

Agenda

Zoom Link: <https://ncsu.zoom.us/j/98870611463?pwd=RYk05WnbGPqR0fU08Ve5z6iNaR8Nxr.1>

- **9:00 am Introduction (Peter and Cliff)**
 - Organization and main goals (Peter)
 - SPLICE Web site and key links (Cliff)
 - [SPLICE Project Home Page](#)
- **9:15 am Community engagement (Peter and Tiffany)**
 - Community engagement mechanisms (Peter)
 - Workshops
 - Working Groups
 - Mini-Grants
 - Summer Schools
 - Broader engagement (Tiffany)
 - Discussion and questions
- **9:30 am Content Hub: Contributing and using smart learning content (Cliff, Peter, Thomas)**
 - Catalog and smart content (Cliff)
 - Community contributions and reuse (Peter)
 - Authoring smart content
 - CodeWorkouts (problems)
 - WEAT for PCX (examples)
 - Using smart content
 - Practice builder
 - OpenDSA
 - Codecheck Assignment Builder
 - Increasing Community contributions
 - DOI for Smart Content packages
 - Protocols and formats - WGs on the content side (Cliff)
 - Lightweight protocols
 - PEMPL
 - “German PEMPL”
 - Parsons’ problems - content and analytics
 - Worked examples
 - Interventions and services (Thomas)
 - Discussion and questions
- **10:00 am Data Hub: Contributing and using learner data (Ken and Thomas)**
 - New datasets

- New workflows
- Data beyond Data Shop
- Increasing Community contributions
 - DOIs for Data Sets
- CS Datasets & Metadata beyond of Datashop (Thomas)
 - CS Datasets from the CompeEd WG:
 - [CompED-Systematic-Analysis of Open Access CSed Datasets](#)
- Discussion and questions
- **10:15 am Final discussion and question session**

Notes / AB Advice

- **Working Groups:**
 - Barb: Are the working groups associated with ITiCSE? I find that a good structure to make sure the group meets and accomplishes something.
 - Do the working groups get any funding?
 - Tiffany: No they do not. We could offer support to leaders/grad student who convene their groups a certain number of times or who define and achieve concrete deliverables (e.g. Paper, dataset). That could be helpful.
 - AP: Or toss some funding at a grad student attached at the project who has responsibility for keeping the group organized / reporting outcomes
- **WG Organization:**
 - Cay: It is helpful for the SPLICE team to be more directive with WG leaders, especially those with non-SPLICE leaders, to communicate clear expectations (e.g. meet with X frequency, propose an outcome to achieve by X date).
 - Andrew: Creating space for the working groups to meet (at the workshops, after the workshops to allow zoom participation) would be good. It also creates a deadline if they know they're going to meet.
 - Andrew: I know you're balancing getting people involved in what they are interested in with making progress ... but there may be too many WGs. Each probably has 1-2 motivated people, but they're going to divide the interested labour pool, which is probably still small.
 - For example: I think the replication and interoperability groups are both exciting, but can they work together? Would the replication group be able to use what interoperability is aiming to put together?
 - TP: **I think this is an important point for us to discuss** if there's time, and certainly within the SPLICE team moving forward.
 - Andrew: I'm seeing lots of content – and some help for deploying them (yay, smart content!) – but do we yet have any infrastructure or groups working on infrastructure for deploying randomized interventions, so we can evaluate content or tools?

- For reference: this is something built into one of the interventions Price is deploying in the replication WG.
- Andrew: Has the snapshots group essentially paused? Does it need a kick to update the standard and to push to get more adoption?
 - Thomas Answer: The group is currently intentionally in a maintenance phase, with the goal of responding to support requests and updating the standard as issues are raised. That said, there could definitely be a 2.0 instantiation that worked to increase adoption. Would that be exciting? The current lead is Ayaan.
 - Andrew: Probably less *exciting* than the replication work you're doing or the infrastructure interoperability problems. The low-level protocols are enabling work, so increasing adoption may simply be "get it into the later stage WGs".
 - TB: Do you mean making sure the newer WGs are using the low-level protocols?
 - AP: I think there are opportunities there. For example, any data generated by the replication group should be in progsnap 2 (and probably will be, thanks to Price) since it's snapshot data.
 - AP: There are citations on PS2 - how do you track that, build it up, etc.
- **Mini-grants:**
 - Barb: Can you increase the amount? The amount of 5K to 15K is pretty low. You might attract more PhD students if you increase the funding.
 - Barb: Have you contacted people with lots of data like Prairie Learn?
 - Andrew: The examples of previously funded mini-grant projects are quite helpful. But if you want more specific info, please add more structure to the application page.
 - Cay: <https://csssplice.org/> could have a link to the mini grant examples and an application form
 - Homepage has a link to mini-grants page, which has those things: <https://csssplice.org/MiniGrants.html>
- **Workshops / Broader engagement:**
 - Barb: Provide a summary of existing work and demos of the types of smart content - not papers - and give time for groups to do work
 - Monica: I like the idea of having a chance for those in the community to share recent work. Really helps in letting everyone know what folks are doing.
 - Provide space (and guidance) for the WG to meet in person at the workshops
 - Cay: +1 for working groups at SIGCSE workshop
- **Catalog and Content Hub:**
 - Barb: Hard to keep these types of things up to date and most people will not spend the time to look at them - Perhaps do short articles about these in Inroads?
 - Cay: There should be full-text search of the content - finding what you actually want is the hardest part. It should include the content of the exercises themselves. [We ought to be able to do search on everything in the collection of

URLs given for the exercise iframes.] Barb: we have full text search in Runestone and it is still hard to find new things that you want to use.

- Cay: Bundling is essential. He already has a tool that does this, but it would be more stable and trustworthy if it was maintained by a group. Peter's student should contact Cay (or just SPLICE/Cay more generally).
- Barb: +1 on authoring/bundling tools as a good idea

- **Questions [in chat & otherwise]:**

- Barb: How are you tracking reuse of the "smart content"?
- Barb: Is all the smart content at the single exercise level rather than at an assignment level? Perhaps the assignment level and/or course level would be more useful and reusable.
- Barb: Need keywords (CS ontology) or some other easier way to search for "smart content" for a particular context. Plain search might be better than an ontology. Might be useful to automatically detect language elements that are in a question.
 - Ken: Could be a good context for testing of LLM tagging of content with skill labels — there have been some promising examples of such recently
- Ken: [Inspired by Thomas' idea for community web site ...] How about a Prolific-like website where instead of paying participants in experimental learning studies, learners get free CS ed for participating in learning studies?
 - David: Reminds me a lot of how psychology scales up many of its studies
 - Ken: Yes. Prolific is super useful for that: <https://www.prolific.com/>
 - Peter: <https://www.socraticmind.com/>
 - Thomas: I love this idea, Ken. Or additionally, we make it useful enough to instructors that they assign it as homework.
- Barb: How do you do data sharing? [Perhaps meaning how to do IRB so data can be shared?]
 - Ken: We could perhaps share approved IRB protocols that allow for sharing
 - Cliff: This is precisely the sort of thing that the "Administrative Support" section of the CSSPLICE website (<https://csssplice.org/Administrative.html>) aspires to do. Anyone with example IRBs to post (or examples of anything else listed there) should let Cliff know so that he can arrange to get it posted.

- **Replicating Interventions - from CHAT**

- David: (regarding each school implementing the tool) Reminds me of some services that have sprung up with a "Bring Your Own LLM" approach, where schools are responsible for procuring and adding their own API key for LLM access as a way of sidestepping the tool being responsible for FERPA-protected data
- Andrew: One layer we added to this experiment: we had a thin layer on top that sent students to either the progress bar intervention OR a different hint-type.
 - Andrew: That randomization layer seems important to me as a capability for the future, and if a group doesn't need it, then they can just set a probability of 1 for their 1 intervention.
 - Thomas: Correct! Though the intervention itself does actually handle condition assignment in the absence of that. That's how we did it in the NCSU course.

- Andrew: Good to have it at a higher level, so that you can deploy two unrelated (originally) interventions, but also nice to have some randomization within a single intervention type.
 - David Joyner: (while Thomas is talking about replication) These sorts of challenges are what we've been facing in finding ways to experiment with Socratic Mind (AI-powered oral assessment tool Thad Starner's lab created) as well
 -
- **DataHub**
 - See Ken's slides [here](#)
 - Monica: I noticed on the SPLICE dataset page that there isn't a filter like this. That might be a great thing to add, especially as the catalog grows.
 - Thomas: Agreed - though part of me wonders if we should just link to the google sheet, since it's easy to update and has sorting and filtering for free: <https://docs.google.com/spreadsheets/u/1/d/1jWdemuCXdjKTVWihuliHlmmIAqZKfx7haehEOev1pBg/edit?gid=0#gid=0>
 - Monica: The sheet is likely more helpful. Could you pull from the google sheet so it is dynamic on the page?
 - Monica: I think the learning rate comment is incredibly interesting. I wonder how many studies are RCT?
 - Ken: A number of the 27 datasets had embedded randomized control trials (RCT) in them ... to be sure, such is not required to measure learning rate. We have seen cases of learning rate varying by experimental condition ... I can point to if interested ...
 - Monica: Yes, I'd be interested in learning more.
 - Ken: https://www.researchgate.net/publication/221570569_Instructional_Factors_Analysis_A_Cognitive_Model_For_Multiple_Instructional_Interventions This IFA approach is a kind of within-participant experimental design that tracks outcomes of instructional interventions in changing learning rate. There have been 3-4 follow-up using it.
 - Here's an analysis of learning rates in two different RCT conditions: [Distinguishing qualitatively different kinds of learning using log files and learning curves](#)
- **Suggestions:**
 - Barb: A [CS ontology](#) would be an important contribution. It should map several common words to one concept in a hierarchy. Work on LLM assignments/exercises is very hot right now and there is a big need for this. Creating [authoring tools](#) is a good direction which might increase use. I also find putting smart content into bundles is a good idea. Have a part of the SIGCSE workshop that [helps people create mini-grants](#). Or do this as a follow-up webinar after. Thomas is [shipping code \(Docker\)](#) to people to try to run experiments in their context - that helps overcome the IRB issues - seems like a good way to try this. I agree with Ken - Infrastructure building requires community building. I think doing follow up work with people who attend the summer training would be good. Do a follow-up survey and/or interviews with attendees from workshops and summer schools to see what attendees did/used.
 - Peter: We would really benefit from an ontology. There are several (including one from ACM), but none is really matching our education-focused needs. Suggestions would be appreciated!

- Andrew: Stefik (UNLV) also does work on replication, and he's also shipping interventions in docker. He and his student might have some expertise to share.
- Barb: be more proactive about followup from summer school and workshops
 - Ken: many subsets of collaborators continue to meet to follow up on the work they started at summer school, but we could be more proactive
 - Barb: You should track the outcomes (papers, collaborations) from the summer school
 - Barb: Suggest Followup-surveys
 - Andrew: Citation analysis on datasets? (From convo: that means the datasets need DOIs.)
- David: is working on "How can we use the data we're gathering to provide more personalization for students..."
 - But the challenge is finding the right assessment measures that align with the goals of the intervention
 - So scaling up intervention is something, but it has to be paired with well-matched assessments
- Monica: Concerns about reidentification in large datasets - as we gather and share, how do we prevent this? Or put in place good standards to minimize the risk, especially of minors.
- **Engaging users (debriefing)**
 - Can we build a community of instructors who use our infrastructure and ready to run experiments?
 - Ensure ACM curriculum coverage or advanced placement testing