Audio paper 7 – MEDIATING SUSTAINABLE CITIES: a mixed-method approach to social media data analysis

Welcome to the series of audio papers: Mediating Sustainable Cities!

I am Paola Monachesi, researcher at Utrecht University. In the previous audio paper, I have shown that the investigation of social media data usually relies on quantitative methods, for example word frequency and social network analysis. Quantitative approaches reveal interesting patterns with respect to the topics of interest of the creative migrants and their communicative interactions. However, in some cases, one might want to go beyond these general patterns and analyze the data at a deeper level to discover relations among these topics and more intricate connections. One would like to get similar levels of information from the analysis of social media data that we gain through focus groups or surveys.

Jingle

In this audio paper, I will discuss an innovative mixed-methods approach to social media data analysis, inspired by recent proposals in health social science, which combines quantitative and qualitative methods.

The method is based on what is called *explanatory sequential emergent design* that consists of a quantitative phase followed by a qualitative one. The qualitative phase is based on a thematic analysis, in which the quantitative results are interpreted to reveal the deeper meaning of the data.

Design can be *fixed* or *emergent*.

Fixed mixed method studies are those in which the use of quantitative and qualitative methods is predetermined and planned at the start of the research process. Emergent mixed methods designs are found in studies in which the use of mixed methods arises due to issues that emerge during the course of conducting research.

For example, when quantitative results ask for further explanation.

This was the case in my work. The quantitative analysis based on hashtag frequency revealed that a specific category of creative migrants, that is architects and designers was particularly interesting to be further investigated. They deal with some of the topics that play a role in the smart city discourse such as data and technology, but they also post messages about sustainability and social innovation. However, the quantitative analysis only highlighted these topics but didn't clarify whether there was a relation among them and what kind of relation.

I wanted to know whether digital technology is employed by these creative migrants to give rise to a more human-centered smart city, concerned with environmental protection and social equity. Therefore, I decided to investigate the relation among these topics in more detail through qualitative analysis.

Jingle

Let's see which are the various steps of the qualitative analysis carried out after the quantitative analysis discussed in the previous audio paper.

First, a data set was created that included all the tweets of the specific category of creative migrants. In my case these were 6000 tweets, so it was feasible to carry out a thematic analysis based on hashtags and tweets reading.

As I discussed in detail in the previous audio paper, hashtags are a way to add metadata to shared content, and they represent a way to highlight topics of interest. It is for this reason that I decided to view them as codes used to tag our data from which themes can be derived by simply collating the hashtags.

This approach is different from that usually employed in the literature that creates themes through manual coding. In the standard approach, codes function roughly as tags. It works in this way: a researcher gets acquainted with the data, reading through it, then he codes interesting features of the data, across the data set. He then collates data relevant to each code, searches for themes by collating the various codes into potential themes and finally assignes names to them.

In my case, I let the codes emerge from the data since I use the hashtags that user assigned to the tweets, as code. This is what hashtags do, they tag tweets.

As in the work of Braun and Clark, I assume that "a theme captures something important about the data in relation to the research question and represents some level of patterned response or meaning within the data set."

I have taken an inductive approach, based on observation of the data, since the themes are strongly linked to the data through the hashtags, but at the same time the approach is theoretically driven since I have focused only on themes that are relevant for our research question. The themes are technology, sustainability, and social innovation. For example, hashtags related to the theme sustainability are: #climatechange, #Parisagreement, #recycling, #Energymatters, #goinggreen, #circularcity, just to name a few.

I have extracted the data coded with these hashtags for each theme and assessed whether the remaining data could be coded with any of the identified hashtags in order to avoid possible omissions given that there is quite a lot of variation in the way users assign hashtags.

I have verified that the collated extracts make a coherent story for each of the three themes and identified sub-themes by further grouping the hashtags. In my case, the theme sustainability comprises 382 tweets which are 6.4% of the total and they include subthemes such as sustainable development, climate change and circular economy.

The themes' division makes possible to quantify how many tweets deal with a specific theme or subtheme and to analyze the tweets in detail in order to assess whether this category of creative migrants contribute to broadcasting an innovative use of technology in support of sustainable urban development.

Jingle

I would like to conclude by summarizing briefly the results of the qualitative analysis from which it emerged that in the communication of Architects and Designers, technology plays a relevant role, both in its technocratic aspects and as a means to serve the environment and citizens's needs. These creative migrants play an important role in broadcasting grassroots initiatives aimed at social inclusion and circular economy.

Quantitative analysis such as the one presented in the previous audio paper shows general patterns and it is a useful first step. For example, it revealed that creative skilled migrants are an interesting community to analyze in the context of the smart city. They are concerned with the theme of sustainability and technology and the group of architects and designers are the most active. However, the qualitative thematic analysis is a necessary step to interpret their message, to discover their attitude and their position. It shows that they contribute to shape a discourse on a more human-centered smart city. By highlighting specific local projects, they also provide insights on how it is put into action. **This is an additional example of the relation between digital and environmental sustainability and it shows the relevance of the methodology employed to reveal the deep meaning of the data.**

Digital technologists, urban planners, and municipalities could be inspired by creative migrants. The community of architecs and designers can play an important role in raising awareness within civil society by broadcasting relevant topics and linked activities through Twitter and by giving rise to a more emancipatory and innovative smart city agenda.

This is the last audio paper of this first series. I would like to continue it in the future when new results of my research will become available. It would be interesting if students and colleagues would join this series with their work. Their audio papers can contribute to make research on digital and environmental sustainability more widely available within the academic community. The ultimate aim is to get back control on our data and their use in urban policies…and not only.

Thanks for listening to the audio papers of this series!