Title: Searching Graphics and Text in Technical Documents: A Brief Overview and Plan

Speakers: Richard Zanibbi

Collection/Series: Theory + AI Symposium

Date: April 08, 2025 - 11:40 AM

URL: https://pirsa.org/25040072

Abstract:

What would effective and usable tools for searching text and graphics in research papers look like? In this talk we sketch a partial answer to this question, with reference to recent work in the Document and Pattern Recognition Lab at RIT. Two multimodal paper search prototypes, one for math (MathDeck) and one for chemistry (ReactionMiner search) will be used for illustration. A simple framework based on 'jars' of available information sources can organize and relate the actions performed by people and automated systems when retrieving, analyzing, and synthesizing sources. We will organize our answer sketch around this framework, and share open questions and research opportunities related to enhancing multi-modal search tools for expert and non-expert users.

Note: ReactionMiner was developed in collaboration with NCSA and the Han lab at the University of Illinois, Urbana-Champaign.
MathDeck demo: https://people.rit.edu/ma5339/mathdeck_landing
ReactionMiner search demo: https://reactionminer.platform.moleculemaker.org/home

Biography:

Richard Zanibbi is a Professor of Computer Science at the Rochester Institute of Technology (RIT, USA) where he directs the Document and Pattern Recognition Lab (dprl@RIT). His research focuses upon the recognition and retrieval of graphical notations, particularly for mathematics and chemistry. He is also a member of the Molecule Maker Lab Institute (MMLI), one of the first NSF AI Centers. He received his PhD from Queen's University (Canada), and was an NSERC Postdoctoral Fellow at the Centre for Pattern Recognition and Machine Learning (CENPARMI) at Concordia University before joining RIT.

dprl
Document and Pattern Recognition Lab

MOLECULE MAKER LAB INSTITUTE

RIT
Rochester Institute of Technology

Searching Graphics and Text in Technical Documents

A Brief Overview and Plan

Richard Zanibbi, Department of Computer Science, Rochester Institute of Technology
Director, Document and Pattern Recognition Lab (dprl@RIT)
Member, Molecule Make Lab Institute (MMLI NSF AI Center @ UIUC)
April 8, 2025

Motivation

What would **effective** and **usable** tools for **searching text and graphics** in research papers look like?

Information Task Framework

Information Tasks: The "Jar Model"

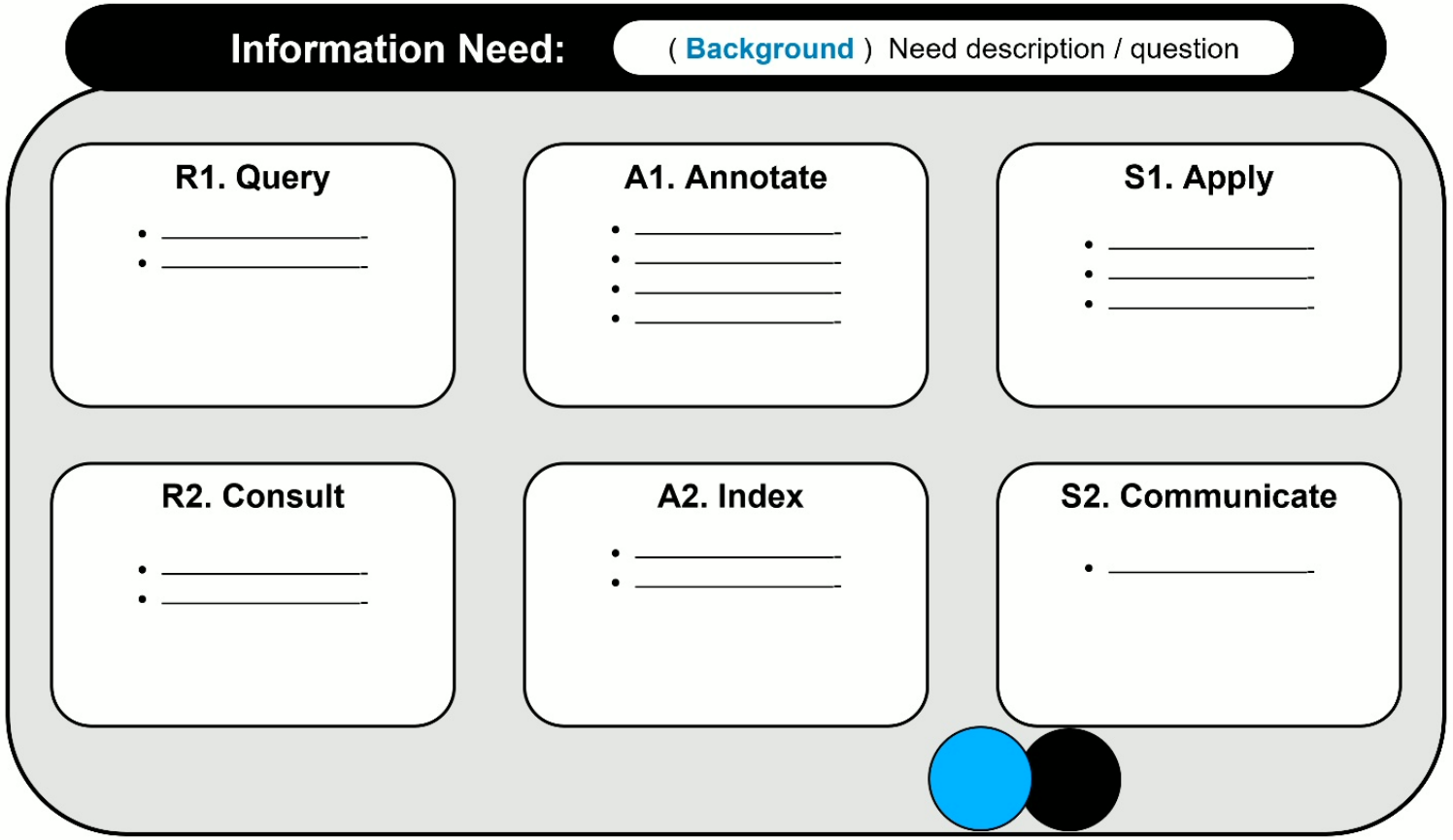
To characterize how sources are used when working with information, we adapt a framework commonly used for ***sensemaking*** tasks with three components.

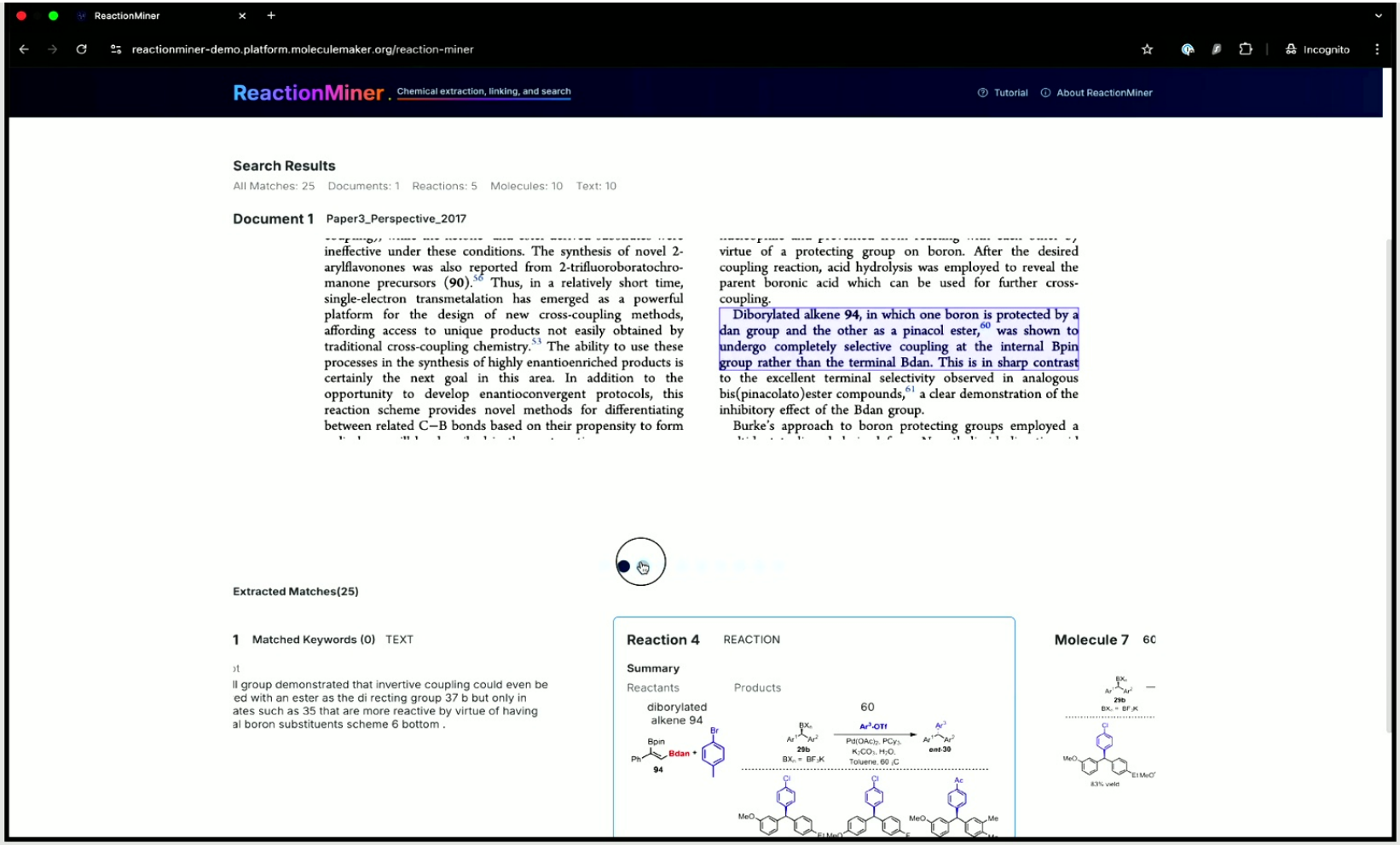
1. **Retrieval:** Locating relevant sources **(search)**, and finding information in sources **(consulting)**

2. **Analysis:** Organizing sources **(indexing)**, and adding information to sources **(annotation)**

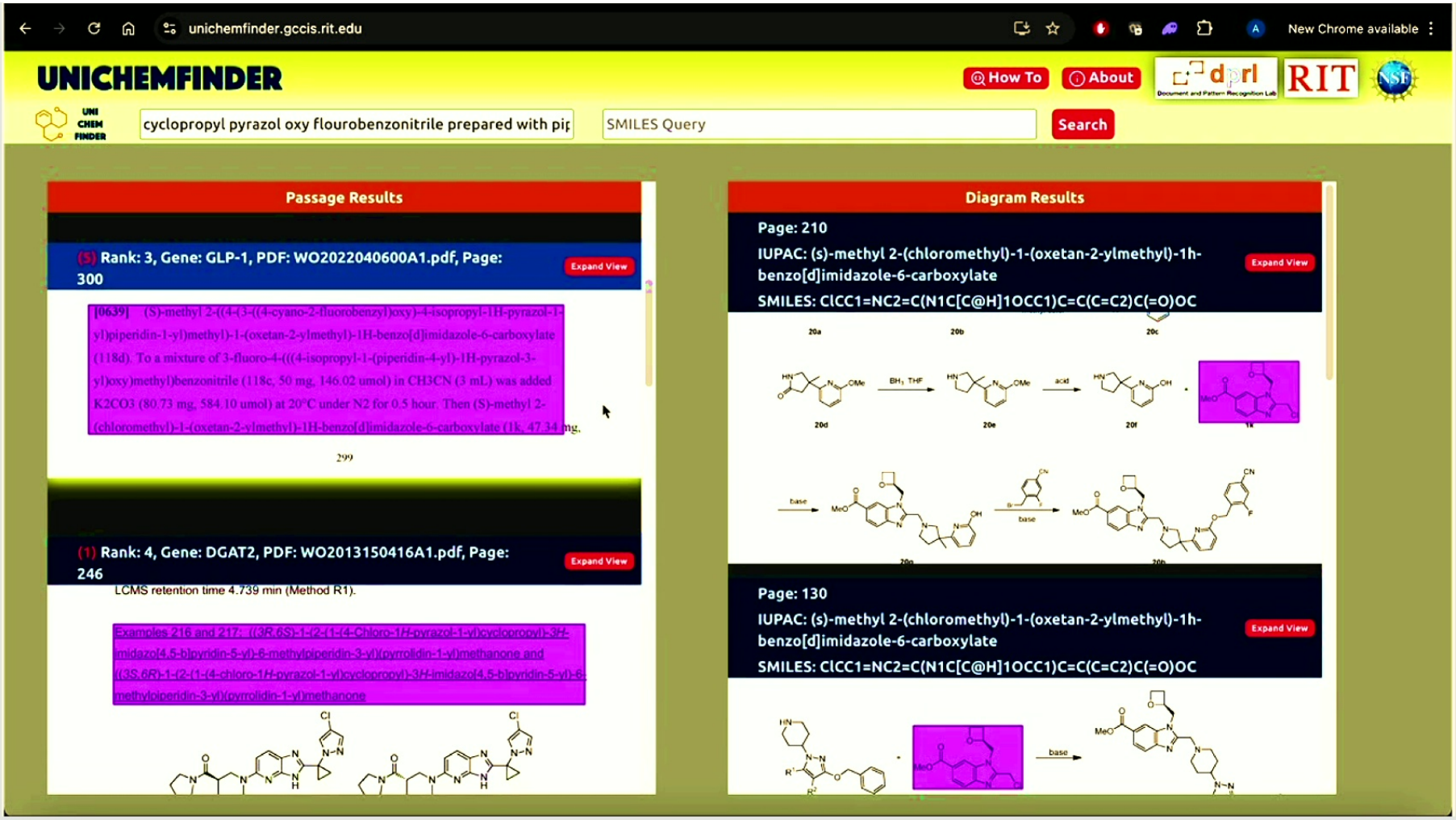
3. **Synthesis:** Using available information from sources **(application),** and creating *new* sources **(synthesis)**

Zanibbi, R., Mansouri, B., and Agarwal, A. (2025) Mathematical Information Retrieval: Search and Question Answering. Foundations and Trends in Information Retrieval: Vol 19, No. 1-2, pp. 1-190.

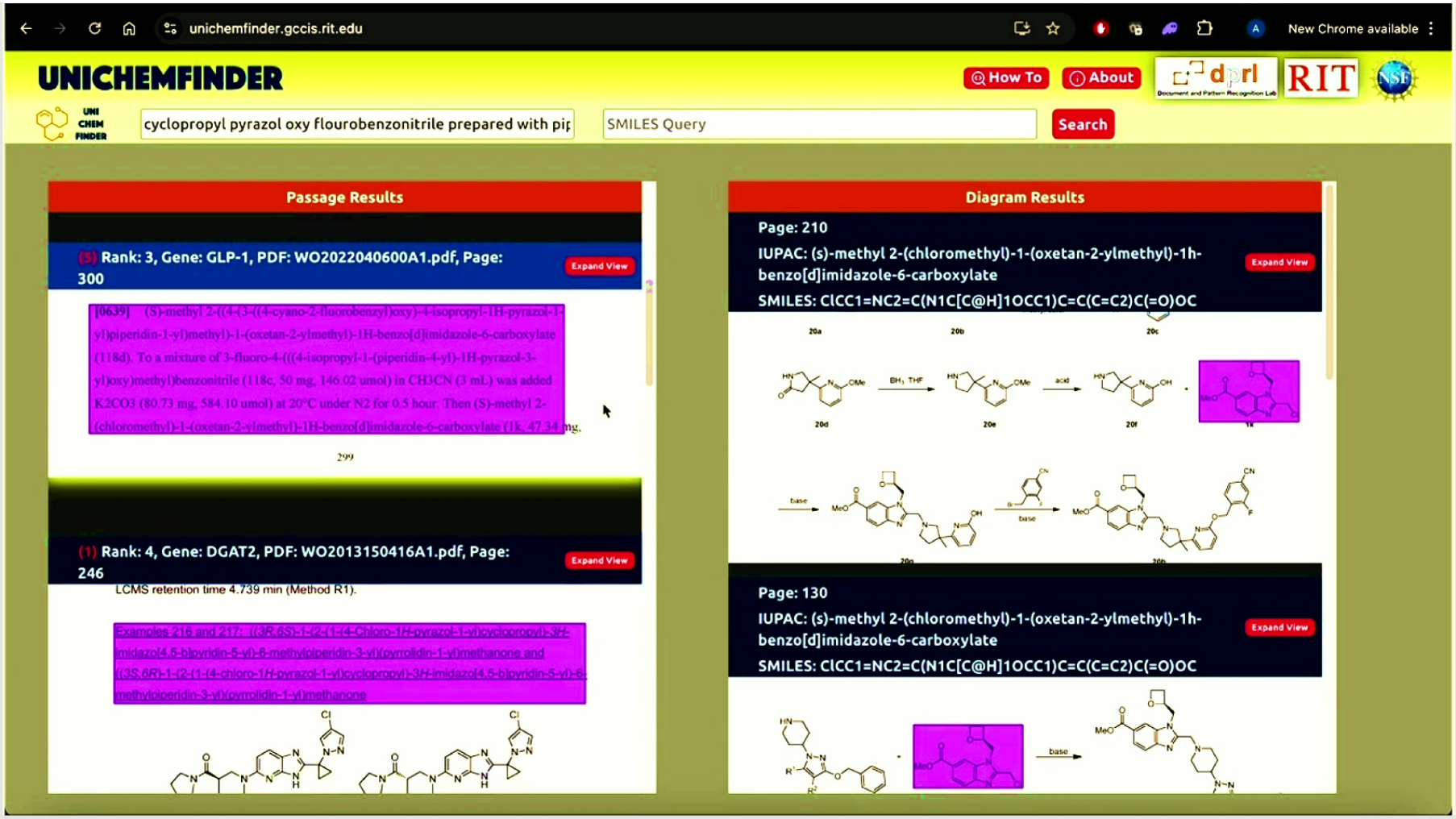

Information Need:
(Background) Need description / question
R1. Query
A1. Annotate
S1. Apply
R2. Consult
A2. Index
S2. Communicate



Shah, A.K., Dey, A., Luo, L., Amador, B., Philippy, P., Zhong, M. Ouyang, S., Friday, D.M., Bianchi, D., Jackson, N., Zanibbi, R., and Han, J. (2025) Multimodal Search in Chemical Documents and Reactions. SIGIR 2025, Padua, Italy (to appear).



Dey, A., Stanley, N., and Zanibbi, R. ***Targeted Multi-Modal Passage Search for Molecules and their Synthesis Pathways***. SIGIR 2025, Padua, Italy (to appear).



Dey, A., Stanley, N., and Zanibbi, R. ***Targeted Multi-Modal Passage Search for Molecules and their Synthesis Pathways***. SIGIR 2025, Padua, Italy (to appear).

Where to Next?

A Sketch — Comments and Suggestions Welcome!

Retrieval

- Improved graphics search (e.g., math formulas), **graphics + text** search, and **graphics ↔ text** search
- Search and navigation of **full models,** even across domains (e.g., for math; in progress w. Nickvash Kani, UIUC)
- Application to **Retrieval Augmented Generation (RAG)** for LLMs
- Conversational search, with interactive query intent clarification and refinement (per Star Trek)

Where to Next?

A Sketch — Comments and Suggestions Welcome!

Synthesis

- **Online collection** for sharing/searching entity card collections (e.g., for MathDeck) or other small information ‘nugget’ sources
  - Q. Particularly useful in smaller community settings, e.g., classrooms and research labs?
- Improved **multi-modal search result pages** and graphics + text **navigation**

Zanibbi, R., Mansouri, B., and Agarwal, A. (2025) Mathematical Information Retrieval: Search and Question Answering. Foundations and Trends in Information Retrieval: Vol 19, No. 1-2, pp. 1-190.

Acknowledgements

Thank you for your attention.

dprl@RIT: Ayush Kumar Shah, Abhisek Dey, Bryan Manrique Amador, Patrick Philippy, Behrooz Mansouri, Wei Zhong, Kenny Davila, Matt Langsenkamp, Yancarlos Diaz, Robin Avenoso, Gavin Nishizawa, Ming Creekmore, Brandon Kirinchich, ***Lee Giles & CiteSeerX group (Penn State),*** and ***Douglas W. Oard (UMaryland, College Park)***

MMLI@UIUC: Jiawei Han, Leo Luo, Blake Ocampo, Scott Denmark, Ming Zhong, Siru Ouyang, Heng Ji, Nick Jackson, David Friday, Cindy Chan, and Huimin Zhao
National Center for Supercomputing Applications (NCSA):
Matt Berry, Kate Arneson, David Bianchi, Bingji Guo, and Sara Lambert

UWaterloo: Frank Tompa, Edward Lank, George Labahn, and Jimmy Lin