David Sathiaraj

davids@trabus.com

http://ds153.github.io

Education

Louisiana State University (LSU)   Baton Rouge, LA
- Ph.D., School of Electrical Engineering & Computer Science
  Research Area: Data Mining, Data Visualization, Big Data Analytics

Louisiana State University   Baton Rouge, LA
- M.S., Department of Computer Science

Louisiana State University   Baton Rouge, LA
- M.S., Department of Industrial Engineering

Osmania University   Hyderabad, India
- B.E., Department of Mechanical Engineering

Work History

Director of Data Science   January 2020 - Present
- Trabus Technologies   San Diego, CA

  – Head of AI/Data Science - applying AI to Environmental Informatics, Transportation, Cybersecurity and Wireless Technologies - https://www.trabus.com/ai. Lead a team of AI/ML engineers and full-stack developers in an Agile environment to develop cutting edge AI and data-driven products that span various data domains.

  – Chief Data Scientist - Develop AI models towards the lifecycle maintenance planning for the COLUMBIA-class submarines (SHAPM project - Shipbuilding Acquisition Program Management).

  – Lead AI and software engineering team for the River Information Services Enterprise (RISE) - https://www.trabus.com/rise. Led projects related to data-driven preventive maintenance planning (Lock Status Analysis Tool, https://lsat.trabusrise.com) at USACE managed lock and dams and water reservoirs and AI-based river level management software to aid US Inland Water Transportation through the Waterway Action Plan (WAP, https://wap.trabusrise.com).

  – Led the design and development of the first AI-based Voyage Planner for the Inland Waterways, RippleGo (https://ripplego.com).

  – Led a TRABUS and Texas Transportation Institute (TTI) research team to study the influence of solar glare and inclement weather on pedestrian crashes.

  – Led the Trabus team in development of *devsecops* and automated provisioning tools in support of the US Army's Labyrinth Cybersecurity program.

  – Successfully led a Trabus - Texas A&M University team to bid for and launch the NOAA-funded Southern Regional Climate Center (SRCC) at Texas A&M University. Led a fast-track IT effort to launch a website, climate data portal and full-stack data operational center for the new SRCC (https://www.srcc.tamu.edu).

  – Lead Deep Learning Algorithm development efforts for the river level prediction tools, bridge air gap predictions, travel time predictions and other marine transportation and logistics' tools.

  – Lead development of the Hydrological and Environmental Data Repository (HEDR) - a large river and environmental data warehouse with GraphQL based API

  – Develop and lead software development efforts for generating analytics for efficient management of Turnarounds in Oil and Gas Industry (https://picassoanalytics.com)

- **Adjunct Professor**   June 2022 - Present

  Dept. of Mathematics, Information & Computer Science (MICS)

  Point Loma Nazarene University   San Diego, CA

  - Developed and taught a senior-level course in Data Science. Topics covered: Data science analysis and techniques using Python, data acquisition and storage (*PostgreSQL, MongoDB, Redis*), development of data science dashboards using web-frameworks such as Django, data analytics using *pandas, numpy and scipy*, classification and clustering (machine learning) algorithms and deep learning using *keras, tensorflow and scikit-learn.*

- **Adjunct Professor**   September 2021 - Present

  Geovisualization Center

  San Diego State University   San Diego, CA

- **Assistant Professor**   April 2015 - August 2020

  Geography & Anthropology, LSU   Baton Rouge, LA

  Associate Director   August 2013 - August 2020

  NOAA Southern Regional Climate Center, LSU   Baton Rouge, LA

  - Developed a Big Data, systems engineering program and curriculum focussed on Computational Geosciences, conduct research at the intersection of Computer Science, Environmental Informatics, Transportation and other Big Data Domains.

  - Manage one of 6, national Regional Climate Centers at LSU - 24/7 operational data warehouse for environmental information.

  - Designed and Deployed LSU's first Science DMZ network (https://fasterdata.es.net/science-dmz/)- a 40 Gbps network to transmit Big Data sets and visualization products between LSU and other partner sites.

  - Developed a data-driven analysis to study impacts of extreme weather on traffic flow in Atlanta, GA http://wxtraffic.srcc.lsu.edu

  - Developed machine learning based predictive analytics platform for rain and river stage interactions. http://rainriver.lsu.edu

  - Developed a water reservoir information system and database. http://reservoir.srcc.lsu.edu

  - Developed AI and healthcare analytics solutions and data mining of clinical/healthcare data in research collaboration LSU Health Sciences Center, New Orleans, LA.

  - Lead software developer at the SRCC for the ACIS (Applied Climate Information System) project - a nationwide effort by the six NOAA Regional Climate Centers in US to provide an integrated climate data services and systems. ACIS is a large scale 24/7 data collection and access engine providing data services via APIs to a host of governmental agencies and industry. ACIS is a multi-layer climate data warehouse comprising of numerous climate data sets, metadata and statistical analysis routines. Information about the ACIS software library can also be found here: http://www.rcc-acis.org, http://data.srcc.rcc-acis.org/doc/

  - Lead a team of software and systems professionals to develop and establish real-time environmental data delivery engineering systems. http://climdata.srcc.lsu.edu and http://hrly.lsu.edu

  - Provide vision and oversight for SRCC's IT infrastructure and climate data products.

  - Lead developer for the Southern Climate Impacts Planning Program (SCIPP) http://southernclimate.org/pages/data-tools, one of NOAA's Regional Integrated Sciences and Assessments programs.

- **Founder and CEO** June 2015 - December 2019
  Pecan Analytics, a Data Analytics Startup.
  http://www.pecananalytics.com

  - Developed a fast paced software start-up from the ground-up with patented and patent-pending technologies.
  - Focus on Healthcare Analytics, Cybersecurity, Political Analytics and Environmental Informatics.
  - Awarded a startup seed grant by the LSU Research and Technology Foundation ($LIFT^2$ grant).
  - Designed and Developed Big Data Systems engineering and Analytical/Quantitative solutions for data-driven political campaigns. In 2014, used the same data analytics in the winning campaign of Senator Bill Cassidy in Louisiana.
  - Developed machine learning algorithms and visual analytics engines for fast mining and analysis of large data sets. http://vis.pecananalytics.com
  - Developed a data-driven enterprise software application platform with automated sensor data collection systems, machine learning algorithms and browser-side analytics and visualization.

- **IT Manager** August 2009 - July 2013
  NOAA Southern Regional Climate Center Baton Rouge, LA

  - Collect, archive, analyze and visualize large climate data sets.
  - Be SRCC's lead in development of climate data products using the ACIS system
  - Lead Developer for SCIPP - developed visualization and data analytic tools for large climate data sets (using Python, Redis, Postgresql, Mongodb and Tornado). Some of the tools developed are hosted here:
    http://www.climate.gov/decision-support/department/decision-support-tools and http://www.southernclimate.org/data.php.
  - Designed data analytics platform combining hazards data sets such as storm surge, extreme events theory and return frequency analysis (Project collaborations with US Department of Energy, Oak Ridge National Lab and Pacific Northwest National Lab, a SCIPP project, http://surge.srcc.lsu.edu). The storm surge data and predictive analytics platform has also led to interactions with several leading insurance companies that are interested in using the predictions to set rates for clients on the US Gulf Coast.
  - Designed and developed parallel computation libraries for generation and visualization of large, gridded climate data sets.
  - Developer of visualization tools for analyzing climate data hazard data sets (such as storm surge, drought, hail, tornadoes, hurricanes, wind) and climate extremes (such as extreme heat and cold conditions, long wet or dry spells). Tools used included Redis and PostGIS.
  - Development of decision support tools for drought planning and understanding historical climate trends.
  - Manage a Linux-based super-computing cluster for data analysis and visualization.
  - Developer of the Datzilla Project - a data reporting tool for the National Weather Service and National Climatic Data Center (a customization of Bugzilla).
  - Developer of map-based and chart-based data analytic tools using spatiotemporal software such as Redis, Mongodb, PostGIS and R.

- **Application Systems Engineer** June 2001 - August 2009
  NOAA Southern Regional Climate Center Baton Rouge, LA

- Developed data ingest procedures using Python for large climate data sets and streams.
- Developed client-server tools for ACIS using Python, PostgreSQL, RPC and NetCDF(a scientific data access library). Implemented web applications that extract and provide climate data using tools such as Twisted, Tornado and Python.
- Developed and deployed web-services and map-based software for Northrop Grumman's (Omaha) decision support systems portal.
- Developed GIS map-based data analytics software using Javascript, R, Python and Postgis.
- Developed and maintained parsers for satellite and network based data streams. Data ingest systems were automated, robust and fault-tolerant.
- Developed pyPloticus - a Python module for the Ploticus graphing library (written in C).
- Migrated legacy software libraries (based on Fortran and C) from an IRIX environment to Python on a Linux environment.
- Conducted several tutorials and talks for using the ACIS library and on the use of Python and R in scientific programming.

Patents and Publications

- RippleGo - An AI-based Voyage Planner for US Inland Waterways, D Sathiaraj, A Smith, E Rohli, C Hsieh, A Salindong, N Woolsey, A Tec, 2023 IEEE Conference on Artificial Intelligence (CAI), 372-373, 2023

- Near-term forecasting of water reservoir storage capacities using long short-term memory, E Rohli, N Woolsey, D Sathiaraj, Environmental Data Science, 2, E30, doi:10.1017/eds.2023.25.

- Artificial-intelligence-based waterway information system, US Patent 11,620,523, https://patents.google.com/patent/US20220245456A1

- A Predictive Analytical System And Method, 2018, https://patents.google.com/patent/US20190205770A1

- Predictive Analytics-based U.S. Inland Waterways Voyage Planning Analysis Tool (VPAT), Joseph Celano and David Sathiaraj, Predictive Analytics-based U.S. Inland Waterways Voyage Planning Analysis Tool (VPAT), Technical Report, Department of Energy, United States, 2021

- Predicting climate types for the Continental United States using unsupervised clustering techniques, D Sathiaraj, X Huang, J Chen, Environmetrics, 2019;30:e2524.

- Data-driven analysis on the effects of extreme weather elements on traffic volume in Atlanta, GA, USA, D Sathiaraj, T Punkasem, F Wang, DPK Seedah, Computers, Environment and Urban Systems, 2018, 72, 212-220.

- Modeling Hydroclimatic Change in Southwest Louisiana Rivers, ZG Xue *et. al.*, Water, 10 (5), 596.

- Deriving Data-driven Insights from Climate Extreme Indices for the Continental US, Workshop on Data Mining in Earth System Science, International Conference of Data Mining, 2017.

- Improving Predictive Accuracy In Elections, Big Data, 5 (4), 325-336.

- A Review of Tropical Cyclone-Generated Storm Surges: Global Data Sources, Observations and Impacts, H. F. Needham, B.D. Keim and D. Sathiaraj, Reviews of Geophysics, 2015.

- System, Method And Computer Program Product For Data Mining Applications (Patent Issued: May 13, 2014) - http://www.google.com/patents/US8725663

- On Identifying Critical Nuggets Of Information During Classification Tasks, David Sathiaraj and E. Triantaphyllou, IEEE Transactions on Knowledge and Data Engineering, 2013, 25(6), 1354-1367.

- A Global Database of Tropical Storm Surges, Needham, H.F., B.D. Keim, D. Sathiaraj, and M. Shafer, EOS Transactions, 2013, 94(24), 213-214.

- Spuriously induced precipitation trends in the southeast United States, Jason Allard, Barry D. Keim, Jessica E. Chassereau and David Sathiaraj, Theoretical and Applied Climatology, 2009, 96(1-2), 173-177.

- Partial Duration Series Rainfall Events at El Paso, Texas, Faiers. G. E., B. D. Keim, K. Jammigumpula, D. Sathiaraj, Pennsylvania Geographer, Summer 2005.

- Common parts grouping heuristic: an iterative procedure to cell formation, David Sathiaraj and Bhaba R. Sarker, International Journal of Production Planning and Control, 2002, 13(5), 481-489.

Conferences

- Prediction of River Stages along US Inland Waterways Using Deep Neural Nets, 101st American Meteorological Society Annual Meeting, 2021.

- Interactive climate data analytics for hazards mitigation, planning and emergency management, David Sathiaraj, Climate Informatics, 2013.

- Storm Surge Return Periods for the U.S. Gulf Coast, Hal Needham, Barry Keim, David Sathiaraj, Mark Shafer, ATC-SEI Advances in Hurricane Engineering Conference, October 2012.

- Storm Surge Return Periods for the United States Gulf Coast, , Hal Needham, Barry Keim, David Sathiaraj, Mark Shafer, World Environmental and Water Resources Congress in Albuquerque, NM, May 2012.

- Storm Surge Return Periods for the U.S. Gulf Coast, Hal Needham, Barry Keim and David Sathiaraj, AAG Annual Meeting, New York, February 2012.

- Building Climate Date-Driven Information Tools Using Python, David Sathiaraj, Joel James, Yixin Luo and Jinwoong Yoo, 92nd American Meteorological Society (AMS) Annual Meeting, January 2012.

- Creating Custom Tiles For Maps, Louisiana Remote Sensing and GIS Conference (LARSGIS), April 2009, Baton Rouge, LA.

- Spatial Analysis and Visualization of Climate Data Using R, useR 2008, August 2008, University of Dortmund, Germany.

Theses

- On Identifying Critical Nuggets Of Information During Classification Tasks, David Sathiaraj, PhD Dissertation, School of Electrical Engineering and Computer Science, LSU, March 2013.

- Evaluation of surface-fitting algorithms for reverse engineering of free-form surfaces, MS Thesis, Department of Industrial Engineering, LSU, October 2000.

Grants and Contracts

- Co-PI, River Information Services Enterprise contract with US Army Corps of Engineers, 2017-Present - $5,000,000

- PI, NOAA Southern Regional Climate Center sub-contract with Texas A&M Transportation Institute, 2021-2026 - $1,400,000

- Co-PI, Dept. of Energy Phase-1 SBIR award - $150,000

- Amazon Research Award, $10,000/year worth of Amazon AWS cloud credits.

- co-PI, NOAA RISA grant, 2018-2022, $2.6 million ($650k/year).

- PI, LIFT^2 grant, PI, Invention to Commercialization grant to develop data science algorithms and visual analytics technologies for the political domain, $43,930.80).

- PI, Lakvold Group - Analyze real-estate data and apply data mining strategies, Jan-Dec 2016, $40,700.

- co-PI, NOAA Southern Regional Climate Center Grant, October 2013-Present, approx. $314,000/year.

- UCAR Equipment Grant, 2010-11 ($7,558).

Technology Stack

Languages: Python, Javascript, C, PHP, Go, R, SQL, bash/zsh

Databases: AWS Redshift, AWS RDS, MongoDB, Redis, InfluxDB, PostgreSQL, MySQL, Memcached

GIS: PostGIS, QGIS

DevOps and Containers: GitHub Runner, Ansible, Docker, Kubernetes

Libraries: Keras, Tensorflow, Weka, scikit-learn, sqlalchemy, jQuery, pandas, requests

Visualization: d3.js, dc.js, mapbox/leaflet.js, highcharts/echarts, matplotlib, plotly, GDAL

Web Frameworks: Django, Flask, Tornado, Twisted, React

API: REST, GraphQL

Web Servers: Apache, Nginx

Cloud: Digital Ocean, AWS (Lambda, RDS, SQS, SES, SNS, S3, EC2)

Operating Systems: Redhat Linux, Ubuntu, CentOS and Mac OS X

Awards and Certifications

- Professional Scrum Master, Scrum.org, January 2022

- Amazon Research Award, AWS Credits for Research, Fall 2018.

- Louisiana Technology and Research Foundation's, LIFT^2 Grant winner - Fall 2015.

- The 2011 LSU Ellinor H. Behre Prize in Science Writing, The Sigma Xi Research Society (LSU Chapter).

- NSF Travel Scholarship for R conference, useR 2008 (Dortmund, Germany).