UC Merced

Title

Towards Capturing Scientific Reasoning to Automate Data Analysis

Permalink

https://escholarship.org/uc/item/85d2d1xf

Journal

Authors

Gil, Yolanda
Khider, Deborah
Osorio, Maximiliano
et al.

Publication Date

2022

Peer reviewed

Towards Capturing Scientific Reasoning to Automate Data Analysis

Yolanda Gil
University of Southern California, Marina del Rey, California, United States

Deborah Khider
Information Sciences Institute, Los Angeles, California, United States

Maximiliano Osorio
University of Southern California, Los Angeles, California, United States

Varun Ratnakar
Information Sciences Institute, Los Angeles, California, United States

Hernan Vargas
University of Southern California, Los Angeles, California, United States

Daniel Garijo
University of Southern California, Los Angeles, California, United States

Suzanne Pierce
The University of Texas in Austin, Austin, Texas, United States

Abstract

This paper describes an initial cognitive framework that captures the reasoning involved in scientific data analyses, drawing from close collaborations with scientists in different domains over many years. The framework aims to automate data analysis for science. In doing so, existing large repositories of data could be continuously and systematically analyzed by machines, updating findings and potentially making new discoveries as new data becomes available. The framework consists of a cycle with six phases: formulating an investigation, initiating the investigation, getting data, analyzing data, aggregating results, and integrating findings. The paper also describes our implementation of this framework and illustrates it with examples from different science domains.