PSLC DATASHOP a data analysis service for the learning science community

What is DataShop?

DataShop is an open data repository and set of associated visualization and analysis tools for fine-grained, longitudinal learner data.

DataShop enables you to:

- Import and export datasets
- Visualize student performance
- Perform learning curve analysis
- Discover cognitive models of student knowledge
- Use data to improve instruction
- Programmatically retrieve data via web services

What data is in DataShop?

DataShop contains vast amounts of free data.

Domain	Datasets	Student Actions	Students	Student Hou
Language	60	2,838,843	3,263	5,66
Math	73	32,201,021	8,887	99,08
Science	30	3,675,770	3,890	19,83
Other	30	4,789,826	11,058	28,69
Total	193	43,505,460	27,098	153,27
t.				

Over 150,000 student hours! 🗲

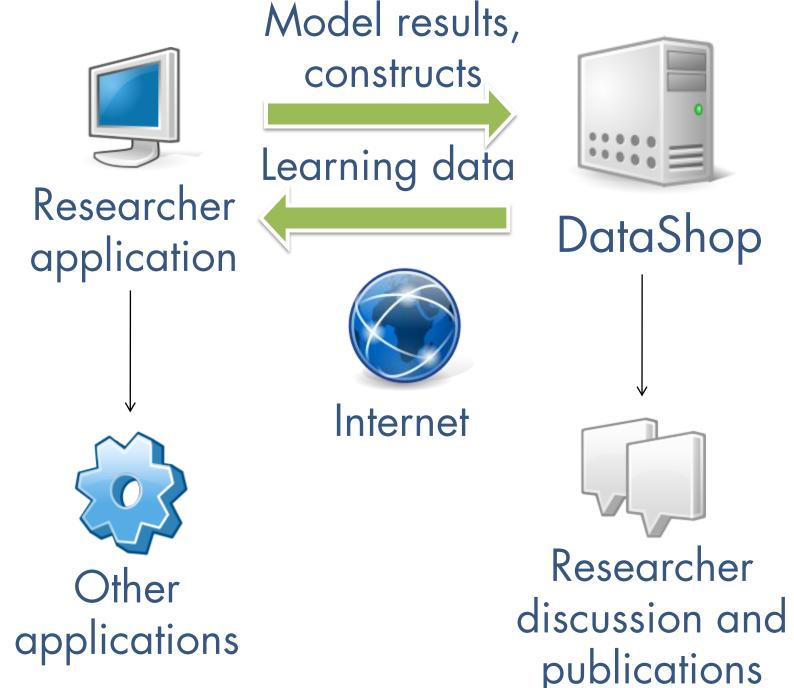
What are some analyses that have used DataShop data?

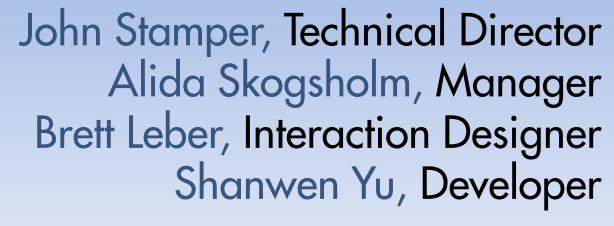
- Adapting to when students game an intelligent tutoring system (Baker et. al.)
- Self-explaining in the classroom: Learning curve evidence (Hausmann and VanLehn)
- Optimizing knowledge component learning of Chinese using a dynamic structural model of practice (Pavlik et. al.)
- Extending learning factors analysis to model reading transfer (Leszczenski and Beck)
- Exploring alternative methods for error attribution in learning curves analyses (Nwaigwe et. al.)



pslcdatashop.org







erformance Profiler	<u>Error Report</u>	Learning Curve	<u>Export</u>	
and Files KC Models	Problem Breakdo	<u>wn</u>		
/ Area (1996-97)				
sholm on 2008-06-10 15:00	10 KCs <u>export rename delete</u> 0:32.0 status: ready to use		Add or modify a KC mode Export (Step 1) Import (Step 2)	el
-to-kc 5161.16 BIC: 5677.64			1. Select a KC model file to imp C.\Documents and Settil Brow	
ect-transaction-to-kc 5642.15 BIC: 6054.03	2 KCs export r status: ready to us	<u>ename delete</u> ;e	DataShop will first try to verify the va file. Information about this process After verifying, you will be given the importing or cancel the process.	will be displayed below.
n 2008-07-10 10:37:57.0 to-kc 5131.29 BIC: 5687	13 KCs export r status: ready to us		* Import New model "Improve If the information above is the "Import" button to cont press "Cancel" to stop the this file. Beginning insertion process Processing file 100%	correct, press inue. Otherwise, import process for
1 nger on 2008-12-06 14:23: to-kc 5085.89 BIC: 5628.53	12 KCs export r 14.0 status: ready to us		Initial import process comp 141 rows in the file were p Running learning factors an building sample information (Note: this may take a very depending on the size of th LFA in progress.	rocessed. alysis (LFA) and for new models.
	1 KCs export r	ename <u>delete</u>	[Import] Cancel	

Coming in future releases

- existing dataset

- Error bars on learning curves

KDD Cup 2010 Educational Data Mining Challenge > http://pslcdatashop.org/KDDCup

Awarded to the PSLC and DataShop, this year's challenge asked participants to predict student performance on mathematical problems from logs of student interaction with Intelligent Tutoring Systems.

The KDD Cup is the premier data mining competition that attracts top data mining and machine learning experts from industry and academia.

The competition ended on June 8. There were:

655 registered participants

130 participants who submitted predictions

3,400 submissions

The datasets used for the challenge were:

Dataset

Algebra I 2008-2009

Bridge to Algebra 2008-2009

The competition addressed questions of both scientific and practical importance. Improved models could be saving millions of hours of students' time (and effort) in learning algebra. These models should both increase achievement levels and reduce time needed to learn.

Easy-to-use API for using web services to add fields to an

Ability to upload models to DataShop for dissemination, making it easier to build on each other's work

Improved scalability and speed: accommodate a planned 100x increase in quantity of data in DataShop

Support for timestamps with millisecond precision

Students	Steps	File size
3,310	9,426,966	3 GB
6,043	20,768,884	5.43 GB

