

This day by day diary was meant to keep track of the process and ongoing discussion in the Data Visualization table during La Cura Summer School 2016. It starts from Tuesday, as Monday was dedicated to introductions.

Tuesday, 24/08/2016

### Morning

We formed groups. We are Giorgio, Ami, Francesca, Martina, Anna, Joanna, Ruggero, Antonio, Jon.

We discussed along with the identity and social listening groups to understand the conceptual framework for the summer school. Identity is meant to represent the process, rather than the output or ideological grounds that started it.

Afterwards, we started working as a group by doing an overview of the possibilities with our data visualization. Interactive/non-interactive, static/dynamic as an installation, with data coming and including the projector as a possible medium for display. The projector would be problematic for two reasons: it would be hard to see anything as the showcase takes place in a glass building at Expogate in Milan. Moreover, it can introduce further complexity as it takes away the possibility to have a touchscreen as an interface. This means we would have to build, connect and map another device to make it interactive.

The possibility of directly and temporarily altering the data also emerges (e.g. I don't like that this theme is not so talked about. By interacting with the visualization on the screen, I will make it bigger).

### Afternoon

We saw how the human ecosystem platform works, and how to set up users and researches, and how to fetch data from the API.

The hard part is to be able to set up a meaningful research from a cultural and meaning producing (produrre senso) way, which is one of the main goal of the social media listening group.

Then we discussed how to proceed, since there is the need for the social media listening group to discuss and experiment to be able to create and refine its research.

We thought we were stuck and started discussing how to overcome this obstacle. Various positions emerged: we could decide which type of data we're interested in (image, text, numbers, etc), whether to go more conceptual or "problem-solving oriented", do research on possible outcomes, experimentation to understand which types of visualization would be possible.

We agreed that the output of human ecosystems is problematic because it already gives structured data that is needed for a particular visualization.

We discussed quantitative and ordinal data vs categorical (qualitative) data to understand which type of visualization to build.

Jon quoted the sensemaker platform and polarity management method (identify poles in a qualitative dataset so to give it some navigation hints through quantitative interpretative scales). Building on that, Ruggero talked about a geography of conflict in a dataset and its archival representation.

Summary of the afternoon to give the next morning at the general assembly:

We had lots of doubts how to get unstuck without a precise dataset. One option was to go help Oriana's group but we figured we have also other things to take care of. So we decided to carry out several experiments with a qualitative visualization, maybe try a quantitative one. If more discussion is needed to converge, we will do some research with a clearer mind set on which are of interest we have (e.g. qualitative). After that, we will explore our tools (D3, visualization instruments, identity creation for the data visualization, possible interactions).

---

## Wednesday, 25/08/2016

### Morning

Things to bring up in the assembly:

Public: summary of the afternoon before

Public: do we put a shared folder for resources?

Private (only for a matter of time, for Salvatore): Ruggero noticed that for data visualization, data is structured in a way that makes it hard to experiment within the framework, so you always imagine what is already visible (that is especially true for the Social Media Listening group, that only access the pre-set visualizations available).

Things brought up during the assembly:

Oriana talked about the kitchen and bar handling as integral part of the summer school, and that situations arising in that area should be dealt with the same methodology as the rest of the working groups.

Salvatore told everybody to keep track of the tools and processes setup and used in the group and to make them apparent to the identity group.

At the end of the assembly, we focused on understanding the data we have. We agreed on the need to look at the contents of the data sources first thing. We mapped the areas of interest: techniques, sources, etc. We then explored the contents of a tweet and the responses from the instagram api endpoints. We understood how facebook api works.

We discussed the possibility to combine at least two data sources, one of which external to Human Ecosystems. An agreement wasn't reached and the discussion moved on.

Before lunch, we had a talk to understand how to proceed. Some people seemed less involved than other in proposing the data, so we discussed the possibility to present D3 as well in the afternoon, so that we could proceed in parallel on the two topics.

## Afternoon

Salvatore showed us the content and schema of the Human Ecosystems database, and setup a user to be able to access it for further explorations.

Afterwards, we started exploring D3, starting from the structure of a JSON file, to html, to the way html and data can be manipulated with Javascript. What emerged from the first presentation is that the tools is probably too complicated to build a complex visualization. So we discussed whether it made sense to work on the visualization, leaving the code for the end, or to break the visualization into several small ones. To decide, we had to look at data. We agreed on meeting the morning after to discuss the issue.

Ruggero talked with Oriana in the evening. What emerged is that there was a misunderstanding on what the Social Media Listening group would provide. Researches are performed by typology, not topic. There are different hypothesis right now: interconnection as empathy (topic without coordinates and languages) and the general idea is to subdivide searches into three "funnels" of information that can be remixed. Those typologies are topics, situations, places.

---

# Thursday, 26/08/2016

## Morning

We met with the Social Media Listening group so to understand better how they are building their researches and how we might use them. They worked on how to create clean data that can be remixed, and have a focus in how people express about a topic. They made three hypotheses:

- H1: search by topic | no geo, 1 elem in several languages. Example: empathy
- H2: search by situation | works with context. Example: earthquake keyword parameters, multiple language, semantic unit
- H3: search by place | different level from micro to macro (geo), include named identity

With that in mind, we started thinking about possible visualizations. Each one worked individually for a little under an hour to get a sense of the data and look for possible visualization proposals, and then presented it to the group.

While discussing the possibilities, we soon realized that the scope of the test researches already set up needed to be broadened, and we worked on a conceptual model of how to do it {pic of the chart}.

In short, we decided to set up several other resources around the semantic area of inhabiting, and its opposite. Inhabiting, living, occupying, dwelling, etc and uninhabiting, leaving, abandoning, etc.

We then talked with the Social Media Listening group to figure out a way to enhance and implement this idea, and agreed to work on it autonomously with an infiltrate from their group.

## Afternoon

In the afternoon, the infiltrate (Eugenio), helped us set up researches. We also started brainstorming to find the final shape of the visualization. It was a long process that took over four hours, with proposals and counterproposals, which ended by voting and choosing 4 possible outputs, and the agreement that the morning after we would decide which one to choose with a fresh mind.

This is because the group wished to have the opportunity to work hands on with D3. So we agreed we would stop working on the visualization and spend the last day trying to implement as much as possible.

---

# Friday, 27/08/2016

## Morning & afternoon

Jon caught us up on what he did the afternoon before infiltrating the Social Media Listening group. He highlighted the fact that they were working more from Human Ecosystems platform user perspective, how to generate a manual and a set of conceptual clustering to build meaningful resources, rather than data visualization perspective, which has been something we've misunderstood the whole time (their focus was more conceptual than practical). Jon proposed a clustering based on crisis, exploring issues, and public events.

First thing in the morning, we had two round of votes to choose the shape of the visualization to use. At the same time, a small general assembly was held, with one representative from each group, to catch everyone up (and be able to coordinate the outputs as a consistent whole) about all the options. Anna was our representative, and then switched to Ruggero. Afterwards, we spent the rest morning working on choosing which data sources to use within the context of the semantic area around inhabit and uninhabit, and how to build the interface.

We subdivided into focus groups to deal with the interface, visualization and interaction aspects, and then come back together in the afternoon to put everything together.

Further discussion was needed to choose the elements of the interface. In the end, we decided to have keywords on the two sides (the original semantic area and its opposite on opposite sides of the interface), and have several emotion filters on top. An interesting point was that, while each tweet could be analysed according to the circumplex model of emotions, many of the images from Instagram cannot, as they don't embed a textual component that can be analyzed from Human Ecosystem at present state. So the need to understand how to explain that, and choose whether to expose it or not arose, and led to further discussion.

Eugenio, from the Social Media Listening group, helped us set up all the researches we needed to collect data for.

We eventually wrapped everything up around 7 pm, and Joanna closed a short sketch presentation with the final visualization to implement.

---