Governance by Design: A Spectrum of AI Teaching Tools for Computer Science

Ethan Dickey

PURDUE UNIVERSITY
Department of Computer Science

A campus-scale portfolio of AI tools that scale teaching and learning responsibly through diverse governance & measurable learning gains

Live Demonstrations

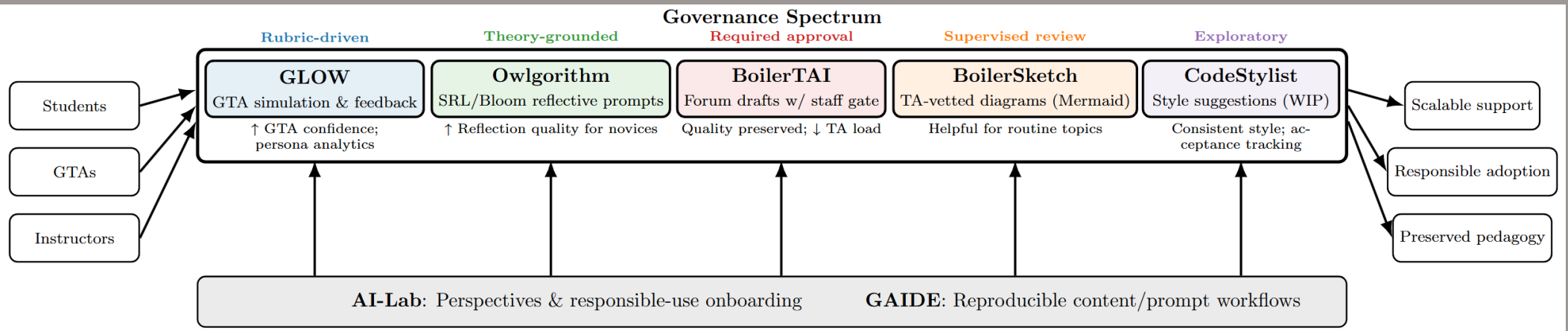
GLOW:

For sample test code, see cs.purdue.edu/homes/dickeye

Owlgorithm:

CodeStylist:

BoilerSketch: tablets below!
BoilerTAI: tablets below!



Why Governance Matters

As AI becomes woven into instruction, the challenge isn't whether to use it, but *how* to guide its use responsibly. Governance decisions shape how students, GTAs, and instructors interact with AI tools, influencing trust, workload, and learning depth.

Challenge	Our Approach
AI use in teaching grows faster than policy	Quickly deploy context-specific governance rather than bans or full automation
Human capacity for feedback is limited (n≥800)	Embed rubrics, review gates, or supervision in AI workflows
Students need reflection, not answers	Scaffold SRL and metacognition through guided prompts
Quality and ethics need visibility	Make each governance choice explicit and measurable

Fig. 1 situates 5 AI teaching tools and two supporting frameworks along a governance-by-design spectrum, ranging from strong governance in rubric-anchored automation to exploratory guidance. Each tool embodies different balances of oversight and autonomy:
GLOW – a rubric-based simulation where GTAs practice office-hour conversations with AI personas and receive structured feedback.
Owlgorithm – a reflection assistant that creates Bloom-aligned, self-regulation questions about a student's code in competitive programming courses.
BoilerTAI – a forum assistant producing draft responses that require staff review and approval before posting.
BoilerSketch – a multimodal tutor generating TA-vetted Mermaid diagrams for common CS concepts.
CodeStylist – a developing tool offering course-specific code-style coaching through LLM suggestions.
AI-Lab and GAIDE – cross-cutting frameworks providing responsible-use training and reproducible prompt/content workflows.
Together, these components illustrate how governance diversity – not uniform control – supports scalable, trustworthy AI integration in computer-science education.

Evidence of Impact

Across six Purdue CS and ENGR courses and more than 1100 participants from 2024-2025, early findings reveal consistent, measurable improvements in both learning outcomes and teaching efficiency. **GLOW** participants improved rubric scores and reported greater confidence handling complex student personas. **Owlgorithm** led to longer, more analytical reflections and richer self-explanations after programming tasks. **BoilerTAI** reduced median forum response times by more than half while maintaining high student satisfaction and lowering TA cognitive load. **BoilerSketch** enhanced clarity for foundational topics, with most staff rating its diagrams helpful when supervised. **AI-Lab** participants showed increased comfort and more mindful, ethical GenAI use across coursework.

Together, these results highlight how matching governance to context – through rubrics, theory alignment, review gates, or supervision – can simultaneously improve readiness, reflection, efficiency, and trust, offering a reproducible model for responsible, scalable AI integration in computing education.

What We've Learned About Responsible Scale

1. **Match governance to context.** Rubrics, review, or supervision, not one size fits all.
2. **Scaffold reflection, not replacement.** AI should assist and accent thinking, not do it.
3. **Keep humans visible.** Accountability increases trust and adoption.
4. **Evaluate continuously.** Governance is an evolving system, not a checkbox.
5. **Integrate seamlessly.** Effective governance feels like good course design, not extra friction.

Next Steps: deployment of CodeStylist; GLOW campus expansion; cross-department replication studies (multiple).

Collaborators and Acknowledgements

Andres Bejarano, Co-PI
Postdoc and Staff collaborators: Bárbara Fangues, Quiondriya Gee, Rhianna Kuperus, Houyame Lkhider-Hudson
Undergraduate Research Assistants:
Shruti Goyal, Erin Kramer, Peter Kurto, Juliana Nieto, Ashok Saravanan, Alex Siladie, Anvit Sinha, Zachary Sy, Vivan Tiwari, Libra Vento
Special thanks is extended to the Innovation Hub, who funded much of this work (IH-AI-23002), David Nelson and Brooke Harris-Thomas, and the instructional specialists and TAs who demoed and provided feedback.
See cs.purdue.edu/homes/dickeye for resources on AI-Lab and GAIDE.