

# A new method measures the integration or segregation of immigrants based on their tweets

March 15, 2018, Spanish National Research Council

An international team led by researchers from the Spanish National Research Council (CSIC) has developed a method to measure the integration or segregation of immigrants based on the messages they write on the social network Twitter. In the work, published in the journal *PLOS ONE*, the researchers developed a method to use Twitter data to analyse the degree of spatial segregation of immigrant communities. "The users' communities of origin are determined by the language in which the tweets are posted, establishing an 'idiomatic algebra' to assign the most likely community to which a tweet belongs," explains the study's director, José Javier Ramasco, CSIC researcher at the Institute for Cross-Disciplinary Physics and Complex Systems, in Mallorca, Spain.

"If all the messages are in the local language, then the user is considered to be a local resident. On the other hand, if some messages are in the language of an immigrant community, it can be assumed that the user knows that language and belongs to that community," he adds.

The language used, together with the location of the messages, make it possible to find the typical residential areas of the different

communities and to study whether they are more or less concentrated in those areas than the local population. "This method has allowed us to analyse immigrant communities in 53 of the world's largest cities. In each one of them we can define a metric that measures the spatial integration capacity of the immigrants living there," Ramasco explains.

By applying this metric, cities can be divided into three categories: those with high integration capacity, those with few or highly segregated immigrant communities, and an intermediate category between both extremes, explains Ramasco. "In the first group (high integration) we find cities such as London, San Francisco, Tokyo and Los Angeles, while at the other extreme (low integration), we see cities such as Detroit, Miami, Toronto and Amsterdam," he explains.

In addition to considering cities, the researchers can analyse how different cultures are integrated within the countries where these cities are located. The best integration is found among nearby cultures, for example, Latin-based language speakers (speakers of Portuguese and Italian) in South American Spanish-speaking countries, or those from European countries within the United Kingdom. The cases of greater segregation occur between extremely different cultures.

This method offers a new source of data to analyse the segregation or spatial integration of immigrants' residences. The online data, which is intended for other purposes, is immense, and constantly updated. These studies offer a significantly reduced-cost form of access to near real-time information, with study areas on a global scale. "We hope that this first work will open up the possibility for future use of this data to study integration. We also hope that it

may be a valuable complement beyond the scientific community for managers and public authorities who are in charge of immigration," concludes Ramasco.

