AI in Education
Integrating AI into Assignments

DATE
NAME
TITLE
UNIT
CONTACT
Today’s primary learning objectives

• Examine strategies and methods for designing assessment activities
• Explore ways instructors and students can use AI chatbots for assignments
• Examine strategies and issues related to using AI for grading
For the sake of transparency and to model best practices, I want to share some of the AI tools I used when developing this workshop. I utilized ChatGPT, Midjourney, DALL-E, and Designer in PowerPoint. Additionally, Grammarly was installed to ensure polished content. If we consider algorithms and machine learning beyond just language models, I should mention that I relied heavily on Google as well. For event planning, Outlook, which has AI features, was invaluable, along with Smartsheet for registration forms. And let's not forget Zoom, which also incorporates AI.

Yes, I might be overemphasizing, but my goal is to highlight, perhaps with a touch of humor, the blurred and ambiguous lines in our understanding of AI tools. This serves as an engaging entry point to our topic, acknowledging the complexities and contradictions we face as we navigate this new landscape.

There may not be clear cut rules and policies that we can simply follow. It is up to us to use our critical thinking skills to collectively decide what is the best fit for our teaching and learning situations.

Acknowledgements

- ChatGPT4 for summarizing, brainstorming, and suggestions on the content.
- MagicSchool.ai for ideation
- Midjourney & DALL-E to generate images
- Designer feature within PowerPoint to suggest slide layouts and stock images.
- Grammarly for spelling, grammar, conciseness, word choice, and so on.
- Smartsheets for automating registration.
- Outlook and Zoom for event coordination.
- Google for web search...
Let’s stay human

We are all here together in this space and time. We all have different perspectives and are coming from different points. Let’s take advantage of that!

I encourage you to share your thoughts and ask questions. Also connect with folks here today.

I’m certainly prepared to be flexible and improvise as needed. So please be engaged. Be human.
There are a lot of reasons to care about AI. For me personally, I’ve been feeling more optimistic. Maybe it’s the spring weather, but I think the feeling of being off-balance and overwhelmed with AI has settled, and now I’m feeling more confident.

For me creating things is incredibly interesting. Using my imagination is satisfying. I think many of us in higher education might feel that way. And to be honest, sometimes a job or work can be a drag. But I think AI adds a whole new layer of creative power to people. It is empowering and motivating. If AI helps us stay engaged in our work and learn about new things and imagine new possibilities, then that certainly worth caring about.

AI skills are valuable. A lot of graduate and doctoral students have been attending these AI workshop series. Several of them have told me that these skills are valuable to them for their career aspirations. The same might be said for instructors too. These skills will be valuable when seeking promotion or to pivot to other kinds of work.

I think similarly, of our own students want to learn about AI. So integrating AI into...
our course activities in pedagogically sound ways is a service to our students.

- Lastly, AI can help us improve how we do our work. Maybe we just want to learn about AI tools so it can do more of the tasks that we hate doing. That has a big impact! It is important for our quality of life. So it makes sense for us to be here.
As a warm-up, let’s explore some of the possibilities of AI. Go to MagicSchool.ai, and log in with your existing Google or other credentials. Spend a few minutes exploring some of the different tools. Our purpose is simply to get a sense of the different types of tasks teachers and students might use AI for.

It is more geared towards K12 audiences, so you may find the actual outputs are not as sophisticated as you would need in your course. But for this warm-up today, for just the sake of brainstorming and inspiration, I think it is perfect.

Many magic tools could be used in assessment related tasks (or at least inspire and inform how you might prompt a more powerful chatbot). For example:

- Exemplar & non-exemplar
- Rubric generator
- Student work feedback
- Clear directions
- Make it relevant

And so on.
Now that you have ideas for some instructional and learning tasks AI can be used for, let’s dive deeper and think more intentionally about how these tools could align with what we know about good pedagogic practices.
We often use these jargon words when talking about assignments without being clear on what we mean. We may be making things more complicated than they need to be by using these jargon words.

For the sake of clarity, I just want to quickly clarify a few terms as we proceed. Because often we are developing an assignment that incorporate many of these things at the same time and we don’t think to unpack them into these specific parts.

**Formative and summative**
Formative/summative are adjectives that can describe assessments, evaluations, feedback, practice, and so on.
Formative is something that during the learning process
Summative is at the end of the learning process

You might hear folks talk about “formative feedback” on the first draft of a student essay, or “summative assessment” for a final essay.

**Assessment vs Evaluation**
Assessments are meant help you understand what someone knows. You might hear
the term “learning assessment” for an activity that is trying to understand what someone learned.

Evaluations are about assigning value. Usually this means grading or giving points. For example, you might have a “performance evaluation” to determine if a student is performing a skill at certain required level. Or a summative evaluation like an exam where a minimum grade is required to earn a certification.

**Feedback**
There are many types of feedback, like expert, Peer, Self, Formal or informal, Oral or written, and so on
Feedback is usually meant to aid in improving something.
Feedback is also a two-way street. Feedback to students is meant to improve their learning. For example, to clarify misunderstandings, identify and reinforce correct understanding, motivate them, and so on.
Feedback from students to the instructor is meant to improve the teaching. It can help you identify where instruction is working or not working, or to better understand student needs.

**Practice**
Practice can be formal or informal and take many forms. It is usually about students improving a specific skill. Practice and feedback go together. It is usually formative, sometimes summative. Often it doubles as an assessment. But usually not an evaluation
What are some guiding principles?

- **Good assessment design comes first**
  Solid pedagogy will guide us in designing activities or assessments that use AI. To me this means having a systematic method for designing assignments. A common one you may have heard of is Backwards Design. This is an instructional design method where you start by defining the instructional goal, what you want students to be able to do after the lesson. And then from there you define how they will demonstrate that, next how you assess that, then last what content or practice they need.

- **AI alone cannot pass the course**
  I believe that we should be teaching students things that an AI cannot easily do. If a generic AI response could pass this course, then I probably need to rethink things. If I am only teaching my students to do things an AI could do, what value does the course have in their broader education? To be fair, there may be situations where you want to teach students fundamental skills that an AI could do. But likely those basic skills build towards more critical thinking or sophisticated skills that an AI couldn’t do, and that a human (or human+AI team) could do better. That connection and rationale should be clearly communicated to students. I heard one instructor say he tells his students, “I don’t care what the chatbot thinks, I want to know what you think.”
Hold ourselves to the high standards of scholarship
Ethics, integrity, and responsible use are important skills that students learn and practice like anything else. We should hold ourselves and teach students how to live up to the standards of our university, discipline, and community. For example, in previous workshops we talked about Transparency, Accuracy, Responsibility, Attribution, Originality, and Quality as guiding values and strategies. These are concrete things that you could evaluate in an assignment.

Expert use requires expertise
I learned this phrase from a talk by Derek Bruff, who was the director of the Center for Teaching at Vanderbilt. He has a great blog about AI in education. This means that in order to use AI like an expert, for example to assist with medical diagnoses, or to write code for app development, we need to have some expertise ourselves. So, when creating assignments that use AI, I want to make sure that however we are using AI it matches our level of expertise in the domain. For example, if I am a novice in chemistry, then I might start by using AI chatbot as a reflection partner or a motivational coach. But if I were an expert in chemistry, then I might use AI to crunch data for me or to recommend solutions, where I have the expertise to critique or evaluate its outputs.
Many of these proven teaching practices can be compatible with students using AI in an assignment.

Learn more about these strategies and the research behind them on the “Integrating AI into assignments” webpage on the Stanford Teaching Commons at: https://teachingcommons.stanford.edu/teaching-guides/artificial-intelligence-teaching-guide/integrating-ai-assignments

---

Teaching strategies that can align with AI

- **Leverage multiple modalities**
  - More in-class activities
  - Multiple modes of expression

- **Clarify grading practices**
  - Robust citation practices
  - Clear assessment criteria

- **Assess learning throughout**
  - Focus on the process
  - Leverage formative feedback

- **Meaningful activities**
  - Personalize assessments
  - Use authentic tasks

- **More advanced learning**
  - Emphasize metacognitive reflection
  - Prioritize higher-order thinking
A method for designing assignments

1. Why should we care about the assignment?
2. What are students supposed to learn to do? (a.k.a. learning outcome).
3. How will they demonstrate what they learned?
4. How will you measure what they demonstrate?
5. What do they need to practice?
6. What tools would enhance that practice?

For example, let’s say you start with “This is a math class, students need to solve complex mathematical equations.” You might then say, “Ok, students will demonstrate that by solving equations in a quiz.” Then “I assess that by checking how many times they get the correct answers.” Easy. Done.

But we could unpack that further to design a more active and effective assignment.

First, perhaps this assignment is meaningful because this particular skill is used frequently in introductory Physics and Chemistry courses. With that in mind, it is important for them understanding the process of solving the equation, not just to simply get the correct answer.

To demonstrate that they understand the process, students will record themselves solving an equation step-by-step while explaining their logic as they go. The instructor evaluates this work based on a rubric that notes the major steps of solving the problem.

To practice, students solve different problems and then explain it to each other in pairs. Students also analyze incorrect solutions to identify where mistakes were made
and propose a correct solution.

To enhance this assignment, students use iPads for digital whiteboards that they can more easily record and share. AI chatbots are used to generate practice problems and incorrect solutions for students to analyze. A more advanced use of AI might have students explaining their problem-solving process to a chatbot that has been prompted to play the role of a novice learner.
This comes from the successive approximations model of instructional design.

Basically, we start by defining the goal and assessing the needs. Perhaps you note down the most important characteristics of your students, or relevant details about the learning environment.

Then design and test a solution. I recommend starting small. Perhaps it is just one new assignment or adapting an existing assignment. Perhaps you test out a solution with former students or even simulated students using a chatbot.

Then we should get feedback to iterate our assignment. Maybe I am reflecting on how a specific lesson went early in the quarter, to inform what happens later in the quarter.

It is powerful to talk to colleagues, students, and support staff throughout the process. Every stage can benefit from you having a good sense of what would best serve the students and from the perspective of other experts. Reach out to CTL to request a consultation any time.
I’ve been talking about a lot of conceptual things here, but ultimately every course is unique. I can’t tell you what exactly to do because it’s different for each discipline or instructor.

Instead, here are some resources that I hope will spark some ideas for you. This video playlist has many short videos of different kinds of instructors talking about ways they use AI.

In small groups, please spend about 5-7 minutes selecting and watching one video from this playlist. Then spend an additional 7 minutes discussing the proposed idea and how it might apply to your situations.
More examples of AI-powered lessons

- [Al Pedagogy Project](https://metallab.harvard.edu/ai-pedagogy) from metaLAB (at) Harvard
- **Exploring AI Pedagogy** from the MLA-CCCC Joint Task Force on Writing and AI
- [TextGenEd: Continuing Experiments, January 2024 Collection](https://wac clearinghouse.org/) from WAC Clearinghouse

If you’d like to see examples of more comprehensive lesson plans, check out these resources.

They are more in-depth so, you may want to save these for exploring later on your own. But I wanted to share them with you as future food for thought.
I gave you some food for thought when thinking about designing an assessment or assignment that use AI: Guiding principles, an ID method, the successive approximations model of design, a bunch of pedagogic strategies, and ideas for calibrating assessment criteria.

What else might you add to that list?
Recall some of the tools in MagicSchool.ai and how AI might be used for grading.

For example, AI chatbots might generate initial feedback on student learning, assist you in developing a rubric, or generate example solutions to a complex problem.
We often don’t think about it explicitly, but grades have many functions.

Instructors are usually aware of how grades functions in their courses. For example, participation points that contribute 15% to the final grade, are used to track progress and motivate students. Or grades like “8 out of 10 correct” on practice quiz are functioning as feedback for students on what content elements or skills they have mastered. Often the grading for one assignment serves multiple purposes.
However, students may not be as aware of the different ways grading functions. You probably have observed that some students might hyper-focus on grades as a checkpoint for advancement and as an evaluation of quality. They might be overlooking the other functions of grading to the detriment of their learning. Perhaps they see assignments as “busy work” or due dates as needlessly strict, rather than seeing assignments as a form of practice or due dates as a way of keeping pace. Consider ways you can communicate to students the purpose of grades in different contexts and how they can make the most of them.
Trust and transparency are foundational
Students care a lot about grades. So, it’s important that they understand why and how you are using AI for grading, how it benefits them, and how it affects their grades. If not, students may begin to have doubts about the course. Does the instructor care about the work students do? Is the grading fair? And so on, which can really erode trust, motivation, and the quality of learning.

Empower humans in the loop
Assure that a human is always in the loop when using AI particularly for grading. Perhaps you check the AI outputs for fairness and accuracy before sharing them with students. Or perhaps you facilitate activities where students compare, critique, and reflect on AI generated grades vs human generated grades, before giving themselves a final self-evaluation. Give students agency in using AI. Perhaps students upload their work to an AI themselves.

Another example is you might use AI to develop a grading rubric for a summative evaluation assignment, like a final paper. You might instruct the AI to evaluate the
paper based on the rubric, but to only identify areas that are not being met and to not assign any grade, just highlight areas to look closer. Then the actual assigning of a grade is done only by you.

Remember, your expertise, acknowledgement, and stamp of approval is uniquely meaningful and should not be replaced by an AI.

**Leverage AI for process-oriented tasks**
Along the same lines as the previous point, think about using AI for tasks within the grading process, rather than for assigning grades themselves. Perhaps AI can help with sorting student submissions into more manageable chunks, making your feedback more concise, or double-checking your scoring calculations.

**Quality and efficiency**
This principle also relates to the previous point. Here think about how AI can improve the quality of assessment and evaluation, not only making things more efficient. For example, maybe you have all the quiz results in a big spreadsheet and use AI to identify the most common misconceptions, then generate explanations that address those misconceptions using a variety of examples that connect to students interests which in turn improves the quality of your feedback.
Calibrating grading criteria for AI

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+</td>
<td>Far exceeding AI outputs, an exemplar</td>
</tr>
<tr>
<td>A</td>
<td>Significantly better than AI</td>
</tr>
<tr>
<td>B</td>
<td>Better than AI, but still could be improved</td>
</tr>
<tr>
<td>C</td>
<td>Some human value added beyond generic AI outputs, needs work</td>
</tr>
<tr>
<td>D</td>
<td>Barely better than generic AI outputs, minimal human value added</td>
</tr>
<tr>
<td>F</td>
<td>Generic AI output, no human enhancement</td>
</tr>
</tbody>
</table>

My intention here is to think more intentionally about what is the new “average” in light of what AI tools can do.

For example, if we can now use AI to generate generic images of most anything we want, then we might now have higher standards on student slide presentations on how they use images to persuade the audience. Or maybe I take a different approach and lean away from visual communication. If an AI can make any slide presentation look visually engaging, then maybe that frees me up to evaluate other skills that are important to the discipline instead. Perhaps I am now evaluating how students respond to live questions in a Q&A or poster session during class.

To calibrate your rubric, you might use a chatbot to generate example student submissions based your assignment instructions. Try prompting the AI with just the basic assignment instructions, then with some additional prompting, then maybe try one where you really work with the chatbot to refine the output. How much human input, or going back and forth with the AI is needed to get an A-level output? Is that back-and-forth with a chatbot a useful learning activity for the student? Those examples and perhaps prior examples of student submissions can help you calibrate an assignment to match what would be a good stretch for students using AI.
When calibrating any rubric or assessment criteria, whether suing AI or not, think about the zone of proximal development, learning is optimal when students stretch and challenge themselves, but not too hard that it overloads them or causes them to give up. Same if an assignment is too easy.
Currently, there are no vetted campus AI tools that are widely available to Stanford students or instructors. For the time being, most instructors and students would likely be using a commercially available AI chatbot. Therefore, when using these commercially available chatbots, there are some issues to consider.

At Stanford as of June 2024:
• Some developer AI tools are approved
• No AI chatbot is centrally supported for general use
• UIT security reviews are under way
• Limited chatbot pilots in SoM, GSB, and other units

See GenAI Tools Evaluation Matrix for latest status updates.
Responsible use of AI

- Privacy
- Security
- Access

Learn more at Responsible AI at Stanford

Privacy
Anything you put into a commercially available chatbot can potentially be used by the company that owns that chatbot service. The terms of service vary, but in general the guidance is don’t put anything private or sensitive into a chatbot and opt out of sharing data where available. This is an important consideration when putting students work into a chatbot.

Security
Chatbots like any web service have varying levels of security. Again, the terms of service vary, and different companies might have better track records than others. But ultimately, most of us here (myself included) and the average student or instructor are not security experts. So again, the guidance is to not put any secure or sensitive information into a chatbot. Again, this is worth careful consideration if you are putting student work into a chatbot.

Access
This is also a challenge because most advanced chatbots require paid subscriptions. Free versions of chatbots are typically have lower quality outputs or limit how often you can use them. This issue of affordability and access can create inequities
especially for students.
What can our guiding principles tell us?

1. Be transparent with students
2. Never put private or sensitive information into a chatbot
3. Get informed consent before you put student work into a chatbot
4. Don’t put copyrighted materials into a chatbot without permission
5. Help students make smart and informed choices
6. Take responsibility for anything you do using AI
7. Use the most accessible or affordable tools whenever possible
8. Provide alternatives where possible for those who can’t or don’t want to use AI

Things are still evolving and there are no clear-cut answers.

But our guiding principles of trust, transparency, empowering humans, expert use, and high standards of scholarship, can inform our decisions here as well.

These are suggestions that we’ve talked about within my academic technology community. What other practices or guidelines come to mind for you?
How might you use AI for grading-related tasks?
Design an AI-enhanced assignment
Now we’re going to practice applying these concepts and strategies to AI in a design sprint.

This is just practice so don’t expect to have something polished. It is more to explore a bunch of AI stuff and share ideas to reinforce this structured way of designing an assessment.
If you want to have a bit more structure, I’ve prepared a worksheet you can use. There are printed copies on the table too.

Does everyone have a computer or can pair with someone who does? A computer or device is useful for exploring ideas and getting fuel for your brainstorm.
What is your initial assignment concept?

What next steps might you take to further refine it?
Continue to engage

• Use AI chatbots for your work tasks
• Read the AI Teaching Guide on the Teaching Commons website
• Attend upcoming CTL workshops on AI
• Talk about AI with colleagues and students
Wrap-up activity

Insert link or QR code to your survey activity here.
Terms for using these slides

This slide presentation was created by Stanford Teaching Commons. It is licensed under Creative Commons BY-NC-SA 4.0 (attribution, non-commercial, share-alike).

We encourage you to organize your colleagues to engage with AI topics. Consider how you might adapt, remix, or enhance these slides for your own needs.