RISHI VERMA

E-mail: (Private) riverma@apache.org,
(NASA) rishi.verma@jpl.nasa.gov
Website: https://riverma.github.io/

EXPERTISE

- Design and implementation of mission-critical, high-availability software data systems
- Computing architectures for Big Data and real-time streaming data
- Software development lifecycle efficiency improvement & innovation

EDUCATION

University of Southern California, Los Angeles, CA, USA
- Master of Science: Computer Science

Indiana University, Bloomington, IN, USA
- Bachelor of Science *with Distinction*: Computer Science
- Bachelor of Science *with Honors*: Computer Information Systems
- Minor: Mathematics

Peking University, Beijing, China
- Int'l Business Course: Guanghua School of Management

EXPERIENCE

U.S. Department of State Fulbright Program **2019 – present**
- Specialist
- Selected by the U.S. Department of State Fulbright peer-review committee as a national specialist in the field of computer science
- Offer expert consultation in the field of information technology for international organizations

NASA Jet Propulsion Laboratory, Pasadena, CA, USA **2008 – present**
Deep Space Network Machine Learning Research
- Principal Investigator (PI) (active)
- Lead a team of two PhD researchers in developing innovative machine learning based algorithms for detecting weather, equipment, and radio-frequency-interference (RFI) anomalies in real-time streaming network data from NASA's Deep Space Network

Advanced Rapid Imaging and Analysis
- Systems Engineer (active)
- Provide systems design guidance and leadership to a team of over ten experts in the fields of computer science, machine learning, and natural hazard detection to rapidly identify global natural hazards and provide near real-time satellite overlay imagery and products to the global scientific and policy communities

Planetary Data System Imaging Node

- Systems Engineer (active)
- Lead a team of five software developers and two system administrators in designing, maintaining, and deploying key NASA imaging and archival software to house imagery data from all NASA planetary missions
- Responsible for overall project schedule formulation and adherence as well as leadership in infusing innovate technological and process improvements including cloud computing and machine learning

Deep Space Network Complex Event Processing

- Cognizant Design Engineer (active)
- Lead a team of three software developers to design, develop, and deliver a revolutionary real-time "Big Data" software system to ingest, filter, store, and visualize all of NASA's Deep Space Network monitor and control data – supporting over 20 active NASA missions
- Selected through NASA's rigorous project lifecycle review to be deployed worldwide to all NASA Deep Space Communication Complexes (DSCCs)

Megacities Carbon Project Portal Data System

- Cognizant Engineer
- Led a team of two software developers and two system administrators in designing, building, and maintaining an innovative real-time climate change focused sensor data management system and web-portal
- Project tapped by United Nations as a top "Big Data project to watch for" (2014)
- Coordinated and built collaborations with Aalborg University (Copenhagen, Denmark)
- Live portal accessible at: https://megacities.jpl.nasa.gov/portal/

Early (Cancer) Detection Research Network Laboratory Catalog & Archive Service

- Cognizant Engineer
- Led engineering activities on a National Institutes of Health (NIH) National Cancer Institute (NCI) and NASA collaboration to build a virtual laboratory environment for nationally dispersed cancer biomarker research data

Time Correlation Service

- Software Engineer
- Designed and implemented the primary means by which select NASA deep-space missions estimate on-board spacecraft clock time
- Actively used by NASA for precisely timing the execution of commands to select spacecraft

CO_2 Virtual Science Data Environment

- Software Engineer
- Led "backend" engineering activities on a comprehensive effort at bringing together models, data, and tools necessary to perform research on atmospheric CO_2
- Accessible at: http://co2.jpl.nasa.gov/

Defense Advanced Research Project Agency (DARPA) XData Initiative

- Software Engineer
- Led software data system engineer activities on a project aiming to meet emerging data challenges by developing computational techniques and software tools for processing and analyzing large, imperfect and incomplete data

Defense Advanced Research Project Agency (DARPA) Memex (Ending Human Trafficking) Initiative

- Software Engineer
- Aided in the development of a software data system aiming to provide search and visualization tools for U.S. law enforcement to help identify human traffickers and human trafficking victims

Airborne Snow Observatory

- Software Engineer
- Aided in the construction of a real-time software data system for an innovative airplane-based LIDAR system for measuring and estimating snowfall snow-water-equivalents and composition in the California Sierra Nevada mountain range

Pasadena Complete Streets Coalition, Pasadena, CA **2014 – 2018**

- Steering Committee Member
- Provided management guidance to coalition organization dedicated to improving street infrastructure within Pasadena, California
- Developed and led effort for open sourcing technology projects within coalition, including websites and other software products

Orvium, Tallinn, Estonia **2018**

- Technical Advisor
- Provided expert technical advisory to a start-up focused on providing a "decentralized social platform for scientific funding, collaboration, and publications management based on Blockchain and Artificial Intelligence"

Apache Software Foundation, Wakefield MA, USA **2011- 2016**
Object Oriented Data Technology (OODT)

- Member of Project Management Committee (PMC)
- Part of select decision making and project leadership team for this open source effort

PUBLICATIONS

4th Planetary Data Workshop, Flagstaff, AZ **2019**
Lead author: *"Archive Manager and Processor (AMP)"*

Planetary Science Informatics and Data Analytics Conference, St. Louis, MI **2018**
Lead author: "*Archive Inventory Management System (AIMS) – A Fast, Metrics Gathering Framework for Validating and Gaining Insight from Large File-Based Data Archives"*

Third Planetary Data Workshop, Flagstaff, AZ **2017**
Lead author: "*Next Generation Parallelization Systems for Processing and Control of PDS Image Node Assets"*

13th International Conference on Space Operations, Daejeon, South Korea **2016**
Co-author: "*Achieving Fast Operational Intelligence in NASA's Deep Space Network Through Complex Event Processing"*

IEEE International Conference on Big Data (Big Data) **2015**
Co-author: "*SciSpark: Applying In-memory Distributed Computing to Weather Event Detection and Tracking"*

IEEE 16th International Conference on Information Reuse and Integration **2015**
Lead author: "*Extending Spark Analytics through Tika-Based Information Extraction and Retrieval"*

International Joint Conference on Artificial Intelligence in Space **2015**
Co-author: "*Radio Array of Portable Interferometric Detectors (RAPID): Development of a deployable multiple application radio array"*

International Conference on Electromagnetics in Advanced Applications, Torino, Italy **2015**
Co-author: "*Radio Array of Portable Interferometric Detectors (RAPID): Development of a deployable multiple application radio array"*

American Geophysical Union Fall Meeting, San Francisco, CA **2014**
Lead author: "*A Distributed, Open Source based Data Infrastructure for the Megacities Carbon Project"*

International Geoscience and Remote Sensing Symposium (IGARSS), Quebec, Canada **2014**
Co-author: "*24 Hour near real time processing and computation for the JPL Airborne Snow Observatory"*

IEEE Computer-Based Medical Systems (CBMS), New York, NY **2014**
Lead author, *"A Laboratory-Targeted, Data Management and Processing System for the Early Detection Research Network"*
ACM Workflows in Support of Large-Scale Science, New York, NY **2013**
Co-author: *"Time-bound analytic tasks on large datasets through dynamic configuration of workflows"*

American Geophysical Union Fall Meeting, San Francisco, CA **2012**
Lead author: *"Developing a GIS for CO_2 analysis using lightweight, open source components"*

IEEE International Conference on Information Reuse and Integration, Las Vegas, NV **2012**
Co-author: *"Developing an Open Source, Reusable Platform for Distributed Collaborative Information Management in the Early Detection Research Network"*

The Federation of Earth Science Information Partners (ESIP), Washington, D. C. **2012**
Co-author: *"Carbon Dioxide and GIS : CO2 Virtual Science Data Environment"*

American Geophysical Union Fall Meeting 2011, San Francisco, CA **2011**
Lead author: *"A virtual science data environment for carbon-dioxide observations"*

PROPOSALS

National Space Technology Applications Advanced Concepts **2017**
PI: DSN Data-Driven Network Monitoring Context for Operations

Earth Science Directorate Spontaneous Engineering Improvement **2016**
PI: JPL Software Market – a software catalog for promoting reuse through efficient search

Strategic Initiative Proposal for the Research & Technology Development Fund **2014**
Co-PI: Archiving, Processing and Dissemination for the Big Data Era

NASA ACCESS Solicitation **2013**
Co-PI: ICARUS: Integrated Climate Analysis distRibUted Services

KEY PRESENTATIONS

Kafka Summit, San Francisco, CA **2019**
Presenter: *"Mission-Critical, Real-Time Fault-Detection for NASA's Deep Space Network using Apache Kafka"*

Robots and Bots Development Talk + Workshop, Pasadena, CA **2019**
Presenter: *"Planetary Data System: Data Access for the Software Development Community"*

NASA/JPL Asian American Council, Pasadena, CA **2019**
Presenter: "The Power and Profoundness of Sound"

Canberra Deep Space Communications Complex (CDSCC), Canberra, Australia **2018**
Presenter: "Complex Event Processing"

NASA/JPL Laboratory Management Council (LMC), Pasadena, CA **2017**
Presenter: "Data Driven Network Monitoring Context for DSN Operations"

Breakthrough Listen / Search for Extra Terrestrial Intelligence (S.E.T.I) Workshop, **2016**
University of California at Berkeley, Berkeley, CA
Presenter: "Supportive Data Architectures"

NASA/JPL Director's Review and Discussion (DRD), Pasadena, CA **2015**
Presenter: "Real-time Big Data Processing for the DSN"

Aalborg University, Aalborg, Denmark **2014**
Presenter: "Megacities Carbon Project"
Co-presenter: "Leveraging the power of Apache for Science"

New School of Design, New York City, NY **2014**
Presenter: "Megacities Carbon Portal"
ApacheCon North America 2013, Portland, OR **2013**
Presenter: *"Searching for cancer biomarkers with Apache OODT"*

NASA Earth Science Data System Working Group Meeting, Annapolis MD **2012**
Presenter: *"Developing a GIS for CO2 analysis using lightweight, open source components"*

NMI Build and Test Workshop, Madison, WI **2008**
Co-presenter: *"Open grid computing environments: building and testing on NMI"*

SKILLS

- Technologies: Apache Kafka, Elasticsearch, Apache Spark, Ansible, Lambda architectures, Apache OODT, Dockerization / multi-app orchestration, Amazon Web Services (S3, Lambda, EC2), Apache Solr, Apache Nutch, RESTful services, CentOS/Ubuntu, Maven, Adobe Illustrator, Linux administration
- Programming: Scala, Java, Node JS, Android, Ruby

LANGUAGES

- English: full fluency
- Hindi: full fluency
- French: basic knowledge
- Japanese: basic knowledge