

HDR² From Harnessing the Data Revolution to Harvesting the Data Revolution

Contribution ID: 1

Type: **not specified**

Keynote talk: Let's make Data Science more like Science

Wednesday 26 October 2022 13:30 (30 minutes)

Data-driven exploration has revolutionized science and led to the establishment of Data Science as a new discipline that integrates approaches from computer science – including data management, visualization, machine learning – statistics, applied mathematics, and many application domains. I will give my perspective of how the field emerged and evolved over the past decade, and the virtuous cycle it has enabled which fuels interdisciplinary research that derives new problems and solutions for multiple areas.

A critical challenge in data science is how to empower domain experts to engage in data-driven exploration. While computing and storage are essentially free and data is abundant, we need humans in the loop to generate insights from data. Toward this end, there have been many efforts that aim to democratize data science, and today it is relatively easy to derive results. But it is also easy to derive incorrect results. I will give examples of common mistakes and problems that can affect results and that are hard to detect, and argue that, akin to natural systems, we must experiment with and observe data science pipelines to understand their behavior, assess the validity and properly explain their results. In essence, we need to make Data Science more like science and work towards democratizing trust and robustness.

Data & Cyberinfrastructure

Education and Outreach

Research

Presenter: FREIRE, Juliana (New York University)

Session Classification: Keynote Speakers