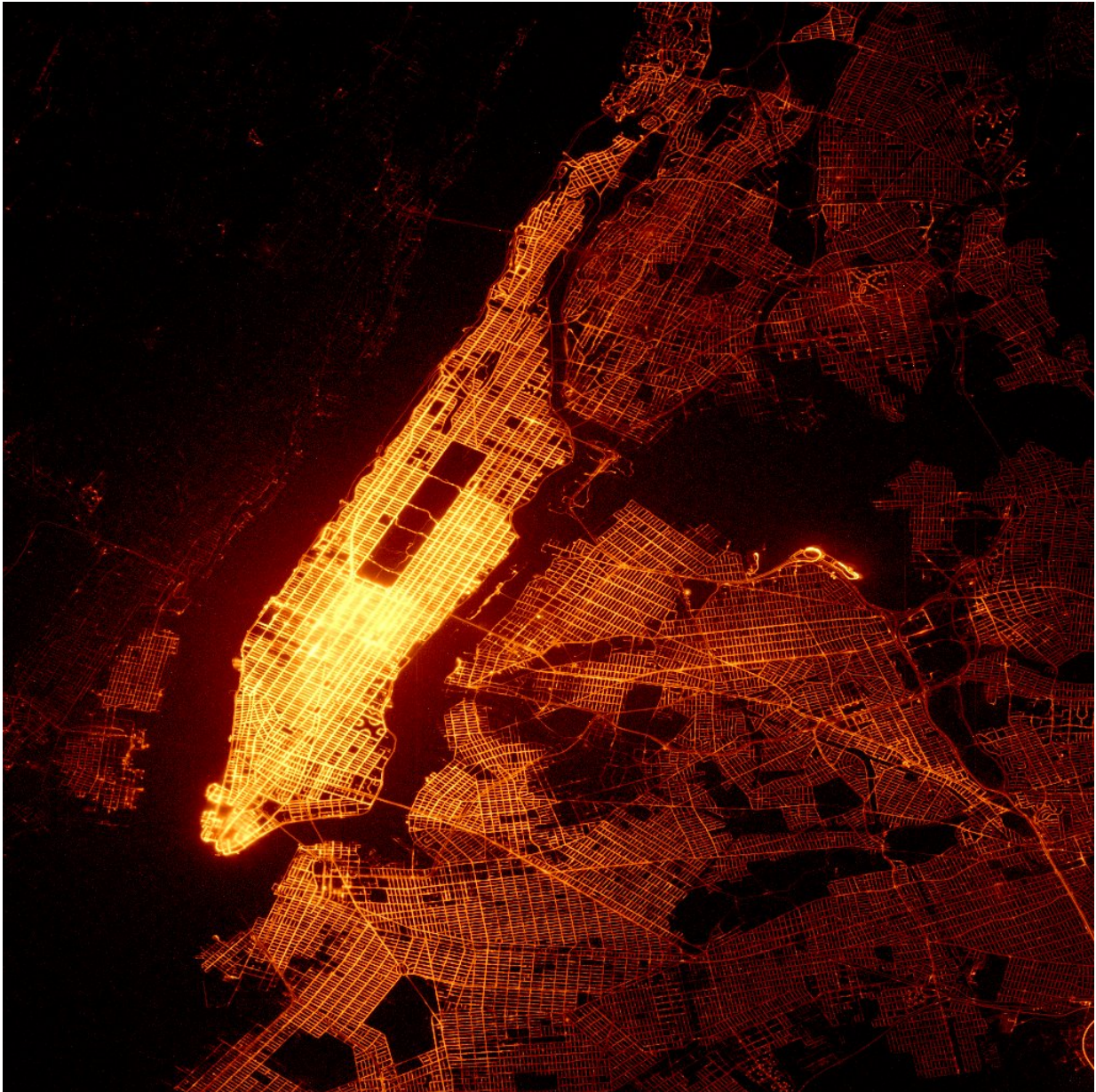# Big Data exploration in the era of Gaia

June 20 2018

This screenshot of an image is based on 1 billion rides of yellow cab users in

New York. Credit: Breddels & Veljanoski (RUG)

Two astronomers from the University of Groningen (The Netherlands) have developed a software library that can effortlessly generate visualisations based on hundreds of millions of data points. Maarten Breddels and Jovan Veljanoski initially developed their software to handle the enormous quantity of data from the Gaia mission. However, the software can also show patterns in other large data files. The software is open source and free to use. The researchers explain the ins and outs in an article that has been accepted for publication in the journal *Astronomy & Astrophysics*.

Breddels and Veljanoski call their [software](#) Vaex, which stands for "visualisation and exploration of big tabular datasets." The interactive software can generate visualisations of billions of data points in only one second. It behaves similarly to Google Maps. When panning or zooming, an updated or more detailed map appears almost immediately. However, Google Maps runs on fast, powerful servers, while Vaex works on a laptop.

The power of Vaex lies in the combination of several smart techniques. First, it uses a smart algorithm that maximises all available computing power. Then, it reads only the required data from the hard disk and sends it directly to the main memory of the computer. Finally, it is extremely memory efficient, and the working memory does not store unnecessary copies of the data.

Breddels has showcased Vaex live at several conferences. As an example, he used a dataset consisting of 1 billion entries related to the Yellow Cab taxis in New York City. He shows which taxi rides are the most lucrative, and where the taxis should wait in any part of the day to

maximise their profit. This example shows how Vaex can be interesting and beneficial for general applications outside of astronomy.

More information: Vaex: Big Data exploration in the era of Gaia. arxiv.org/abs/1801.02638