PSLC DataShop: A Data Analysis Service for the Learning Science Community

John Stamper[1], Ken Koedinger[1], Ryan S.J.d. Baker[2], Alida Skogsholm[1],
Brett Leber[1], Jim Rankin[1], and Sandy Demi[1]

[1] Carnegie Mellon University, Human-Computer Interaction Institute
`{jstamper,krk,alida,bleber,jimbokun,sdemi}@cs.cmu.edu`
[2] Worcester Polytechnic Institute, Department of Social Science and Policy Studies
`rsbaker@wpi.edu`

The Pittsburgh Science of Learning Center's DataShop is an open data repository and set of associated visualization and analysis tools. DataShop has data from thousands of students deriving from interactions with on-line course materials and intelligent tutoring systems. The data is fine-grained, with student actions recorded roughly every 20 seconds, and it is longitudinal, spanning semester or yearlong courses. Currently over 188 datasets are stored including over 42 million student actions and over 150,000 student hours of data. Most student actions are "coded" meaning they are not only graded as correct or incorrect, but are categorized in terms of the hypothesized competencies or knowledge components needed to perform that action.

DataShop provides a number of features to facilitate data analysis including a data schema that allows researchers to import data into DataShop or export data from the repository in order to perform additional analysis. DataShop offers a number of online analysis tools to perform functions, such as visualizing student performance and analyzing learning curves. Researchers can export cognitive models, make changes, and upload the changed model for further analysis. One new feature that has been added to DataShop is an easy-to-use API for using web services to access the repository. These web services allow developers to identify data sets in the repository and directly export data from them at the transaction or student step level. In the near future, developers will be able to add new fields back into the repository with the use of our web services for custom fields.

Researchers have analyzed these data to better understand student cognitive and affective states and the results have been used to redesign instruction and demonstrably improve student learning [1]. Researchers can find out more and sign up for access to DataShop from our website: **`http://pslcdatashop.org`**

Reference

1. Koedinger, K.R., Baker, R.S.J.: A Data Repository for the EDM commuity: The PSLC DataShop. In: Romero, C., Ventura, S., Pechenizkiy, M., Baker, R.S.J. (eds.) To appear in Handbook of Educational Data Mining. CRC Press, Boca Raton