

## Data Science Seminar Series

### Large-scale graph representation learning and computational biology through sparse matrices



#### **Aydin Buluc, Ph.D.**

**Staff Scientist and Principal Investigator  
Lawrence Berkeley National Laboratory  
Adjunct Assistant Professor  
University of California, Berkeley**

**Date:** Wednesday, April 14th, 2021

**Time:** 4:00 PM – 5:00 PM EDT

**Location:** Zoom Virtual Room

**Web Link:** <https://njit-institute-for-data-science.eventbrite.com>

Solving systems of linear equations have traditionally driven the research in sparse matrix computation for decades. Direct and iterative solvers, together with finite element computations, still account for the primary use case for sparse matrix data structures and algorithms. These sparse "solvers" often serve as the workhorse of many algorithms in spectral graph theory and traditional machine learning.

In this talk, I will be highlighting two of the emerging use cases of sparse matrices outside the domain of solvers: graph representation learning methods such as graph neural networks (GNNs) and graph kernels, and computational biology problems such as de novo genome assembly and protein family detection. A recurring theme in these novel use cases is the concept of a semiring on which the sparse matrix computations are carried out. By overloading scalar addition and multiplication operators of a semiring, we can attack a much richer set of computational problems using the same sparse data structures and algorithms. This approach has been formalized by the GraphBLAS effort. I will illustrate one example application from each problem domain, together with the most computationally demanding sparse matrix primitive required for its efficient execution. I will also cover novel parallel algorithms for these sparse matrix primitives and available software that implement them efficiently on various architectures.

Aydin Buluç is a Staff Scientist and Principal Investigator at the Lawrence Berkeley National Laboratory (LBNL) and an Adjunct Assistant Professor of EECS at UC Berkeley. His research interests include parallel computing, combinatorial scientific computing, high performance graph analysis and machine learning, sparse matrix computations, and computational biology. Previously, he was a Luis W. Alvarez Postdoctoral Fellow at LBNL and a Visiting Scientist at the Simons Institute for the Theory of Computing. He received his PhD in Computer Science from the University of California, Santa Barbara in 2010 and his BS in Computer Science and Engineering from Sabanci University, Turkey in 2005. Dr. Buluç is a recipient of the DOE Early Career Award in 2013 and the IEEE TCSC Award for Excellence for Early Career Researchers in 2015. He was a founding associate editor of the ACM Transactions on Parallel Computing.