

# DATASIM Conceptual UX/UI Design Report

Data and Training Analytics Simulated Input Modeler

10 July, 2019

This work was supported by the U.S. Advanced Distributed Learning (ADL) Initiative (HQ003419C0061). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the ADL Initiative or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes.



**Distribution Statement A**

Approved for public release: distribution unlimited.

**REPORT DOCUMENTATION PAGE**

*Form Approved  
OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

**PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YYYY) 10-07-2019		2. REPORT TYPE Research Paper		3. DATES COVERED (From - To) 31-03-2019 to 10-07-2019	
4. TITLE AND SUBTITLE DATASIM Conceptual UX/UI Design Report				5a. CONTRACT NUMBER HQ0034-19-C-0061	
				5b. GRANT NUMBER N/A	
				5c. PROGRAM ELEMENT NUMBER 0603769D8Z	
				5d. PROJECT NUMBER UPR_030603769D8Z776	
6. AUTHOR(S) Shelly Blake-Plock				5e. TASK NUMBER 20_0400_0603769D8Z	
				5f. WORK UNIT NUMBER HQ0642922724	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Yet Analytics 201 E. Baltimore St., Ste. 630 Baltimore, Maryland 21202				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) OUSD Personnel & Readiness Advanced Distributed Learning Initiative 13501 Ingenuity Drive, Suite 248 Orlando, Florida 32826				10. SPONSOR/MONITOR'S ACRONYM(S) OUSD/P&R/FE&T/ADLI	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Distribution A					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT DATASIM is an open-source application that will provide a valid means of producing the data sets necessary to benchmark and stress test the Total Learning Architecture (TLA) and distributed learning acquisitions. Additionally, DATASIM can help learning scientists, ISDs, IT and technology acquisition, and decision-making stakeholders to determine the effectiveness of xAPI data design and implementation across the TLA. This document provides a draft of the conceptual experience and user interface necessary to enable the researcher of DATASIM to accomplish key tasks.					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 19	19a. NAME OF RESPONSIBLE PERSON Nick Armendariz
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (Include area code) 407-381-5550

# DATASIM Conceptual UX/UI Design Report

## Project DATASIM

Data and Training Analytics Simulated Input Modeler

Prepared by Shelly Blake-Plock, PI, Yet Analytics, Inc.

Submitted to ADL TPOC on July 10, 2019

---

## Purpose of this Document

DATASIM is an open source application that will provide a valid means of producing the datasets necessary to benchmark and stress test the Total Learning Architecture (TLA) and distributed learning acquisitions. Additionally, DATASIM can help learning scientists, ISDs, IT and technology acquisition, and decision-making stakeholders to determine the effectiveness of xAPI data design and implementation across the TLA.

This document provides a draft of the conceptual experience and user interface necessary to provide the researcher of DATASIM to accomplish the following key tasks:

1. DATASIM leverages xAPI Profiles to define the learning environment. The end user will need to be able to point to the xAPI Profile(s) upon which to base their simulation.
2. The end user will need to set parameters in order to configure the elements of the environment with which each artificial agent can “perform”.
3. The end user will need to be able to assign a distribution of personae across the population of simulated learners.
4. When DATASIM is run, it will first create a baseline scenario against which subsequently generated valid xAPI datasets of simulated learning activity will be compared for purposes of establishing good fit.

# Table of Contents

<b>Purpose of this Document</b>	<b>1</b>
<b>Table of Contents</b>	<b>2</b>
<b>Introduction</b>	<b>3</b>
<b>Definitions</b>	<b>4</b>
<b>DATASIM Conceptual UX/UI Designs</b>	<b>5</b>
<b>DATASIM Homepage</b>	<b>5</b>
<b>Step 1: Import Your xAPI Profile</b>	<b>7</b>
Alpha Release	7
Future Release	10
<b>Step 2: Confirm Your Environment</b>	<b>11</b>
Alpha Release	11
Future Release	11
<b>Step 3: Select Your Personae</b>	<b>12</b>
Alpha Release	12
Future Release	13
<b>Step 4: Review Simulation Settings</b>	<b>14</b>
<b>Simulation Processing Page</b>	<b>15</b>
<b>Your Simulation Results</b>	<b>16</b>
Alpha Release	16
Future Release	16
<b>Conclusion</b>	<b>17</b>

# Introduction

The Data and Training Analytics Simulated Input Modeler (DATASIM) intends to provide a solution to the problem of the availability of relevant xAPI test data — at scale — which can be used for the purpose of testing and evaluating the TLA.

The project began by considering the following challenge:

- Test data gathered from isolated time-constrained human-input scenarios can not begin to approximate the scale and scope of real-world longitudinal training and lifelong learning.
- The manual creation of test data both comes at great cost and introduces the challenges of human error and limited capacity to the development of complicated and comprehensive xAPI data sets that actually portray reasonable, meaningful, and readable real-life scenarios.

The solution is to build the specifications and reference model for a software application which could be used to simulate learning activity at any scale and generate the activity's event data as valid and meaningful xAPI. Combined with the TLA Reference Model, DATASIM represents an application of the Digital Twins paradigm applied to the learning technology ecosystem. DATASIM provides a means of generating the simulated data of a learner population without the cost of expensive live exercises or interruptive training events. The reduced need for human subjects allows for the repeated testing of tools, technologies, processes, and technical standards at scale and provides the foundation for data interoperability and conformance testing across the TLA.

DATASIM simulations are comprised of an environment made up of one or more xAPI Profiles. Within this environment reside artificial agents that have characteristics that define them and describe their behavior and activity. Conceptually, they may be thought of akin to non-player characters (NPCs) in a video game. DATASIM leverages the inputs to produce a median baseline dataset. That dataset provides a model against which to compare and validate user-defined simulations. The end user has the opportunity to re-model a simulation to bring its component parts closer in adherence to the baseline model or to highlight anomalous behavior within the activity of the agents (such as might be the case if the end user were trying to create a dataset reflecting a population that was highly sleep deprived during an assessment).

The mockups below demonstrate how an end user would interface with DATASIM to build their simulation, review the data, and execute the final dataset.

# Definitions

For the purposes of this document, the following are working definitions.

- Researcher:** This is a general term indicating the end user of DATASIM. In reality, the end user could be a learning scientist, an ISD, an IT stakeholder, or other TLA stakeholders involved in data design and/or the testing and evaluation of TLA components, specifications, and conformance.
- Simulation:** The modeled activity of artificial agents within an environment.
- Agent:** This is a simulated learner whose activity takes place within a defined DATASIM simulation.
- Personae:** This is the typology of characteristics that defines the Agent.
- Population:** This is the universe of all agents in the simulation.
- Environment:** This is the combination of platforms, courses, activities, and events, as defined by a taxonomy. The researcher has chosen the environment for the simulation.
- Baseline Dataset:** A system-generated mean dataset against which the researcher can compare their actual simulation output.

# DATASIM Conceptual UX/UI Designs

The following illustrations present the conceptual user experience and user interface designs for the DATASIM software application.

## DATASIM Homepage

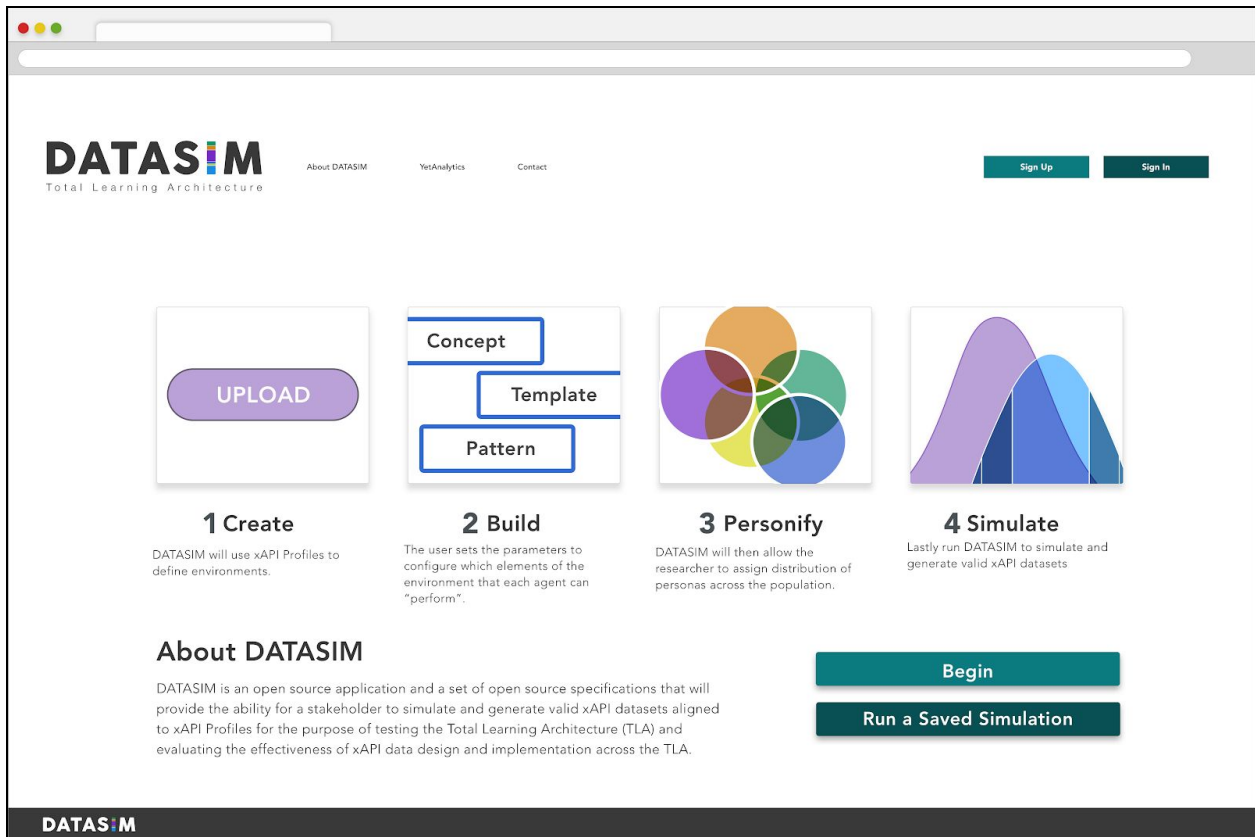


Figure 1: DATASIM homepage

Figure 1 demonstrates the first page the researcher will see in the DATASIM application. The DATASIM homepage includes a description of DATASIM and provides an overview of how to use the DATASIM application. The researcher can create a new account from this page or log into an existing account. As shown on the bottom right of Figure 1, a researcher can either choose to run an existing saved simulation or create a new simulation. The researcher must be logged in and have saved a simulation previously in order to choose the former.

If the researcher meets these requirements and clicks on the “Run A Saved Simulation” button, the application will present the researcher with a list of saved simulations to choose from. Once the researcher selects a simulation from this list, the application will populate the remaining

steps in DATASIM with the saved simulation settings. If the researcher chooses the “Begin” button from the homepage, the application will take the researcher to Step 1 where they can begin the process of building their simulation components.

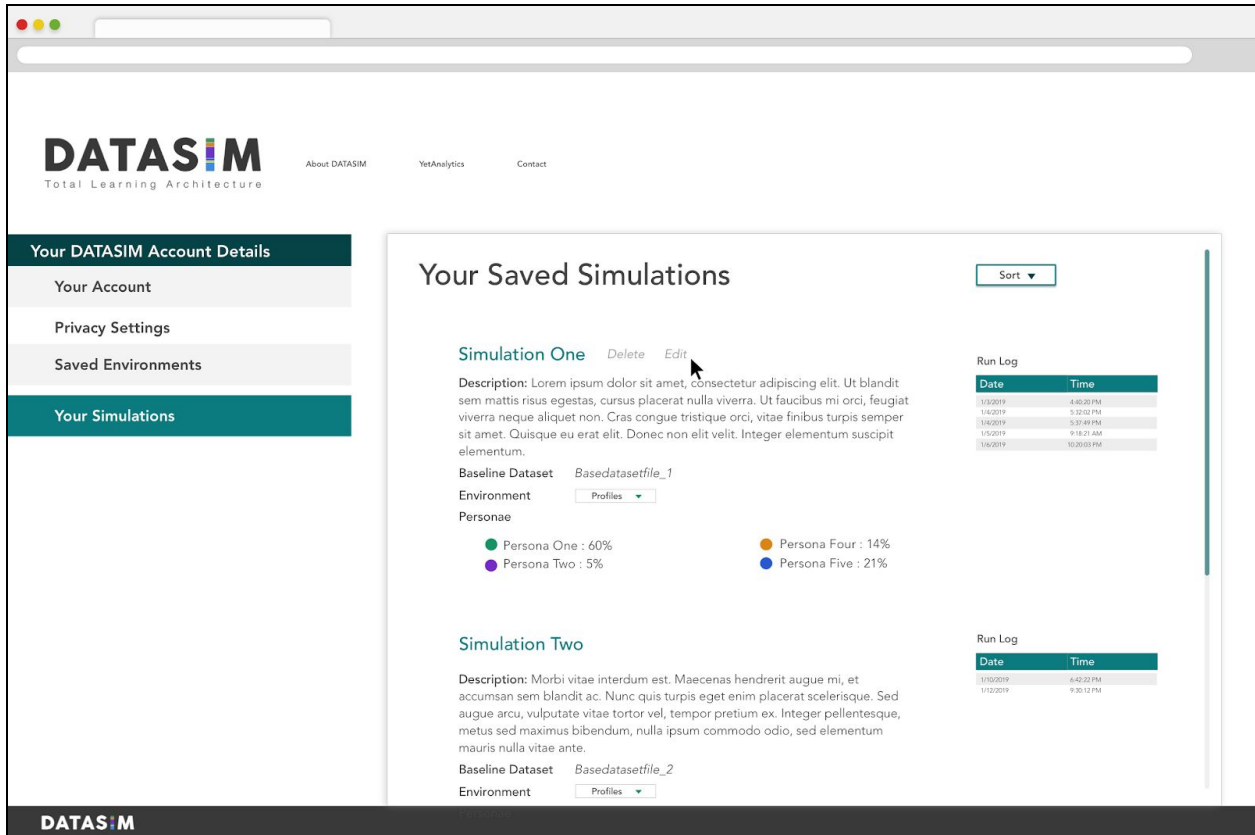


Figure 2: DATASIM account view for a registered researcher



## Step 1: Import Your xAPI Profile

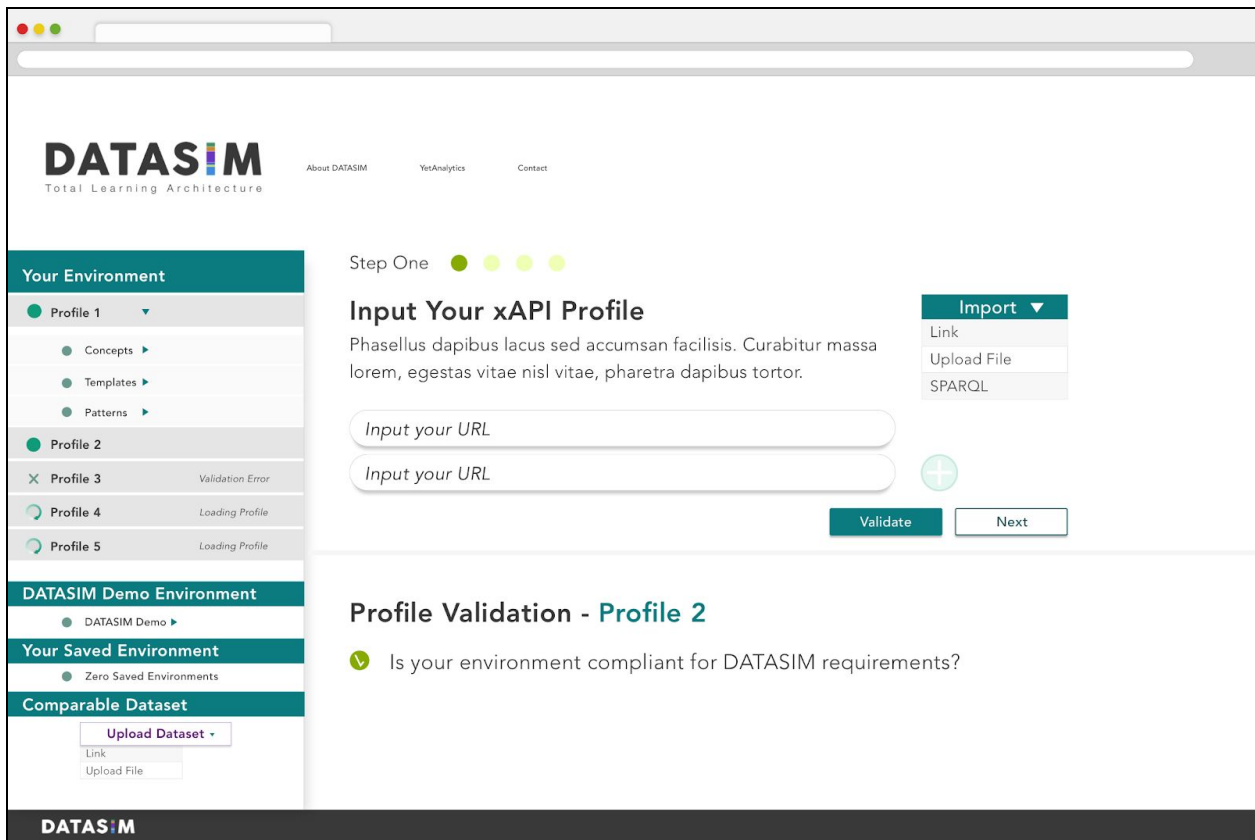


Figure 3: xAPI Profile input

### Alpha Release

Figure 3 demonstrates the first step in building a simulation, where the researcher can input various xAPI Profiles or components of xAPI Profiles which will compose the simulation environment. The application allows a researcher to choose from pre-configured environments or to configure a new environment.

Pre-configured environments include:

- Demo environments (available from the panel on the left side of Figure 3)
- Previously saved environments (the panel on the left side of Figure 3)

Researchers would choose a pre-configured environment by clicking on the appropriate option on the left panel and then clicking the “Next” button.

New environments can be built from:

- Referencing existing xAPI Profiles via URL(s)
- Uploading existing xAPI Profile JSON-LD file(s)

- Executing a SPARQL query

The three options are offered on the right side of Figure 3 under the “Import” menu.

Figure 3 specifically demonstrates how the researcher can input a URL reference to one or more existing xAPI Profiles. The researcher will enter the URL(s), then click on the “Validate” button. The profile will then be retrieved from the URL and sent through a profile validation service. If the profile is validated as conformant to the xAPI Profile specification and deemed sufficient to be used in DATASIM (based on constraints defined by the DATASIM specifications), then the application will indicate that the profile has been validated for DATASIM use.

If an xAPI Profile does not pass validation, the application will indicate that the profile did not pass validation and will suggest possible reasons why (see Figure 4 below).

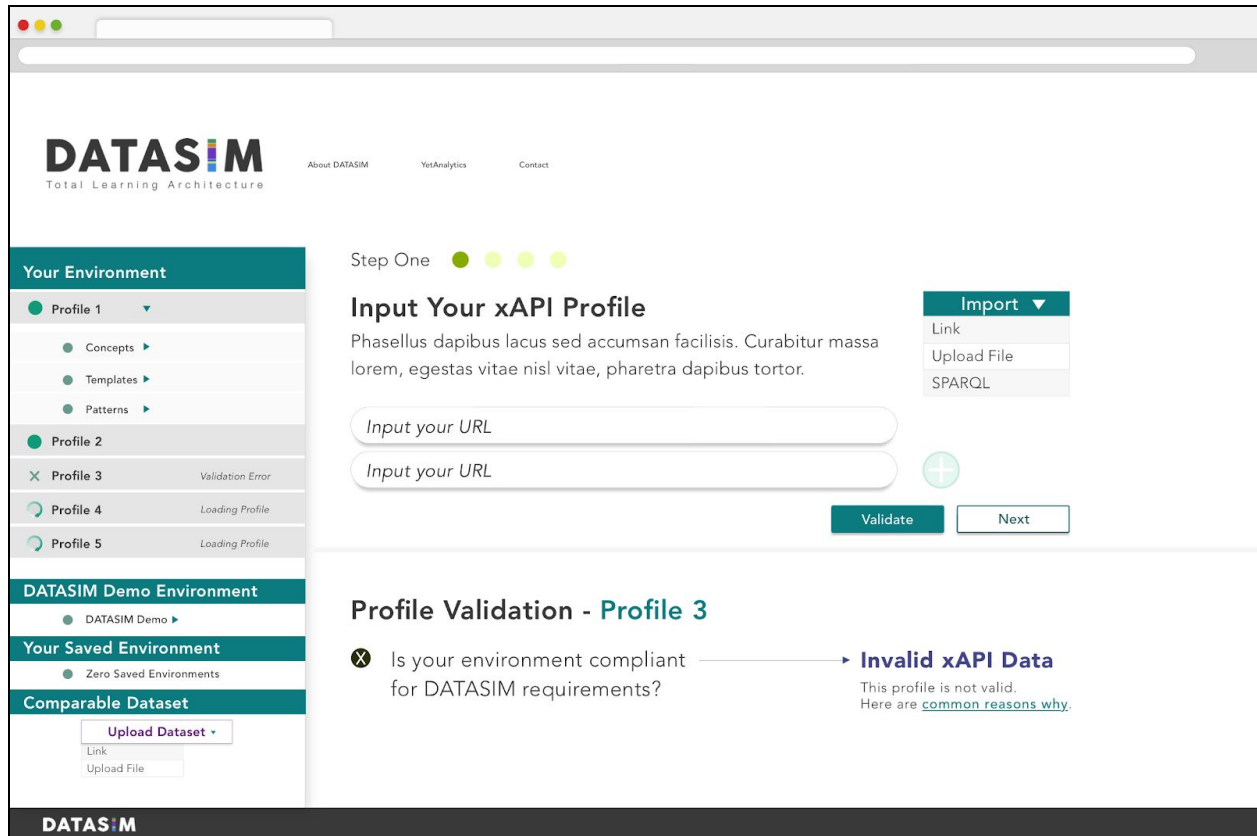


Figure 4: Indication of invalid xAPI data

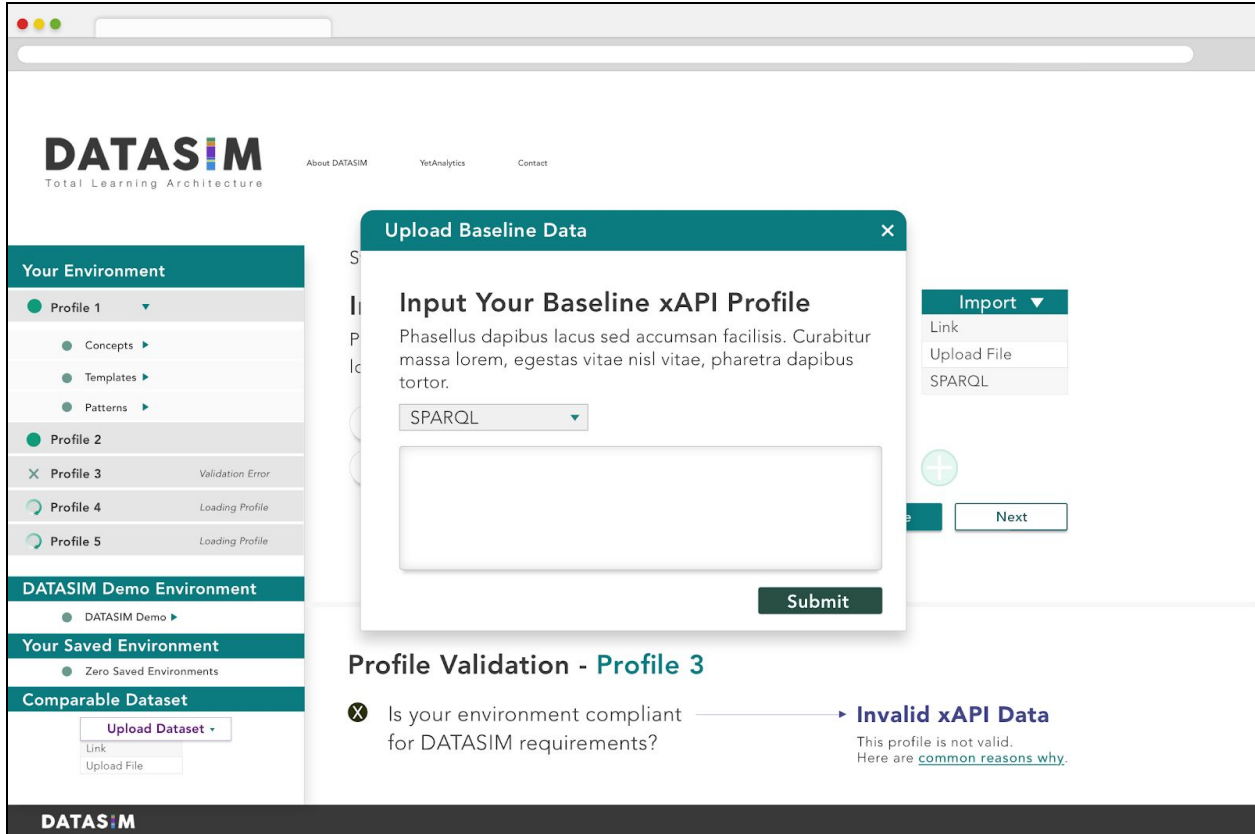


Figure 5: Executing a SPARQL query

Figure 5 demonstrates how a researcher would import xAPI Profile components by executing a SPARQL query. SPARQL is the query language for RDF data that is called out in the xAPI Profiles specification. The advantage of SPARQL generally is that it provides a full set of analytical query operations and is built for the RDF information model. The value proposition to the DATASIM user is the ability to query and input data from an xAPI Profile while retaining all of the inherent schema of the xAPI Profile. This contributes to the quality of resultant data to be used by DATASIM in generating simulations. This feature will be scheduled for a post-Alpha release.

The researcher would input the query in the text box and click on the “Submit” button. The query results will then be sent through the profile validation service. If the results are validated as conformant to the xAPI Profile specification and deemed sufficient to be used in DATASIM (based on constraints defined by the DATASIM specification), then the application will indicate that the query results have been validated for DATASIM use. If the query results do not pass validation, the application will indicate that the results did not pass validation and possible reasons why.

Note that there is no limit to the number of xAPI Profiles or xAPI Profile components that can be uploaded and validated in Step 1. As the researcher uploads various xAPI Profiles and xAPI



Profile components to build their simulation environment, the inputs will populate at the top of the Environments panel on the left side of Figure 5. Profiles that have been validated will be indicated with a solid circle. Profiles that have not been validated are indicated with an X. xAPI Profiles that have been input and are currently undergoing validation will be indicated with an “in process” spinner.

Once the researcher is finished uploading and validating their environment components, the researcher will click the “Next” button to move on to Step 2.

Note that the DATASIM application will automatically produce a baseline dataset by which the researcher can compare their simulation output each time the simulation is executed. The DATASIM application also enables the researcher to upload additional comparable dataset(s) if they so choose. The researcher may do so by uploading comparable dataset(s) via URL, file upload, or LRS credentials (bottom left of Figure 5). Note that while this is not a required component, the application allows the researcher to upload comparable dataset(s) at any step.

#### Future Release

In future releases, DATASIM may allow researchers to build their own xAPI Profiles directly in the application.

## Step 2: Confirm Your Environment

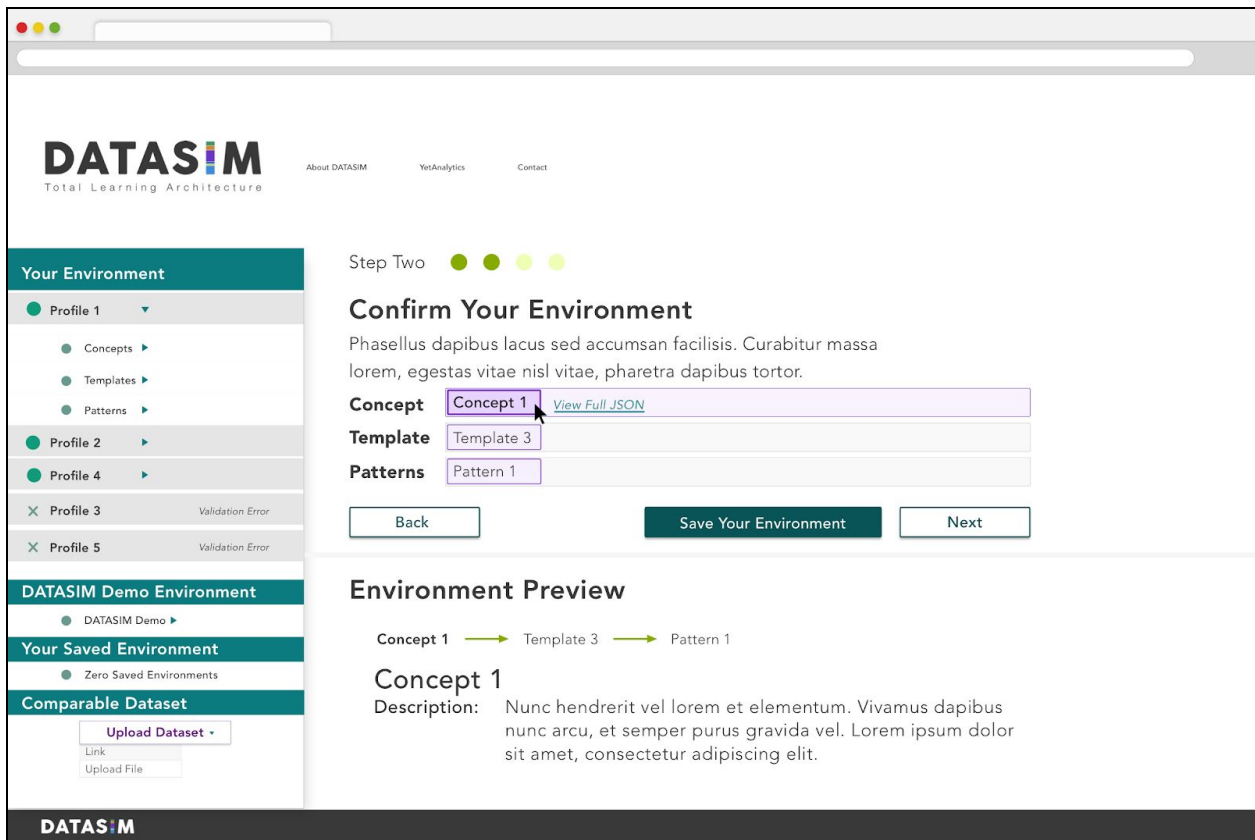


Figure 6: Interface for environment design

### Alpha Release

Once the researcher is done uploading and validating xAPI Profiles and xAPI Profile components, they will need to review and confirm the simulation environment. Figure 6 demonstrates how the researcher can click on the various concepts, templates, and patterns imported in Step 1 to explore their imported environment. If there are any changes that the researcher wants to make to the environment, the researcher will need to navigate back to Step 1, reconfigure and re-upload the imported components. As demonstrated in Figure 6, Step 2 also allows the researcher to save their environment for future use. If the researcher desires to save the current environment for future use, they would click on the “Save Your Environment” button, which will allow them to name and save the environment. Note that the researcher must be registered and logged into the application to save an environment. Once the researcher has reviewed and confirmed the simulation environment, they will need to click on the “Next” button to proceed to Step 3.



## Future Release

In future releases DATASIM will allow for the researcher to reconfigure their environment directly in the application.

## Step 3: Select Your Personae

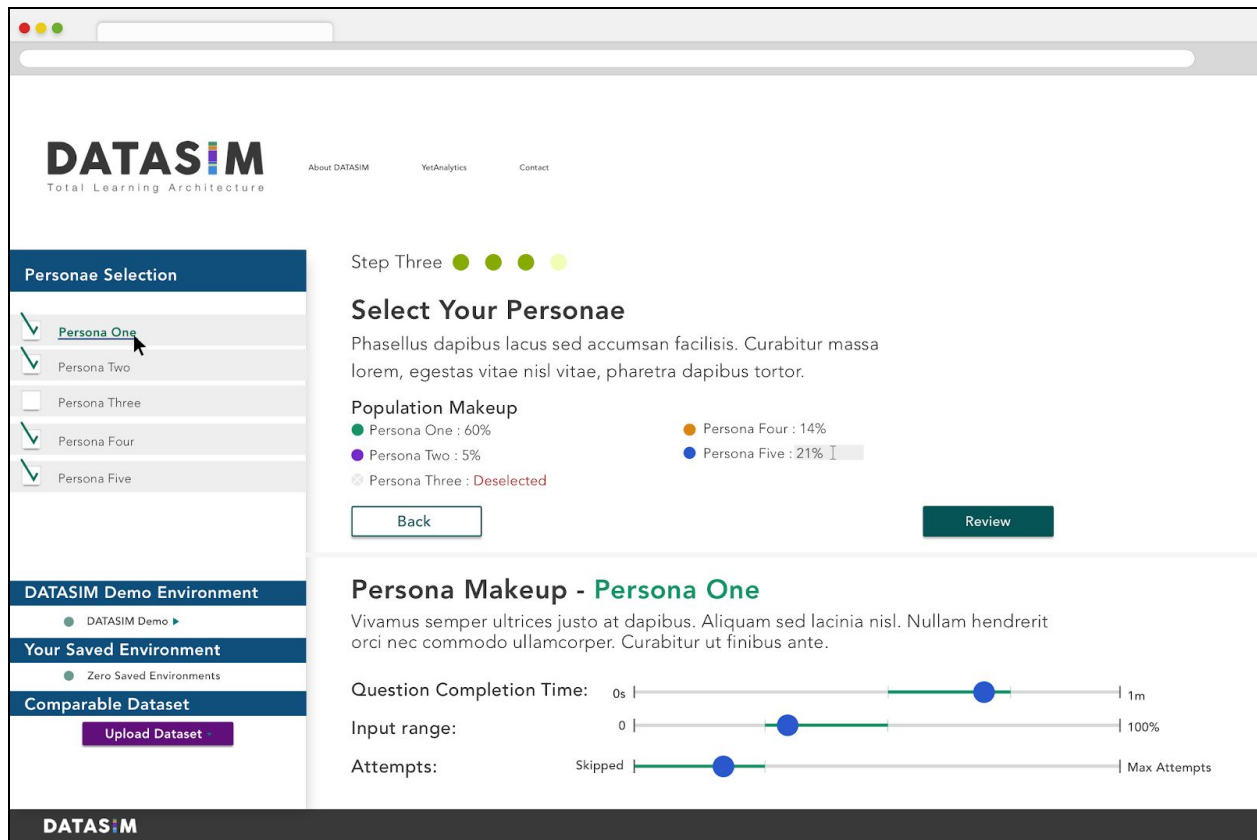


Figure 7: Persona configuration

### Alpha Release

Personae are what define the characteristic range of possible agent activities within the simulation on an event-by-event basis. Each persona is built from multiple characteristics; each characteristic is made up of different behaviors. In any given persona there is one dominant characteristic and multiple latent characteristics. Each persona and the associated characteristics of that persona will dictate an artificial agent's response event-by-event in the simulation.

The Alpha release will include 3-5 pre-configured personae in the DATASIM application. As demonstrated in Figure 7, the application will present the researcher with the list of pre-configured personae in the panel on the left side of the screen. To configure the agent population for the simulation, the researcher will select one or more personae by clicking the check box(es) of the appropriate personae in the left panel. As the researcher selects personae, the agent population makeup will be presented in the middle panel of Figure 7, where the researcher will also input the relative weight of each persona on outcomes of behavior within the overall agent population. Note that personae weights must equal 100%.

If the researcher clicks on a persona from the agent population list, the application will present associated characteristics and behaviors of the selected persona at the bottom middle of the screen (see Figure 7). Once the researcher has selected all desired personae and weights, they will then click the “Review” button to process to Step 4.

### Future Release

In future releases of DATASIM, the application will allow researchers to configure and save their own personae as well as identify ranges of behavior associated with each environment template and/or pattern.



## Step 4: Review Simulation Settings

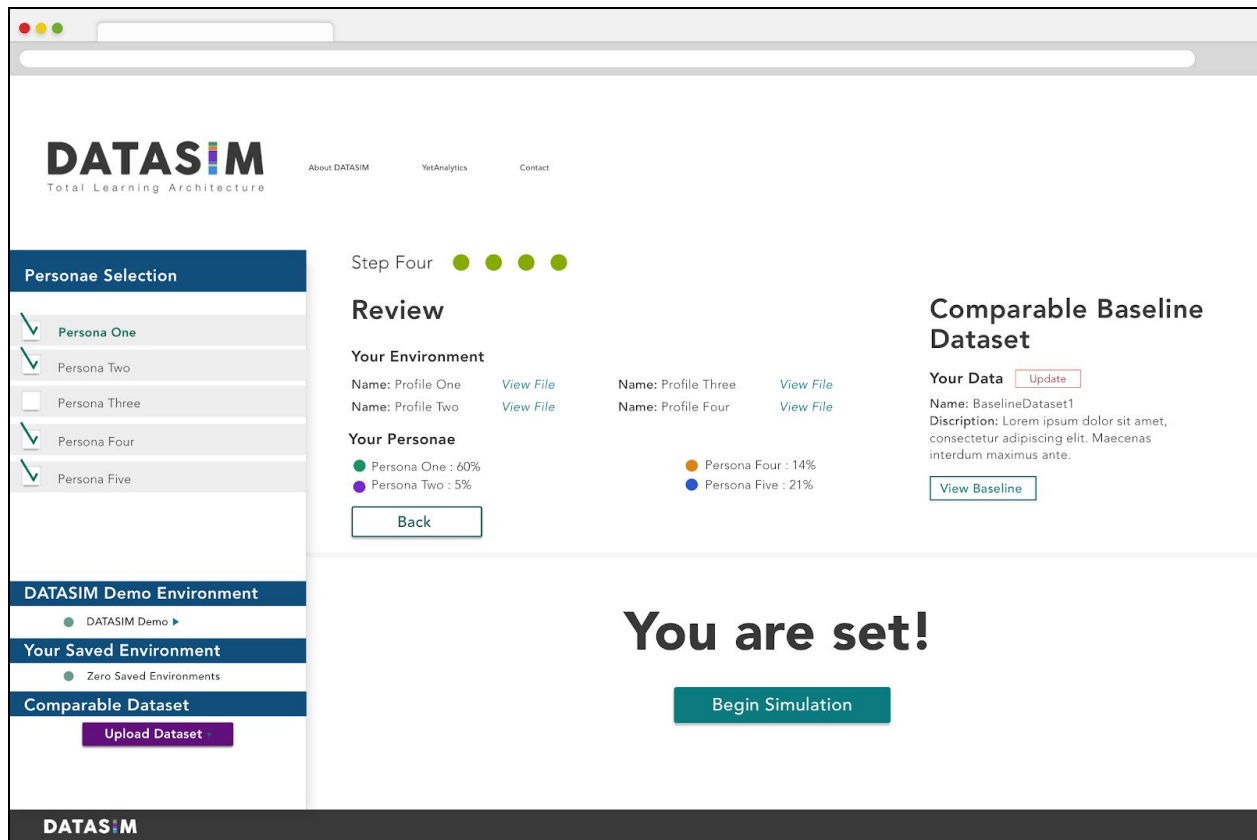
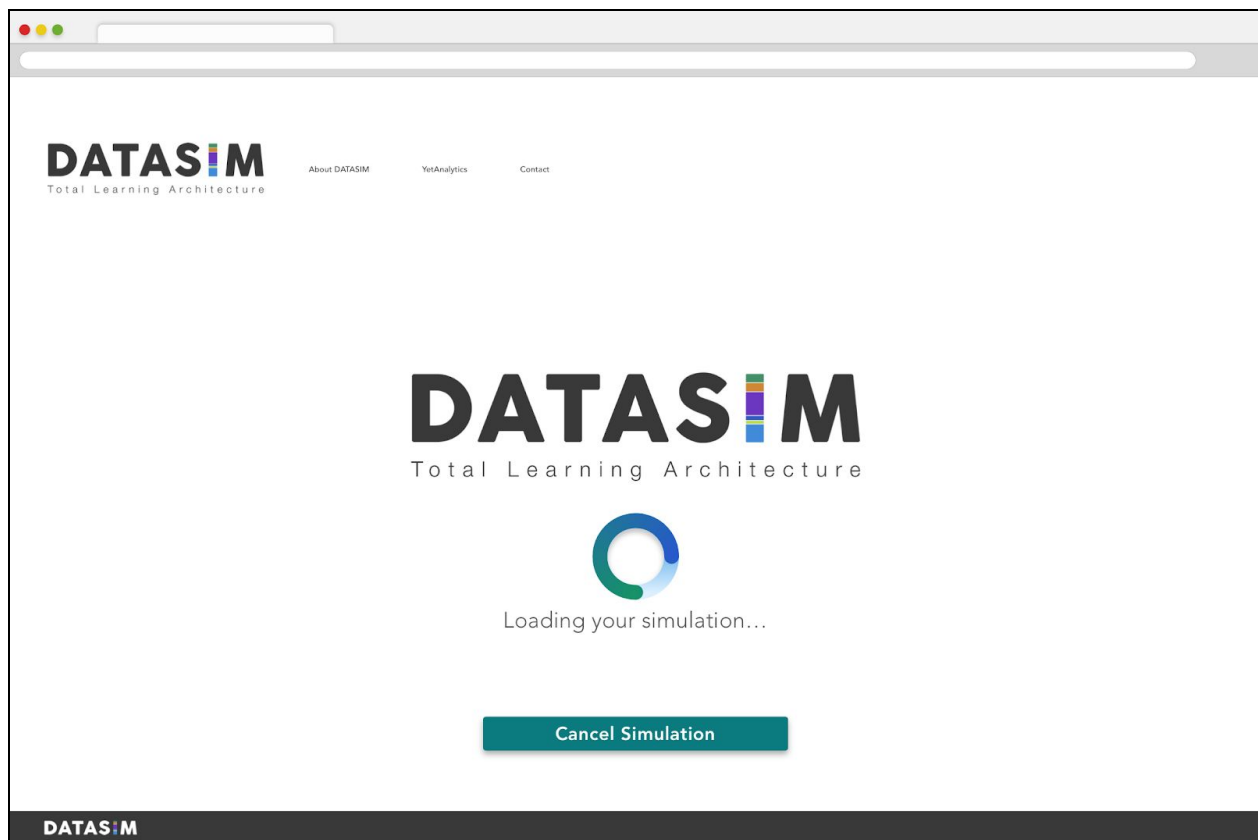


Figure 8: Review parameters

Once the researcher has confirmed their environment (Step 2) and configured their agent population (Step 3), they will need to review their simulation settings as demonstrated in Figure 8. If the researcher wishes to make any changes to their simulation settings, they will need to navigate back to the appropriate step and make the desired changes. At this point, the researcher can click the “Begin Simulation” button to execute the simulation.

## Simulation Processing Page



*Figure 9: DATASIM indicates that a simulation is being generated*

Figure 9 demonstrates the researcher experience while the simulation is executing. The researcher can cancel the simulation on this screen by clicking the “Cancel Simulation” button.

## Your Simulation Results

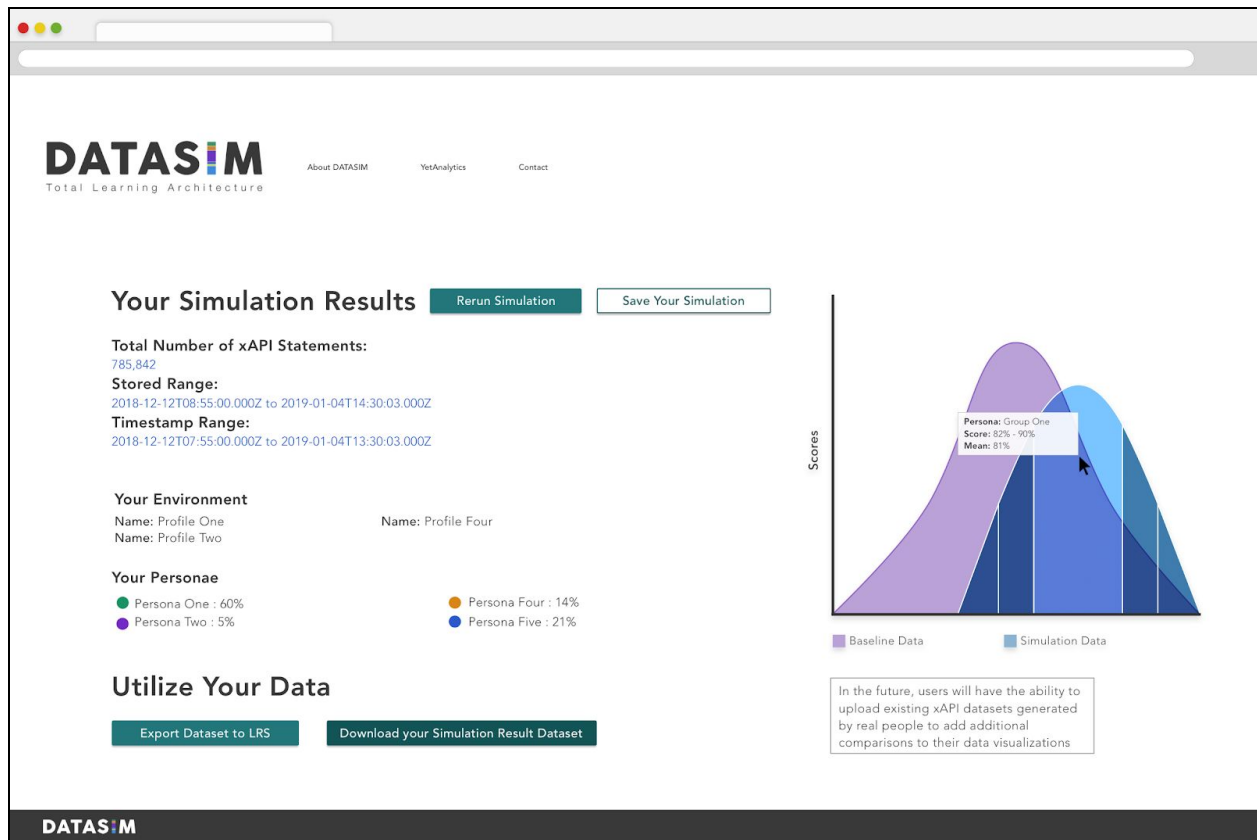


Figure 10

### Alpha Release

The final screen of DATASIM presents summary statistics of the simulated data set (top left of Figure 10), options for the researcher to export the dataset to an LRS and/or download the simulation output for future use (bottom left of Figure 10), and a graph comparing the simulated dataset to the system-generated baseline dataset (bottom right of Figure 10). If the researcher has chosen to upload additional comparison datasets, these comparable datasets will also present on the graph. The researcher can rerun the simulation or save the simulation from this screen as well.

### Future Release

In future releases, DATASIM will allow the researcher to import additional comparable datasets at this stage.

## Conclusion

The purpose of the UX/UI of DATASIM is two-fold. First, it must provide the end user with the ability to generate the appropriate datasets. Secondly, because DATASIM's end users may be of a variety of types — learning scientists, ISDs, IT and decision-making stakeholders — the interface has to be generally accessible. The goal is that the design of the Alpha implementation of DATASIM will follow the design and iterations of design represented by this document and subsequent stakeholder conversations.