Using and Evaluating LLMs in Academic Work

Session 1: Factual Accuracy, Responsibility, and the Limits of Fluency

Moses Boudourides

Faculty, Graduate Program on Data Science
Northwestern University

Moses.Boudourides@northwestern.edu
Moses.Boudourides@gmail.com

instats Seminar

Wednesday, February 25, 2026
4:00 PM – 5:30 PM UTC

Northwestern | SCHOOL OF PROFESSIONAL STUDIES

Session 1: Factual Accuracy, Responsibility, and the Limits of Fluency

1. Challenges of LLMs in Scholarly Practices
2. Factual Accuracy, the Allure of Fluency, and Hallucinations
3. Governance and Validation

The LLM Revolution in Academia

- LLMs are becoming everyday collaborators in academic life.
- They assist in tasks ranging from:
  - Drafting manuscripts and literature summaries.
  - Designing syllabi and educational materials.
  - Supporting student supervision and feedback.

Technical vs. Epistemic Questions

- Adoption raises technical questions (efficiency, speed).
- More importantly, it raises fundamental epistemic and ethical questions.
- Key concern: Responsibility to preserve scholarly reliability and fair attribution.

Responsibility of the User

- The responsibility for verification rests squarely with the human user.
- Using an LLM does not absolve a researcher of responsibility; it heightens it.
- Researchers must verify all outputs.

Limitations of Current Discourse

- Current focus is often on preventing plagiarism and fact-checking.
- This focus on "local accuracy" is necessary but insufficient.
- It fails to address deeper systemic challenges to scholarly communication.

Beyond Isolated Errors

- The most significant risks are not simple factual mistakes.
- Risks involve systemic, structural distortions of knowledge.
- These distortions can misrepresent the entire landscape of a field.

The Primacy of Factual Accuracy

- Factual accuracy is a foundational moral requirement in academia.
- It is not optional or negotiable.
- No level of productivity justifies the introduction of error into the record.

The Mirage of Fluency

- LLMs generate highly fluent and stylistically appropriate text.
- This "polish" can mask underlying inaccuracies.
- It leads to a false sense of confidence in the model's reliability.

How LLMs Function: Probabilistic Prediction

- LLMs are designed to generate plausible sequences of text.
- They are NOT retrieval systems for verified facts.
- They predict the next token based on statistical patterns.

Defining Hallucinations

- Statements that are entirely fabricated but presented with confidence.
- They arise from the model's goal of maintaining surface-level plausibility.
- These errors are often indistinguishable from truth without domain expertise.

Synthesis vs. Accumulation

- Academic work is more than just a list of facts.
- It requires organization, interpretation, and synthesis.
- LLMs can fail at this global level while succeeding at a local level.

Local vs. Global Integrity

- **Local Accuracy**: Are the individual sentences true?
- **Global Integrity**: Is the overall structure of the field represented correctly?
- Distorting relationships between concepts is a failure of global integrity.

Intro to Structural Hallucination

- Distorts the relational structure of knowledge.
- Misrepresents bibliographic landscapes.
- Undermines intellectual attribution.

Conceptual Re-centering

- Assigning disproportionate importance to peripheral concepts.
- Marginal ideas are presented as foundational.
- This reshapes the perceived "core" of the intellectual field.

Accountable Research Partnership

A Necessary Paradigm Shift

- Move from "Uncontrolled Artificial Generators".
- Transition to an "Accountable Research Partnership".
- Governance must be proactive rather than reactive.

The Danger of Uncontrolled Generators

- Systems that produce text without user-side validation.
- Lacks checks for structural or bibliographic coherence.
- Leads to the spread of "polished" misinformation.

The Accountable Partnership

- A collaborative process between human and machine.
- Human actively governs LLM contributions.
- Use of explicit methodological and evaluative protocols.

Academic Inquiry Values

- The partnership is grounded in intellectual honesty.
- Requires rigorous argumentation.
- Committed to a shared, verifiable body of knowledge.

The Governance Framework

- Researcher- and instructor-centric.
- Suite of lightweight computational protocols.
- Individually implementable (not just for AI specialists).

The Multi-layered Validation Pipeline

- Translates text into structured representations.
- Quantitative comparison against trusted sources.
- Ensures structural and bibliographic integrity.

Epistemic Responsibility

- The framework is not meant to replace human judgment.
- It is a set of instruments to discipline human responsibility.
- Focus is on governance, not just performance optimization.

Application in Research

- Validating literature reviews and conceptual frameworks.
- Ensuring intellectual lineages are correctly positioned.
- Shifting from textual polish to epistemic alignment.

Application in Teaching

- Assessing if foundational works are emphasized in curricula.
- Detecting thematic gaps in LLM-generated syllabi.
- Maintaining expert control over curricular integrity.

Application in Student Assessment

- Distinguishing engagement from hollow reproduction.
- Using "validation reports" as assessment tools.
- Moving beyond crude detection to assessing critical use.

Summary: The Governance Shift

Reactive (Old)	**Proactive (New)**
Fact-checking	Structural Validation
Plagiarism detection	Attribution Integrity
Outright prohibition	Methodological Governance

Role of Network Science

- Uses established techniques from network science.
- Provides transparent, quantitative toolkits.
- Makes relational structures explicit and inspectable.

Bibliometric Foundations

- Grounding LLM outputs in real scholarly ecosystems.
- Comparing generated content to bibliometric baselines.
- Ensuring thematic coherence through data.

Reaffirming Academic Values

- Accountability is normative, not just technical.
- Requires a willingness to engage critically with new tools.
- Governance enables trustworthy scholarly contributions.

Session 1 Summary

- LLMs are powerful tools for academic work, but they also pose significant challenges.
- The most critical challenge is not factual error, but **structural hallucination**.
- We need to move from a model of uncontrolled generation to one of an **accountable research partnership**.
- This requires a proactive approach to methodological governance, based on a multi-layered validation pipeline.

Seminar Roadmap

Across sessions we will:

1. Looking Ahead to Session 2: **Diagnose structural hallucination**
   - Deep dive into structural hallucination.
   - We will explore failure modes in detail.
   - Understanding the complex, hierarchical nature of academic knowledge.
2. Translate text into networks
3. Apply network diagnostics
4. Validate citation integrity
5. Reframe attribution and assessment

Questions and Discussion

Thank you!

Questions?

Moses.Boudourides@northwestern.edu
Moses.Boudourides@gmail.com