An international team led by a researcher from IMDEA Networks counts more than two million current symptomatic cases of Covid-19 in Spain

Researcher Antonio Fernández Anta has launched a study to estimate the actual incidence of the pandemic in eleven countries

The method used by the scientific team consists of conducting open and anonymous surveys of the population, promoted through social networks

In order to properly manage the Covid-19 pandemic, governments and the scientists who advise them need the most accurate data possible on the situation. The problem is that such data do not exist. The only data they have are those provided by laboratory tests. In other words, governments know how many people have tested positive, how many are in hospitals and how many are being treated over the phone. But by now it is more than evident that these are not the real data of the pandemic. Scientists around the world are now devising formulas to estimate figures that are more realistic. An international team led by researcher Antonio Fernández Anta of the IMDEA Networks institute is one of them.

This team has launched the @CoronaSurveys project to "Measure the iceberg". This objective is totally graphic: the tip of the iceberg can be seen and that is the number of confirmed cases, but below that tip, submerged beneath the waters, is the rest of the block of ice, of uncertain dimensions, and which could sink any ship that does not know its real size. To achieve this, the team led by Antonio Fernández Anta launched the first surveys on Twitter more than two weeks ago.

Currently, a daily survey is being carried out, open to the participation of all. It asks only three questions: from which region can you provide data (understanding that region can mean a country or a smaller administrative area); how many people do you know in that area; how many people in that area do you know who have symptoms of Covid-19 or that may be associated with that disease? One of the objectives when designing the data search method was to scrupulously respect data protection laws. Consequently, the survey is anonymous and does not collect any information that could lead to the identification of either the participants or the patients counted.

At first, the survey was launched only in Spain but over the days since then more countries have been added: Portugal, Argentina, Chile, Cyprus, France, Germany, Greece, Italy, Japan, United Kingdom and the United States. The form of dissemination of the surveys was also expanded so now, in addition to using Twitter, they are also available through Facebook and LinkedIn.

The results are collected every day. In this way, the study not only provides an estimated daily
total number of symptomatic infected people, but also makes it possible to observe the evolution of the disease. The data provided by the surveys of Spain estimate that last Monday 30th March the number of people with symptoms of Covid in the country was 2,361,650, a much higher figure than the official one, which on that day was 85,195 confirmed Covid-19 patients in the country.

To make the estimate, the researchers take the answers obtained in the survey with some adjustments. First, they eliminate the answers that are very possibly incorrect, such as someone who knows a thousand people and all of them have symptoms. Then they calculate the proportion of people with symptoms in relation to the total. Finally, this ratio is applied to the entire population of the country. The research team is working to develop more sophisticated methods to compensate for possible biases in the data, but what has been observed is that the estimates obtained are close to those obtained with other indirect methods.

About Antonio Fernández Anta

Research Professor at IMDEA Networks since December 2010. Antonio Fernández Anta holds a PhD in Computer Science from the University of Louisiana at Lafayette and completed a postdoctoral stay at the Computer Science Laboratory of the Massachusetts Institute of Technology (MIT). His research is focused on networks and communication, parallel processing, algorithms and discrete mathematics.