

# UCR-STAR: The UCR Spatio-temporal Active Repository

<https://star.cs.ucr.edu/>

Saheli Ghosh   Tin Vu   Mehrad Amin Eskandari   Ahmed Eldawy  
Department of Computer Science and Engineering, University of California, Riverside, USA  
{sghos006,tvu032,mamin021,eldawy}@ucr.edu

## Abstract

*This article describes the UCR Spatio-temporal Active Repository (UCR-STAR). UCR-STAR is a visual catalog for big spatial datasets. Rather than a boring tabular listing of datasets, it provides an interactive map interface that allows users to explore these datasets to assess their coverage, quality, and distribution. This article describes both the functionality of UCR-STAR as well as the underlying system architecture. We believe that this article can help the research community by explaining how to realize research ideas into a workable product.*

## 1 Introduction

Recently, there has been a tremendous growth in spatial data collection from various sources such as satellites, IoT sensors, smartphones, autonomous cars, and others. At the same time, there is a move for open data led by governments, non-profit organizations, and industry which makes hundreds of thousands of datasets publicly available. For example, Data.gov [4], which is maintained by the US government, hosts more than 140,000 datasets that are tagged as *geospatial*. Similarly, other governments, non-government organizations (NGOs), and companies keep releasing data on the web as part of the open data movement [3, 5, 13, 6, 11]. To browse these datasets, existing data repositories provide a plain listing of the datasets with references on how to access and download them, e.g., see Figure 1(a). Users will either have to *guess* what the dataset really contains or download these datasets, figure out how to import them into their favorite tool, before they can interact with the data.

To break from this old interface, this article describes UCR-STAR which provides an interactive web-based interface that allows users to interact with the datasets to assess their coverage, quality, and distribution before even downloading them. Figure 1(b) shows a screenshot of how UCR-STAR looks like. The majority of the screen is devoted to the map-based visualization of the dataset while a smaller part is used to list the datasets and its details. The map portion provides the standard map interactions such as pan and zoom. When a dataset is selected on the left, the map is updated in a fraction of a second to visualize the selected dataset no matter how big it is. Users can also search for datasets using a keyword search or advanced search based on the size or type of the dataset.

UCR-STAR is currently hosting more than 100 datasets with a total size of nearly one terabyte. We welcome requests to add additional datasets to the archive to further facilitate the access to these datasets and help the research community. The rest of this article gives more details about the architecture of UCR-STAR and the research behind it.











