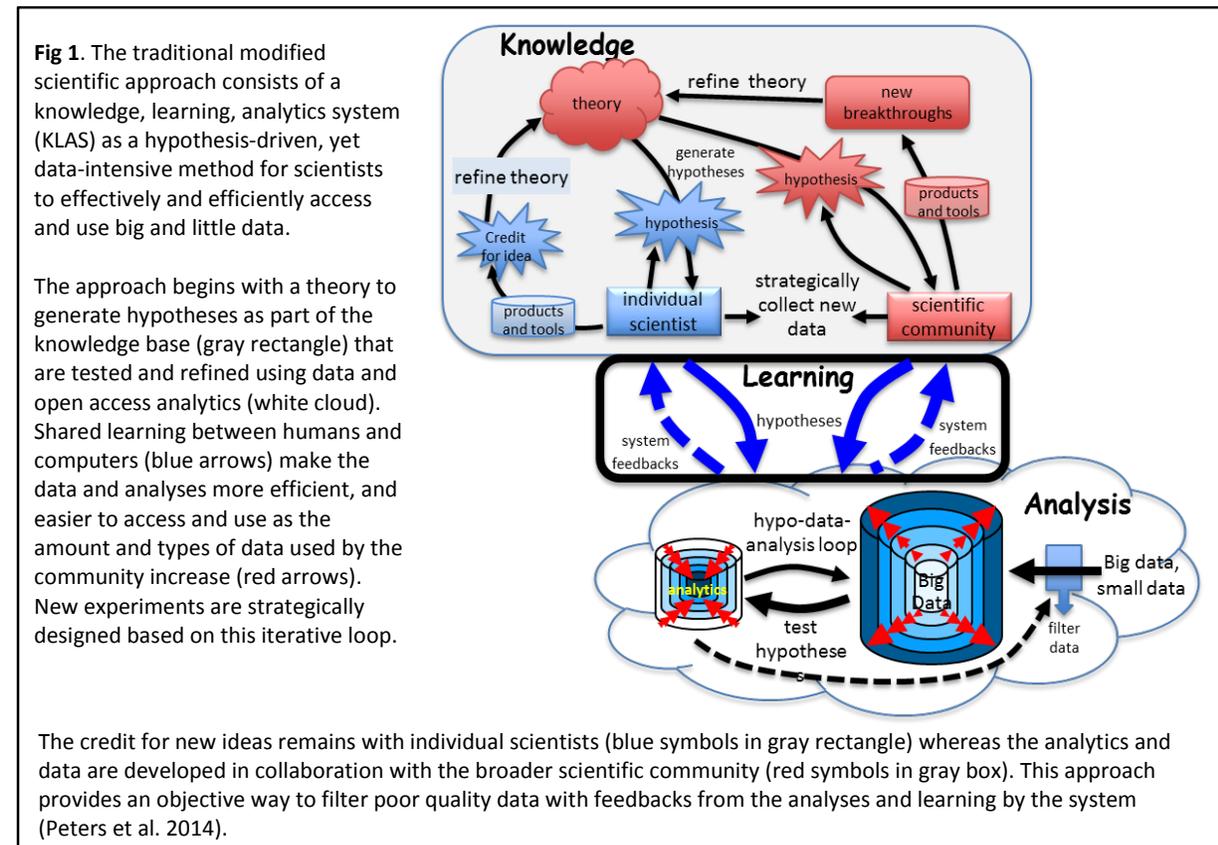# Knowledge, learning, analysis system (KLAS)

Contact(s): Geovany Ramirez, Debra Peters, Kris M. Havstad, Judy Cushing, Craig Tweedie, Olac Fuentes and Natalia Villanueva-Rosales

Abstract: The goal of KLAS is to develop a new scientific approach that takes advantage of the data deluge, defined here as both legacy data and new data acquired from environmental and biotic sensors, complex simulation models, and improved technologies for probing biophysical samples. This approach can be implemented as a knowledge-driven, open access system that "learns" and becomes more efficient and easier to use as data streams increase in variety and size. We are in an early stage of KLAS, but we expect that it will adapt the scientific method to accommodate vast amounts of data, and make them accessible to a broad range of user via open access with an iterative learning process. Also, we expect KLASS to transform ecology and environmental sciences by shortening the time lag between individual discoveries and leaps in knowledge by the scientific community. New experiments can be strategically designed based on feedback from the hypothesis-data-analysis loop. The scope of KLAS is not limited to ecology and environmental sciences; it can be used to manage and process data for different fields. KLAS will be able to interconnect and reuse hypotheses, methodologies, data, processes, and knowledge gained from previous experiments from one domain to another domain, smoothly and transparently to the user.

Fig 1. The traditional modified scientific approach consists of a knowledge, learning, analytics system (KLAS) as a hypothesis-driven, yet data-intensive method for scientists to effectively and efficiently access and use big and little data.

The approach begins with a theory to generate hypotheses as part of the knowledge base (gray rectangle) that are tested and refined using data and open access analytics (white cloud). Shared learning between humans and computers (blue arrows) make the data and analyses more efficient, and easier to access and use as the amount and types of data used by the community increase (red arrows). New experiments are strategically designed based on this iterative loop.

The credit for new ideas remains with individual scientists (blue symbols in gray rectangle) whereas the analytics and data are developed in collaboration with the broader scientific community (red symbols in gray box). This approach provides an objective way to filter poor quality data with feedbacks from the analyses and learning by the system (Peters et al. 2014).

Relationship with the LTER VI proposal: This project is funded by an EAGER award to Peters and Havstad (DEB-1440166), and addresses the data accessibility goals of the LTER VI proposal.

Literature Cited: Peters, DPC, Havstad KM, Cushing J, Tweedie C, Fuentes O, and Villanueva-Rosales N. 2014. Harnessing the power of big data: infusing the scientific method with machine learning to transform ecology. Ecosphere 5:art67. doi:10.1890/ES13-00359.1.