

# Equitable AI for Cognitive Accessibility: Designing CourseCore, a Trustworthy Study Companion for Students with ADHD

Shrutigya Uprety, Junior | Advisor: Dr. Nan Wang, Department of Computer Science, William Paterson University of New Jersey

Benchmark results completed May 11, 2026. Student pilot scheduled for May 13-19, 2026.

## Abstract

CourseCore addresses a gap in AI study tools: ChatGPT and NotebookLM are designed for users with intact executive function, but students with ADHD disproportionately struggle to verify AI outputs, exactly the labor these tools require. We built CourseCore, a study companion that structurally refuses to answer outside the professor's uploaded course materials and filters fabricated citations before display. Across 100 benchmarked prompts, CourseCore produced 0 hallucinations; a baseline LLM produced 50 (50%).

## Research Problem

### Largest disabled-student population in higher education

- ~17% of U.S. undergraduates self-report ADHD (ACHA-NCHA, 2023).
- ADHD is a major accessibility-services population in higher education.
- At WPUNJ, ADHD is a dominant accessibility-services population.
- Generic AI tools were not designed for the students most burdened by verification labor.

**17%**  
U.S. undergraduates

Goal: move trustworthiness from something the student must do to something the system structurally provides.

## How CourseCore Differs

No existing study tool combines course-bounded refusal, citation verification, and executive-function scaffolding in one workflow.

Capability	ChatGPT	NotebookLM	CourseCore
Refuses outside materials	No	Partial	Yes
Filters hallucinated citations	No	No	Yes
Page-as-provenance	No	No	Yes
Forced attempt	No	No	Yes
Tracks mastery	No	No	Yes
Designed for ADHD	No	No	Yes

## Contribution

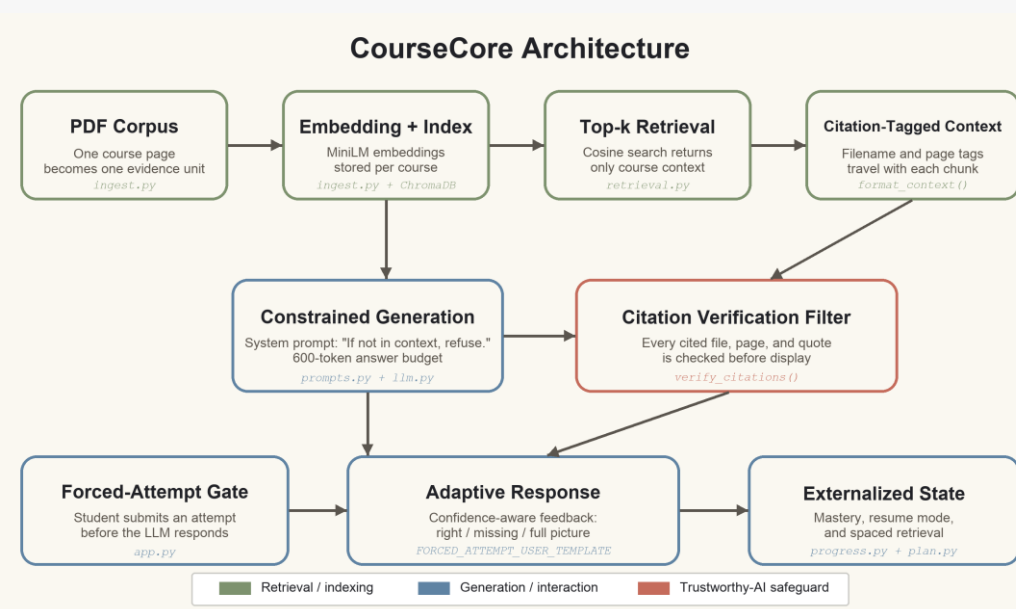
CourseCore is the first AI study system combining corpus-bounded refusal, post-hoc citation verification, and executive-function scaffolding designed for ADHD.

**Bounded Verified Provenance Scaffolding**

Measured outcome: 0 hallucinations across 100 prompts, with 95% out-of-corpus refusal and comparable in-corpus usefulness.

## Method and System Design

### Architecture + benchmark protocol + scoring rule



Bounded

Verified

Provenance

Scaffolding

### Benchmark protocol

- Run both systems on the same retrieved context for each prompt.
  - Judge each response against the retrieved pages, not against world knowledge.
  - Count REFUSED, ANSWERED\_FROM\_CORPUS, and HALLUCINATED outcomes.
  - Compute rates from those counts: refusal, coverage, hallucination, citation coverage, and trustworthiness.
- Adversarial prompts = near-miss AI topics plus trick questions with real course terms but outside content.

### Scoring + pilot design

INFLU: low response R, corpus index C  
FOR each citation in R:  
IF cited page not in C: REMOVED citation  
ELSE IF quoted text not found on cited page: REMOVE quote  
ELSE: KEEP citation  
RETURN: verified response

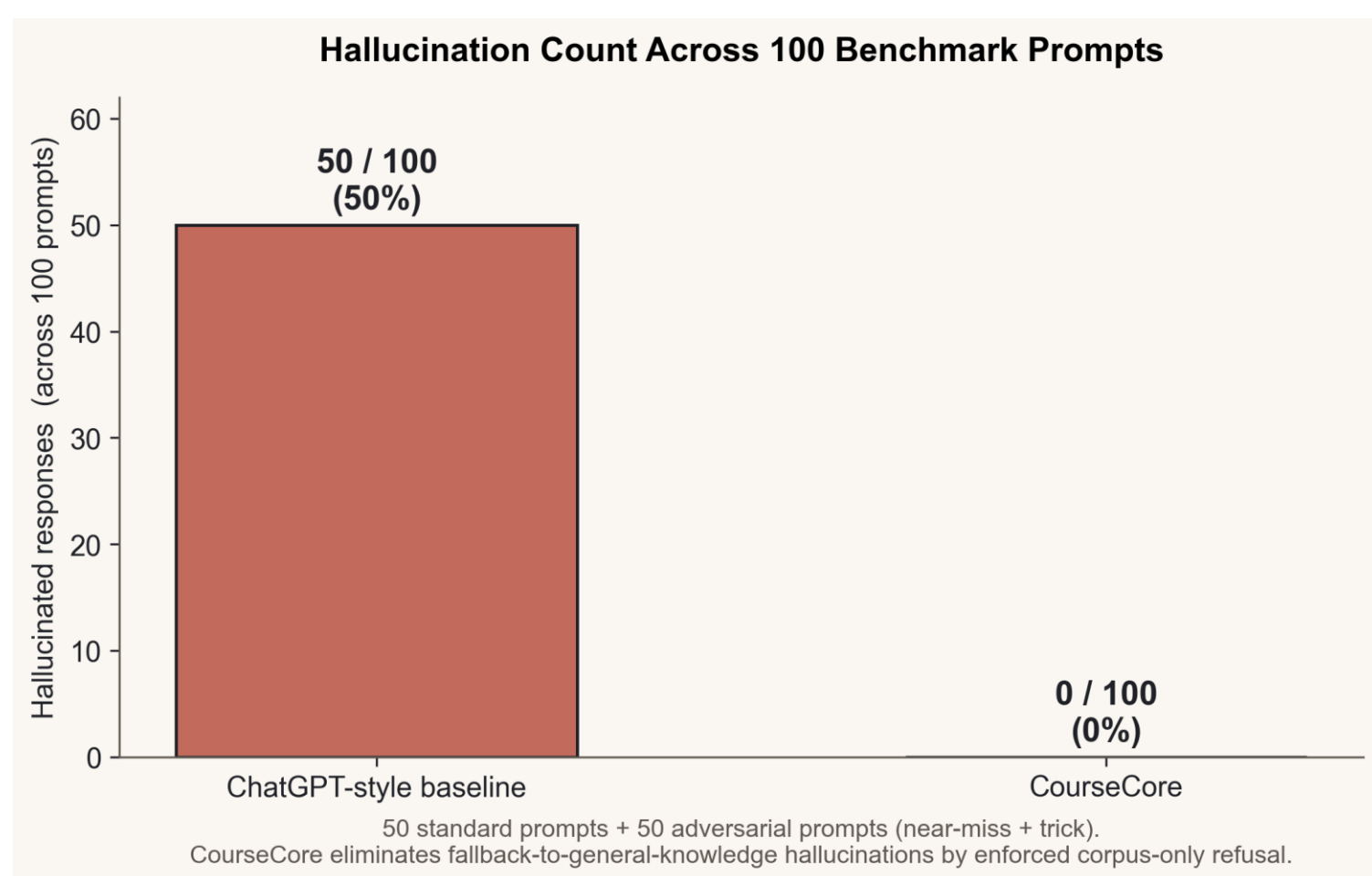
Within-subjects pilot (N=14 WPUNJ students with ADHD): NASA-TLX cognitive load + 7-day free recall + SUS usability, ChatGPT vs CourseCore counterbalanced.

Trustworthiness counts safe refusals and grounded answers; coverage tracks in-corpus usefulness; hallucination cost counts unsupported answers.

## Results

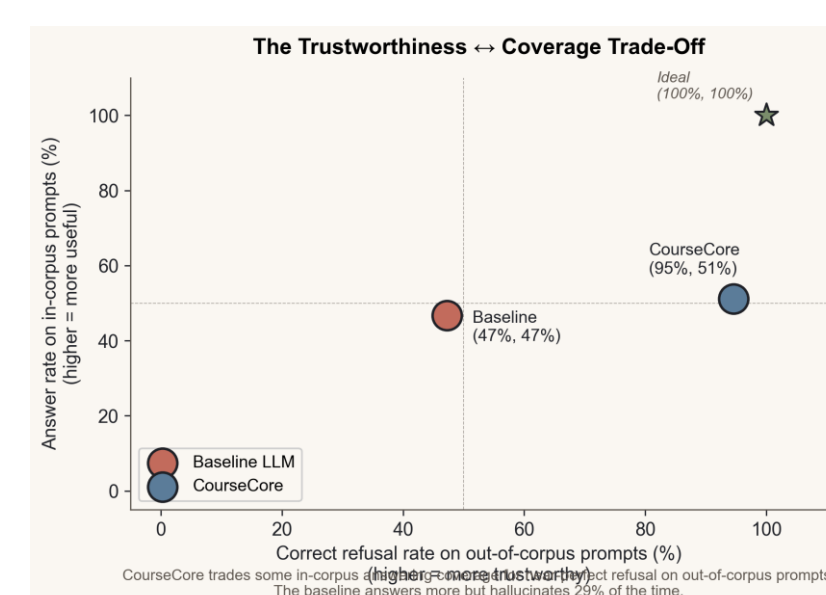
Methods: 100 prompts (50 standard + 50 adversarial) x 2 systems with identical retrieved context. Rates are calculated from REFUSED / ANSWERED\_FROM\_CORPUS / HALLUCINATED labels.

### 1. Hallucination trajectory



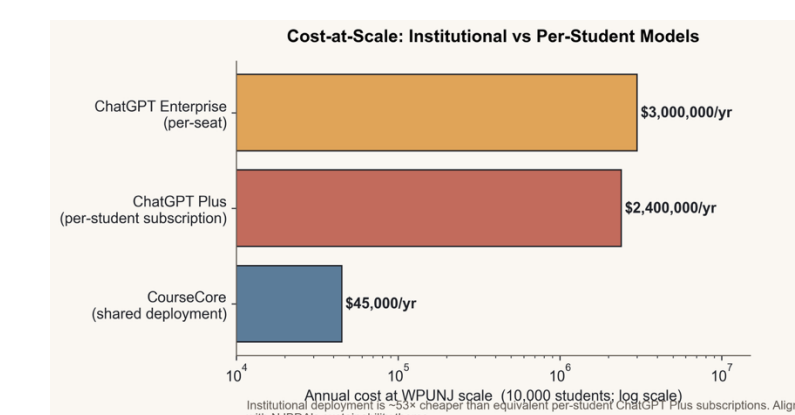
Across 100 benchmark prompts, CourseCore records 0 hallucinations while the baseline reaches 50 out of 100. This is the poster's central result.

### 2. Trustworthiness vs coverage



Higher trustworthiness without losing in-corpus usefulness. The top-right direction is the win condition.  
Interpretation: CourseCore preserves nearly the same in-corpus usefulness while sharply increasing safe refusal on outside content.

### 3. Cost at institutional scale



Measured cost is \$0.0045 per query. At WPUNJ scale, shared deployment is ~54x cheaper than per-student ChatGPT Plus subscriptions.

15/15 WCAG 2.2 cognitive SCs meet

Interpretation: trustworthiness also has an infrastructure side. The system is cheap enough to deploy institutionally and accessible enough to use under real cognitive load.

0 / 100 hallucinations

95% out-of-corpus refusal

23 / 45 in-corpus answers

54x lower annual cost

Interpretation: CourseCore gives up some answer coverage on genuinely in-corpus questions, but it eliminates hallucinations, sharply reduces verification burden, and makes the refusal boundary explicit to the student.

## ADHD Accessibility Rationale

### Failure mode -> design response

Accessibility challenge	CourseCore response
Decision fatigue	Decide-for-me mode
Working-memory drift	Resume and mastery state
Time blindness	Pomodoro and timing cues
Verification burden	Inline citation filtering
Hyperfocus / boundary blindness	Soft-stop detection at 30+ min
Initiation paralysis	Concept browser kills blank-page freeze

## Limitations and Next Steps

Pilot design: counterbalanced within-subjects comparison of ChatGPT vs CourseCore in 14 WPUNJ students with ADHD.

- Benchmark-first evidence; pilot data collection is the next validation step.
- CourseCore still over-refuses some in-corpus questions.
- Current corpus is one AI course only.
- WCAG audit is self-coded, not third-party reviewed.

Planned measures: NASA-TLX, free recall, citation-checking behavior, and SUS usability.

## References

- Antshel, K. M., et al. (2014). Cognitive function in college students with ADHD. *J. Attention Disorders*, 18(2).
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions. *Psychological Bulletin*, 121(1).
- Brown, T. E. (2013). A new understanding of ADHD in children and adults: Executive function impairments. *Routledge*.
- Dunlosky, J., et al. (2013). Improving students' learning with effective learning techniques. *Psych. Sci. Public Interest*, 14(1).
- Fricker, M. (2007). Epistemic injustice. *Oxford University Press*.
- Ji, Z., et al. (2023). Survey of hallucination in NLP. *ACM Computing Surveys*, 55(12).
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning. *Science*, 331(6018).
- Knouse, L. E., et al. (2005). Accuracy of self-evaluation in adults with ADHD. *J. Attention Disorders*, 8(4).
- Lewis, P., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *NeurIPS*, 33.
- Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning. *Psychological Science*, 17(3).
- Volkow, N. D., et al. (2009). Evaluating dopamine reward pathway in ADHD. *JAMA*, 302(10).
- Wing, J. M. (2021). Trustworthy AI. *Communications of the ACM*, 64(10).
- WCAG 2.2. <https://www.w3.org/TR/WCAG22/>
- W3C. Making Content Usable for People with Cognitive and Learning Disabilities. <https://www.w3.org/TR/coga-usable/>
- NCEES. Students with Disabilities, Condition of Education. American College Health Association. NCHA Fall 2023 Reference Group.

