Abstract: Huge amounts of real-time information are generated from a great variety of sources such as mobile/web applications, e-commerce, social networks, or sensor networks. Currently popular data management approaches and frameworks to extracting and understanding patterns from multiple streams take a simplistic view of the data, which ignores multi-relational and multi-dimensional characteristics associated with these streams. Furthermore, the visualization and user-interaction components of these systems are limited to visualizing the outcome of stream processing results. Visual analysis represents a new form of analysis where the user has more control and interactive capabilities either to dynamically change the visualization, analytics or data management processes. Applying visual analysis to such data streams presents many benefits such as automated analysis and interactive visualization. VAStream, established at UL Lafayette via an NSF- MRI award, is a big data system that provides both a big data stream processing platform along with more interactive visualization capabilities. The system environment consists of hardware and software modules to optimize streaming data workflow (that includes data ingest, pre-processing, analytics, visualization, and collaboration components).

Topic detection methods, when applied to text streams from social media, can automatically recognize the onset of new events by modeling them as emerging topics of interest in these conversations. This is very useful for situations such as public safety incidents or mass emergencies, where events are dynamic and early detection is desirable. Similarly, during emergencies, like hurricanes, information is flooded from a wide variety of sources and decision makers should be able to identify emerging trends and patterns, such as water level in river streams, for effective response. Recent research at the NSF-sponsored Center for Visual and Decision Informatics (CVDI), a multi-institutional Industry University Cooperative Research Center (IUCRC), takes a robust strategy by modeling such complex data as time-evolving graph data streams. In this presentation, we introduce the system architecture of VAStream in terms of hardware, and software modules. We also elaborate on our work in the context of the two above-mentioned streaming applications, to demonstrate our overall visual analytics approach and to present experimental findings relating to the system's performance.

Bio: Raghavan is the Alfred and Helen Lamson Endowed Professor of Computer Science at the Center for Advanced Computer Studies, University of Louisiana at Lafayette, the Director of the Center for Visual and Decision Informatics, a multi-institutional, NSF-sponsored, Industry University Research Center (IUCRC), and a co-director of the Laboratory for Internet Computing. His research interests are in information retrieval and extraction, data and web mining, multimedia retrieval, data integration, and link discovery. He has published around 300 peer-reviewed research papers, has served as major adviser for 32 doctoral students and has garnered $15 million in external funding. Raghavan is an IEEE Life Senior Member and an ACM Distinguished Scientist.