	KEY	VALUE	DESCRIPTION	...	Bulk Edit
☑	per_page	100			
☑	type	master			
☑	page	{{page_number}}			
☑	genre	electronic			
☑	year	{{year}}			

4) Define the necessary *Collection Variables*, which specifies which API to call, the authorization required to run queries and by creating our own application; and set other variables (*page number* and *year*) across all queries.

	VARIABLE	INITIAL VALUE	CURRENT VALUE
☑	url	https://api.discogs.com	https://api.discogs.com
☑	user_agent	PostmanDiscogs/1.0	PostmanDiscogs/1.0
☑	consumer_key		OJpqHVzrTditFkLWwNJ
☑	consumer_secret		iRnCkDkUDeKAdfKUCRI
☑	oauth_token		dedehkCQeLGdTUqyoxI
☑	oauth_token_secret		HtyxzdueZuUPhwliLwxy
☑	oauth_verifier		gGlfpSLBZM
☑	username		
☑	page_number	0	30
☑	year	0	1970

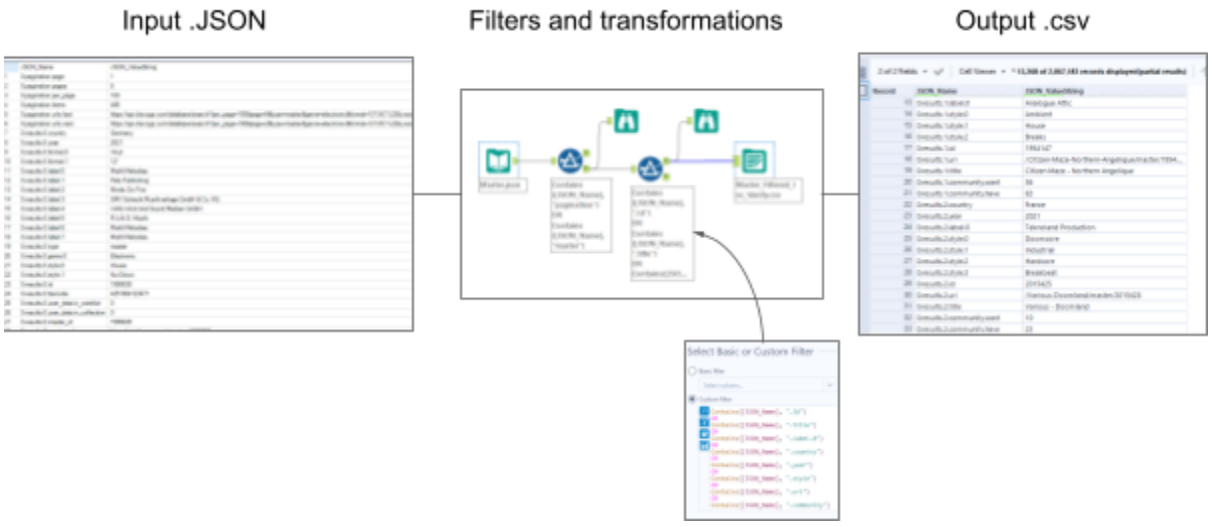
5) Use the *Runner* function, which allows for running one or more API queries multiple times. In this case, 30 runs (that is 30 pages of a 100 releases each) was the maximum size feasible, due to that being the maximum amount of data capable of being stored in a variable.



6) Finally, you get each query results in a JSON format data file containing information on each master release. Each file also contains information on the number releases as distributed via *pagination*. Apart from the *artist* and *title* on the same line (more about why that proved troublesome later) a query result also contains useful metadata about a release's attributes; such as *country, year, format, label* and *style*. Additionally, the result JSON also contains other useful information, such as a release's unique *master id*, the *uri* to its entry on Discogs or even community data such as how many people *want* or *have* the release in question.
7) After each *collection run*, each *result* JSON (from "r") can then manually be strung together to one big Master file, containing all Electronic Music 12'' vinyl master releases from 1970 to 2021 .

Data Filtration via Alteryx

The goal of using Altery was two-fold. First, we could filter through the information in the data file, and select only what is needed in the output CSV file. Secondly, Alteryx also allowed us to index all the releases and their metadata separately - to make the process of creating the network files easier. In retrospect, however, using Alteryx was in many ways a remnant of our efforts to figure out how to extract data from Discogs data dumps. As such, having to do the this process again, it would probably have been easier to either do the entire data processing in Alteryx, *or* in Python - not a combination of both.



Data Processing in Python

After having obtained the filtered and indexed files from Alteryx, the goal of using Python was to create the node- and edge list files to put into Gephi. In all actuality the easiest way of getting the data into Gephi, turned out to be creating a master node file that contained all the meta data, as well as the unique release-id from Discogs itself. Then, creating two seperate *adjacency list* files, one for label and one for artists, that only contained the names of their artists and label, as well as the unique id of the release they connect to. The final file, with both edges and nodes, could then be made by combining all three files.

0) Importing necessary libraries

```
#import libraries
import pandas as pd
import re
```

1) creating the *nodelist* file and creating the proper headers (*title, artist, recordlabels,* etc.).

```
#df for DataFrame
df = pd.read_csv('Master_Filtered.csv')

num_rows = df.shape[0] #number of rows

#create masternode file
node_file = open("masternodeslist_split_no_regex.csv","wb")

#creates header for masternode file
node_file.write("Label; title; artists; recordlabels; year; country; styles; url; want;
have\n".encode("utf-8"))
```

2) Creating the data frame, corresponding to the correct attributes from the header, and explicitly formatting to *utf-8*, to avoid encoding errors when Gephi has to read the file.

```
def write(id,a):
    try:
        # write to masternode file
        node_file.write("{}; {}; {}; {}; {}; {}; {}; {}; {};
{}\n".format(a[9],a[0],a[1],a[2],a[3],a[4],a[5],a[6],a[7],a[8]).encode("utf-8"))

    except:
        print("error: release skipped")
```

3) Using the indexes from the filtered file to separate information from one release to another, and to write/clear a new list for each release.

```
size = 10
res_data = [None] * size
res_id   = 0
cur_id   = 0

for i in range(num_rows):
    x = df.iloc[i,0].split('.') #split result number info

    # set id for current release
    cur_id = int(x[2]) + int(x[0])*100 #corresponds to centi / deca

    # check if new release
    if(cur_id != res_id):
        #save res_data to file
```

```
        write(res_id,res_data)
        res_id = cur_id
        res_data = [None] * size
```

4) Putting information from the appropriate attribute in the data, into the correct position in the array.

```
    # replace / remove symbols that breaks the .csv format
    val = val.replace("\"","")
    val = val.replace("'","")
    val = val.replace(",","/")
    val = val.replace(";","/")
```

(And remembering to remove or replace symbols that would mess with the output CSV format)

```
    # determine attributes (title, year etc.)
    attr = x[3]
    val  = df.iloc[i,1]
  # do something depending on which attribute
    if attr == "title":
        title_val = val.split(" - ")
        # release title
        res_data[0] = title_val[1].strip()
        res_data[1] = title_val[0].strip()
    if attr == "label":
        res_data[2] = val
    if attr == "year":
        res_data[3] = val
    if attr == "country":
        res_data[4] = val
    if attr == "style":
        if res_data[5] is not None:
            res_data[5] =  res_data[5] + ", " + val
        else:
            res_data[5] = val
    if attr == "uri":
        res_data[6] = str('https://www.discogs.com'+val)
    if attr == "community":
        if x[4] == "want":
            res_data[7] = val
        else:
            res_data[8] = val
```

```
    if attr == "id":
        res_data[9] =val
```

5) For creating the *adjacency list* needed to make edges for the network, it is the same procedure as with the *node list* file - except this time we are creating three separate files, each with the *id* of a release, and a specific attribute (*label, artist* or *style*). In retrospect, we ended up using the *style* edge files more as a curiosity, than a part of the final network.

```
#create label edge file
edge_file_label = open("adjacencylist_label_only.csv","wb")

#create artist file
edge_file_artist = open("adjacencylist_artist_only.csv","wb")

#create style file
edge_file_style = open("adjacencylist_style_only.csv","wb")def write(id,a):
    try:
        # write to edge files
        edge_file_label.write("{}; {}\n".format(a[9],a[2]).encode("utf-8"))
        edge_file_artist.write("{}; {}\n".format(a[9],a[1]).encode("utf-8"))
        edge_file_style.write("{}; {}\n".format(a[9],a[5]).encode("utf-8"))

    except:
        print("error: release skipped")
```

6) The script for creating the adjacency lists also has two important additions. To get each artist as a separate entity in the network, a long list of *regular expressions* were needed to properly separate one artist from another. This list was obtained experimentally, by looking through a multitude of different releases on Discogs.

```
# artist(s) by splitting at regex
        artist_regex =
'(?i)\s+&\s+|\s+and\s+|\s*,\s*|\s+vs.?\s+|\s*/\s*|\s*\+\s*|\s+a.k.a.\s+|\s+ft.?\s+|\s+featuring\s+|\s
+presents\s+|\s+w/\s+|\s+feat.?\s+|\s+x\s+|\s*:\s*'
        artists = re.split(artist_regex,title_val[0])
```

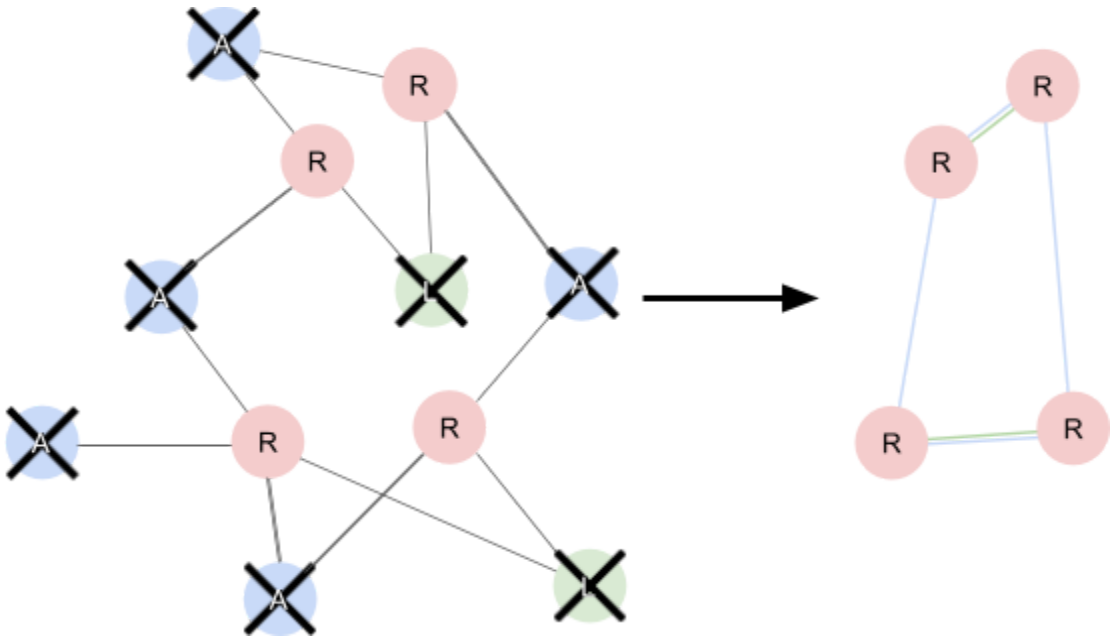
Additionally, each attribute needed for the three different adjacency list files as well as the title of the release, had to have a have a prefix put in (e.g. “title:” before the release title) in order to be able to separate different types of nodes in the final network.

```
res_data[0] = "title:" + title_val[1].strip()

...
```

```
res_data[1] = "artist:"+"; artist:".join(["{}"]*len(artists)).format(*artists)

if attr == "label":
      res_data[2] = "label:" + val
...

if attr == "style":
      if res_data[5] is not None:
         res_data[5] =  res_data[5] + "; style:" + val
      else:
         res_data[5] = "style:" + val
```

Network Projection with NetworkX

To produce a more “readable” network, we opted to “project” the label and artist nodes of our network, onto the release nodes. To achieve network projection, we needed a bi-partite network. Luckily, our network was already partitioned after type, due to the prefixes we put in.



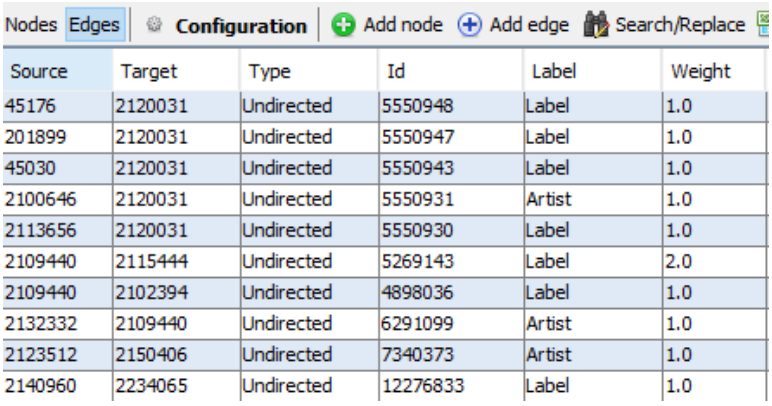
This script uses NetworkX to do network projection on the networks created via the *adjacency lists* only. One set of nodes is projected onto another set of nodes (called *top*- and *bottom-nodes* or *0* and *1* in the script), by referring to the “bipartite” attribute of the network. This attribute was made manually in Gephi, by making a new column and calling all nodes with the “title:” prefix (that is releases) 0, and the other type of node (in the example below: styles)

1.

```
import networkx as nx
from networkx.algorithms import bipartite

B = nx.read_gexf("Style_Release_final.gexf").to_undirected()

top_nodes = set(n for n,d in B.nodes(data=True) if d['bipartite']==0)
# releases
bottom_nodes = set(B) - top_nodes
# label/style/artist etc.

G = bipartite.weighted_projected_graph(B, top_nodes)
#project "1" onto "0", weighted


nx.write_gexf(G, "bipartite_style.gexf")
```

Creating the final network visualization in Gephi

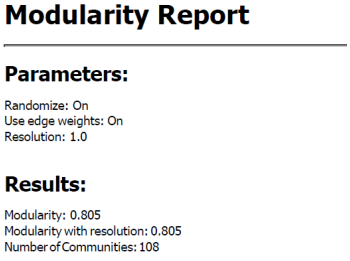
Creating the final network file was then a matter of combining the two projected networks created from the adjacency list files with the master node list file containing all the metadata. Before this could be done however, we also went back to the projected network files, and labeled the edges themselves, so that when all the files where combined, it would be possible to view whether an edge came from an artist o a label,



Nodes | Edges | Configuration | Add node | Add edge | Search/Replace

Source	Target	Type	Id	Label	Weight
45176	2120031	Undirected	5550948	Label	1.0
201899	2120031	Undirected	5550947	Label	1.0
45030	2120031	Undirected	5550943	Label	1.0
2100646	2120031	Undirected	5550931	Artist	1.0
2113656	2120031	Undirected	5550930	Label	1.0
2109440	2115444	Undirected	5269143	Label	2.0
2109440	2102394	Undirected	4898036	Label	1.0
2132332	2109440	Undirected	6291099	Artist	1.0
2123512	2150406	Undirected	7340373	Artist	1.0
2140960	2234065	Undirected	12276833	Label	1.0

Another addition to the file we wanted to make, was to find the difference between how many people wanted a certain release and how many of that release that was actually for sale (via the Column Calculator plugin). While it didn't change much in terms of ranking, it would,



Modularity Report

Parameters:

Randomize: On
Use edge weights: On
Resolution: 1.0

Results:

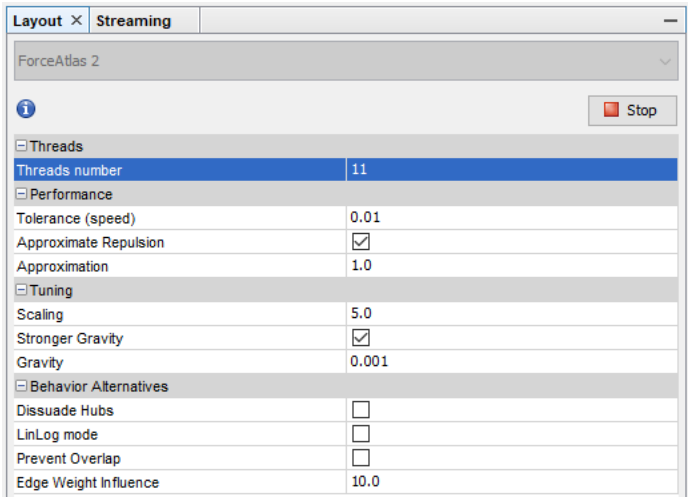
Modularity: 0.805
Modularity with resolution: 0.805
Number of Communities: 108

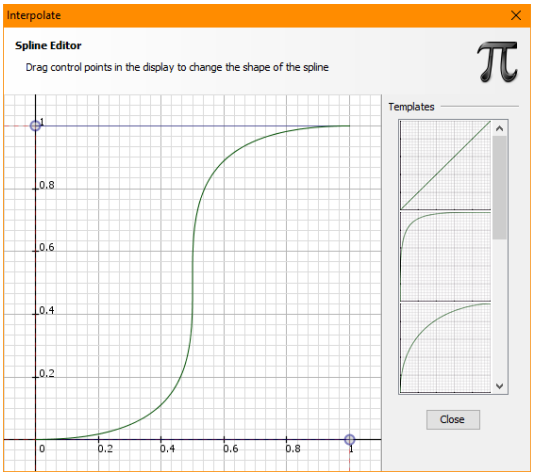
however, seem a better metric for how *coveted* a release is - compared to just looking at how many people want it.
Two final adjustments were done before subjection the network to the ForceAtlas2 algorithm: First, running the “Modularity Class” algorithm to find clusters in our network based on structure, rather than layout. Secondly, we filtered the network down to only the “Giant Connected Component” to avoid “satellites” in the network.
In the end, the final network file looks like this in the data laboratory:

Nodes | Edges | Configuration | Add node | Add edge | Search/Replace | Import Spreadsheet | Export table | More actions

Id	Label	...	title	artists	recordlabels	year	country	styles	want	have	Popularity Metric	url	Modularity Class
2211510	29410		Music Sound...	Stardust	Roulé	1998	France	House	33257	23682	9575.0	https://www.discogs.com/Star...	88
193276	80049		The Age Of ...	Age Of L...	DiKi Records	1990	Belgium	Trance, Tec...	29052	16527	12525.0	https://www.discogs.com/Age-...	28
190099	94963		Gypsy Wom...	Crystal W...	Mercury	1991	US	House, Gar...	22529	13062	9467.0	https://www.discogs.com/Crys...	74
208350	49387		Spacer Woman	Charlie	Mr. Disc Or...	1983	Italy	Italo-Disco, ...	16504	5428	11076.0	https://www.discogs.com/Char...	41
201107	5108		Acid Tracks	Phuture	Trax Records	1987	US	Acid House,...	16117	4907	11210.0	https://www.discogs.com/Phut...	75
2132619	350827		Big Fun	Innercity	KMS	1988	US	House, Tec...	16106	14147	1959.0	https://www.discogs.com/Inne...	90
189192	79590		Go	Moby	Outer Rhythm	1991	UK	Trance, Tec...	16095	12385	3710.0	https://www.discogs.com/Mob...	4
190138	6424		Papua New ...	The Futur...	Jumpin & P...	1991	UK	Breakbeat, ...	15973	14120	1853.0	https://www.discogs.com/The-...	27
2130405	5568		The Promise...	Joe Smoo...	D.J. Intern...	1987	US	House	15775	7983	7792.0	https://www.discogs.com/Joe-...	17
2208811	401133		Deep Inside	Hardrive	Strictly Rhy...	1993	US	House, Dee...	15253	6568	8685.0	https://www.discogs.com/Hard...	0
173804	100665		The Bomb! (...	Kenny Do...	Henry Stre...	1994	US	House	14662	11610	3052.0	https://www.discogs.com/Ken...	90
210379	134087		Keep The Fir...	Gwen Mc...	Atlantic	1982	US	Soul, Disco,...	14239	2439	11800.0	https://www.discogs.com/Gwe...	21

In terms of spatializing the network, we tried to promote the formation of clusters as much as possible. As such, we increased the edge weight from 1 to 10, set the scale relatively low (compared to the size of the network) and specifically turned off “Prevent overlap”.
Similarly, we set size by our “Popularity Metric”, but with a range from 0.1 to 100 and a spline setting to promote making the big nodes visible, and the rest small enough to still clusters





Workshop Preparation

In preparation for the workshop, we have allied ourselves with Poppy's laptop, the software Gephi and a vinyl shop owner named Anders to provide an empirical grounding of our efforts in *making sense* (APA) of our network visualizations. To accommodate the goal of producing tangible design outcomes, we've made a workshop plan. In doing so, we take inspiration from reviewed literature (Law & Lin, 2020; Munk, 2019; Latour et al., 2012; Jensen et al., Accepted/In press; Björgvinsson et al.,2012). From this, three phases have been devised: present, explore, and discuss design possibilities. Each step is intended to both structure the way in which we interact with the visual networks in Gephi as well as maintaining a reasonable timeframe.

Present

At first we present the visual networks in Gephi, explaining what the data consists of (releases, artists, and labels) and where it has been taken from (Discogs). Further, to explain some of the choices we've made regarding selecting data (master releases of electronic music on 12" vinyl) - showing how it can be viewed in the data laboratory of Gephi. Additionally, we ensure to make it clear that the relational structure the visual networks have, don't carry any ontological meaning - in terms of that it's a product of the properties of the data rather than a representation of phenomena 'out there'. Finally, we make it clear what our end goals are: to gain designerly insights through the visual networks acting as "low-fidelity mock-ups" (Jensen et al., Accepted/In press) as a step towards conceptualizing a navigational tool.

Explore

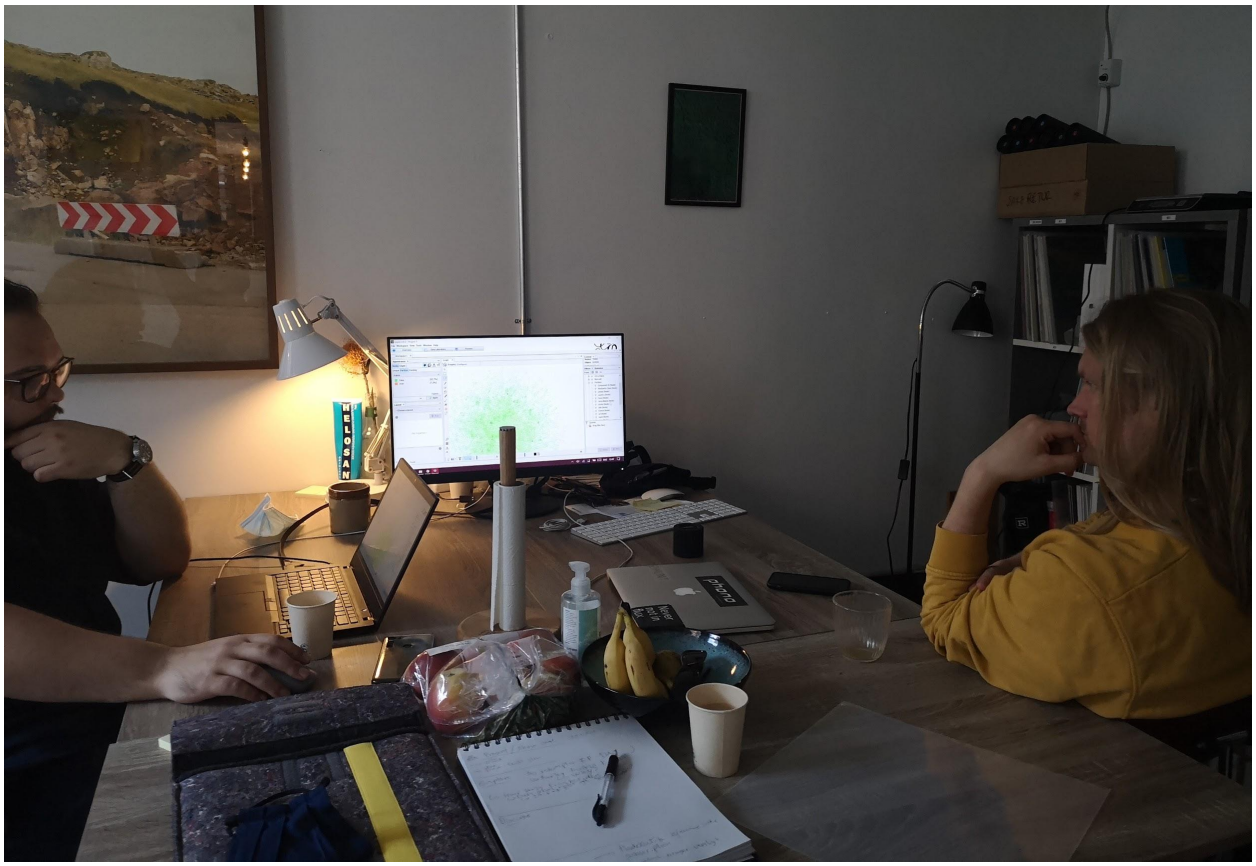
In this phase we strive towards facilitating an explorative and inductive mindset with our participant. A mindset that leaves pre-existing assumptions behind and allows for diving into the data set, by following the idea of "circulating back and forth" (Latour et al., 2012) between different clusters and specific data in the data laboratory in Gephi. With this, we strive towards an approach of "design-by-doing" and "design-by-playing" as intertwined language-games which fosters playful learning (Björgvinsson et al., 2012). Essentially, we want to find music releases, which our participant knows and has a relation to - to then trace how it is connected in the network. In a sense this can be seen as a chase after 'aha-moments'. The basic idea is to get mild confirmation of the idea that the visual network can mediate how music

releases are structured in a new way - different from browsing the Discogs website itself at least. In a sense we hope to make a *multiplicity* more apparent. In doing so our role is to facilitate the iterative processes of visualizing, filtering and labelling - collectively leading up to a discussion of the “design object” (Björgvinsson et al. 2012) or “data practice” (Jensen et al., Accepted/In press) at hand.

Discuss design possibilities

In this phase, we aim to facilitate discussion about design possibilities in the visual network. We’ve decided to directly incorporate discussion because we’ve found that the visual networks have had persuasive qualities that have *urged us* to describe the contents of them. Therefore we would like for the contents of the discussion to become a matter of evaluating what parts of the network data that can say interesting things in relation to the matters of concern of our participant as well as discussing possible use cases and requirements. As leading questions for framing this discussion, we take inspiration from anti-programs (Latour 1990; 1992) - in which we make the rhetorical question of: what would you need to make this ‘Thing’ work if we were to remove X from the situation?

Workshop Evaluation



The above is a picture of the workshop situation at the headquarters of Proton Rec. On the left is Alfred showcasing the networks in Gephi. On the right is Anders, the co-owner of the recordstore.

Abridged: Technical issues meant that it went rather poorly. Not a complete catastrophe, but not the greatest success either.

Elaborated

Gephi ran slowly on the laptop and was crashing all the time, which was a major hindrance for both the process, but also the laissez faire atmosphere, which Anders was fully onboard upon. This meant that the network graphs didn't go much further than being an elicitation device yielding good discussion about genres and the field, but not as much tinkering with the technical parts of it. Eventually, when we made Gephi show the type of visualizations we wanted to pursue, Anders did seem to have lost interest a bit. Further, we found a couple of points of impact in the networks, which could be interesting examples (and is to be found in the github repository). At this point 1:30 hours had passed and everyone was exasperated by all the technical issues - to an extent that the process didn't get fully leveraged. In spite of this, we got

a couple of interesting take-aways which can be seen as design outcomes (and which can outline a future prototype).

Takeaways

If the networks are to work in their explorative function, they have to be user friendly in terms of both speed, how straightforward they are technically speaking - but also that they should not be overwhelmingly large or customizable (as the network is its current state). Rather customization of parameters should be more readily visible and available - not messing about in the data laboratory of Gephi.

The primary design idea: To make an extension for the search engine on discogs, which allows for the use of many of the same metrics we have (genres, styles, year, country, wants / haves and so on). In contrast, the point is to explore related music in a way that allows for 'exploration' more so than using the existing categories and metrics by themselves. This way it's a matter of narrowing the 'search area' by means of a search engine, but to then make available things like the structure or topology, clusters, modularity classes and so on. A variation of this idea is to integrate the music catalogues of music of sellers on discogs for either exploration of it or to explore related music based upon it.

Concrete proposal for Anders: A tool to find new music for the Proton store, by exploring current 'hot' releases for reprint and find related music to them. Alternative: a tool for customers to explore their music by themselves - possibly of parts of the stock that is on remote storage.

Bonus: their(Proton) game is to find interesting and desirable music for reprint beyond selling local and secondhand music. The challenge is that it's manual labour to 'crate-dig' for records on different databases and websites based on their knowledgeability and expertise. The flipside is that there is no guarantee that their customers will buy the records they find - even though it might be hyped abroad - this way, their job is also to frame the music in interesting narratives to persuade customers to buy it.

Idea for a format: Sigma.js webpage, in which search terms and further parameters can be set up and with the addition of the ability of the nodes themselves to be directly hyperlinked to certain stores or just discogs itself for listening and a purchase.

Field notes

1) We cannot circumvent genres, since artists group themselves in ways that make the variety of music very large - even in small or close knit ‘network communities’. This ties into the notion that the same artists might have many different musical ambitions at the same time. In a sense this is reflected in the networks we have produced - genres seem evenly distributed in the network. It’s possible to ‘zoom’ in at an arbitrary part of the network and it would look similar.
2) Having all the music at the ready seems to lead back to the all time classics rather than branching out into the more peculiar or unknown releases. This way we discussed the opportunity to be able to look ‘wide’ rather than ‘deep’, with regard to year filters (for example) in ego networks.

Literature reviews

This section contains literature of the majority of STS-literature used in the project. It's ordered in such a way, that we firstly present our condensed takeaways from each literary source to then have the more elaborate review come in succession below.

The 'Takeaway Bundle':

The condensed interpretations of selected articles presented below affects our choices and participatory efforts as described in the previous chapter of our general approach. We have placed one review of (Whilsther, 2016) that sheds light on the research topic of electronic music before unfolding our side project on Wikipedia.

(Björgvinsson et al., 2012)

To support an environment that enables a participant's (issue expert's) existing skills and tacit knowledge to become resources of the *Thinging*.

- It is an infrastructuring of the socio-material assembly with an attention towards the process in itself as a "design-after-design".
- Be aware of the three layers of activities that in with overarching words can be described as an enactment, a mediation, and an adaptation.

(Akrich et al. 2002)

To (openly) negotiate alliances, to adapt, and mobilize interests.

- To enroll socio-material alliances through trial and error often involve accusations and resistance that stabilizes the design through a continual re-shaping of what *it could* and *could not* become.

(Jensen et al., Accepted/In press)

PDD is inspirational in regards to what activities that can be performed in data practices with participants.

- We *do not* have prosperous circumstances of doing data sprints because of technical and physical circumstances.
- "Open the blackbox", incorporate a large degree of flexibility, and guard against overinterpretation in the co-creation of multiple perspectives.

- Make us question what type of “materialization” we want to perform?

(Munk, 2019)

When applying Single Layer Analysis, we should be attentive towards our account of the “interconnectivity” embedded in the data set and the possible quali-quantitative methods of analysis that can be done.

- Think critically and reflectively about the appropriation of digital entities by qualitatively “situating” quantitative patterns.
- Ensure understanding of technical choices and explain the formatting as an “open blackbox”.

(Latour et al., 2012)

The concept of “monads” directs our attention towards the interconnectivity, once again, and “ways” elements are connected and could be associated differently in terms of aggregates and “starting point”.

- We are to circulate “differently conceived wholes” as partial totalities that make trajectories emerge - as associated “successive attributes” (imitative rays).
- With addition of (Latour, 2014):

(Law & Lin, 2020)

As a matter of being sensible in terms of our Thing not forming a coherent ‘whole’, we are to embrace resistance and alternative ways the course of action could unfold.

- Be sensitive to “otherness”, take into account uncertainty and iteration when translating empirically situated concerns.
- Be care-ful when presenting findings or insights containing “multiple realities” and “non-coherence”.
- With addition of (Law, 1987) and the concept of “heterogeneous engineering”.

Review of: Akrich, M., Callon, M., and Latour, B. (2002). *The key to success in innovation*. International Journal of Innovation Management, 6(2), 187–225. https://doi.org/10.1142/S1363919602000550

Authorship

The following is a review of the texts by Madeleine Akrich, Michel Callon and Bruno Latour with the enticing common title: 'the key to success in innovation' (Akrich et al., 2002a; 2002b). Due to their stature, further introduction of these French scholars seems superfluous, yet, for the sake of good order, they are all rooted in the field of Science, Technology and Society studies and have produced a plethora of influential works regarding socio-technical analysis. For example, concepts of Script Language for describing designer-user interaction (Akrich, 1992), ideas of trials of strength and enrollment (Callon, 1984) and science in the making (Latour & Woolgar, 1979; Latour, 1990; Latour 1999a).

Overview

At the onset, the main reason for choosing this text was to get a perspective on how to forge alliances effectively in innovation processes. So, additionally we chose these texts, because the authors are scholarly progenitors of Actor-Network Theory. The articles are dedicated to presenting some lessons learned about successes and failures in innovation, making the underlying mechanisms intelligible and ultimately making them more manageable (Akrich et al., 2002a, p. 191).

Unpredictable innovation processes

The way in which the lessons learned are presented, is via a collection of case examples, which showcase how innovation, led by experts according to their best knowledge, fails. In this way, Akrich et al. set out to dispel misconceptions about how and why successful technical innovation happens. Namely, they argue that it is a fallacy to emphasize the intrinsic qualities of technology as the determining factor for its implementation (the so-called 'diffusion model') (Akrich et al., 2002a, p. 203; 2002b, p. 208). Said in other words: technical innovations will prevail if they have superior intrinsic qualities compared with the technologies they aim to replace. Other fallacies include, descriptions of innovation as a result of coherent chains of rational decisions (Akrich et al., 2002a, 191-194) and notions of predictability and apriori certainty

while innovation happens (pp. 194-201). What all these arguments amount to is a critique of analysis of technical transformation that investigates the 'social' and the 'technical' separately (Akrich et al., 2002a, pp. 203-205). Alternatively, it is suggested that one should do both together in socio-technical analysis (Akrich et al., 2002a, 205). With this foundational argument in place, the authors present the 'model of interessement ' as an alternative to the 'diffusion model'. What this encompasses, is that successful innovations depend on the active participation of all those who have decided to develop it (Akrich et al., 2002b, p. 208). While this statement is rather obtuse in its most reduced form, in practise it is not. Rather, what it entails is the success of an innovation is based on the innovators' ability to build social and material allegiances by mobilizing interests (Akrich et al., 2002a, p. 205). Said in other words: 1) if you choose the *right* actors to ally with (social and material), your project will be successful, and 2) choosing allies is the same as choosing spokespersons for your cause. This line of thought showcases how agency is dispersed among the actors the innovator manages to enroll.

Takeaways

A point for discussion is that it is argued that mobilizing interests happens via adaptations, series of trial and error and countless negotiations (both social and technical) (Akrich et al., 2002b, p. 207, p. 212), and that these are all based on discourses of accusation (Akrich et al., 2002b, p. 223-224). What this means is that the authors argue that the innovator has to contend with a stream of accusations, when enrolling their socio-material alliances. This line of argument evokes the 'machiavellism critiques' of ANT and Latour in particular (Latour, 1999a; Vikkelsø, 2007). Wel would argue that this is due to the subject matter of 'innovation in the making', given that this type of story is about how power relations shift. In any case, this is expertly showcased in the case examples brought forth in the articles. On a different note, the articles also invoke critiques regarding the performative multiplicity of technical innovation as seen in the case of the Zimbabwe Bush Pump (de Laet & Mol, 2000). Yet, we would argue that the articles at hand actually contain notions of multiplicity. This is apparent in the case example of the photovoltaic kits, as it shows that openness to negotiation in design of technology (adaptability) can be the source of its ability to adhere (Akrich et al., 2002a, pp. 202-205). This type of argument is on the same line as those of 'fluidity' (de Laet & Mol, 2000), the difference being that 'the art of interessement' is an instrumentalization of the notion.

Literature:

Akrich, M. (1992). The De-Scription of Technological Objects. In Law, J., Bijker, W., & Law, J. (Eds.), *Shaping technology/Building society : studies in sociotechnical change* (206-222). MIT. https://hdl.handle.net/2027/heb.01128

Akrich, M., Callon, M., Latour, B., & Monaghan, A. (2002a). The key to success in innovation part I: The art of interessement. International Journal of Innovation Management, 6(2), 187–206. https://doi.org/10.1142/S1363919602000550

Akrich, M., Callon, M., Latour, B., & Monaghan, A. (2002b). The key to success in innovation part II: The art of choosing good spokespersons. International Journal of Innovation Management, 6(2), 207–225. https://doi.org/10.1142/S1363919602000562

Callon, M. (1984). Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay. The Sociological Review (Keele), 32(1_suppl), 196–233. https://doi.org/10.1111/j.1467-954X.1984.tb00113.x

de Laet, M. & Mol, A. (2000). The Zimbabwe Bush Pump: Mechanics of a Fluid Technology. *Social Studies of Science*, *30*(2), 225–263. https://doi.org/10.1177/030631200030002002

Latour, B. & Woolgar, S. (1979). Laboratory life : the construction of scientific facts . Sage Publications

Latour, B. (1990). Technology is Society Made Durable. *The Sociological Review (Keele)*, *38*(1_suppl), 103–131. https://doi.org/10.1111/j.1467-954X.1990.tb03350.x

Latour, B. (1999a). On recalling ANT. The Sociological Review, 47(S1), s. 15–25. https://doi.org/10.1111/j.1467-954X.1999.tb03480.x

Latour, B. (1999b). *Pandora's hope : Essays on the Reality of Science Studies.* Harvard University Press

Vikkelsø, S. (2007). Description as Intervention: Engagement and Resistance in Actor-Network Analyses. Science as Culture, 16(3), 297–309. https://doi.org/10.1080/09505430701568701

Review of: Björgvinsson, E., Ehn, P., & Hillgren, P. (2012). *Design Things and Design Thinking: Contemporary Participatory Design Challenges*. Design Issues, 28(3), 101-116. Retrieved April 7, 2021, from http://www.jstor.org/stable/23273842

Authorship

Erling Björgvinsson (Professor at the School of Design and Craft, Gothenburg University), Pelle Ehn (Professor of Interaction Design, Malmö University), and Per-Anders Hillgren (Associate Professor in Design and Social Innovation, Malmö University) are all occupied with the democratization of innovation processes. They build upon the Scandinavian "school" of Participatory Design and look towards the socio-technical dynamics of especially design materials that enable, restrict, and 'manipulate' humans and non-humans.

Overview

The article departs from addressing former design approaches of "design-for-use" and introduces the notion of "design-after-design" by incorporating Participatory Design (PD). Instead of being attentive to a physical product and an economic bottom line, it moves the center of attention towards the processual aspects of innovation. The authors concern themselves with how to establish a participation in socio-material assemblies described as *Things* (which is borrowed from Latour).

From things to Things

With inspiration from PD, they are taking on the responsibility of staging and "infrastructuring" a space for co-creation (co-design and participant design) within a field of 'developers, users, and markets' characterized by great complexity and diversity in both perspectives and interests (pp. 102-103). The "user" and "market" involvement should ensure that "existing skills" are made a resource in the design process and still allow the participants' tacit knowledge - and not just their formal and explicit competencies - to come into play. They are striving towards an approach of "design-by-doing" and "design-by-playing" as intertwined language-games which fosters playful

learning (pp. 106). The authors' view of PD is as an entanglement that aligns heterogenous matters of concern and language-games by using "design objects" (early prototyping) and "presenters" performing as boundary objects. The envisioned use is not the same as the "actual use", which means that the appropriation of the design is unintelligible (pp. 107). Instead of more 'fixed' entities they present a notion of *Thinging*, which entails designing for the process and not the final product.

Exploring design opportunities

When not designing for a certain use of things but rather a chain of "one design Thing after another", we can engage multiple "presenters" and stakeholders by infrastructuring the design phase with potential boundary-objects supporting future design Things applicable in use-situations (pp. 108).

> *"However, the relations between these design Things, rather than being clear-cut, form a web of interwoven language-games over time"* (Ibid).

The term directs our attention to how public socio-material assemblies (Things in use) are aligned with the design of Things in the temporal dimension of the projects. It is a relationship being shaped in an extended timeframe, as they put it, between the activities at "use time" and at "project time". It seems as though a *Thing* is entangled with other *Things*.

Infrastructuring

The activities of *selection, design, development, deployment, and enactment* on one hand, and the professional activities of *mediation, interpretation, and articulation* on the other, are entangled and intertwined through infrastructuring. These two activities are furthermore mixed up with the performances of *adaptation, appropriation, tailoring, re-design, and maintenance* when thinking further of a design in use (pp. 108). One should enable heterogenous partners to *"bring forth the issues or possibilities they want to explore and see if their vision or issue makes sense and matches with other partners' concerns."* As a result, participants are allowed to set up their own infrastructuring processes and Things while the facilitator is ensuring that the objects designed allow "design-after-design and have at least elements of Thinging" (pp. 114).

Takeaways
In our case of cooperative exploration of data visualizations, would we then enable the “diverse and practical skills” (as they put it) of navigating electronic music and be assisted in the sense-making of these maps as “socio-material assemblies''? Instead of designing for a certain use, we are therefore, according to the authors, to shift our focus as designers from striving towards providing useful products and services, to supporting environments that stimulate the movement of ongoing “design for design Things”. In this way users can be considered mediators that emerge from our infrastructuring process. We are designing a *Thing* by harvesting data, filtering it and curating different ways of visualizing it. This *Thing* is built from a process of *Thinging*, which means it is continually re-appropriated because of many other “design objects” and (non)human actors in micro-processes of the innovation stage.

Review of: Munk, A. K. (2019). *Four styles of quali-quantitative analysis: Making sense of the new nordic food movement on the web*. *Nordicom Review*, *40*(s1), 159-176. https://doi.org/10.2478/nor-2019-0020

Authorship
Anders Kristian Munk (Associate Professor in Techno-Anthropology, Aalborg University) is concerned with data intensive research methods of social science. He is the Lab Director of a research group that specializes in the use of Digital Methods (i.e. data sprints with visual network analysis) as means for participatory approaches and Controversy Mapping.

Overview
The article deals with anthropological investigations that involve quantitative data processing and how an analysis can achieve a qualitative richness. To do this, Munk presents styles of analyses that combine ethnographic sensibilities with Digital Methods in order to situate macro-patterns. Thus, he intends to reintroduce the ambition of understanding “tribal life” or “the native language” by showcasing qualitative interventions in the data practice that informs the researcher of the construction of digital phenomena.

The four styles

1) *Complementary:* A mode of analysis in which the quantitative insights inform additional qualitative investigations that situates these onlife traces. Each is allowed to unfold on their own methodological premises, that is, to co-exist undisturbed from each other. The task is therefore to establish a "program" as in the right conditions that ensures this unfolding (pp. 165ff).

2) *Single Level Analysis:* This style is concerned with how quantitative patterns emerge from qualitative rich interactions on the micro level. It is an attempt to overcome the distinction between macro and micro levels of analysis - as inspired by ANT. In this way it becomes a matter of tracing macro phenomena by inquiring the construction of these in qualitative contexts of deployment. The two levels are contingent (pp. 168ff).

3) *Curation:* By taking into account Richard Rogers' idea of developing "critical analytics", it is required as onlife researchers through qualitative intervention to repurpose the common "vanity metrics" of the Web. In this mode of analysis we are to evaluate relevance and quantifiability of digital performances. What is the nature of these digital traces? The style concerns itself with understanding the technical choices and data tools through social inquiry (pp. 171ff).

4) *Algorithmic Sensemaking:* Instead of the ethnographer discovering patterns it becomes a matter of computation, network analysis and algorithmic community detection and recognition. It is an explorative data analysis approach that in many ways has similarities with ethnographic work in how it does not need to know the "theory of the world" in order to generate hypotheses or questions. A "native worldview" is built from the bottom up - "exploratively and inductively" (pp. 174).

Ethnographic work and anxieties

Taking inspiration from George Marcus' set of methodological Munk attends to the isomorphic conflation of "place, culture and field". The key anxieties are described as how [1] the limits of ethnography are tested when trying to understand a wider context or system; how [2] a researcher's mobility and lack of embeddedness weakens the power of fieldwork; and [3] that the

former privileged position of being placed outside of “hegemonic power structures of the world system” is eroded by attempts of studying up and completing the picture (pp. 170).

Takeaways
Anthropologists are in the case of Digital Methods no longer accounting for “spatially confined, physically locatable, single-site notions of the field” - partly because of the world’s interconnectivity. There is a limit to one’s explanatory capability and it is urgent to be reflexive and attentive to the use of data tools in layout and exploration. We need to accommodate the understanding of these technical choices and ensure situating quantitative ‘patterns’ qualitatively in the ‘context’ of its development. One is obliged to think critically and reflectively about the appropriation of such digital entities through qualitative interventions. How can we repurpose the “vanity metrics” of electronic music to explore and build new perspectives inductively? Inspired by Single Layer Analysis it could be interesting to see whether or not the relationships between artists, labels, and releases in electronic music emerge differently.

Review of: Jensen, T. E., Birkbak, A., Madsen, A. K., & Munk, A. K. (Accepted/In press 2021). *Participatory Data Design: Acting in a digital world*. In G. Downey, & T. Zuiderent-Jerak (Eds.), Making and Doing STS MIT Press.

Authorship
Torben Elgaard Jensen (Professor of Science and Technology Studies, Leader of the Techno-Anthropological Research Group), Anders Kristian Munk (Associate Professor, Lab Director at TANTlab), Anders Koed Madsen (Associate Professor, Executive Committee at TANTlab), and Andreas Birkbak (Associate Professor, Executive Committee at TANTlab). They are all affiliated with Aalborg University at the Department of Culture and Learning. As members of the Techno-Anthropological Research Group and teaching programme, they examine “innovation, knowledge construction and user

involvement” and concern themselves with “Digital Methods” and “the retooling of humanistic methods”.

Overview
The authors present Participatory Data Design (PDD) as a methodology of how to involve knowledgeable stakeholders - “issue experts” as they call it - in participatory approaches to data design and practices. It is often organized as data sprints in workshops spanning over 3-5 days which entails three key processes of “datafication”, “flexible visualizations” and a “materialization” of the collectively generated knowledge.

Iteration
The collaborative performance involves competent stakeholders - often experts or figures of notable agency - within a particular field of inquiry in which relevant data is found and made sense of. An outcome of such a workshop strives towards (co-)creating a useful digital device or object (pp. 117). It entails iterative processes of visualizing, filtering, labeling, and collectively discussing the material at hand (pp. 119). These participatory processes are different every time, but always concern themselves with how versions of data practices can be negotiated so that they improve rather than “diminishes life”. Therefore it is necessary to involve stakeholders who are the potential future users.

Datafication
This is part of the three key processes of PDD. In this first movement it is a matter of collecting substantial data in relation to a specific problem and through these data sets investigate ontological and normative commitments. It is in this regard important not to enter this process with presumptions of what the central issues or matters of interest are (pp. 124). Instead, datafication has the ambition of operationalizing the stakeholders' concerns in respect to the specific data set. It is a mode of pragmatic inquiry that seeks to redefine problems through empirical experiments (pp. 125).

Flexible visualizations
As the second key process it draws resemblance to low-fidelity mock-ups. It is a matter of withholding determinative and definitive representations in relation to the field of inquiry beforehand. Data visualizations are potentially

very persuasive under certain circumstances and can direct the participants' attention towards a shared visual object to which the discussions are to relate to. But, on the other hand, data visualizations also open up for participatory design through investigation and sensemaking. "The craft of producing images" enables large amounts of information to be embedded into condensed formats by organizing and sorting material (pp. 126). The persuasiveness is nevertheless not a predictable and controlled effect. The participative dimension springs from the realization that data visualizations are flexible enough to be played around with because of the ability to quickly generate, display and enhance them in various ways.

Materialization

The goal is to create some material that articulates the participants' knowledge. In the article it is discussed whether or not this materialization should be on the basis of "interference" or "constitutionalism" - with reference to John Law. Again, it differs from project to project, but the first is a matter of preventing the manifestation of any singular or pre-existing ontology, and the latter is concerned with taking into account the multiple positions present and entangling them into a craft on the basis of a common world. This mode is much more an ambition that impacts the course of action than a predefined process. The stakeholders have a more active role in the knowledge production which makes us reconsider the amount of multiplicity invoked, situational ethics to be applied and authorship/ownership retained.

Takeaways

It is troubling for us to incorporate the methodology of a collaborative data sprint since the COVID-19 situation prevents us from using the facilities at the University and meeting with a "larger" group of people. Nevertheless, we are inspired to establish participatory processes that inform the "structuring" and selection of our data. It should allow a data practice that operationalizes the participant's concerns, while guarding everyone against "overinterpretation" by explaining the formatting (p. 126) when dealing with the persuasiveness of such data visualizations. To do this it is important to incorporate a large degree of flexibility to ensure authorship being taken by the participant in the co-creation of electronic music perspectives. Also, we come to consider what materialization we should perform, since we will not be present in "future use cases" to take ownership and direct the course of action.

Review of: Latour, B., Jensen, P., Venturini, T., Grauwin, S., and Boullier, D. (2012). 'The whole is always smaller than its parts' – a digital test of Gabriel Tardes' monads. The British Journal of Sociology 2012 Volume 63 Issue 4, 590-615. (25 pages)

Authorship

Bruno Latour is a french philosopher, anthropologist and sociologist famous for his contributions to the field of STS and his Actor Network Theory. He joins forces with his colleague, Tommaso Venturini, from the médialab of Sciences Po Paris as well as Pablo Jensen (a physicist at Ecole Normale Supérieure), Sebastian Grauwin (a data analyst and Chief Scientific Officer of AUM Biosync), and Dominique Boullier (Scientific Coordinator of the médialab of Sciences Po Paris). This ensemble concerns themself with innovation and data science - amongst many other things.

Overview

In short, the article translates the conceptualization of "monads" by Gabriel Tarde to the exploration of data practices and structures. It builds upon the idea of not distinguishing between the micro and macro level as seen in ANT-approaches through 'generalized symmetry' in dealing with subjects and objects. The matter becomes one of moving back and forth from the entity and its attributes - following successive chains.

Social connections

When dealing with "social connections", earlier definitions entailed divisions into two levels: individual or society; the element and the aggregates - and instead it is posed in this article not to make distinctions between the level of the individual component and that of the aggregated structure when navigating and exploring datasets. Credit is given to Tarde's notion of 'monads' (p. 590) and the authors' ask: "Is there a way to define a longer lasting social order *without* making the assumptions that there exist two levels?" (p. 591).

A monadological standpoint

Because of new digital techniques it is possible to learn how to navigate through overlapping 'monads'. It is described as a notion of a "circulation of differently conceived 'wholes' that are always smaller than its parts" (p. 591f). Circulating back and forth from the entity and its list of attributes (p. 592, 595)

without any starting point or anchor (p. 603). The entity and its 'associates' make up each other. A monadological standpoint is to gather the richness of associations while balancing the complexity of agents with specificity and their intersecting features (p. 606).

Partial totalities of particulars

Context-independent rules cannot 'capture' human behaviour as seen attempted in atomistic approaches (p. 597). Associations characterize an entity, and these differ as point of views turned towards all the other entities one at a time (p. 598). When applying this chain of thought one goes from "particular to more particulars" (p. 599) and might obtain "partial totalities" along the way while remembering that "there is no individual agent; they don't interact; [and] there is no whole superior to the parts" (p. 600). There is no being a part of a structure, because there is no second level common to a collective (p. 604). Surrounding a list of features in data sets with a circle is just making another monad defined by the "provisional end of the expansion of their content" as an envelope of individualized attributes (p. 607).

A 'collecting' activity

Each entity should always have its own trajectory through "successive attributes'' (p. 608) when navigating as Tarde would call it by 'imitation' with 'imitative rays' (p. 609). It is all about looking at monads sharing attributes [translations] that are modified by each sharing becoming something repeated with variations [recognizable transformations] (p. 610). In general, the authors' state that we should "stop talking about collective phenomena distinct from individual ones" and instead focus on the "many different types of *collecting* phenomena" (p. 612).

Takeaways

When exploring a dataset (or visualization) the complexity rises as we investigate 'nodes' in the network. The 'node' is made up from the links (connectivity) streaming from it and from other 'nodes' sharing (some of the) same attributes. The collecting activity consists of these many partial totalities of 'nodes' and their trajectories. How do they intersect or associate themselves with each other?

Review of: Law, J., & Lin, W. (2020). Care-ful Research: Sensibilities from STS. Retrieved April 7. 2021 from http://heterogeneities.net/publications/LawLin2020CarefulResearchSensibilitiesFromSTS.pdf

Authorship

John Law (sociologist and an STS-scholar) is a key proponent of Actor Network Theory and therefore positions himself close academically to Michel Callon, Madeleine Akrich and Bruno Latour. Wen Yuan Lin (Professor at China Medical University Hospital) has worked together with Law on multiple occasions in writing sociological articles that concerns the act of doing science and what sensibilities that needs to be attended to.

Overview

In social science and especially STS-studies it is presented by the authors how certain sensibilities need to be enforced. A reflective stance needs to be applied in discerning when to change the course of the research or resources at hand. What are the possible applications or limitations of a given knowledge production? Because 'things' are conceived as being parts of "never-ending" relational structures and inhabit multiple logics, it is important to consider how the research can be sustained after the point of determination.

No general rules

STS-research is claimed to be done through iterative and uncertain down-to-earth practices. The authors' introduce an alternative *care*-ful research that takes its departure from the concerns present and from cultivating appropriate sensibilities towards an articulation of researchable questions. What is presented is a material-semiotic STS-version of sensitivities that include "material heterogeneities; webby relations and logics; non-coherence; otherness; normativities; and politics".

There are no rules - only research-relevant questions to be asked (p. 1f).

Looking for concerns

Concerns shape our view of the world. What there is to see, the urgency to some perceived entities, and the projected potentials. Educating these sensibilities becomes a matter of technical and disciplinary training; social skills; the art of observing and record-keeping; the organisation and selection

of appropriate bits from piles of material; and how to write and read well or "beautifully" (p. 2). To be care-ful in one's research is according to the authors' to work iteratively with concerns by adapting and sharpening them into questions. Ask what difference the research is supposed to make and how the concerns can be articulated and educated. Question the resources at hand and whether or not to change the performing concerns (p. 3).

A particular heterogeneous materiality

With reference to Haraway (1988) it is stated how knowledge is situated and works in "particular ways in particular circumstances". Different disciplines deploy different sensibilities (p. 4). A core sensibility in STS is nonetheless one of heterogeneous materiality (textual, architectural and technological as well as 'people') even though what to think of as 'materials' still is debated (p. 5). Law and Lin underlines the values, normativity and/or politics embedded in materials. The interest is therefore directed towards 'how materials *do* this' and what alternatives that could be imagined (p. 6).

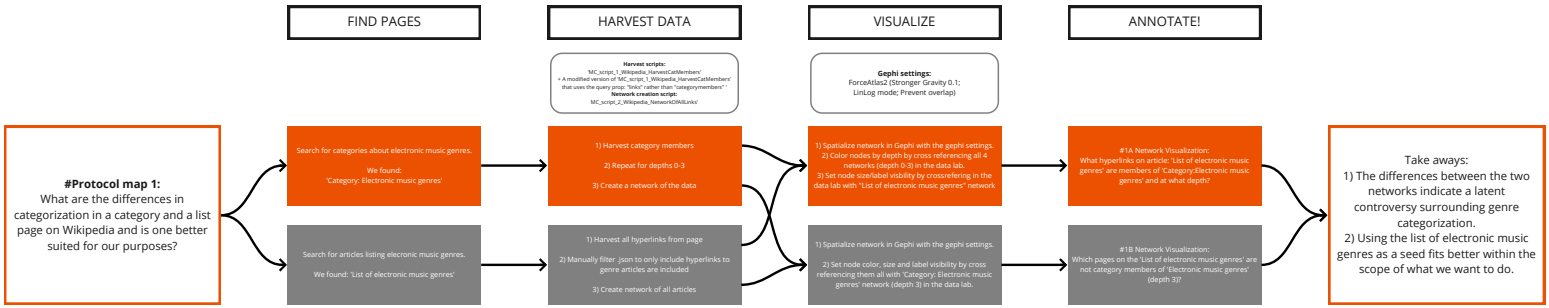
Webby relations

Things *are* webs of relations - described as the notion of relationality (p. 7). These webs are never-ending and the research stops when run out of time [pragmatic] and resources or when the questions have been answered [respectable]. Webs are fragile and are only sustained by processes still going (p. 8). Ask "how" and not "what". Social webs are not necessarily coherent but entails several 'logics' and the interaction between them (p. 9f). What is noticed becomes a matter of one's partial perspective and the *insensibilities* being established (i.e. a researcher not desiring a mess will probably not include it).

Takeaways

It is a matter of being sensible about things not forming a coherent whole while being sensitive to otherness and resistance (p. 11) Especially since we are entering the roles of 'designers' and thus directs or facilitates the matter of course. Embrace the resistance and think of alternative ways the course of action could unfold. By taking into account the empirically situated concerns we need to turn our gaze inwards in considering what "difference" our research is supposed to make. The care-ful work is concerned with "multiple realities, with double vision and non-coherence" (p. 12) by imagining the uncertain and iterative nature of the research (innovation) process (p. 13).

Getting aquainted with the field



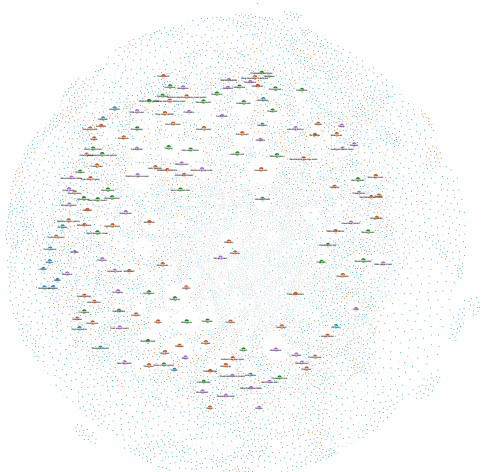
#1A Network Visualization: What hyperlinks on article: 'List of electronic music genres' are members of 'Category:Electronic music genres' and at what depth?

Nodes: 14980
Edges: 262923

Depth at which articles appear

Depth 3 (78.36%)
Depth 2 (24.92%)
Depth 1 (1.38%)
Depth 0 (0.34%)

Label & large node size indicate titles of genre pages from 'List of electronic music genres' that are members of 'Category:Electronic music genres' (depth 3)



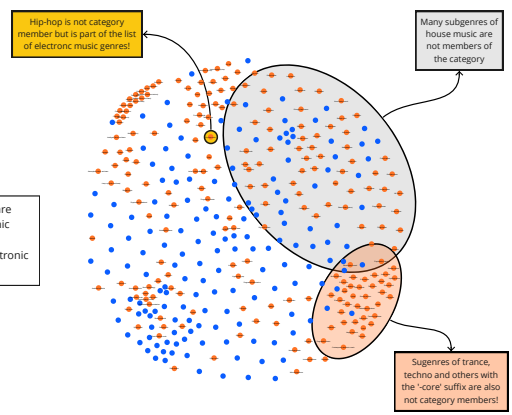
#1B Network Visualization: Which pages on the 'List of electronic music genres' are not category members of 'Electronic music genres' (depth 3)?

Nodes: 369
Edges: 24982

Part of network 1?

Yes (44.44%)
No (55.56%)

Labels indicate titles of genre pages from 'List of electronic music genres' that are not members of 'Category:Electronic music genres' (depth 3)



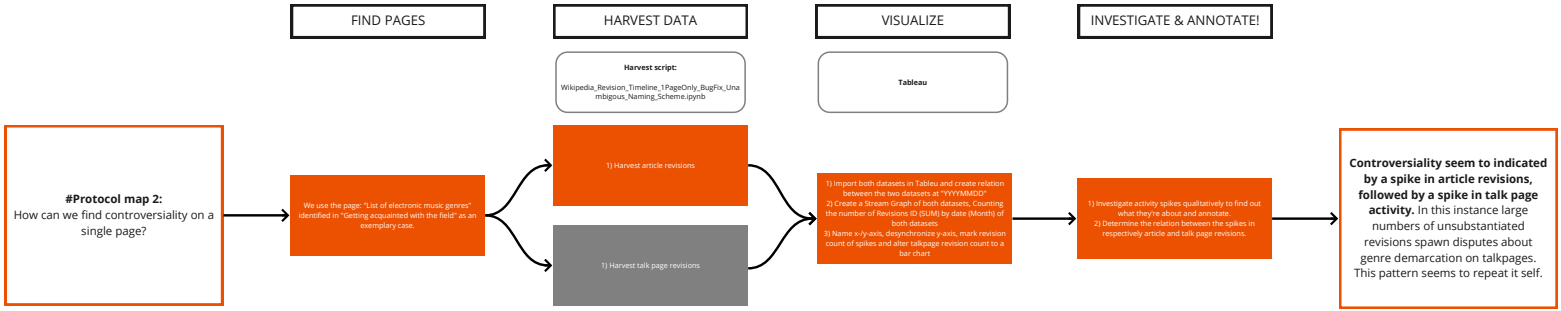
#1A Findings:
1) There are different conventions for categorization of genres / sub genres on the list vs the category. This is interesting because the two networks are dealing with the same topic. This might indicate that the topic of genre categorization might be controversial as the two networks are so far apart in scope and content.
2) Genre articles are found at various depth, which also results in non-genre articles quickly showing up. They might be instruments, festivals, artists or albums.
3) It would thus require a mucher higher depth (and many more nodes) to get to certain sub-genres that are in the list of electronic music genres.

The inconsistency in categorization and relative size needed to gain saturation, would make the "Category:Electronic music genres" a bad fit fro studying genre controversies across all (or most) of electronic music

#1B Findings:
1) The presence of Hip Hop (not a member at all) and House sub-genres (present below depth 6), again indicates a strong difference between communities in how music genres are categorized.
2) While we can not be certain that some genre articles are excluded, the list format of the pages presents an opportunity to study, not only the genre articles hyper-linked, but also the "List of electronic music genres" page it self.

While a list page is arguably more subjective than going by category alone, the "List of electronic music genres" seems to be a good fit for studying genre controversies on genre pages, while still keeping the data amount managable

Finding a genre controversy on a single page



#2 Stream Graph: What edits trigger debate among editors of the page: 'List of electronic music genres'?

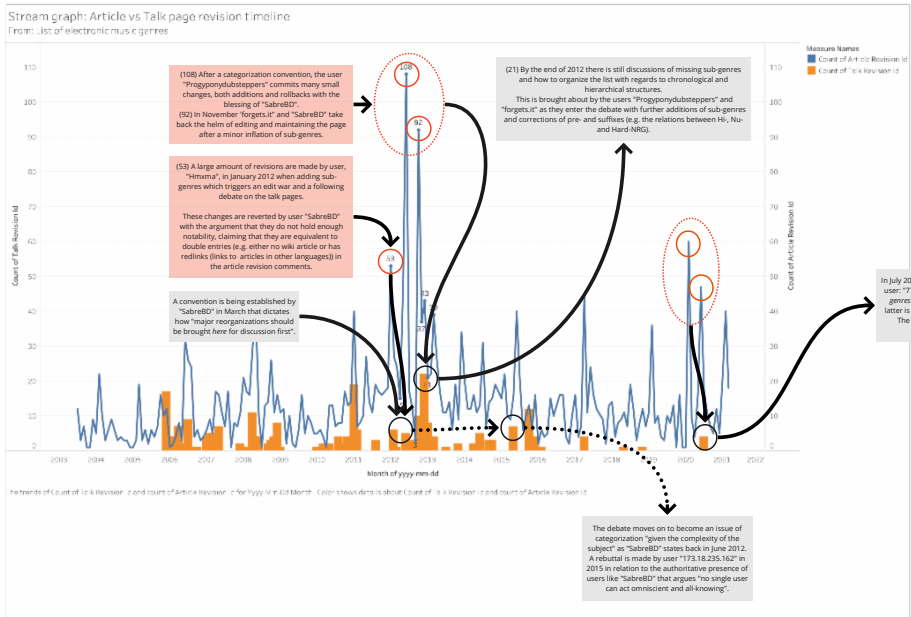
Hypothesis:
That spikes in revisions with subsequent talk page activity on the timeline will be controversial.

#2 Findings:
Unsubstantiated changes to the genres and sub-genres spawn controversy on talkpages. Particularly if they are done by single users not conferring the changes in the talk pages before hand.

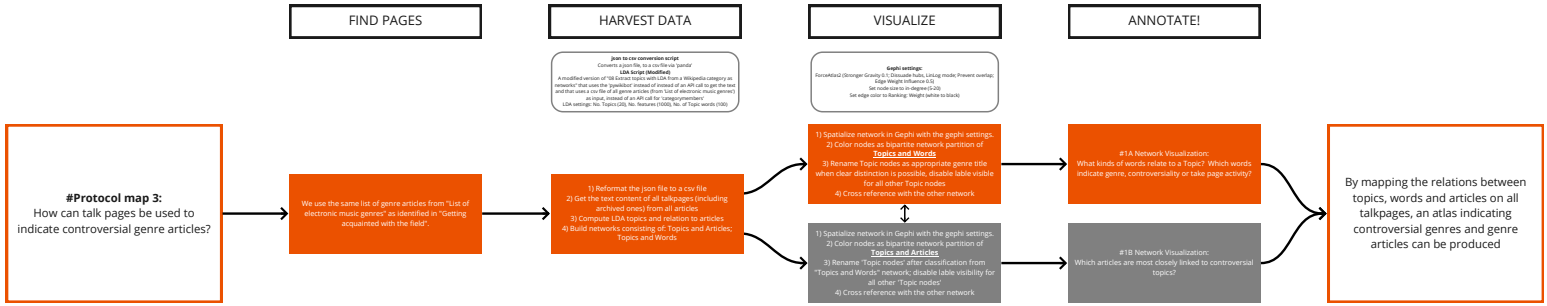
Controversies are (usually) settled by reffering to whether a claim is substantiated by external links or a Wikipedia page (from 2012 this includes 'redlinks').

Additionally the above mentioned spawn debate about naming and categorization conventions and intervals between replys can span over many years (e.g. 2012-2015).

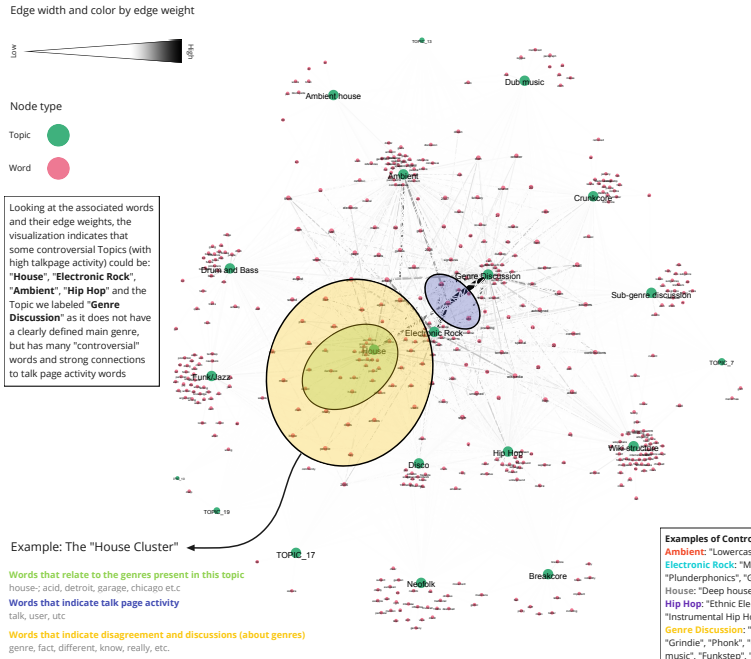
This pattern still seem to be repeating it self...



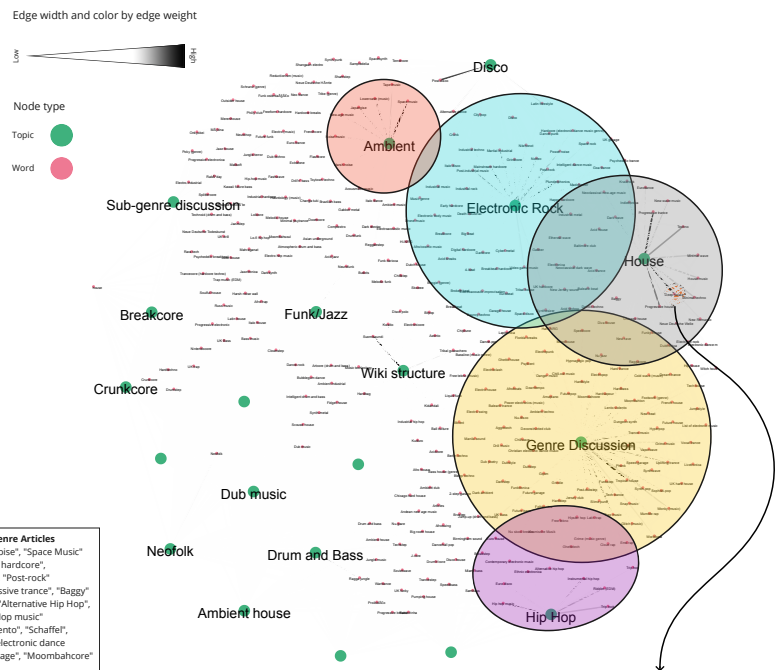
Finding Controversial Genre Pages



#3A Network Visualization: What Topics relate to which words and how do these indicate controversiality or talkpage activity?



#3B Network Visualization: How can the relation between Topics and Words from Network Visualization 3A be used to identify possible controversial genre articles?



Examples of Controversial Genre Articles
Ambient: "Lowercase", "Japanoise", "Space Music"
Electronic Rock: "Mainstream hardcore", "Plunderphonics", "Grindcore", "Post-rock"
House: "Deep house", "Progressive trance", "Baggy"
Hip Hop: "Ethnic Electronica", "Alternative Hip Hop", "Instrumental Hip Hop", "Hip Hop music"
Genre Discussion: "Lento violento", "Schaffel", "Grindie", "Phonk", "Christian electronic dance music", "Funkstep", "Speed garage", "Moombahcore" etc.

#3A Findings:
1) The LDA map of Topics and Words, indicate that certain words relating to discussions and talk page activity, can be used to gauge the relative controversiality of a Topic and (by association), it's genre pages.
2) The map also suggests that some genres (primarely House and Electronic Rock, but also Ambient and Hip Hop) are more controversial than others.

By looking at the relations between words and topic, we can get a measure of which Electronic Music genres might be more controversial than others

#3B Findings:
1) The LDA map of Topics and Articles, indicate that certain genre articles are more controversial than others.
2) Interestingly, the "Genre Discussion" Topic/cluster also poses a question: Are the genre specific clusters (tied together by both genre terms **and** controversial words) *more* or *less*, controversial than the "Genre Discussion" Topic/cluster (which is presumably more tied together by the presence of controversial words/words indicating timestamp activity).

By looking at the article nodes with the strong edge weights to controversial Topics, we get an indication of which Genre Articles are more controversial than others.

Story example:
When taking a closer look at the talk page of the page **'Deep House'** (indicated controversial by the edge weight / proximity to the 'house' topic), lots of similar arguments appear to the initial page. Interestingly the debate on this page is different in nature. The posts are more extensive and have more outward references. Furthermore the users take particular interest in debating how platforms have an effect on how the sub-genre is being defined. Namely that there has been a 'beatport debacle', refering to how www.beatport.com has been mislabling/overusing 'Deep House' and that this has had an effect on the use of the term in general - and most importantly, confusion and mislabeling on this article.