Keith S. Beattie

(510) 205-3889
KSBeattie@lbl.gov

SCIENTIFIC INTERESTS

My interests lie at the intersection of scientific research and software engineering. How software is developed and used in a scientific context, who develops it, and how the research, software and people can all mutually achieve their respective scientific mission, engineering impact and personal career goals. My personal goal is to see scientific software and those who develop and support it, be seen throughout the research community, as first class citizens in the scientific mission.

SOFTWARE ENGINEERING COMMUNITY ENGAGEMENT

ASCR Workshop on the Science of Scientific-Software Dec 2021
Participated in ASCR led workshop exploring the need and potential for future funding for sustainable scientific software

Better Scientific Software (BSSw) 2019 - Present
Fellowship Honorable Mention, 2021

United States Research Software Engineer Association (US-RSE) ***2019 - Present***

LBNL-RSE Working Group ***2019 - Present***
Founded a CSE "Cabal" to to bring together lab employees doing primarily software engineering or computational related work into a common lab-wide community with the goals of supporting their individual careers and advocating for the importance of developing sustainable software with respect to the research mission of the lab, the DOE and the larger scientific community.

EXPERIENCE

Lawrence Berkeley National Laboratory, Berkeley, CA **— Dec 2022 to Present**
Group Lead: Sustainable Software Engineering

Founded and leads a new group within the Scientific Data Division focusing on the tools, processes, people and careers that are required to create and maintain software that supports scientific achievement with an emphasis on stability and reproducibility.

Lawrence Berkeley National Laboratory, Berkeley, CA **— Aug 2002 to Present**
Computer Systems Engineer II - IV

> **Lux Zeplin - Aug 2020 to Present**
> USDC Deputy, Data Movement Software Development and Release Management
> As the USDC (US Data Center) deputy, represent the LZ collaboration to NERSC on project compute, network and archival needs. As a software engineer, help develop and monitor offline data movement systems, and manage releases for various analysis packages for LZ - an international dark matter experiment of 250 scientists across 35 institutions.
>
> **IDAES - Feb 2016 to Present**
> Task Lead for Software Development and Release Management
> Lead 10-20 developer/contributors for the software development, user support and release management for the Institute for the Design of Advanced Energy System (IDAES) Integrated Platform. This IDAES-IP is an entirely open-source product providing capabilities to solve complex design and optimization problems using equation-oriented solvers. The success of this project has led to the creation of the following dependant projects, for which I also lead the SW Dev and Release Management efforts:

PARETO - Jun 2021 to Present
Lead 5-7 developers for PARETO, an open-source, optimization-based, decision-support application for produced water beneficial reuse by upstream operators, midstream companies, technology providers, water end users, research organizations and regulators.

WaterTAP - Aug 2020 to Present
Lead 10-15 developers for WaterTAP, a NAWI funded open-source library of water treatment models to assess various treatment trains, such as desalination, through simulation, optimization, and other advanced methods.

DISPATCHES - Jul 2020 to Present
Lead 10-15 developers for DISPATCHES, a GMLC funded open-source platform used to identify and optimize integrated energy systems for operation within the bulk power system via energy market signals.

CCSI/CCSI^2 - Jan 2011 to Present
LBNL Co-PI, Task Lead for Software Development and Release Management
Lead the software development and release management for a toolset of over 30 modeling and simulation products for the Carbon Capture Simulation Initiative (CCSI) and its successor project CCSI^2 - a partnership of national labs, industry and academic institutions applying chemical engineering computational models and simulations to accelerate the commercialization of carbon capture technologies in power plants.

SPOT - Feb 2015 to Sep 2020
Software Engineer
Co-maintainer of first LDRD funded “Superfacility” system tying ALS and other beamline data sources to NERSC compute, analysis, presentation and archival systems via the SPADE distributed data movement application.

Berkeley Drosophila Genome (Fruitfly) Project - Mar 2013 to Feb 2015
Software Engineer
Developed an integrated system to enable automated, high content screening for microscope imaging. Worked with a small team programming a microscope, a slide loader and an image processor module to build a database of embryonic drosophila region of interest (ROI) images to discover gene expression patterns. Patent issued Oct 2019.

The STAR Experiment - Dec 2009 to Apr 2016
Software Engineer
For STAR experiment, built out 1PB XRootd data storage cluster using unused space on ~200 PDSF compute nodes. Wrote various tools for populating, maintaining cluster and querying STAR file catalog for derived datasets for analysis on NERSC supercomputer.

NetLogger - Apr 2008 to Oct 2009
Software Engineer
Contributed to NetLogger - both a methodology and set of software tools for debugging and performance analysis of complex distributed applications - by writing, updating parsers, applying it to several projects and migrating DB to Mongo backend.

Icecube - Jun 2004 to Dec 2010
Software Engineer, DAQ Release Manager
Developed and managed releases of the surface data acquisition (DAQ) software subsystem for Icecube - NSF funded, international high-energy neutrino detector installed in the ice below the South Pole. Introduced to a geographic, scientific, and organizationally diverse team of scientists: version control, issue tracking, automated testing, and regular release processes.

Akenti - 2004
Software Engineer
Ported the C/C++ side of Akenti - a PKI based authorization policy engine written in Java, C/C++ - to gcc v3, eased development and distribution by reworking GNU Autoconf configuration and split out Java side to use Ant for building.

PyGridWare - 2002 to 2004
Software Engineer
Wrote an initial implementation of the XMLDigSig specification in Python to provide message-level security via XML digital signatures as part of the WS-Resource framework for interoperability with the Globus Toolkit.

Bigstep, San Francisco, CA — **Jan 1999 to Oct 2001**
Sr Release Engineer, Build & Configuration Manager - **Sept 1999 to Oct 2001**
Solely managed all aspects of source code management, builds, release, and training for a development team of up to 40 contributors in two geographic locations. Coordinated development and release planning between Product Management, Product Development, QA and Network Operations groups.

Software Engineer - **Jan 1999 to Oct 2000**
Core team engineer contributing to architecture and subsystem specifications, design and implementation of Java servlet engine and back-end DB schema. Team employed Java Beans, JDBC, XML, SQL, Dynamo 4.5, Solaris and Oracle 8i to create one of the first web sites for small businesses to create and run their own on-line stores.

The Baan Company, Santa Clara, CA — **May 1997 to Jan 1999**
Technology Engineer II
Developed Java middleware components, including C++ and COM interfaces, for use in creating custom thin-client GUIs to the Baan ERP system. Lead the documentation, testing and distribution to internal and external Baan developers.

Lawrence Berkeley National Laboratory, Berkeley, CA — **April 1994 to May 1997**
Graduate Research Assistant, ALS Distributed Collaboratory Experiment Environment
Developed an object-oriented interface to a fault-tolerant, ordered multicast communication protocol and a distributed resource management system for control and coordination of ALS project components.

General Services Administration, San Francisco, CA — **Sept 1991 to April 1994**
Computer Programmer, Traffic Management Division, Federal Supply Service Bureau
Developed, documented, distributed and supported a FoxPRO freight carrier database application obviating the manual collection, storage and retrieval of commercial freight carrier rates resulting in a dramatic increase of Traffic Management Division's productivity.

PUBLICATIONS, PRESENTATIONS & PATENTS

PARETO: An Open-Source Produced Water Optimization Framework
Markus G. Drouven, Andrés J. Caldéron, Miguel A. Zamarripa, Keith Beattie,
Accepted for publication in Optimization and Engineering, 12-Sept-2022, OPTE-2022-3071-R1

Software Engineering Challenges and Best Practices for Multi-Institutional Scientific Software Development
Invited webinar presentation for the DOE Exascale Computing Project (ECP) series on Best Practices for HPC Software Developers, Aug 2021

Lessons learned working with protected assets in an open-source collaborative scientific software project
K. Beattie and D. Gunter, 2021 IEEE/ACM International Workshop on Body of Knowledge for Software Sustainability (BoKSS), 2021, pp. 13-14, doi: 10.1109/BoKSS52540.2021.00014.

High content screening workflows for microscope imaging
Frise, Erwin; Booth, Benjamin; McParland, Charles P.; Beattie, Keith S.; Fisher, William W.; Hammonds, Ann; Celniker, Susan E., US Patent No. 10430955, Issued: 01-Oct-2019

Real-Time Data-Intensive Computing
Parkinson, D. Y., Beattie, K., Chen, X., Correa, J., Dart, E., Daurer, B. J., Ushizima, D. (2016). In PROCEEDINGS OF THE 12TH INTERNATIONAL CONFERENCE ON SYNCHROTRON RADIATION INSTRUMENTATION (SRI2015) Vol. 1741. doi:10.1063/1.4952921

Making advanced scientific algorithms and big scientific data management more accessible
Venkatakrishnan, S. V., Mohan, K. A., Beattie, K., Correa, J., Dart, E., Deslippe, J. R., Parkinson, D. Y. (2016). IS and T International Symposium on Electronic Imaging Science and Technology. doi:10.2352/ISSN.2470-1173.2016.19.COIMG-155

Community Access to MODIS Satellite Reprojection and Reduction Pipeline and Data Sets
Hendrix, V., Li, J., Jackson, K., Ramakrishnan, L., Ryu, Y., Beattie, K., Agarwal, D. Poster session presented at AGU 2012.

Enabling petascale science: data management, troubleshooting, and scalable science services
Baranovski, A., Beattie, K., Bharathi, S., Boverhof, J., Bresnahan, J., Chervenak, A., Tierney, B. (2008). SCIDAC 2008: SCIENTIFIC DISCOVERY THROUGH ADVANCED COMPUTING, 125. doi:10.1088/1742-6596/125/1/012068

The Icecube Data Acquisition Software: Lessons Learned During Distributed, Collaborative, Multi-Disciplined Software Development
K S Beattie, C T Day, D Glowacki, K D Hanson, J E Jacobsen, C P McParland and S J Patton, Presented at CHEP 2007, Victoria, Canada

TEACHING

College of Notre Dame, San Mateo, CA – **Summer 1999**
Lecturer, Mathematics and Computer Science Department
Instructor for 2nd Semester Java course covering advanced Java programming topics: Object-Oriented programming & design, Exceptions, Standard utilities & data types, Threads, Network & File I/O and AWT.

San Francisco State University, San Francisco, CA – **Jan 1991 to Jan 1993**
Lecturer, Mathematics and Computer Science Department
Instructor of introductory and intermediate algebra courses.

EDUCATION

San Francisco State University, San Francisco, CA – *MS Computer Science* **1997**
Thesis: "*An Object-Oriented Interface To Reliable, Ordered Multicast Communication*"
Advisors: Prof. Jozo Dujmovic, Prof. James Wong, Dr Deborah Agarwal

San Francisco State University, San Francisco, CA – *BA Mathematics* ***1990***