

163. The KBase Narrative User Interface

Stephen Chan¹, Jason Baumohl¹, Aaron Best², Jared Bischof², Ben Bowen¹, Tom Brettin², Tom Brown², Shane Canon¹, John-Marc Chandonia¹, Dylan Chivian¹, Ric Colasanti², Neal Conrad², Brian Davison³, Matt DeJongh⁶, Paramvir Dehal*¹ (psdehal@lbl.gov), Narayan Desai², Scott Devoid², Terry Disz², Meghan Drake³, Janaka Edirisinghe², Gang Fang⁷, José Pedro Lopes Faria², Mark Gerstein⁷, Elizabeth M. Glass², Annette Greiner¹, Dan Gunter¹, James Gurtowski⁵, Nomi Harris¹, Travis Harrison², Fei He⁴, Matt Henderson¹, Chris Henry², Adina Howe², Marcin Joachimiak¹, Kevin Keegan², Keith Keller¹, Guruprasad Kora³, Sunita Kumari⁵, Miriam Land³, Folker Meyer², Steve Moulton³, Pavel Novichkov¹, Taeyun Oh⁸, Gary Olsen⁹, Bob Olson², Dan Olson², Ross Overbeek², Tobias Paczian², Bruce Parrello², Shiran Pasternak⁵, Sarah Poon¹, Gavin Price¹, Srivdya Ramakrishnan⁵, Priya Ranjan³, Bill Riehl¹, Pamela Ronald⁸, Michael Schatz⁵, Lynn Schriml¹⁰, Sam Seaver², Michael W. Sneddon¹, Roman Sutormin¹, Mustafa Syed³, James Thomason⁵, Nathan Tintle⁶, Will Trimble², Daifeng Wang⁷, Doreen Ware⁵, David Weston³, Andreas Wilke², Fangfang Xia², Shinjae Yoo⁴, Dantong Yu⁴, **Robert Cottingham³**, **Sergei Maslov⁴**, **Rick Stevens²**, **Adam P. Arkin¹**

¹ Lawrence Berkeley National Laboratory, Berkeley, CA, ² Argonne National Laboratory, Argonne, IL, ³ Oak Ridge National Laboratory, Oak Ridge, TN, ⁴ Brookhaven National Laboratory, Upton, NY, ⁵ Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, ⁶ Hope College, Holland, MI, ⁷ Yale University, New Haven, CT, ⁸ University of California, Davis, CA, ⁹ University of Illinois at Champaign-Urbana, Champaign, IL, ¹⁰ University of Maryland, College Park, MD

<http://kbase.us>

Project Goals: The KBase project aims to provide the capabilities needed to address the grand challenge of systems biology: to predict and ultimately design biological function. KBase enables users to collaboratively integrate the array of heterogeneous datasets, analysis tools and workflows needed to achieve a predictive understanding of biological systems. It incorporates functional genomic and metagenomic data for thousands of organisms, and diverse tools for (meta)genomic assembly, annotation, network inference and modeling, allowing researchers to combine diverse lines of evidence to create increasingly accurate models of the physiology and community dynamics of microbes and plants. KBase will soon allow models to be compared to observations and dynamically revised. A new prototype Narrative interface lets users create a reproducible record of the data, computational steps and thought process leading from hypothesis to result in the form of interactive publications.

The Narrative interface enables users to easily customize, execute, and share a set of ordered KBase actions that target their specific systems biology hypotheses. The datasets, analyses, and thought processes underlying the execution of these actions are captured in the form of an interactive, dynamic publication called a Narrative. Within each Narrative, users can interleave text and commentary with workflows, so that hypotheses and conclusions can be captured as well as raw results and procedural notes. The Narrative, with the help of Workspaces, maintains provenance and metadata for all data, thereby providing a virtual reference list for all Narratives. A sophisticated social framework will soon allow members of a research team to share, execute, modify, and comment on Narratives at multiple levels of granularity enabled by the KBase infrastructure.

Intrinsically reproducible, Narratives will serve as a new type of publication by (1) explicitly capturing the parameters associated with various algorithms, (2) clearly showing how users applied the algorithms to selected datasets, and (3) transparently recording the process by which resulting output was used to derive biological insight and conclusions. Users who access a Narrative that another researcher has created and shared in KBase not only will be able to see a complete inventory of the data and algorithms underlying a conclusion, they will be able to repeat the computational experiment with the push of a button, even altering parameters to achieve different or improved results.

The Narrative interface is built on top of the open source iPython platform and deeply integrated with the KBase infrastructure. For those seeking to develop new KBase functionality, this allows for rapid integration of KBase-compliant services and KBase data, and leverages the existing community support. Detailed documentation and support is available on the kbase.us website explaining how to use the Narrative interface and the process for adding new features.

KBase is funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research.