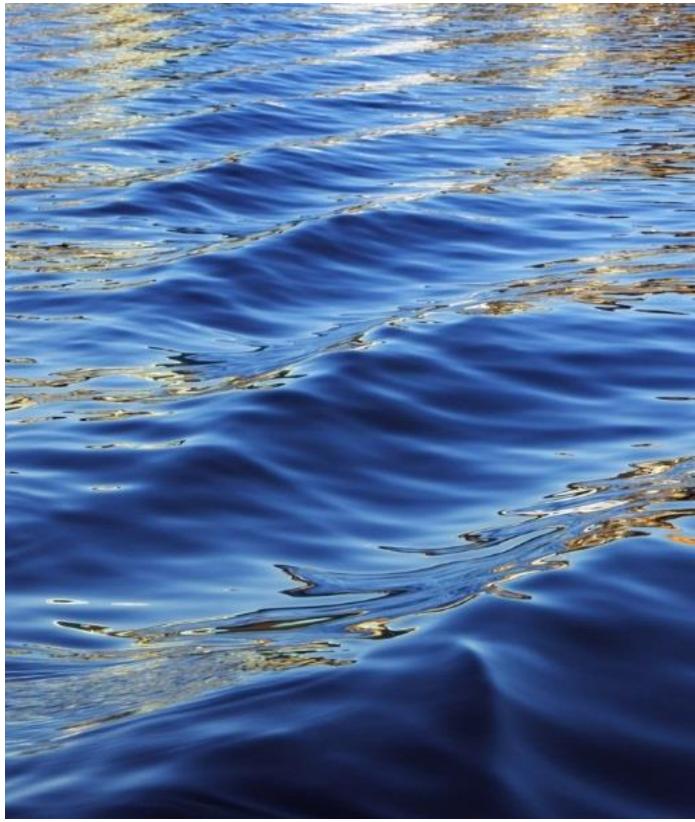
TURI

TOXICS USE REDUCTION INSTITUTE

Evaluating the Effectiveness of AI Tools in Extracting Regulatory-Relevant Data from Carbon Nanotube Studies

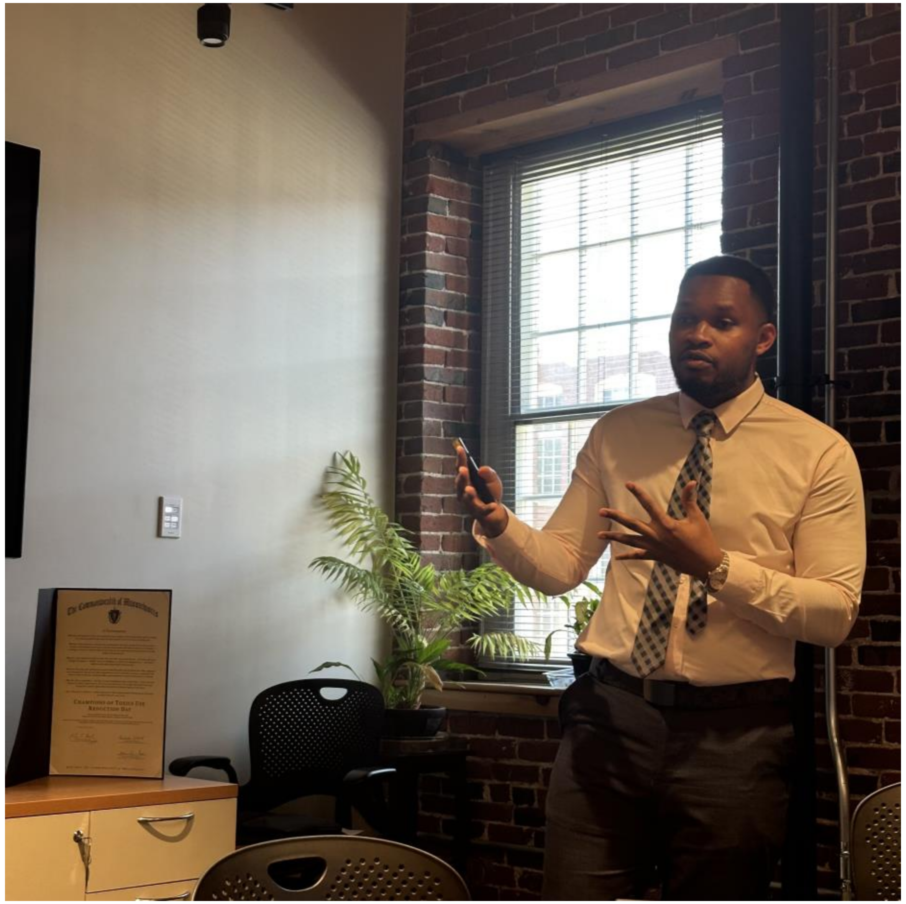
Presented by:
Alicia McCarthy
Senior Safer Chemicals Scientist
Toxics Use Reduction Institute (TURI)



UMass Lowell Capstone Project

Graduate MPH Practicum Project
Arthur Musinguzi, MPH-Epidemiology
UMass Lowell

Project Purpose
Evaluate how Large Language Models (LLMs) could be used as a tool in answering Toxics Use Reduction Act (TURA) Science Advisory Board (SAB) follow-up questions that require rapid, large-scale literature review.



TURI
TOXICS USE REDUCTION INSTITUTE

What Is the TURA Science Advisory Board (SAB)?

SAB Overview

- Independent scientific advisory body established under the Toxics Use Reduction Act (TURA)
- Provides technical and scientific recommendations to MassDEP and the TURA Administrative Council
- Reviews chemicals proposed for listing, delisting, or categorization

Who Serves on the SAB?

- Appointed experts in toxicology, public health, epidemiology, chemistry, industrial hygiene, and risk assessment
- Members serve in an independent role to ensure evidence-based recommendations

Challenges

- SAB questions often require reviewing dozens of studies quickly
- Manual extraction is time-consuming
- Industry/Lobbyist stakeholders often have more analysts and can submit extensive comments faster

Methods Overview

- Collected CNT studies, SAB questions, and staff responses
- Reviewed journal licensing & AI-use policies
- Selected five LLMs & review data-handling policies
- Completed prompt-engineering education
- Developed strict exact-text prompts
- Randomly sampled 30 peer-reviewed CNT studies from 120 total
- Processed each study through all five LLMs using the same prompt set
  - Eight prompts per study (one per variable)
  - 240 extraction outputs per LLM tool
- Compared outputs & analyzed error pattern
  - Accurate (match)
  - Omission (missing data)
  - Incorrect extraction
- Assessed feasibility, operational constraints, and implications for staff understanding of AI capabilities and limitations

Variable	**Description**	**Examples**
Study Type	Study design or format	Experimental, review
Nanomaterial Type	Specific carbon nanotube studied	SWCNT, MWCNT
Biomarkers	Biological markers measured	Inflammatory cytokines, fibrotic markers
Key Nanomaterial Properties	Physicochemical properties of nanotubes	Surface area, diameter, length, purity
Toxicity/Health Effects	Adverse outcomes observed	Pulmonary inflammation, genotoxicity
Exposure Assessment	Exposure routes or conditions	Inhalation dose, duration, pathways
Study Findings	Main conclusions related to nanotube toxicity	Pulmonary toxicity
Funding Source	Source of research funding	For bias assessment

Table 1: Target Variables for Information Extraction from Carbon Nanotube Studies

TURI
TOXICS USE REDUCTION INSTITUTE

Key Findings

All five LLMs showed generally reliable accuracy, with omissions being the most common error.

Best performance: clear, structured information.

Challenging fields: nanomaterial properties, exposure details

LLMs were most useful for screening large literature sets and spotting repeated patterns.

Journal licensing limits AI use: cloud tools may store uploaded PDFs unless they guarantee ephemeral, non-retained processing.

Paid versions (Enterprise subscription)

Local network/offline LLM option

Operational constraints (file size limits, memory, data-handling policies) influence tool suitability for regulatory or research contexts.

Final Takeaways from This Project

- LLMs are most useful **after** literature is identified
  - Use literature-mapping tools to speed up article discovery: **LitMaps, Connected Papers, Research Rabbit** (non-generative, AI/Machine Learning to map citation networks)
  - Reference managers for automated organization
  - Helpful for early-phase synthesis or preparing summaries for SAB responses.
- LLMs help organize information and spot repeated patterns across studies.
  - Struggle most when information appears in multiple sections or uses inconsistent terminology
  - Human interpretation is needed for nuanced toxicology
- Licensing, storage, and file-size limits shape when LLM can be used.
- This project builds internal understanding of AI's capabilities and limitations.
- AI tools evolve rapidly (Results reflect the capabilities *as of early 2025 models)*
- Limited to a theoretical evaluation: no real-world deployment.
  - This project assessed *feasibility*, not implementation. No AI was used to support active SAB decisions.

TURI

TOXICS USE REDUCTION INSTITUTE

http://Turi.org

126 John Street, Suite 14
The Offices at Boott Mills West
Lowell, MA 01852

Alicia McCarthy

Senior Safer Chemicals Scientist

978-934-3889

Alicia_McCarthy@uml.edu