

*Paola Monachesi*  
*P.Monachesi@uu.nl*  
*PhD, Faculty of Humanities*  
*Utrecht University, Utrecht, The Netherlands*

## **DUTCH ELDERLY AND GEOGRAPHICAL INFORMATION IN TWITTER DATA**

### **Introduction**

In the Grage (Gray and Green in Europe: elderly living in urban areas - [www.grageproject.eu](http://www.grageproject.eu)) project we investigate the role that information and language technology can play in transforming cities into environments that support green and healthy lifestyles for the elderly. The goal is to be able to analyze and visualize social media data to achieve a better understanding of people's behavior with respect to green and healthy lifestyles and eventually develop mechanisms based on personal technologies to inform citizens on best practices. The Netherlands is at the forefront in the use of internet and social media by the elderly: In 2016, almost 60% of those between 65 and 75 years old were active on social media while those above 75 years old were 20%. As for the use of social networks, in the years 2014-2016, it has registered an increase from 70% to 83% in the age group 45-65 years old while in the age group above 65 years old, the increase has been from 26% to 39% (CBS, 2017).

Since a growing number of Dutch elderly are active on social networks, we have decided to collect Twitter data from Dutch users for our analysis. They have been divided in three groups based on chronological age and life stages that relate to working life. The youngest group is that of those under-55 years and it includes those people that constitute the workforce. The second group involves those that are between 55 and 67 years, this group is characterized as pre-retirement. The

third group involves those over 67 years and it includes the post-retirement individuals. Detecting users of the desired age is quite challenging in social networks, since their age is not publicly available and it requires a manual step to categorize them in the three groups on the basis of the pictures provided and the information in their profile.

In the project, we mainly focus on a language analysis of the interaction of the users, however, in this paper we explore the possibility of using geographical data to assess whether we can get useful behavioral insights. More specifically, we analyze the location of the users and their network as well as their relation with the environment they tweet from by means of an analysis of location hashtags.

### **Geographical data**

The data of the selected users is archived in a database and it is possible to identify the location for the person belonging to the network of our users. More specifically, we have selected one test user per age group with 600-900 users in his network and exported the results in Excel. Furthermore, the program Tableau (<https://www.tableau.com/>) has been used to relate place names to locations on a map and get a visualization of the most important places where the relations of the users are located. The different colors are related to the type of relation while the size reflects the number of users that come from a given place. In Fig. 1, we have a visualization of the network of a user younger than 55 years and we notice that he has several followers from the same area he comes from, as well as from the main cities in the Netherlands. On the other hand, the followees are more scattered and there is very little overlap among the location of the followers and the followees. In fig. 2, the network of a user in the 55-67 age group is visualized, in this case there is quite some overlap between the location of the user and that of the followers and followees which are all located in one place.

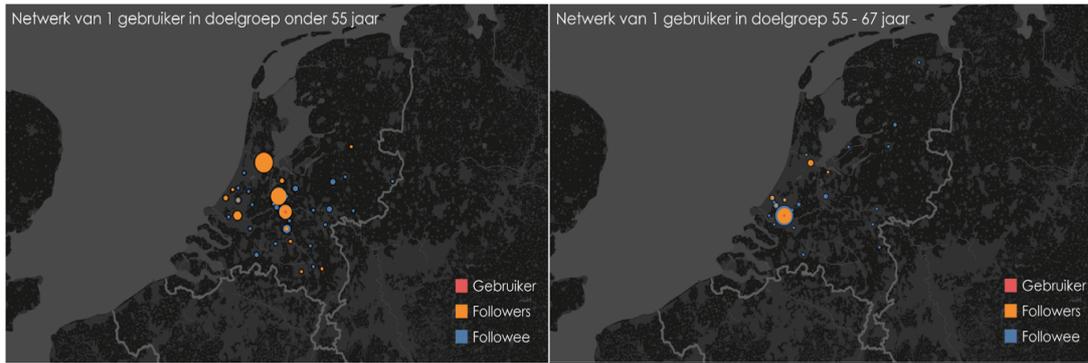


Figure 1

Figure 2

In fig. 3 we find a configuration where the user doesn't have any relations in the place where he is located which is in between two big cities in the Netherlands, that is Amsterdam and Utrecht where instead most of his followers and followees are located. Finally, in fig. 4 we see another example (exploiting a similar methodology as in the previous ones) of a visualization of the locations where all the relations of the users come from. In this case, it appears that the followers and followees come mainly from the big Dutch cities, with Amsterdam, Utrecht, Rotterdam and Den Haag belonging to the top 4 locations.

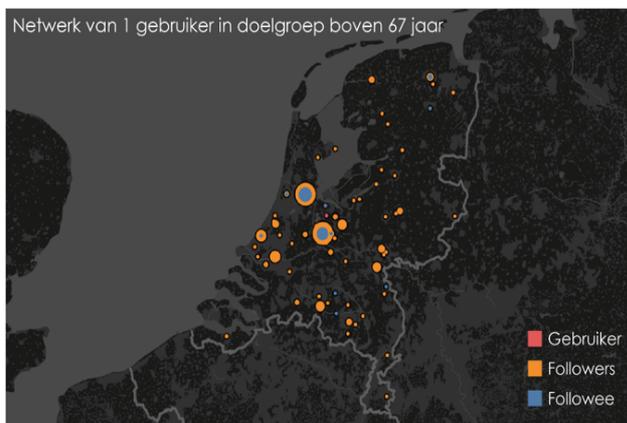


Figure 3

Top 20 steden met meeste twitter gebruikers uit user\_ff



Figure 4

Another interesting source of geographical information is provided by the use of location hashtags by the various age groups. In Carella and Monachesi (2018), we have presented a behavioural analysis of the elderly on the basis of hashtags, which are a social media specific feature. They are a relevant source of information given that they are used to index keywords or topics in Twitter and

more generally to emphasize something. We have selected the top 100 most frequent hashtags in each group which we have then divided into categories, one of which is *location*. It is interesting to note that the over-67 group shows an extensive use of location-tags (20.2%): more than three times as much as the younger group (6.6%). A detailed analysis reveals that they use location hashtags to promote the city they tweet about revealing thus a strong attachment with the environment.

## **Conclusions**

We have shown that Twitter data allow for a geographic visualization of the various relation patterns of the users in the different age groups. We suggest that it should be combined with a language analysis of the tweets of the users in order to shed light on the nature of these communication patterns and relations (i.e. private, work, interests, etc.) as well as on the motivation behind the use of Twitter.

***Acknowledgements:*** This paper is based on results of a project that has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 645706. This article reflects only the author's view and the REA is not responsible for any use that may be made of the information it contains. We would like to thank the students of the HAS Geo-trends project for creating the maps.

## ***References:***

- Centraal Bureau voor de Statistiek (CBS). Available online: <http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=83429NED&D1=0,12-14,25-45&D2=0,11-14&D3=0&D4=2-4&HDR=T&STB=G1,G2,G3&VW=T> (accessed on 06 March 2017)
- Carella, V. and P. Monachesi (2018) Greener through Gray: Boosting Sustainable Development through a Philosophical and Social Media Analysis of Ageing. *Sustainability*, **10**, 499.