# VAL: Virtual ASSISTments Laboratory

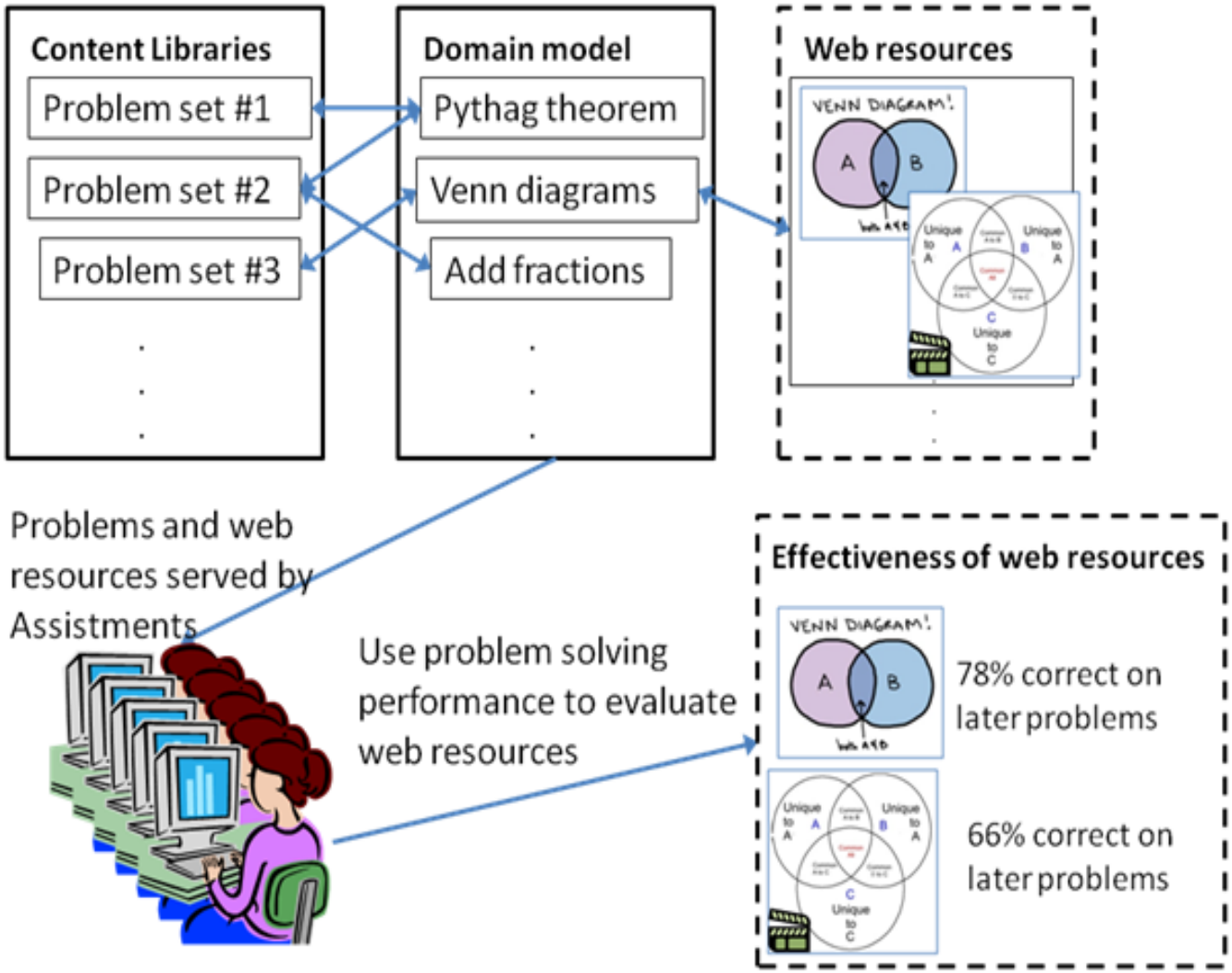
Joseph E. Beck and Neil T. Heffernan

# ASSISTments + web pages = WEBsistments

- ASSISTments:  web-based ITS
  - Tens of thousands of student users
  - Primarily used for math education, but is independent with respect to curriculum

- Have augmented it with web pages to teach students

# Automatically evaluate pages



# ASSISTments + web pages = WEBsistments

- ASSISTments:  web-based ITS
  - Tens of thousands of student users
  - Primarily used for math education, but is independent with respect to curriculum

- Have augmented it with web pages to teach students
  - **Which web pages?**

# How we select web pages

- Have been searching the web to find good ones
  - We don’t want to be in the content creation business
- Why not create web pages?
  - Too expensive, most ITS designers lack the knowledge

# Ok, who said ITS designers were creating the pages?

- There are other content creators out there

- E.g., cognitive and educational psychologists
  - They often have ideas they want to test

# Problems with running RCTs (Randomized Controlled Trial)

- Need to acquire subjects

# Need to acquire subjects

- Want representative sample of your population
  - Much psych research done with undergraduates
  - Should be a diverse population
- Want to get several subjects
  - Costs money, or competing with others for subject pool (can get nasty)

# Problems with running RCTs

- Need to acquire subjects
- Authenticity of learning task

# Authenticity of the learning task

- Many lab studies on learning use a synthetic task
  - Goal is to control for subject's prior knowledge
- So they work on a meaningless task that has no relevance to anything later
  - Anyone see a problem?

# Problems with running RCTs

- Need to acquire subjects
- Authenticity of learning task
- Longitudinal is hard

# Longitudinal is hard

- Expensive to bring subjects back in later
  - More than just twice as hard, you'll lose subjects so need to overrecruit
  - Also more variability in subjects over time → more subjects needed
- Need to keep contact info for subjects
- Reserve lab facilities again
  - Easier to run 2 studies

# Problems with running RCTs

- Need to acquire subjects
- Authenticity of learning task
- Longitudinal is hard
- Collating the data

# Collating the data

- Subject pre-test / questionnaire
- Performance during experimental condition
- Post test / questionnaire
- Have to create one file for analysis

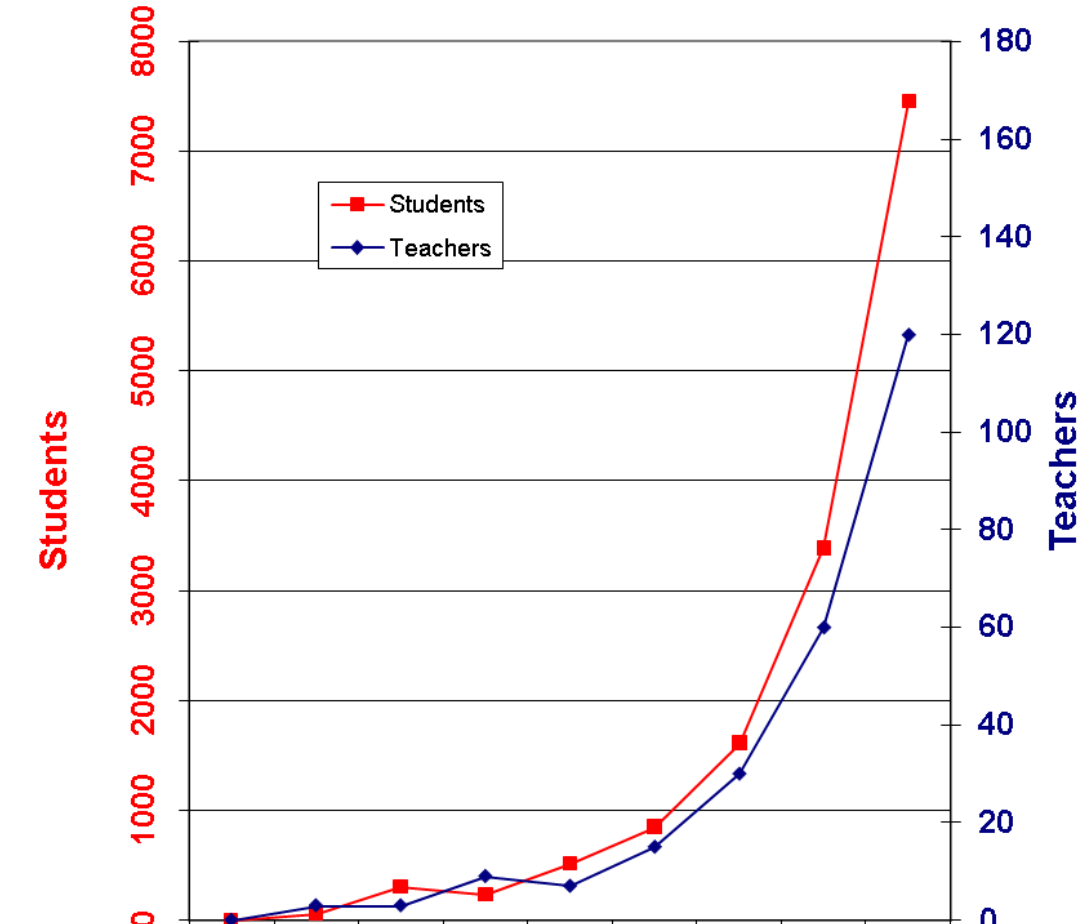
# That’s a lot of problems!

- Need to acquire subjects
- Authenticity of learning task
- Longitudinal is hard
- Collating the data

# Can ITS make this easier?

- Need to acquire subjects
- Authenticity of learning task
- Longitudinal is hard
- Collating the data

# You want subjects? (>20k students now)




| | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 |
|---|---|---|---|---|---|---|---|---|---|
| Students | 0 | 60 | 300 | 230 | 514 | 848 | 1612 | 3386 | 7451 |
| Teachers | 0 | 3 | 3 | 9 | 7 | 15 | 30 | 60 | 120 |

Year

# Authentic learning task

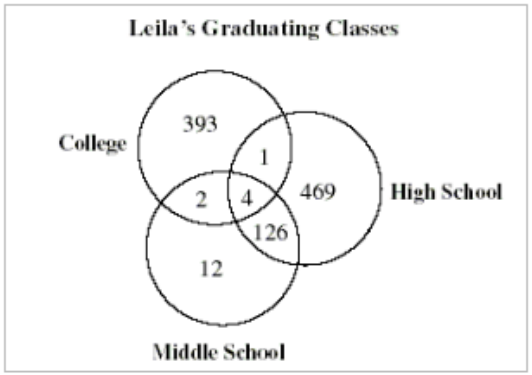
Assistment - Previewing Content - Mozilla Firefox

http://www.assistments.org/preview/assistment/4517

ASSISTments

You are previewing content. Problem 30 - 2002 (#4517)

The Venn diagram below shows Leila's graduating classes from middle school, high school, and college. How many students graduated together from **both** Leila's middle school and high school?




Comment on this question

Break this problem into steps

*Type your answer below (mathematical expression):*

Submit Answer

Done

# Longitudinal is easy

- Ming Feng predicted student test scores on a state test
  - Two years after students finished with ASSISTments
- We store information in our database by student

# Collating the data

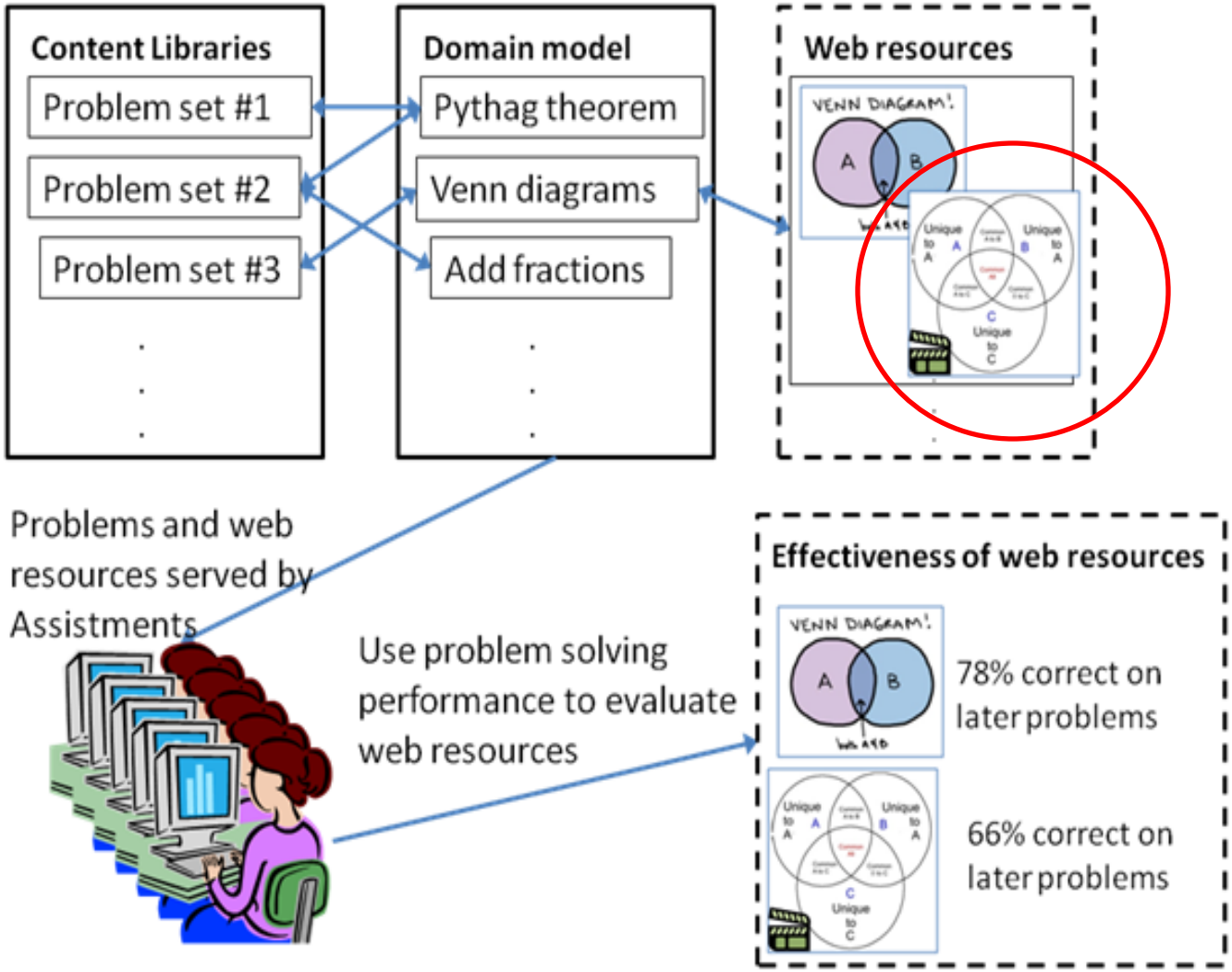
- That’s why we have databases :-)

- Everything logged, relatively straightforward to assemble

# Summary

- Running good RCTs on human learning is hard
- ITS have several aspects they can simplify
- Big question: how do researchers get their experiment into ASSISTments?
  - **???**

# Can evaluate an *arbitrary* web page



# The model

- Researcher has idea for how to teach
- Instantiates experiment as two competing web pages
- Add pages to WEBments
- Collect data on their performance

# Simplifies experimental cycle

| Step in experimental process | Manual | Semi-automated | Automated |
|---|---|---|---|
| Generate research idea | • | | |
| Acquire subjects | | • | |
| Intervention | | • (design) | • (administer) |
| Record & collate data | | | • |
| Export data to tab-delimited file | | | • |
| Analyze and write up results | • | | |

# A useful contribution

- Makes experimenter's life easier
- Possibly gets better content into our systems
- Connects us with domain researchers
- A win-win!

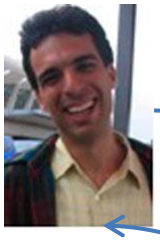
# Concerns?

# Concerns

- What if experiments are badly conceived?

- Do teachers get control over what happens in the classroom?
- At WPI we have some trust (and they're using our system)
  - Does not necessarily transfer to others
  - We don't want to burn our goodwill

# Thinking of a much bigger process



Proposes / revises

Receives feedback and improves

Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study

Provide feedback

Computes pedagogical effectiveness

**Computational Engine**

Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# Ido gets an idea…

Proposes / revises

Receives feedback and improves

Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study

Provide feedback

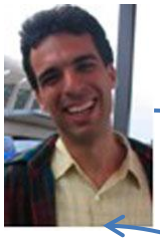
Computes pedagogical effectiveness

**Computational Engine**

Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# Ido creates project pitch (text, video…)



Proposes / revises

Receives feedback and improves

Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study

Provide feedback

Computes pedagogical effectiveness

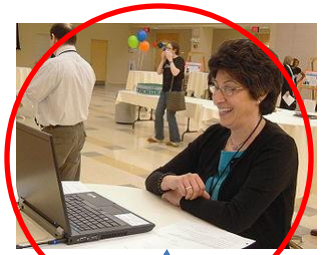
**Computational Engine**

Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# Gets feedback from teacher

Proposes / revises



Receives feedback and improves

Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study

Provide feedback

Computes pedagogical effectiveness

**Computational Engine**

Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# If approved gets feedback from students

Proposes / revises

Receives feedback and improves

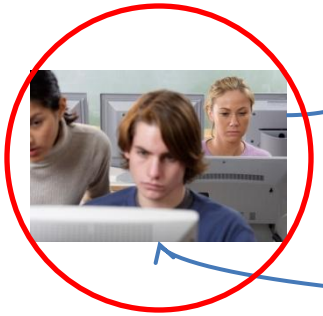
Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study



Provide feedback

Computes pedagogical effectiveness

**Computational Engine**

Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# Comments are useful

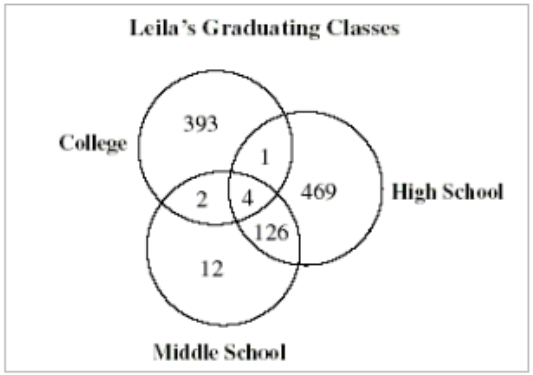
Assistment - Previewing Content - Mozilla Firefox

http://www.assistments.org/preview/assistment/4517

ASSiSTments

You are previewing content. Problem 30 - 2002 (#4517)

The Venn diagram below shows Leila's graduating classes from middle school, high school, and college. How many students graduated together from **both** Leila's middle school and high school?




Comment on this question

Break this problem into steps

*Type your answer below (mathematical expression):*

Submit Answer

Done

# Gets automated analysis of experiment

Proposes / revises

Receives feedback and improves

Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study

Provide feedback

Computes pedagogical effectiveness

**Computational Engine**

Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# After experiment finishes, adds data to his experiment's entry

Proposes / revises

Receives feedback and improves

Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study

Provide feedback

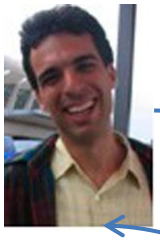
Computes pedagogical effectiveness

**Computational Engine**

Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# Didith comes along and checks if anyone has a similar idea



Proposes / revises

Receives feedback and improves

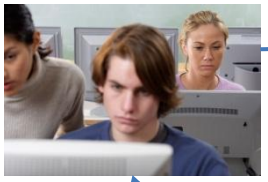
Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study



Provide feedback

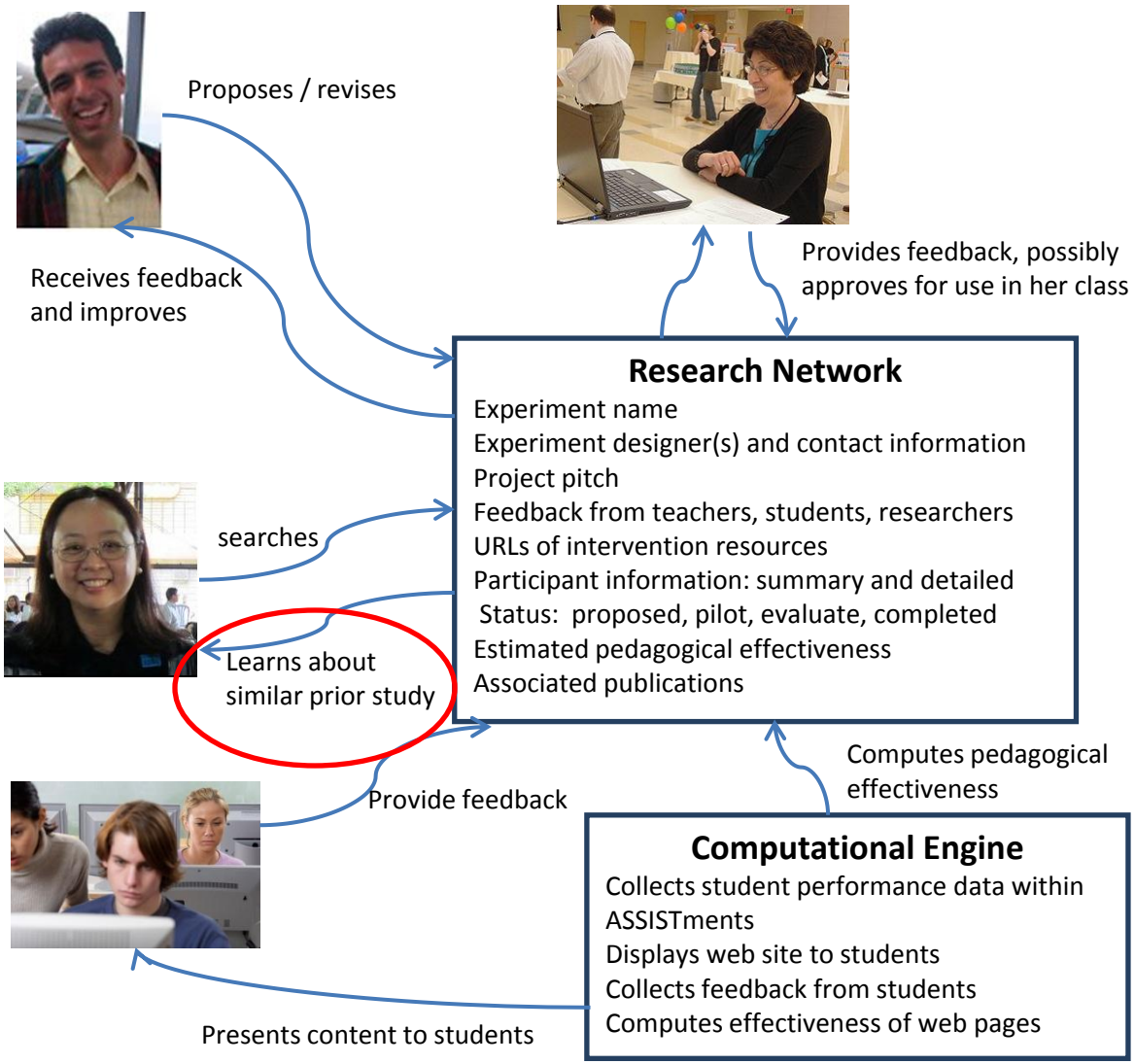
Computes pedagogical effectiveness

**Computational Engine**

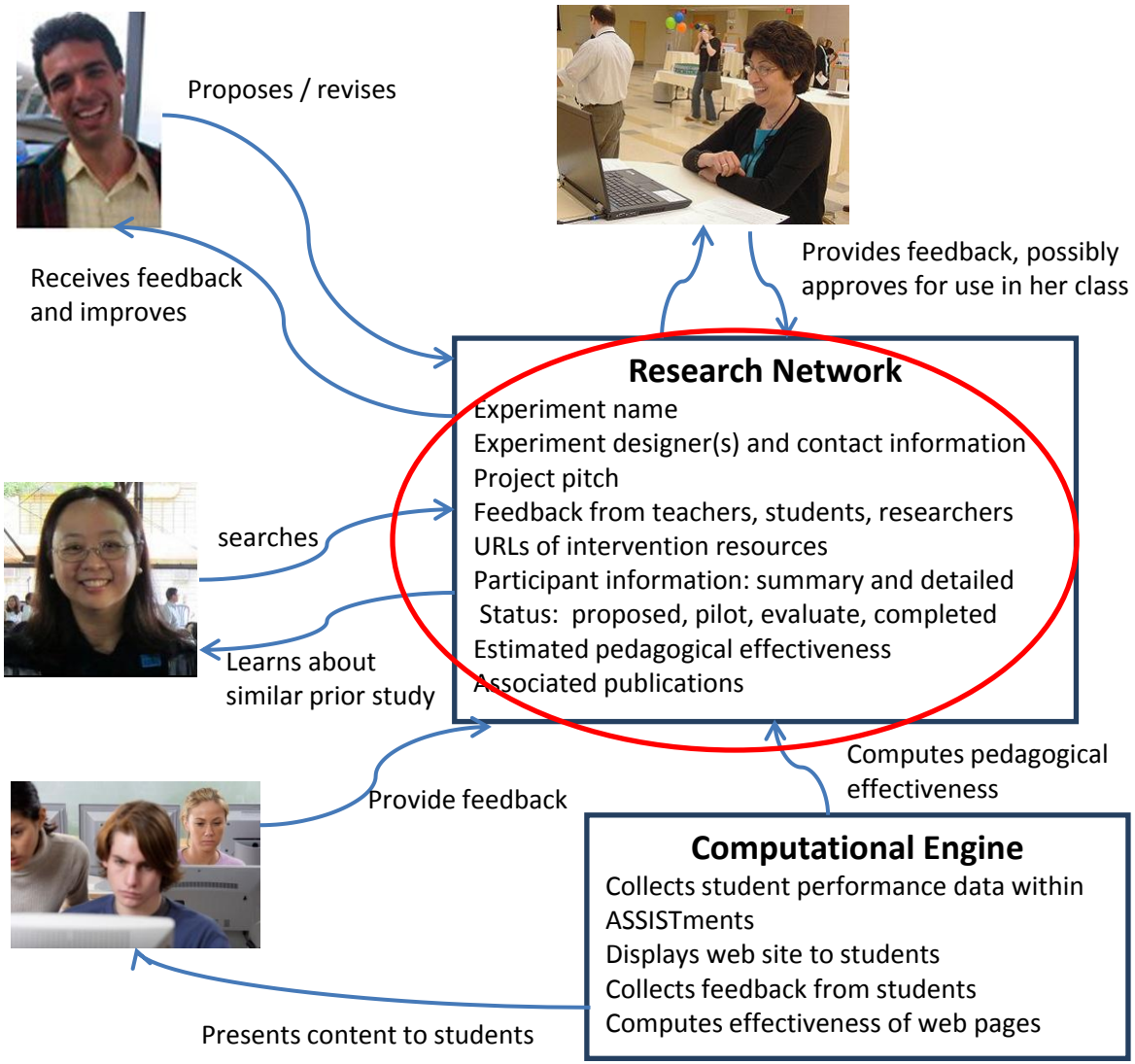
Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# Can see outcomes of prior study



# Much better than a paper abstract



# Has all the data you will need to *replicate*

Proposes / revises

Receives feedback and improves

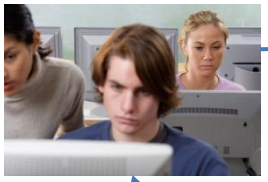
Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study



Provide feedback

Computes pedagogical effectiveness

**Computational Engine**

Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# Replication

- Essential for science to happen
- Ionnidis:  medical researcher
  - Estimated 50% of medical research reached the wrong conclusion
  - Tested on the very best work, and found overturn rate was 40%
- Why EDM has replication as a listed goal

# An interesting ecosystem

Proposes / revises

Receives feedback and improves

Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study

Provide feedback

Computes pedagogical effectiveness

**Computational Engine**

Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# Easier to generate ideas

| Step in experimental process | Manual | Semi-automated | Automated |
|---|---|---|---|
| Generate research idea | | • | |
| Acquire subjects | | • | |
| Intervention | | • (design) | • (administer) |
| Record & collate data | | | • |
| Export data to tab-delimited file | | | • |
| Analyze and write up results | • | | |

# Future work

- Almost all of it :-)

# Computational Engine mostly exists

Proposes / revises

Receives feedback and improves

Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study

Provide feedback

Computes pedagogical effectiveness

**Computational Engine**

Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# Ido and Didith exist

Proposes / revises

Receives feedback and improves

Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study

Provide feedback

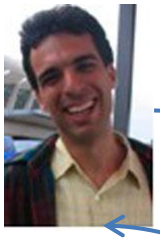
Computes pedagogical effectiveness

**Computational Engine**

Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# Need to create network and get teachers more involved



Proposes / revises

Receives feedback and improves

Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study

Provide feedback

Computes pedagogical effectiveness

**Computational Engine**

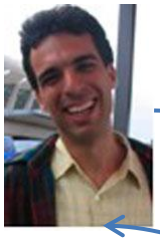
Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# Why I think this work is neat

- Makes ITS much more useful
  - Experimental platform might be best use (Chuck Perfetti, over a decade ago)
- Addresses authoring tool problem
  - Lots of tools proposed and built, none took off
  - Authors use whatever tool they want to create a web page
- Some meaty EDM challenges

# Tough EDM challenges



Proposes / revises

Receives feedback and improves

Provides feedback, possibly approves for use in her class

**Research Network**

Experiment name
Experiment designer(s) and contact information
Project pitch
Feedback from teachers, students, researchers
URLs of intervention resources
Participant information: summary and detailed
Status: proposed, pilot, evaluate, completed
Estimated pedagogical effectiveness
Associated publications

searches

Learns about similar prior study

Provide feedback

Computes pedagogical effectiveness

**Computational Engine**

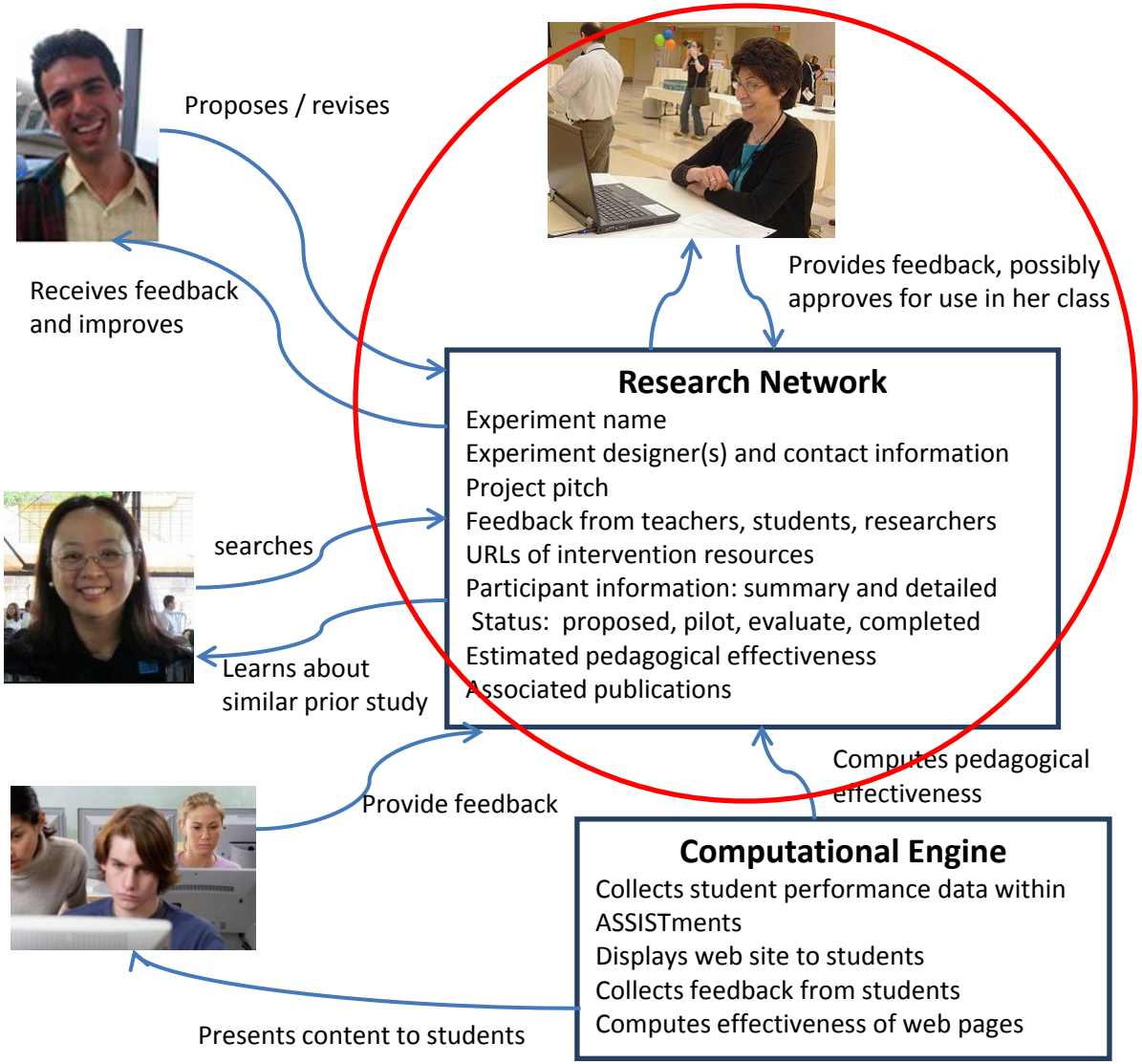
Collects student performance data within ASSISTments
Displays web site to students
Collects feedback from students
Computes effectiveness of web pages

Presents content to students

# Neat EDM work

- Automatically evaluate an arbitrary intervention
  - Do we know enough to automatically create outcome variables
- Automatically output features for researchers
  - Which ones are important

# Also social computing work



# Conclusions

- We're trying to make this happen
  - Trying to acquire funding
- In the meantime, working on evaluating known web pages (WEBsistments → EdRank)
- One of our current "big ideas"
  - WEBsistments, EdRank, VAL