Product Discovery Report
for a Visual Data Management System

December 2018

A Report by Caktus Group
for the University of North Carolina School of Education
derry@email.unc.edu


© 2018 by Caktus Group. Material in this paper is based upon work supported by a National Science Foundation Cyberlearning grant (Award no. 1217027) called "EXP: Implementing a Workflow Visualization System for Design-Based Research." All rights reserved. Any opinions, findings, or conclusions expressed in this paper are those of the author and do not necessarily reflect the views of the funding agencies, WCER, or cooperating institutions. Readers may make verbatim copies of this document for noncommercial purposes by any means, provided that the above copyright notice appears on all copies. WCER working papers are available on the Internet at https://wcer.wisc.edu/publications/working-papers
Abstract

This report is a companion piece to WCER working paper 2018-14, “A Cyberinfrastructure for Design-Based Research: A Collaborative Design Project Proposal.” That paper stressed that the educational research community would be well served by a mutually created cyberinfrastructure that encourages and supports engagement by multiple design-based researchers in working toward answers to important theory-driven research questions, moving the field toward a “bigger science” approach.

This current report, prepared by the Caktus Group for the University of North Carolina School of Education, outlines a process by which a single digital tool specific to design-based research (DBR) methodologies could be built. The goal in commissioning this report was to discover the most effective and suitable system to recording and visually representing learning environment workflows, building searchable collections of well-catalogued assets, and securely sharing data and analysis with the community of peers.

With that goal in mind, the Caktus Group, in collaboration with Dr. Sharon Derry, built and administered a survey to better understand the needs of the DBR community. Following the survey, and informed by its results, the Caktus Group facilitated a discovery workshop to identify design and technical requirements for the system, outline its desired architecture, and gather information for an estimate of the cost of building such a system.

This report proposes a full-scale design for a workflow visualization system to support design-based research in the learning sciences. It builds on participant experiences with prototypes and incorporates outcomes of survey research and a two-day intensive design workshop led by Caktus Group that took place on July 31–August 1, 2017, in Durham, NC. The work was conducted under direction of Sharon Derry. Invited participants in the workshop were Janice Anderson, Kelly Barber-Lester, Lana Minshew (all from University of North Carolina at Chapel Hill), and Vanessa Svhila (University of New Mexico).
Product Discovery Report for a Visual Data Management System

Design-based research (DBR) methodologies for studying student outcomes associated with specific educational interventions create a wealth of multi-modal data (video, assessment data, curriculum materials, student work) for researchers to study. Searching through large, multi-modal datasets, researchers must be able to rapidly identify and locate specific items needed for research on effectiveness of educational interventions. Often, specific data must be selected, accessed, and analyzed during a short turnaround time between design iterations. The multi-modal qualitative and quantitative data sources leave research teams with an untenable data management task.

These problems can be tackled by categorizing collected data, organizing and storing data in the context of visual representations of learning environments’ workflows, and employing search and query tools to access data easily within those contexts. This ability to collaborate and share data and analysis with peers would build a strong foundation for the entire DBR community to improve and standardize research processes, facilitate immediate knowledge sharing, and speed the rate at which effective new educational applications are created. Particularly, building a single digital tool specific to DBR methodologies to tie all of these aspects together (representation, organization and storage, search, and sharing), in a way that makes each step easier, would ultimately improve the DBR approaches that practitioners can apply in various learning settings.

I. Methods

The University of North Carolina (UNC) School of Education engaged the services of the web experts at Caktus Group to discover the most effective and suitable system solution to recording and visually representing learning environment workflows, building searchable collections of well-catalogued assets, and securely sharing data and analysis with the community of peers.

In collaboration with Dr. Sharon Derry, Caktus Group built and administered a survey to better understand the needs of the DBR community. Following the survey, and informed by its results, Caktus Group facilitated a discovery workshop to identify design and technical requirements for the system, outline its desired architecture, and gather information for an estimate of the cost of building such a system.

1. Survey

Survey questions were constructed by Dr. Derry in collaboration with her team and the Caktus team. The survey was built, distributed, and responses were collected using Qualtrics. The survey was administered among members of the International Society of the Learning Sciences. There was a total of 81 respondents, of whom 41 completed the survey. In total, 50 of the respondents indicated clear understanding of design workflows. Of those 50, 30 provided responses to all 25 survey questions, and the remaining 20 respondents answered questions selectively.

The survey results provided information about ways researchers carry out design projects, elements that are important to be captured in design workflows, ways researchers collect and access data, types of data, types of software currently in use, capabilities that researchers would
look for in a new digital system, and more. See Appendix 1 for a detailed summary of the survey results.

2. Discovery Workshop

A discovery workshop is a design-thinking based, user-centered approach to problem solving that is leveraged in the context of software design and development. It allows all stakeholders to explore various aspects of the system under consideration, including user perspectives, user flows, and contexts in which the system and its parts will function.

Caktus Group conducts discovery workshops structured around three questions:

- What is the problem that must be solved?
- For whom is that problem being solved?
- How will the problem be solved?

The specific workshop goals included:

- refining the problem statement
- identifying user flows and tasks
- defining the list of most valuable features
- defining system architecture
- identifying possible approaches to user interface design
- generating information for a technical estimate.

Dr. Derry’s and Caktus’ teams reached workshop goals through collaboration on:

- formulating the problem statement for the project
- identifying project constraints
- identifying the key target user persona
- mapping out user flows and detailed user tasks that the system should support
- prioritizing the system’s features
- diagramming the system architecture.

2.1. Problem Statement

Researchers who pursue DBR are in need of a web-based application that:

- allows for diagramming learning environments workflows (henceforth, “workflows”),
- captures the hierarchical structure of workflows,
- assists in managing and easily accessing large numbers of data files and associated analyses,
- is searchable through the use of workflows,
- enables secure sharing of workflows with or without associated data file and analysis with peers, and
- captures the structure of the overarching research project.
2.2. Constraints

Constraints are any limiting requirements or circumstances under which the system would be built and released. We identified the following constraints for the visual data management system:

- must comply with security requirements of individual academic institutions
- must comply with IRB guidelines
- must be easy to use
- may meet resistance to adoption by seasoned researchers who have established processes and tool sets
- may be difficult to adopt widely because of lack of standardized DBR procedures
- must be maintained over time
- may require some level of customer support
- must have allocated sufficient time and resources to be developed, tested, and iterated upon.

2.3. Key Target User Persona

Seasoned researchers have established procedures and tools they might not be willing to give up for a new tool. For that reason, and based on information from interviews that Dr. Derry’s students conducted with a number of researchers with varying levels of experience in the field, the workshop participants concluded that a researcher with fewer years of experience might be a key target user persona. Such a researcher is likely to struggle more with lack of standardized DBR processes and the amount of data that DBR generates. She/he may seek a single tool that would allow her/him to organize their work and data, and access it easily for analysis. Here is a hypothetical target user profile:

My name is Emily Sharp. I’m an assistant professor at Pineapple Under the Sea University. I am passionate about improving public education, particularly in STEM disciplines.

Over the last three years, using DBR methodologies, I have been studying a variety of approaches to teaching science in high school. I want my work to have real impact on practitioners in the field and on students, but I cannot analyze data from my research fast enough. There is simply too much data to analyze and no good way to organize and keep track of it.

I need a digital tool that would allow me to capture the structure of my research projects, diagram learning environment workflows, and visualize levels of their hierarchy, manage large amounts of data and associated analysis, search and filter my data in the context of learning environment workflows, and securely share data and analysis with peers.

2.4. User Story Mapping

User story mapping is a visualization technique popularized by Jeff Patton. It is frequently leveraged in the product discovery process for software. During the workshop Dr. Derry’s and
Caktus’ teams mapped out the anticipated user flows and associated with them user tasks. Design implications are discussed in detail under section II: Resulting Design Considerations.

2.5. Feature Prioritization

The user story mapping activity revealed user flows and tasks that the system would ideally support. However, some features needed to support those user flows and tasks could be developed at later stages of the system’s life cycle. Workshop participants identified features that could be relegated to later releases of the software.

2.6. System Architecture

Once user flows and tasks were identified through user story mapping, the workshop participants discussed relationships between discrete system entities. System entities identified for the visual data management system included:

- projects
- users
- conjecture Maps
- workflows / Workflow Iterations
- nested workflows / nested workflow iterations
- workflow nodes
- folders / files
- participant groups / participants
- settings
- tags.

A system entity relationships diagram is provided as Appendix 2.

II. Resulting Design Considerations

1. Overview

For the proposed visual data management system to support management of large amounts of data, be scalable, and serve as a vehicle for sharing knowledge, analysis, and data, we at Caktus Group recommend building a web-based application. We also recommend using the Django web development framework due to its excellent suitability for data-driven web applications. The sections below discuss the features proposed for the application.

2. Most Valuable Features

The following list of most valuable features is a direct result of user story mapping and feature prioritization conducted during the discovery workshop. It includes features deemed necessary for the system to deliver on user value. Based on results from the survey and the discovery workshop discussions, these features seem essential for the target audience to adopt the visual data management system.
2.1. Home Page
The application will need a home page. The essential features that need to be supported by the home page include:

- Link to log in:
  - User can click on the link to log in.
- Security statement:
  - User can read a security statement to access the application’s compliance with security and IRB guidelines.

2.2. User Accounts and Projects
We propose that within the first release of the application, users will be able to create accounts by invitation only. This approach will ensure that only users with serious interest in the project will be granted access to the application, and the application’s storage capabilities will not be overwhelmed by excessive amounts of data.

When creating an account, users will be asked to provide only the most essential information such as name, user name, email address, and password. Users will be able to add optional information to their accounts at any time once the account has been created. That information may include institution, highest degree earned, address, and phone number.

Within their accounts, users will be able to create projects. Projects will allow users to organize work and data pertaining to research projects they run.

It is imperative that users collaborate with team members on projects. We propose that users be able to grant permissions to team members to view; view and edit; or view, edit, and share projects.

The following are the user flows and user tasks that pertain to user accounts and projects within the application:

- Create account:
  - User can create an account by invitation only.
  - User must enter required information to create an account (Name, Username, Email address, Password).
- Manage account:
  - User can enter optional account information (Institution, Highest degree earned, Phone number, Address).
  - User can add / edit / delete account information (only optional information can be deleted).
  - User can delete an account.
- Log into an account:
  - User must enter email address.
  - User must enter password.
  - User can reset password.
• Create project:
  o User can create a project.
  o User must name a project to create it.
  o (Project metadata such as date and owner will be added to the project automatically upon its creation.)

• View project:
  o User can view who has access to project.
  o User can view conjecture maps associated with the project.
  o User can view the list of workflows, their levels, and iterations within a project.
  o User can view the list of folders and files in the project.
  o User can view the list of participant groups / participants associated with the project.
  o User can view project’s metadata (descriptions, tags, owners, date created etc.).

• Manage project:
  o User can edit project name.
  o User can add / edit / remove project description.
  o User can add / remove project tags.
  o User can give team members project permissions to view and edit; view, edit, and share; or view-only.
  o User can back up a project.
  o User can recover a project from backup.
  o User can delete a project.

2.3. Conjecture Mapping

Conjecture mapping is a process of coming up with a learning environment design that for theoretical reasons should accomplish a certain set of goals. For the visual data management system to support a researcher’s work, it must not only allow for diagramming learning environment workflows, but also their theoretical underpinnings (the conjecture maps).

The following are the user flows and user tasks that pertain to conjecture mapping within the application:

• Create conjecture maps:
  o User can create a conjecture map.
  o User must name a conjecture map to create it.
  o User must use a conjecture map template to create it.
  o User can add text fields to template columns to build the conjecture map.
  o User can use arrow connectors to connect text fields across columns.

• View conjecture maps:
  o User can view a conjecture map.
  o User can see all connections of the conjecture map to other entities within the system.

• Manage conjecture map:
Report for a Visual Data Management System

- User can connect / disconnect conjecture maps and their iterations to workflows, workflow levels, workflow iterations, and workflow nodes.
- User can link / unlink files to conjecture maps and their iterations.
- User can edit a conjecture map.
- User can delete a conjecture map.
- User can add / remove conjecture map tags.
- User can add / edit / remove conjecture map description.
- User can copy and rename a conjecture map to create an iteration.
- User can copy a conjecture map to another project.

2.4. Workflow Diagramming

Learning environment workflows are a visualization tool that researchers use in DBR approaches. In the context of the visual data management system, they would also serve as a visual search and filtering tool to allow researchers to access data and files quickly within the context of workflows, their levels, iterations, and nodes.

**Workflow anatomy.** Workflows would be built as a series of nodes connected to one another with connectors derived from workflow visual language. Each node would be represented by an icon. The application would provide a set of core icons, plus a generic icon that could be customized by the user to represent any node that cannot be represented by a core icon.

A user could zoom into each node in a workflow by expanding it into its own (lower level) workflow. The user would be able to navigate into lower level workflows and back. Workflows could be marked as intended or enacted.

Workflows, their levels, and nodes could be annotated with description and tags. They could be linked to folders / files, participant groups / participants, and assigned particular settings.

**Workflow iterations.** At any level of their hierarchy, workflows could be iterated upon and saved as iterations.

**Core icons.** The core list of icons to represent workflow nodes is derived from the results of the survey and includes:

- assessment
- tools/technology
- expected learning outcomes
- learning activity
- text-based learning material
- multimedia learning material
- scaffolding
- stations/exhibitions.

We recommend that designing icons be treated as its own project, separate from the development effort. It would require hiring a visual designer, and iterating on the design through a process of design and user testing until visual representations are found that align with users’ mental models.
The following are the user flows and user tasks that pertain to learning environment workflows within the application:

- **Create a workflow:**
  - User can create a workflow.
  - User must name a workflow to create it.
  - User can build a workflow by adding nodes.
  - User can connect workflow nodes with connectors derived from visual language for workflows (a final list of connectors must be decided upon).
  - User can create nested workflows (workflow levels).

- **View a workflow:**
  - User can view a workflow and its iterations.
  - User can zoom into a lower level workflow (possibly by double clicking on a workflow node).
  - User can back up from a lower level workflow into a higher level workflow.
  - User can see and access conjecture maps associated with workflow, workflow level, workflow iterations, or workflow nodes.
  - User can view and access files / folders associated with workflows, workflow levels, workflow iterations, or workflow nodes.
  - User can view and access participant groups associated with the workflows, workflow levels, workflow iterations, or workflow nodes.
  - User can view the library of core icons that can be used to represent nodes.

- **Manage a workflow:**
  - User can edit names of workflows, workflow levels, workflow iterations, or workflow nodes represented by a generic icon.
  - User can add / edit / delete description of workflows, workflow levels, workflow iterations, and workflow nodes.
  - User can add / remove tags associated with workflows, workflow levels, workflow iterations, and workflow nodes.
  - User can delete workflows, workflow levels, workflow iterations, or remove workflow node(s) from workflows.
  - User can copy and save workflows at any level as a new workflow.
  - Users can copy and save workflow at any level as an iterations (iterations are copies of the same workflow that remain connected).
  - User can mark a workflow, workflow level, or workflow iteration as intended or enacted.
  - User can manually add / edit / remove date and time of implementation (enactment).
  - User can attach / detach files to workflows, workflow levels, workflow iterations, and workflow nodes.
  - User can link / unlink to conjecture maps from workflows, workflow levels, workflow iterations, and workflow nodes.
User can link / unlink to participant groups / participants from workflows, workflow levels, workflow iterations, and workflow nodes.

- User can assign / unassign “Setting” to a workflow iteration (e.g., urban, rural etc.).
- User can assign / unassign another user to workflows, workflow levels, workflow iterations, and workflow nodes to communicate to a colleague which entity they should look at.
- User can add / remove time span to a workflow node.

2.5. File Management

A main goal of a visual data management system is to provide an easy way for researchers to organize, store, and access their files (including research data) in the context of the project architecture and learning environment workflows.

The proposed system would use workflows as a primary means of organizing and accessing files. Users would be able to attach folders and files to workflows, their levels, iterations, and nodes. These folders and files could then be searched for in the context of and across workflows.

The following are the user flows and user tasks that pertain to file management within the application:

- Upload files:
  - User can create folders.
  - User can upload files into folders.
  - A file naming convention is enforced by the system upon file upload (the file naming convention needs to be defined).
  - User can manually name the file if they do not want to follow the system’s naming convention.
  - User can enter who collected data in the case of research data files.

- Securely store files:
  - User must store data on a secure server.
  - The system must comply with IRB guidelines.
  - The system must comply with security requirements of individual academic institutions.

- View files:
  - User can view folders and files.
  - User can view files inside folders.
  - User can see what workflows, workflow levels, workflow iterations, workflow nodes, conjecture maps, participant groups, or participants are linked to a given file / folder.
  - User can see the contents of some files within the visual data management systems (e.g. the system may allow streaming time-based media in the browser for file formats that are easily handled in supported browsers).
• Manage files:
  o User can delete files / folders.
  o User can manually overwrite the system-generated filenames.
  o User can add / remove tags from files.
  o User can add / edit / delete file descriptions.
  o User can copy / move files between folders and projects.
  o User can link / unlink files to workflows, workflow levels, workflow iterations, workflow nodes, conjecture maps, participants groups, other files.
  o Assign / unassign a user to files / folders to communicate to a colleague which file they should look at.
• Download files:
  o User can download a file and open it automatically in its native application (e.g. by double-clicking on the file within the visual data management system).

2.6. Participant Groups
Researchers must be able to associate the workflows they build (together with the relevant data and files) to groups of participants and to individual participants. Participant groups could be nested within one another and could include:

• school districts
• schools
• classes
• teachers
• students
• parents.

The following are the user flows and user tasks that pertain to participant groups and participants within the application:

• Create participant groups:
  o User can create a participant group.
  o User must name a participant group to create it.
  o User can create nested participant groups.
  o User can assign / unassign type to a participant group.
  o User can import a list of participants into a group (e.g. an Excel spreadsheet with student unique IDs and pseudonyms).
• View participant groups:
  o User can view participant groups.
  o User can view nested groups within their parent groups.
  o User can view participants lists in participant groups.
  o User can see what workflows, workflow levels, workflow iterations, workflow nodes, conjecture maps, folders, and files are linked to a given participant group or participant.
• Manage participant groups:
  o User name edit a name of a participant group.
  o User can edit a participant group’s type.
  o User can add / edit / delete a participant group description.
  o User can add / remove participant group tags.
  o User can add / delete individual participants in a list of participants
  o User can delete participant groups.
  o User can copy participant groups across other participant groups (e.g. copy a group of students to another class).
  o User can enroll / unenroll participant groups (e.g., students) in a workflow.
  o User can link / unlink participant groups to files, conjecture maps, workflow nodes
  o User can assign a “Setting” participant group.

2.7. Search and Filtering

A primary goal for building a visual data management system is to offer researchers an easy way to access data and files. It is especially important that researchers be able to search for and filter data and files across multiple dimensions that include:

• projects
• conjecture maps
• workflows, workflow levels, workflow iterations, workflow nodes
• participant groups and individual participants.

Researchers must also be able to search and filter data and files by a number of criteria:

• keywords (terms found in user entered descriptions)
• tags (user entered words or short phrases that are used to label system entities)
• date / time
• setting
• data format (e.g., .mp3, .mp4, .txt, .jpg etc.)
• data type (e.g., research data, learning object, study design document, design material, etc.)
• who collected data.

The following are the user flows and user tasks that pertain to search and filtering within the application:

• Search:
  o User can search
    ▪ projects;
    ▪ workflows, workflow levels, workflow iterations, and workflow nodes;
    ▪ conjecture maps;
    ▪ files (by filename and metadata, not file contents) / folders;
    ▪ participant groups and individual participants.
Report for a Visual Data Management System

- User can search by
  - tags;
  - keywords;
  - date / time;
  - data type (if there are many possible values; for fewer possible values, this could be handled by filtering);
  - who collected data.
- User can search within the constraints of applied filters.
- Filter:
  - User can filter
    - projects;
    - workflows, workflow levels, workflow iterations, and workflow nodes;
    - conjecture maps;
    - files (by filename and metadata, not file contents) / folders;
    - participant groups and individual participants.
  - User can filter by
    - date range;
    - file format (if there are few possible values; for many possible values, this would be handled by search);
    - who collected data.
  - User can apply multiple filters within and across entities (e.g., by multiple workflow iterations and participant groups).

2.8. Sharing

During the discovery workshop, we decided that sharing capabilities for the first release of the software will be limited to granting levels of permissions for projects to team members as discussed under section 2.2: User Accounts and Projects.

2.9. Other

In addition to user flows and tasks associated with the system’s core functionalities, Dr. Derry’s and Caktus’ teams identified a few auxiliary requirements. Those include:

- Working offline:
  - User can add / edit description to any entity in the system.
- Technical support:
  - User can report bugs by sending an email to an email address provided within the application.
- Host application:
  - User’s parent institutions can purchase managed hosting to host an instance of the application.
• Analytics:
  o A report from the application logs or database can be generated by developers building the system.
  o A word cloud from tags used within the application can be generated by developers building the system.

3. Entity Relationships

Relationships between entities of which the proposed system would be comprised have been capture in the entity relationships diagram (see Appendix 2).

4. Interface Design Ideas

A set of 24 wireframes is attached as Appendix 3 to illustrate possible interface design ideas to support the identified user flows. An example wireframe is included in Figure 1.

Figure 1. A workflow view in the context of the system navigation (left-hand panel).
Workflow attributes such as the description and tags are displayed in the right-hand panel. The three panels in the lower part of the screen show entities attached to or associated with the workflow in view. The “Files” panel features a collapsible “Filter” overlay.
III. Estimate

Caktus will build the visual data management system using Python, Django, and supporting technologies. Caktus builds and deploys Django-powered projects every day, and we continually refactor and improve our internal processes to ensure the best possible outcomes for our clients. Django, by design, is perfectly suited for sites that have complex needs and unique data models. It is highly flexible and easy to develop.

Iterative phases of development allow for increased collaboration between Caktus and the Visual Data Management System stakeholders. Every two weeks, the Caktus team will produce artifacts or production-ready code for review and deployment.

1. Proposed Technology Stack

Subject to review and approval, Caktus will deliver the website redesign project with the following technologies and toolsets:

<table>
<thead>
<tr>
<th>Type</th>
<th>Application/Toolset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programming Language</td>
<td>Python 3.6</td>
</tr>
<tr>
<td>Web Framework</td>
<td>Django 1.11 LTS</td>
</tr>
<tr>
<td>Developer languages and tools</td>
<td>Git, CSS, JavaScript, HTML</td>
</tr>
<tr>
<td>Provisioning, Deployment &amp; Servers</td>
<td>Project to include deployment tools using Ansible (and/or other deployment methods as needed)</td>
</tr>
<tr>
<td>Methodology</td>
<td>Agile and Scrum</td>
</tr>
<tr>
<td>Database</td>
<td>PostgreSQL</td>
</tr>
</tbody>
</table>

2. Development Methodology

Caktus strongly believes in close collaboration and consistent communication, demonstrated by following Agile project best practices. Agile development is characterized by the division of tasks into short phases of work and frequent reassessment and adaptation of plans. Caktus performs two-week iterations (also called sprints) using Agile principles. The goal of each iteration is to create working, tested software. To facilitate this, a Caktus project manager (PM) will coordinate the efforts of developers, ensure timely deliverables, and manage communications between Caktus and stakeholders. The backlog of tasks for the project will be reviewed by the PM and client contact prior to each sprint in order to ensure that the most valuable tasks are worked on for each iteration. Feedback from the client and end users as applicable is encouraged at the end of each iteration. The PM will also manage the following:

- Regular meetings with client team. Client will receive an agenda the day before.
- Typically covered topics include updates, software improvements, new features, a technical analysis (including a summary of technical improvements and/or results from benchmark tests), and risk identification. Action items are identified and assigned in or after the meeting.
• Daily internal meetings with the Caktus team. The PM meets daily with the development team to ensure that each sprint’s deliverables are on track and to discuss any challenges that may need to be brought to the client’s attention.
• Project schedule and prioritization. Caktus runs two-week development sprints for prioritizing tasks and accomplishing them quickly. The PM prioritizes work into sprints with the goal of completion within the project timeframe.
• Maintain transparent communication. The Caktus team is available to the client for any questions or comments, and we ensure a prompt reply. Caktus can adopt the preferred communication tools of our client. If there is no preference, Caktus will use Basecamp.

3. Server Infrastructure and Security
Caktus has created and released as open source the django-project-template. We utilize this as a starting point for all new Django-based projects, as it embeds best practices and provides boilerplate code and scaffolding to start development immediately.

The template includes:
• developer setup and configuration documentation to streamline onboarding new developers
• use of virtual environments to manage Python project dependencies
• scripted server provisioning and deployment using Fabric and Ansible to Virtual Machines on Amazon Web Services
• support for industry standard web technologies out of the box: PostgreSQL, PostGIS, RabbitMQ, Celery, Nginx, Gunicorn, Memcached, NodeJS, Less CSS.

The django-project-template enables Caktus to provision servers and deploy testing environments efficiently. A fully deployed staging environment will be up and running within the first few days of development.

Caktus also takes security seriously and follows best practices, including the rigorous standards of the Django project. Native Django security features are considered highly robust with protection from cross-site scripting, cross-site request forgery, SQL injection, clickjacking, and more. Caktus performs regular reviews and patch updates to maintain security.

4. Schedule
Caktus can begin scheduling this work effort as soon as the client signs all agreements. We anticipate that the project will take 29 2-week Sprints.

As an example, if the first development sprint begins on January 1, 2018, the expected end date would be February 4t 2019.
5. Budget

The total price for this effort is $1,225,000. Caktus will invoice on a monthly basis for work completed. The actual invoicing schedule will be confirmed once the contract has been awarded and the work has been scheduled.

IV. Conclusions

As is apparent from this report, the proposed visual data management system would be a complex application that would require a significant amount of time and resources to develop. Designed and built in the spirit of collaboration by gathering feedback from the DBR community through field testing, such a system would bring enormous benefits. It would allow researchers to understand the data they collect in the context of an overarching project structure, as well as in the context of conjecture maps and learning environment workflows with increasing levels of granularity (from high level to low level workflows, across workflow iterations, and against individual workflow nodes). If adopted broadly, the system could support guiding DBR methodologies in the use of best practices and help standardize DBR approaches.

The list of most valuable features that has been detailed and estimated in this report constitutes a set of essential requirements that would deliver sufficient value to potential users to ensure that the system would be tried by researchers in the field within the constraints of a limited and controlled release. For example, initially users could only create an account by invitation, and projects would only be shared among members of the same team. That would ensure a level of control over who is using the system, and would initially eliminate a need for certain security measures such as requiring a user to attach an IRB certification to their profile before a project can be shared with them.

Dr. Derry’s and Caktus’ teams also discussed enhancements to the most valuable features as well as additional features that could be added to the system later on to improve its functionality and increase its user value.

A few selected examples of enhanced user flows and tasks are listed below:

- Create an account:
  - Anyone can create a trial account by clicking a link to “create account” from home page.
  - User with a trial account has access to a limited amount of storage space.
  - User with a trial account can request more storage space.

- Manage account:
  - User can attach an IRB certificate to their account/profile.

- Manage projects:
  - User can archive a project.
  - User can assign an analysis status to a project (e.g. in progress, done, published, etc.).

- View conjecture maps:
  - User can view conjecture maps side by side to compare them.
• View workflows:
  o User can see where exactly workflows, workflow levels, or workflow nodes differ across iterations.
• Manage workflows:
  o User can assign an analysis status to a workflow (e.g. in progress, done, published, etc.).
• Share:
  o Users can make projects, workflows, files / folders public.
  o Users can request permissions from other users to access their projects, workflows, files / folders.
  o Users can grant permissions to any qualified user for their projects, workflows, files / folders.
V. Appendices

Appendix 1: Summary of Survey Results
Appendix 2: Entity Relationships Diagram
Appendix 3: Wireframes of Select Interfaces
Appendix 1: Summary of Survey Results

The following is a summary analysis of the survey’s 50 respondents who indicated a clear understanding of design workflows. The survey begins with question 3.

Question 3: Which elements are important to your design and there should be as icons included in icon pallet? (abbreviated)

Responses to Question 3

Number of responses: 38

Rank order of suggested icons with distribution of responses (very important, somewhat important, unimportant):

1. Learning activity (38, 0, 0)
2. Assessment (32, 5, 1)
3. Expected learning outcomes (31, 5, 2)
4. Scaffolding (27, 10, 0)
5. Tools/technologies (26, 12, 0)
6. Text learning materials (13, 22, 2)
7. Multi-media materials (14, 22, 2)
8. Stations/exhibits (7, 20, 8)
9. Other 1 (10, 2, 0)
10. Other 2 (8, 2, 0)

Comments for “Other”: Other 1:

- Context of design
- Learner characteristics/demographics (part of context?)
- Level of activity—whether individual, group, plenary
- Process data icons to visualize what type of data collected
- User needs—empathy based “pain points” in design
- Human resources required
- Mediating processes as in Sandoval’s conjecture maps (also other icons from conjecture mapping)
- Types of scaffolding

Other 2:

- Curriculum being used
- Technical support
- Movement in room
- Intro and debriefing material
- Theory/principles that informed design (similar to Sandoval’s conjecture maps)
- Collaboration structures
- Teacher reflection points
• Forms of data collected/available
• Feedback (not sure what this meant but might be assessment of how design worked or feedback to students)

Suggestion from survey taker: Either represent learning resources as more specific categorizations or represent them in more general way to be annotated.

**Researcher Reflections on Question 3 Results**

Most icons listed in survey considered important although biased by classroom researchers, non-traditional or informal environments less represented. Learning activities, assessments, expected learning outcomes and scaffolding unequivocally important although these need to be divided into types and it is not clear how much of this should be done with icons and how much with check boxes or annotations to icons. Users are clearly calling for more types but (later in survey) one user notes that learning too many icons is too hard, like learning a new language. I agree.

It is unclear how to break down learning resources as icons or sub-categories represented as check-box or written annotations. Are tools and technologies part of broader category of learning resources? One user was dissatisfied with text versus multi-media classification of learning resources.

Clearly there is interest/need in being able to include theoretical basis for design, conjectures Sandoval’s conjecture maps are mentioned but don’t think this alone would permit adequate archiving and retrieval of data, certainly not for multiple purposes. Including conjecture maps capability in system seems important.

There are some good suggestions to consider such as add icon to quickly access context (including student demographic information, debriefing material?), process data icons to visualize where and what types of data (and analyses?) are available, user based pain points in design, human resources needed here, technical support needed there, teacher reflection points (this may be a type of data collection).

User experience design discipline offers some evidence that there is a difference between users’ stated needs and their actual needs as revealed through their interactions with software. For that reason testing designs with users is important. It might be good to treat icons design as a distinct effort during which various options could be tested and an optimal approach could be identified in a process of systematic design and user testing.

**Question 4: Which of the following is typical of your design projects?**

*Responses to Question 4*

Number of responses: 35-36

Rank order of design projects types with distribution of responses (typical, somewhat typical, not typical at all)

1. One evolving design, several iterations, one setting (21, 14, 0)
2. One intended design, evolving differently in several settings (21, 14, 1)
3. Studies begin with idea but very different designs evolve in different settings (12, 11, 13)
4. Other (2, 0, 0)

Comments for “Other”:
- Like option 2 but with comparison/control group with different design for contrast.
- Within major iterations there are small iterations, say changes to day 4 based on experience with day 3.

**Researcher Reflections on Question 4 Results**

Single and multiple setting design implementations with some design evolution in each are typical. Less typical are “messier” studies in which designs are allowed to change drastically in different settings, but this is still a practice. Should it be or might tool scaffold more systematic design or control?

Comments seem important: There are micro-iterations within design projects in which design changes are made from day to day and need to be represented; and that control/comparison groups must be represented as contrasting designs.

The software should allow for creating iterations based off of an initial design with an ability to make a distinction between and intended and implemented design across a variety of settings.

**Question 5: What annotations/elaborations would you likely want to add to or embed in your design at the time of design creation?**

**Responses to Question 5**

Number of responses: 35-36

Rank order of annotations/elaborations to add to or embed in designs with distribution of responses (very important, possibly important, unimportant)

1. Comments on specific segments or elements of design (30, 6, 0)
2. Comments on overall design, describing settings and populations, for example (28, 7, 0)
3. Theoretical conjectures related to design/design elements (26, 10, 0)
4. Assessment materials (26, 10, 0)
5. Learning materials for use in specific parts of the design (26, 9, 1)
6. Teacher/docent guidance/professional development materials (18, 18, 0)
7. Other ideas (explain): (5, 2, 0)
8. Other ideas (explain): (1, 1, 0)

Comments for “Other”:
- Survey/interview questions
- Description of main learning activities or tasks
- How student groups will be decided
- Comments for those responsible for conducting the study what to do when and how and why
• Learner needs (empathy based understanding of learner pain points)
• Notes of different kinds of implementations
• Comments on intended outcomes/understandings
• Links to types of data and/or methods

**Researcher Reflections on Question 5 Results**

All types of notations listed in the survey are regarded as important enough to support. Most of the ideas provided by responders under “other” can be accommodated in the categories we already had in mind. For example, we already planned to allow researchers to annotate learning activities and this could also include explanation of how groups in collaborative activities were formed. Intended outcomes will be an element in the workflow that can be annotated. Overall comments on design could include the “take away” messages desired by one responder. It is possible that the system could prompt the researcher for types of notation that especially inexperienced researchers might not consider, this is scaffolding for good research practice, although prompts should not be a nuisance and it should be possible to easily ignore what is not wanted. The request to make comments on the learning science that informs design is included under theoretical conjectures and it is possible that some responders were not familiar with the ideas of theoretical conjectures and conjecture mapping to know that is where the learning science would be included.

Interesting ideas we had not previously considered: A set of suggestions relate to methodology and implementation notes, which could serve other researchers or remind a researcher of exactly what they did with a project (when returning to dataset years later it is often not easy to recall or find notes on what exactly the data collection and research procedures were). Methodologies could include both implementation and data analysis information. This seems like a good idea.

Another unique idea we had not considered is an icon for “pain points” that learners experience. Teachers might also experience “pain points. To indicate where design pain points occurred and annotate these seems a good suggestion.

The challenge will be to decide what categories of notations to iconize and within those, how to provide non-obtrusive hints or prompts to provide. Using the logic of scaffolding, these hints could be “faded” or maybe just turned off by the researcher. Or maybe researcher could customize the system to include the scaffolding wanted. This might be helpful if, for example, a lead researcher is supervising students and post-docs. The researcher might pick the categories of notations/data that are needed to scaffold junior researchers.

The tool could allow for annotating workflows at various levels of information hierarchy, starting at a project level, down to workflow design, workflow iterations, and icons that make up the visualization of the design. Some of that information could be captured by the tool as metadata pertaining to a specific design. Some of the information the users reported as desired seems better suited for file linking than for annotating (e.g. linking for files with survey/interview questions rather than annotating a design with such questions).
Question 6: Which of the following tool capabilities would best help you manage the work following the design as it progresses in situ and keeping up with changes that occur?

Responses to Question 6

Number of responses: 29

Indicated 1st, 2nd, 3rd, 4th, and 5th choice. Listed in rank order (not in order of answers):

1. Modify intended to get enacted, save two versions, original & enacted (16, 11, 2, 0, 0)
2. Representing enacted design with annotations on original intended design (6, 14, 8, 1, 0)
3. Modify original to express enacted design only, original fades (5, 4, 19, 1, 0)
4. Other (see comments) (2, 0, 0, 27, 0)

Comments for “Other”:

- A workflow saved for each teacher who enacts to capture the overall implementation of that specific teacher
- A place to annotate reasons for implementation changes + the two versions
- Versioning (multiple, not just intended/enacted) with annotations for reflection on design process.

Researcher Reflections on Question 6 Results

The preferred is two versions: original and enacted. There should be capability of making notes for why there were design changes when they occur. However, there is also need to have different versions for different implementations (by different teachers, for example). It should be possible by allowing researcher to create multiple intended versions, a different one for each setting, then modify each of those to create enacted version. This would also take care of the need to have ability to compare experimental to control or comparison groups. The person requesting multiple versions might have something more complex in mind but there is a tradeoff and we have to do what is possible/practical to accomplish.

Since each user would have their own account for the tool, each user would be able to save and capture the iterations of their workflow designs in their intended and enacted versions.

Question 7: What forms of assessment might be included in your designs?

Responses to Question 7

Number of responses: 26-34

Rank order of assessment categories with distribution of responses (very important, somewhat important, unimportant)

1. Evaluation of learner artifacts rubrics (28, 5, 1)
2. Analysis of video data rubrics/coding schemes (26, 7, 1)
3. Questionnaire/survey (24, 9, 1)
4. Interview tasks (22, 10, 2)
5. Written tests (21, 12, 1)
6. Analysis of online conversation rubrics (10, 18, 6)
7. Online assessments (9, 10, 7)
8. Other (explain): (3, 1, 0)
9. Other (explain): (2, 0, 0)

Comments for “Other”:

Other 1:
- Learner descriptions of actions taken/why (this could fall under interview)
- Classroom observation protocols
- Concepts maps, Teacher observation, Self-assessment, Peer assessment etc.
- Evaluation of tasks with digital interfaces and physical objects/manipulables

Other 2:
- Evaluation of multiple drafts/iterations of an artifact
- Analysis of video-records of joint discourse and collaborative problem-solving

**Researcher Reflections on Question 7 Results**

Evaluation of learner artifacts rubrics and analysis of video data rubrics/coding schemes, questionnaire/survey, interview tasks, and written tests are the forms of assessments most frequently used in workflow designs. About half fewer respondents indicated online assessment forms as very important compared to respondents who considered offline assessment forms as very important. This indicates that an ability to link to assessment files stored locally in the user’s account would be of more importance than an ability to link to online assessments. However, both capabilities should be offered within the tool as one would anticipate an increasing use of online assessment resources.

Additional comments imply there is a large variety of assessment types that are employed in workflow designs, many of them qualitative in nature, probably best suited for annotations and descriptions.

The tool could support this variety by offering a base icon to represent assessments and allowing users to annotate or tag the icon to reflect the specific assessment types used. Additionally, an ability to link from an assessment icon to assessment files or student artifacts would give users a quick access to data relevant to each assessment.

A possible interface design approach is to allow users storing data associated with their design to check boxes associated with an assessment icon they entered into workflow to indicate what types of assessments the researchers used and whether instruments, data or analyses are stored (analyses may involve multiple assessments but links still connect assessments with related analyses). Info from these related items 7-X could be used by us to identify frequently used categories of assessments. Researchers could check categories they used then annotate them and attach instruments/data, or create new assessment categories and annotate. I imagine a visual interface where you later during search mouse over assessment icon and easily see at a glance that assessments and data are available and what types.
Questions 8 and 9

Questions 8 and 9 were presented to respondents who, in question 7, indicated that written assessments or interview tasks were very important or somewhat important.

Question 8: Which types of written tests might be included in your designs?

Responses to Question 8

Number of responses: 21-31

Listed in a descending order of received number of responses:

1. Short answer (31)
2. Multiple choice (22)
3. Essay (21)
4. Other (15)

Comments for “Other”:

Other 1:
- Gist statements
- Concept maps
- Numerical answer
- Sketches
- Descriptions of actions taken/choices made/explanations for why
- MCQs plus open ended explanation responses
- Drawings
- Reflection
- Recording answers to digital tasks
- Application of material to real-world setting

Other 2:
- Summary writing
- Drawing the correct relations between different terms
- Questions about multiple drafts/iterations of an artifact

Question 9: Which types of interview tasks might be included in your designs?

Responses to Question 9

Number of responses: 27-29

Listed in a descending order of received number of responses:

1. Video (29)
2. Audio (27)
3. Field notes (27)
4. Other (7)
Comments on “Other”:

Other 1:
- Online survey
- Stimulated recall
- Artifacts produced during interview
- Collect responses to open-ended questions in a survey
- Sketches
- Questions about multiple drafts/iterations of an artifact

Other 2:
- Task driven

Researcher Reflections on Questions 8 and 9 Results

In addition to the most frequent types of written assessments and interview tasks, respondents listed a variety of methods through which tests and interviews are delivered. Those constitute interesting and creative ways practitioners employ in the search for conclusions on the efficacy of instruction.

The rich assortment of approaches could be supported within the tool by allowing users to customize or tag base icons representing written tests and interviews. In addition, an ability to link out to students artifacts and other relevant files (e.g. notes or analysis) could provide means for quick and easy access to detailed and complete data pertaining to any given instance of a written assessment or an interview.

Question 10: What basic types of learning activities are typical in your work?

Responses to Question 10

Number of responses: 32-33

Rank order of learning activities with distribution of responses (very important, somewhat important, unimportant)

1. Small group collaborations (31, 2, 0)
2. Whole class discussions (18, 12, 3)
3. Demonstration/Model (15, 16, 2)
4. Study/reading (14, 13, 6)
5. Out of class explorations (10, 16, 6)
6. Whole class presentation (9, 18, 6)
7. Other (explain): (3, 0, 0)
8. Other (explain): (2, 0, 0)

Comments to “Other”:

Other 1:
Researcher Reflections on Question 10 Results

A variety of learning activities from small group collaborations to out of class explorations are typical in the work of many practitioners. In additional comments, the respondent pointed out learning activities conducted with individual students. Some interesting additions to the provided set of responses include guided reflections and guided goal-setting. Activities associated with digital design and prototyping indicate that learning activities are being considered in the context of modern-day skills learners will likely need to develop.

It seems that separating the type of activity from the learner constellation pattern (individual, small group, whole class etc.) could provide a more flexible interface design for the tool that would allow users to account for a variety of learning activities as well grouping contexts. It may also be worth considering to allow users to create custom representations of learning activities to accommodate new types of learning activities reflecting the needs of modern-day learners.

Questions 11–16

Questions 11 through 16 were presented to respondents who, in question 10, indicated that small group collaborations, whole class discussions, whole class presentations, study/reading, demonstration/model, or out of class explorations were very important or somewhat important in their work.

Question 11: Which of the following small group collaborations are among learning activities typical in your work?

Responses to Question 11

Number of responses: 20-28

Listed in a descending order of received number of responses:

1. Solve a problem (28)
2. Discussions (27)
3. Design together (24)
4. Do an inquiry or experiment (21)
5. Research together (20)
6. Other (7)

Comments for “Other”:

- Close reading exercises to achieve comprehension in order to build knowledge
- Contribute and evaluate ideas
- Create a product (infographic) together
- Team case studies
- Writing workshop / peer review
- Critique

**Question 12: Which of the following whole class discussions are among learning activities typical in your work?**

**Responses to Question 12**

Number of responses: 22-25

Listed in a descending order of received number of responses:

1. Teacher/Docent/Expert led (25)
2. Learner led (22)
3. Other (3)

Comments for “Other”:

- Debates
- Teacher or expert giving feedback/comments when needed
- Group led

**Question 13: Which of the following whole class presentations are among learning activities typical in your work?**

**Responses to Question 13**

Number of responses: 19-21

Listed in a descending order of received number of responses:

1. Small groups (27)
2. Teacher/Docent/Expert (21)
3. Individual learners (19)
4. Other (1)

Comments for “Other”:

- All of them. It depends on the content of the class.
Question 13: Which of the following types of study/reading are among learning activities typical in your work?

Responses to Question 14
Number of responses: 18-23
Listed in a descending order of received number of responses:

1. Collaborative (23)
2. Silent/individual (18)
3. Other (2)

Comments for “Other”:  
- Whole class

Question 15: Which of the following types of demonstration/model are among learning activities typical in your work?

Responses to Question 15
Number of responses: 21-28
Listed in a descending order of received number of responses:

1. By teacher/docent/expert (25)
2. By learner(s) (21)
3. Via technology (28)

Question 16: Which out of class explorations are among learning activities typical in your work?

Responses to Question 16
Number of responses: 14
Listed in a descending order of received number of responses:

1. Field trips (14)
2. Explorations on grounds (14)
3. Other (10)

Comments on “Other”:  
- Virtual field trip
- Outdoor tasks
- Search additional information when needed
- In students’ own homes and neighborhoods
- Individual or collaborative exploration at home
- Tech-based (skype, zoom, etc.) interview with experts
- Exploration using simulation models
Online research
Museum exhibit exploration

Researcher Reflections on Question 11-16 Results

Respondents comments suggest that all types of small group collaborations, whole class discussions, whole class presentations, study/reading, demonstration/model, or out of class explorations are typically employed depending on the context of instruction. This suggests that the context itself is an important attribute and should be captured within the tool for workflow designs.

Both context and the type of learning activity along with its specific applications could be reflected within the tool as annotations or tags to be associated with a base set of icons. As in the case of assessments, it would also be important to enable linking from icons representing specific learning activities to relevant files.

Question 17: What basic forms of learning technologies are included in your educational design work?

Questions 17

Responses to Question 17

Number of responses: 29-33

Rank order of learning technologies with distribution of responses (very important, somewhat important, unimportant):

1. Software (28, 5, 0)
2. Devices (21, 11, 1)
3. Technical or lab equipment (7, 11, 11)
4. Other (4, 0, 0)
5. Other (1, 0, 0)

Comments to “Other”:

Other 1:
- Network connectivity required
- Sensors
- Analog/materials
- IoT physical/digital objects
- Furniture

Other 2:
- Immersive simulations
**Researcher Reflections on Question 17 Results**

Software and devices are commonly used in educational design, more so than technical or lab equipment. The additional comments indicate that some of the technologies used require Internet connection. A couple of innovative examples include IoT (Internet of Things) objects and immersive simulations, both suggesting that cutting edge technologies find their way to educational instruction.

Given that practitioners leverage latest technologies to design and deliver instruction, it is only fitting that they should have at their disposal a digital tool to store, organize, search and access data they collect through administration of workflow designs. Additionally, with the growing an inevitable reliance of the Internet, such a tool should be web-based to allow easier sharing of data and analysis among peers.

**Questions 18-19**

Questions 18 and 19 were presented to respondents who, in question 17, indicated that devices or software were very important or somewhat important in their work.

**Question 18: Which of the following devices are included in your work?**

**Responses to Question 18**

Number of responses: 29

Listed in a descending order of received number of responses:

1. Laptops (29)
2. Tablets (23)
3. Phones (smart or cell) (15)
4. Tabletops (12)
5. Handheld other than phones (8)
6. Other (4)

Although some respondents indicated “Other” devices being included in their work, they did not provide any details.

**Responses to Question 19**

Number of responses: 25

Listed in a descending order of received number of responses:

1. Search engines such as Google, Safari, etc. (25)
2. Specialized apps (21)
3. Other (17)
4. Social media (11)
5. Games (10)
6. Other (5)

Comments to “Other”:
Other 1:

- Bespoke Coding Systems
- Mapping
- Online databases and data visualization tools
- Software for analyzing data and creating graphs
- Online learning environment developed for the project
- R
- Knowledge building software
- Moodle courses
- Custom Apps
- Modeling software
- We develop our own custom simulations which use motion tracking technology, and scientific annotation tools.
- Collaborative writing & review software (e.g. Google docs)
- Online textbook
- Simulation models
- Word, excel, LMS
- Research databases

Other 2:

- Automatic Live Data Collection
- Images/video
- Interactive simulations
- Quiz apps
- Google drive or other online workspace for students

**Researcher Reflections on Questions 18 and 19 Results**

In response to questions 18 and 19, respondents listed devices and software they use to deliver instruction to learners as well as those they use to manage and analyze their data. Although some examples represent stand-alone software installed on an operating system, several of the applications listed are web-based.

The fact that the number of responses to “Other” is relatively high indicates a wide variety of software respondents use.

Any new digital tool should draw inspiration from the types of software already in use. A web-based application that builds off of existing user mental models for data storage, organization, search, and sharing would be easier to learn and have a greater chance for adoption.
Questions 20: What types of learning/performance outcomes are the goals of your designs?

Responses to Question 20

Number of responses: 33

Rank order of learning/performance outcomes with distribution of responses (very important, somewhat important, unimportant):

1. Critical thinking (26, 5, 2)
2. Conceptual understanding (26, 7, 0)
3. Learner-designed artifacts (25, 5, 3)
4. Disciplinary practices (24, 7, 2)
5. Performance skill (19, 14, 0)
6. Self-regulation (18, 9, 6)
7. Other (5, 0, 0)
8. Other (1, 0, 0)

Comments to “Other”:

Other 1:
- Belief/attitude change
- Constructs related to motivation
- Three-dimensional science learning performances
- Drafting/revision
- STEM engagement/interest

Other 2:
- Constructs related to emotions

Researcher Reflections on Question 20 Results

Because of the qualitative nature of learning/performance outcomes, the new tool should support an ability to annotate workflow designs at various levels of information hierarchy (including workflow iterations and icons) with descriptions and tags. That would allow users to search for data associated with workflows and to analyze learning/performance outcomes in the context of those workflows.

A couple of additional comments deserve mention since they allude to changes in learners’ psychological makeup. Those include belief/attitude change, constructs related to motivation, and constructs related to emotions. Together with a suggestion from question 3 regarding an icon that represents user pain points, these comments indicate a desire to analyze learner’s mental-emotional states alongside the measurement of outcomes understood in terms of performance. The tool should therefore allow for associating learners’ mental-emotional states with stages of instruction.
Questions 21: What types of scaffolding are most important in your design work?

Responses to Question 21

Number of responses: 29-33

Rank order of scaffolding with distribution of responses (very important, somewhat important, unimportant):

1. Designed discussion prompts, questions or question starters for learners (23, 8, 2)
2. Scaffolds embedded in instructional materials (22, 10, 1)
3. Scaffolds embedded in technology (22, 7, 4)
4. Formative assessment (21, 8, 0)
5. Teacher/docent as guide of the side (20, 11, 2)
6. Designed discussion prompts, questions or question starters for teacher/docent (20, 10, 3)
7. Establishing norms or rules for discourse (19, 11, 3)
8. Assigning defined learner roles (10, 14, 9)
9. Scripts for learners (9, 9, 15)
10. Scripts for teacher/docents (8, 8, 17)
11. Other (1, 1, 0)

Comments to “Other”:

- Technology-enabled feedback to individuals during the learning process
- Teacher as learning partner

Researcher Reflections on Question 21 Results

A range of types of scaffolding are employed in educational design workflows. The tool we propose to design should allow users to indicate which types of scaffolding were used at the various levels of workflows’ information hierarchy. This could be accomplished by providing users with a base icon for scaffolding that could be customized by tagging or annotations.

Questions 22: Which of the tool’s capabilities would you care about?

Responses to Question 22

Number of responses: 33

Rank order of the tool’s capabilities with distribution of responses (very important, somewhat important, unimportant):

1. Find data within and across iterations associated with researcher assigned tags/keywords (for example, energy, science practices, self-regulation) (28, 5, 0)
2. Archive and locate data associated with workflow elements within a single implementation or design iteration (27, 5, 1)
3. Archive and locate data from across multiple design iterations that have identical or very similar workflow elements (23, 10, 0)
4. Other (4, 0, 0)
Comments to “Other”:

- Archive and locate data from different experimental groups within the same and across different iterations
- IRB requirements for working with data
- Locate data that represents one collaborative group or one individual participant
- Find all data for a single person

**Researcher Reflections on Question 22 Results**

The responses to provided tool capability options and the additional comments indicate that users would like an ability to search for data within and across workflows and their iterations. Respondents also reported a need archiving data. Additionally, an ability to search for data against learner groups and individual learners would be desired in the context of workflows and their iterations. Based on responses to some previous questions, it can be inferred that data searches at the level of icons that constitute a workflow might also be helpful.

The tool could provide an archiving capability or could simply allow users to create multiple iterations off of the initial design workflow. A robust search engine could allow users to search for data across multiple dimensions within the application, including workflows, workflow iterations, icons, and learner groups or individual learners.

**Questions 23: Which of the following types of data do you need to archive in easily accessible form to support and speed up your research process?**

**Responses to Question 23**

Number of responses: 33

Rank order of data types with distribution of responses (very important, somewhat important, unimportant):

1. Learner artifacts (for example, drawings, test responses) (29, 4, 0)
2. Transcripts of recordings (26, 7, 0)
3. Video recordings (for example, of classroom interaction) (25, 8, 0)
4. Written data from assessments (22, 10, 1)
5. Numerical data from assessments (22, 9, 2)
6. Audio recordings (for example, recorded focus groups) (21, 10, 2)
7. Teacher/docent artifacts (for example, lesson plans or reflections) (17, 12, 4)
8. Other (4, 0, 0)
9. Other (2, 0, 0)

Comments to “Other”:

Other 1:

- Student assessment data from district
• Numerical social network data (just stating this here as SNA data needs to be stored differently than other numerical data)
• Design thinking/empathy data such as pain points
• Teacher feedback, surveys, interviews, and implementation notes

Other 2:
• Pre/post assessments (in addition to assessments as part of the intervention)

Researcher Reflections on Question 23 Results

Based on the provided responses, it can be concluded that the types of data formats to be stored by the tool would include text, PDF, spreadsheets, images, video, audio, and other specialized data formats such as those required for numerical social network data.

The tool would have to allow for upload and storage of a wide range of file formats. More importantly, the server where the tool is hosted would have to afford robust storage capabilities given the importance of video and audio data that carry large amounts of information.

Questions 24: For what purpose do you need to be able to easily archive and access video recordings?

Responses to Question 24

Number of responses: 22-28

Listed in a descending order of received number of responses:

1. To find specific segments from large video repository (28)
2. To find specific clips from smaller library of previously selected video (22)
3. Other (6)

Comments to “Other”:

Other 1:
• CA of group interactions
• To reference global phenomena from large video repository
• Interaction analysis
• To share videos among coders
• To find specific segments from large audio repository

Other 2:
• To find specific segments from drafts/written artifacts

Researcher Reflections on Question 24 Results

Respondents indicated that they access video (and audio) recordings in order to identify segments from large video repositories as well as clips from smaller libraries of previously selected video.
The need for accessing source footage further highlights that the tool would have to support storing large amounts of data. Storage capabilities of the tool could be extended by provisioning a dedicated media server separate from the hosting server for the tool itself.

While the tool should support linking various levels of workflow information hierarchy to a variety of file formats, retaining a connection between the large source repositories and the small libraries derived from them may require a certain level of organizational discipline on the part of the users. Filing data in logically named and organized folders as well as file naming conventions would be important to facilitate the tool’s search capabilities and easy access to data.

**Questions 25: When archiving and accessing data from within the tool, what relationship would you expect the transcripts to have with the respective sample video clips and source recordings?**

**Responses to Question 25**

Number of responses: 1-17

Listed in a descending order of received number of responses:

1. Linked to sample video clips and to the source video in video data corpus (with accompanying search and storage costs) (17)
2. Linked only to sampled video clips (11)
3. Not linked to any recordings, transcripts are data (4)
4. Linked only to searchable video data corpus (with accompanying search and storage costs) (1)

**Researcher Reflections on Question 25 Results**

Responses to this question provide further evidence for high storage capacity needs. Additionally, an ability to retain connections between transcripts and their source recordings is of great importance. Such connection could be supported within the tool by linking both the transcripts and their source recordings to various levels of workflow information hierarchy. While the tool would support creating such connections, the need for organizational discipline on the part of the users cannot be underestimated. Creating logical folder/file structure and following a consistent naming convention would go a long way in retaining relationships between source data and their derivatives.

**Questions 26: What types of technology tools are you currently using in your design work?**

**Responses to Question 26**

Number of responses: 20-30

Rank order of design projects types with distribution of responses (very important, somewhat important, unimportant):

1. Spreadsheets: (24, 5, 1)
2. Statistical software: (21, 4, 3)
3. File search/organization capabilities provided by your computer’s operation system: (24, 2, 2)
4. Qualitative analysis programs or services: (15, 7, 4)
5. Data management software: (8, 6, 6)
6. Project management software: (6, 3, 12)
7. Other: (3, 0, 0)

Comments to “Other”:
- Hierarchical Linear Modeling
- Bespoke coding software
- IRB-approved data management software

Respondents also provided examples of software they use in their work. Spreadsheets:
- Excel
- Google Sheets

Statistical software and programming languages:
- SAS
- SPSS
- JMP
- Excel
- R
- Gephi
- Python

File search/organization capabilities:
- Mac OS Spotlight and Finder
- Windows
- Baidu, Google
- Dropbox
- Box

Qualitative analysis programs or services:
- InqScribe
- FinalCut Pro
- SNA
- NVivo
- Dedoose
- Atlas
- MAXQDA
- Atlas.ti
- Transana
• ELAN

Data management software:
• SQL
• R
• Excel
• Qualtrics
• Google Drive
• Filemaker

Product management software:
• Google docs/Bespoke
• Trello
• Asana
• Basecamp
• Google Drive

**Researcher Reflections on Question 26 Results**

Respondents reported using a range of software products from file/data management, through statistical and analysis applications, to project management solutions.

The proposed tool would offer data/file organization and management, data search, and possibly data sharing capabilities. Given that data collected in design-based research can be either qualitative or qualitative, with qualitative data making up a significant share, it would be hard to design one tool that would also handle all types of data analysis. There are great many sophisticated tools on the market already that specialize in quantitative or qualitative analysis. There aren’t, however, data organization and management tools that would allow users to tie their data to design workflows. We believe that equipping researchers with a powerful tool to access large amounts of data quickly and easily in the context of their design workflows is the first and crucial step in facilitating analysis that would occur outside of the tool, and could be stored and accessed as records from within the tool.

**Questions 27: What criteria would a new tool have to meet for you to try it in your work?**

*Responses to Question 27*

Number of responses: 22-31

Rank order of criteria with distribution of responses (very important, somewhat important, unimportant):

1. It would allow me to handle and analyze large amounts of data easier and faster than I can do now. (26, 5, 0)
2. It would pass security requirements at my institution. (26, 3, 2)
3. It would be free or affordable on a modest budget. (25, 6, 0)
4. It would be easy to use and learn. (24, 7, 0)
5. It would be reliable with good technical support. (22, 9, 0)
6. It would allow for more systematic synthesis of findings across projects. (18, 11, 2)
7. It would permit the application of analytics and big data approaches in design research. (14, 11, 6)
8. It would be compatible with other tools I use. (12, 4, 6)
9. It would allow broader sharing of data because of standardization across researchers. (10, 15, 6)
10. Other: (2, 0, 0)

Comments to “Other”:
- It would be customizable/programmable to a certain extent
- Better organization and data audit options

Some respondents also provided a list of tools they currently use with which they would like the new tool to be compatible:
- Excel, SPSS, R
- Output of jpegs for models, and csv files for data
- Video formats, transcripts from InqScribe
- Qualtrics
- We currently use Box / Atlas quite a bit and would hope to continue to. I also don’t know what future tools we’ll want.
- Excel, MAXQDA
- I want to link to other tools I use, not replicate them in another tool
- Excel, SPSS, R
- Excel
- Python, R

**Researcher Reflections on Question 27 Results**

Many respondent indicated a desire for the tool to assist them in handling and analyzing large amounts of data. As alluded in the discussion of the previous question, a distinction needs to be made between the tool’s capability to enable fast and easy access to data for the purpose of analysis and conducting analysis itself. The tool we are proposing would facilitate data analysis by providing quick, easy, and efficient access to large amounts of data across various dimensions (e.g. workflows, workflow iterations, instruction settings, time, learner groups, etc.). It would not, however, be a tool to conduct data analysis *per se*. Qualitative and quantitative data analysis are best conducted using specialized tools available on the market.

Another requirement many respondents deemed very important is security. That requirement would likely have to be addressed at an institutional level.

Low-cost and ease of use were other criteria ranked by many respondents as very important. Providing a complex, powerful, and comprehensive piece of software at low or no cost would
require a funding source or a social entrepreneurial model that could be explored. Achieving optimal and easy to learn user experience will require conducting usability testing built into the development process.

A digital tool requires continuous maintenance. Software inevitably comes with bugs that are only discovered through repeated use. Additionally, any web-based application requires ongoing security patches and updates to keep up with the development of browsers in which it is running. A requirement of good technical support raises the question of long-term maintenance and ownership of the tool. Unless there is an organization behind the tool that continues to maintain it, it would be hard to make the tool be reliable long term, and it would be impossible to provide any technical support.

The question of the new tools compatibility with existing tools in use would require further investigation. Any compatibility with specific software that goes beyond an ability to upload and manage files generated by the software would be a major undertaking.

Sharing data was ranked as very important by the least number of respondents. Given potential security challenges associated with data sharing, it may be prudent to exclude data sharing from the first iteration of the tool. The authors of this report would hope, however, to garner more support for data sharing in the future and to tackle security challenges associated with it. One of the benefits of building a tool to support faster and easier access to data for the purpose of analysis would be an ability to share that data and associated analysis with peers in order to contribute to the collective body of knowledge. But the security challenges associated with it may require a dedicated effort to design and build that capability after the tool’s most essential features are build and optimized.

Questions 28: Are there any specific needs or suggestions you have for data organizing/retrieving that have not been covered in this survey?

Responses to Question 28

List of responses to the open-ended question:

- There’s nothing mentioned about anonymization of data. Would the tool require it? If so, how would it enforce it, specifically with video/audio data? If not, how would it protect privacy?
- Linking of multiple video/audio channels from the same event
- It’s important to have a hierarchical data structure that allows for multiple classes taught by a teacher and/or multiple teachers within a school, etc.
- The data stored in one computer can be easily packed and copied to another computer.
- Chinese version
- Provide an online training component; there are so many technological options for doing some tasks that you named. It is confusing. Besides, I have no budget; so, all needs to be free or very low costs.
• Honestly, I’m confused about how you envision what you call “design workflow” because you’re not clearly distinguishing research design from learning environment design. Compare Sandoval’s conjecture maps – they allow you to map out prospective learning environment functions and the evidence you would need to see those, and how those, if seen, link to outcomes. That is a useful tool for organizing research workflow. That’s a tool that the field needs.

**Researcher Reflections on Question 28 Results**

Respondents offered additional suggestions regarding data organization and retrieval. A need for data anonymization has been pointed out. Data security and privacy would not doubt have to be thoroughly considered in the design and development of the tool. An ability to link media sources to workflow designs and their elements, a need for a hierarchical structure of data within the tool, as well as ease of use and cost considerations have been reiterated as important aspects of the tool. Copying data from computer to computer would not be necessary if the toll is built as a web-based application. It would allow users to access data from multiple locations and devices without a need to copy data and to keep it in sync.
Appendix 2: Entity Relationships Diagram
Appendix 3: Wireframes of Select Interfaces
Visual Data Management System
Interface Design Ideas

The 24 wireframes included in this appendix show possible approaches to selected views and key user flows.
Workflow Preview and Navigation

Within a project, a user can
- use the left-hand panel to navigate workflows and their iterations
- see a description of the project in the right-hand panel when the "Workflows" label is highlighted in the navigation panel
- see a description and other properties of a workflow the name of which is highlighted in the navigation panel
- see a preview of a workflow highlighted in the navigation panel
- see workflow levels in the preview panel when a single workflow (that has multiple levels) is highlighted in the navigation panel
- access icons to edit, duplicate, or delete a workflow, as well as to attach file and/or participant groups / participants to a workflow
- create a new workflow
- see if a workflow is "intended" or "enacted"
- see conjecture maps, files, and participant groups / participants (three lower panels) associated with a workflow highlighted in the navigation panel
- apply filters (lower panel) to files attached to a workflow that is highlighted in the navigation panel
- select a node within a workflow
- see a description and other properties of a selected node
- see conjecture maps, files, and participant groups / participants (three lower panels) associated with a selected node
- create a lower level workflow "within" a selected node
Energy Iteration 1

Create New Workflow

Conjecture Maps
- Workflows
  - Energy Iteration 1 (Intended)
  - Energy Iteration 2 (Enacted)
  - Energy Iteration 3 (Enacted)
  - Energy Iteration 4 (Enacted)
  - Nervous System (Intended)

Files
- Participant Groups
  - Rhinecliff District
    - Abernathy Middle School
      - Teachers
      - Students
        - 8th Grade
        - 7th Grade

Workflow Description
Click on description text to start editing
Nullam quis risus eget urna mollis ornare vel eu leo. Curabitur blandit tempus porttitor. Cras justo odio, dapibus ac facilisis in, egestas eget quam. Integer posuere erat a ante venenatis dapibus posuere velit aliquet. Donec id elit non mi porta gravida at eget metus. Duis mollis, est non commodo luctus, nisi erat porttitor ligula, eget lacinia odio sem nec elit.

Setting
Click on text to edit
None

Tags
Click a tag to start editing
Commodo × Nullam × Tontor × Nibh ×

Add new

Conjecture Maps | Date Created | Tags
--- | --- | ---
C-Map Magna | 02/02/2017 | Vehicula...
C-Map Etiam | 12/15/2016 | Aenean...
C-Map Malesuada | 11/02/2016 | Vehicula...
C-Map Dolor | 11/12/2016 | Aenean...
C-Map Tristique | 09/15/2016 | Aenean...

Files | Date Created | Tags
--- | --- | ---
Sit.doc | 02/15/2017 | Vehicula, Sit, Vul...
Lorem.doc | 02/05/2017 | Parturient, Elit.
Mollis.txt | 01/20/2017 | Vestibulum, Pha...

Participant Groups | Participants
--- | ---
District Rhinecliff | 2799
Abernathy Middle School | 2790
Teachers | 2791
Students | 2792
7th Grade | 2793
7th Grade Period 2 | 2794
7th Grade Period 1 | 2795
2796
2797
2798

Filter Files
Edit Workflow and Workflow Nodes

Within a workflow edit view, a user can
- view a library of icons in the right-hand panel (the icons must be designed)
- add nodes to workflow, possibly by dragging icons into the workflow
- remove nodes from a workflow
- select icons in the icon library to view their description and other properties
- access a link to edit an icon
- connect nodes within a workflow
- apply ending types to node connectors (the list of endings must be determined)
Conjecture Maps Preview and Navigation

Within a project, a user can

- use the left-hand panel to navigate the conjecture maps
- see a description of the project in the right-hand panel when the "Conjecture Maps" label is highlighted in the navigation panel
- see a description of a conjecture map the name of which is highlighted in the navigation panel
- see a preview of a conjecture map highlighted in the navigation panel
- create a new conjecture map
- access buttons to edit a conjecture map, attach files to a conjecture map, link a conjecture map to participant groups / participants, delete a conjecture map
- see workflows, files, and participant groups / participants (three lower panels) associated with a conjecture map highlighted in the navigation panel
Home Screen for a Logged in User

A logged in user
- can see all their projects in chronological order, based on the date created
- can see summary of information about a project on its card
  - project name
  - excerpt from project description
  - date created
  - link to manage users
    - clicking on the link to manage users opens up a modal or a dialogue where all users with access to the project are listed and can be managed (added, deleted, edited permissions)
- can search their projects by keywords
- can create a new project
- can delete an existing project
Project Rhinecliff 2017
Etiam porta sem malesuada magna mollis euismod. Sed posuere consectetur est at lobortis. Etiam porta sem malesuada magna mollis euismod. Morbi leo risus, porta ac consectetur ac, vestibulum at eros.
Date Created: 03/12/2017
Manage User Access

Project Wake County
Date Created: 01/05/2017
Manage User Access

Project East NC
Nunc quis risus eget urna mollis ornare vel eu leo. Curabitur blandit tempus porttitor. Cras mattis consectetur purus sit amet fermentum. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Cras justo odio, dapibus ac facilisis in, egestas eget ...
Date Created: 01/05/2017
Manage User Access
Participant Groups Preview and Navigation

Within a project, a user can
- use the left-hand panel to navigate the nested participant groups
- see a description of the project in the right-hand panel when the "Participant Groups" label is highlighted in the navigation panel
- see a description and other properties of a participant group or a participant whose name is highlighted in the navigation panel
- see a preview of participant groups or participants inside a participant group highlighted in the navigation panel
- create a new participant group
- upload a participant list into a group
- edit a participant group type
- see conjecture maps, workflows, and files (three associated with a participant group highlighted in the navigation panel)
Participant Groups:

- Rhinecliff District
- Wake District

Project Description:
Nullam quis risus eget urna mollis ornare vel eu leo. Curabitur blandit tempus porttitor. Cras justo odio, dapibus ac facilisis in, egestas eget quam. Integer posuere erat a ante venenatis dapibus posuere velit aliquet. Donec id elit non mi porta gravida at eget metus. Duis mollis, est non commodo luctus, nisi erat porttitor ligula, eget lacinia odio sem nec elit.

Tags:
- Commodx
- Nullam

Add new
Within a project, a user can
- use the left-hand panel to navigate folders and files
- see a description of the project in the right-hand panel when the "Files" label is highlighted in the navigation panel
- see a description and other properties of a folder or a file the name of which is highlighted in the navigation panel
- see a preview of a folder or a file inside a folder highlighted in the navigation panel
- create a new folder
- upload a files into a folder
- see conjecture maps, workflows, and participant groups/participants (three lower panels) associated with a folder or file highlighted in the navigation panel