



NEW MEXICO TECH · NMCCoE · NSF EPSCoR

From Documents to Discovery

Advancing Research Collaboration in New Mexico with AI

A collaboration-discovery infrastructure for connecting researchers across New Mexico institutions

Principal Investigator: **Dr. Lorie Liebrock**

Team: **Raisa Islam · Noah Fry**

NMCCoE Contact: **Jessica Trujillo Torres**

AGENDA

What we'll cover



Introduction

The team, NMCCoE, & EPSCoR, and why this project exists



The IHLLM

Using and administering the in-house LLM



Two ways to participate

Run a local IHLLM or just contribute documents to the statewide instance



The Research-LLM

Querying and uploading to the statewide instance



Access & support

Getting lite-admin access and where to get help

Who we are & what funds this



The Team

NEW MEXICO TECH

Led by PI Dr. Lorie Liebrock with researchers Raisa and Noah, building the platform end-to-end.



NMCCoE

HOST & COORDINATOR

The NM Cybersecurity Center of Excellence hosts the statewide instance, supports onboarding, and provides technical support.



NSF EPSCoR

FUNDING PROGRAM

Supported by NSF grant #2435071 E-CORE RII: Research Infrastructure Optimization for New Mexico (RIO-NM) initiative and collaborating research partners.

Finding collaborators is challenging



“Who else is working on this?”

Expertise that should connect rarely does — especially in new, fast-moving research areas where networks haven't yet formed.

- Relevant work is scattered across many institutions and departments
- Expertise lives inside published papers, not in any central index
- Emerging fields have no established collaborator network to tap into
- Manual searching is slow, partial, and quickly out of date

THE SOLUTION

A Research-discovery LLM



A retrieval-augmented LLM that indexes published research and lets anyone find potential collaborators by topic and expertise — in plain language.



IHLLM — In-House LLM

Runs locally at each institution. Indexes that institution's own published papers and answers queries privately, on-premises.



NM-LLM — Statewide LLM

An aggregated instance hosted by NMCCoE. Combines contributions from participating institutions for cross-institutional discovery.

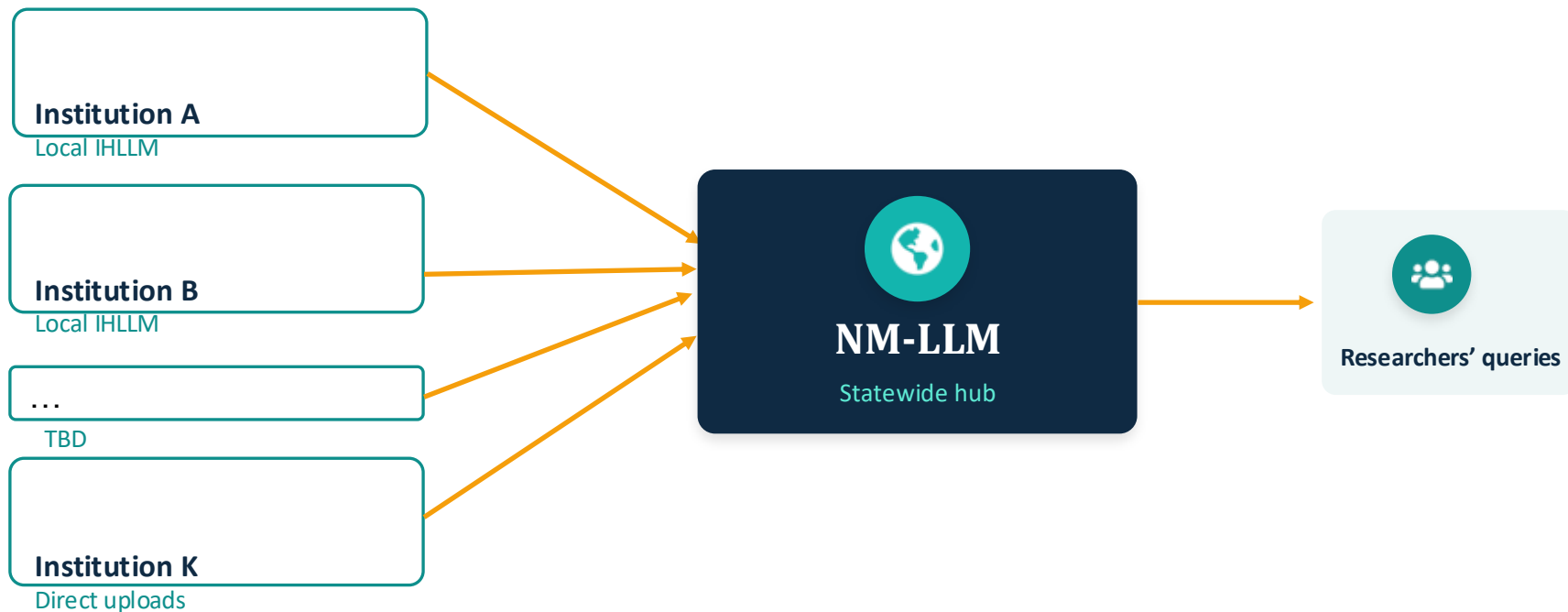
Architecture of Research LLM



Under the hood: FastAPI service · RAG pipeline · SentenceTransformer embeddings · Ollama-hosted LLM

HOW IT WORKS

From local to statewide hub



Set up your own Institutional IHLLM Instance



1

Get the code

Obtain the IHLLM codebase to deploy at your institution.

2

Set up for your org

Configure a new instance following the LLM Quickstart Guide.

3

Collect — search engines

Supplement the collection by pulling papers via search engines.

4

Collect — faculty assisted

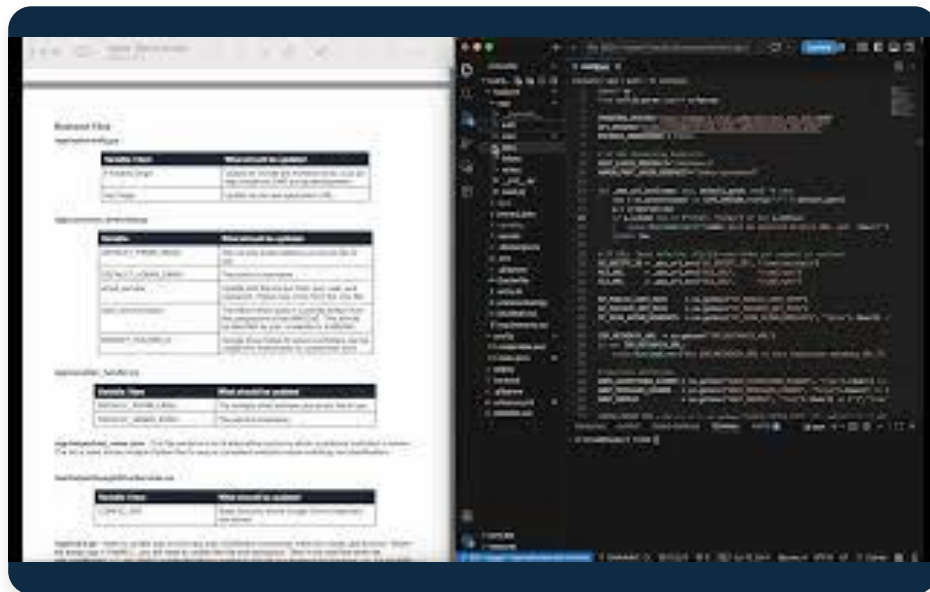
Gather published papers with help from your faculty.

5

Export & import

Export your institution's data and import it into the NM-LLM.

Quick Setup Guide for IHLLM

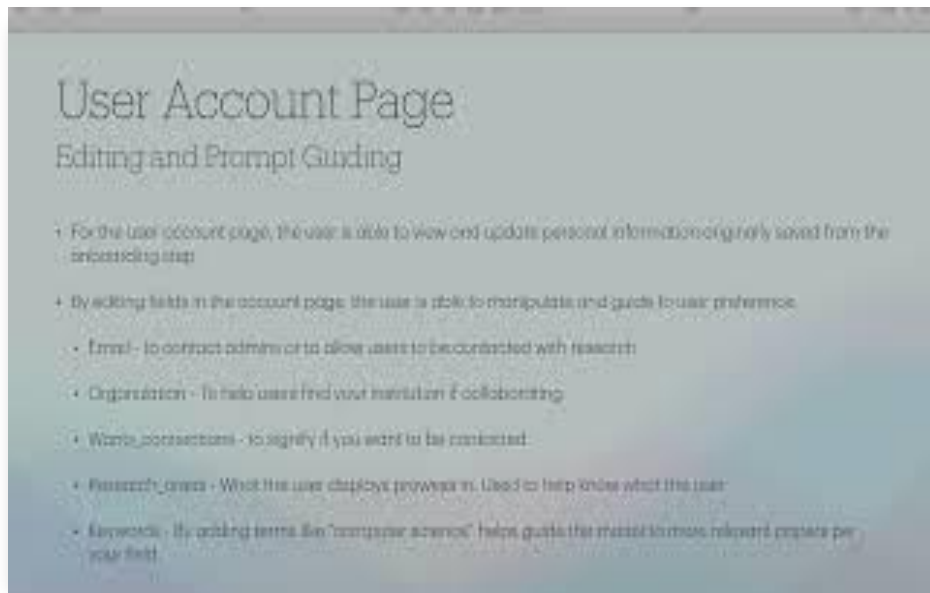


Content

- Frontend Updates
- Secret Files requirement
- Backend Updates

Reference: <https://www.youtube.com/watch?v=bHByMczJoO0>

Using the IHLLM

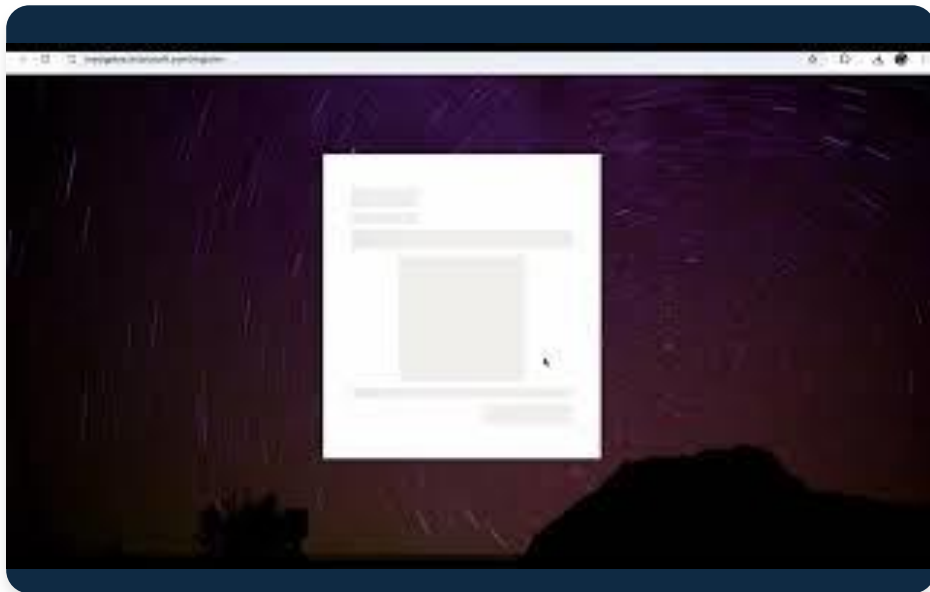


Content

- How to create account
- Ask questions in natural language over the indexed paper collection
 - Refine by topic to scope down to a specific research area
 - Get information about a specific paper
- User account information update

Reference: <https://www.youtube.com/watch?v=BspUpt0d0Ug>

Administering the IHLLM



Content

- Upload new document
- Pre-process documents
- Retrain model
- Download pacers (automated)
- Handle the database
- Set time for automated download retry
- Copy files to drive for statewide version
- List JSON tilte

Reference: <https://www.youtube.com/watch?v=mtINTf1KIM0>

IHLLM Faculty Paper Uploads



Researchers receive an email listing papers we couldn't access publicly — each gets a private Google Drive folder created just for them.



Upload

- Searchable PDFs of the listed papers
- Final published versions preferred
- Preprints (arXiv) OK if needed



Don't upload

- Drafts or incomplete manuscripts
- Sensitive or confidential or restricted access files
- Slides, posters, or datasets

How to upload



Open the Drive link in your email



Sign in with your institutional email



Drag & drop your PDFs



Wait until uploads finish

One PDF per paper · no duplicates · upload-only — please don't remove existing files or reshare · contact the admin with any issues.

Contribute Documents Directly to Statewide LLM



Collect searchable PDFs

Gather published research papers as text-searchable PDF files.



Bundle into a ZIP

Package the PDFs into a single ZIP archive for transfer.



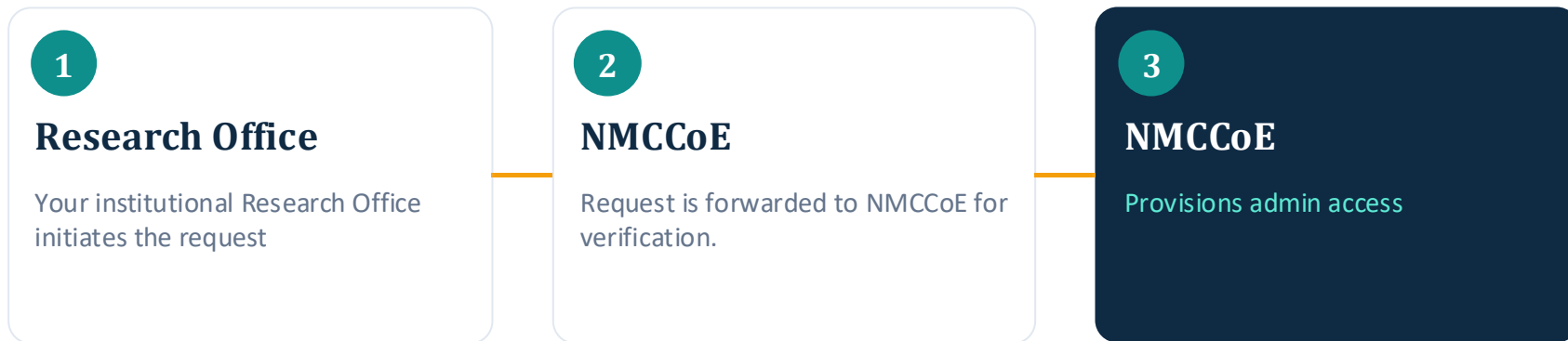
Import into the Research LLM

Upload the archive; the system ingests, embeds, and indexes it.



Getting admin access

Institutions that want to upload data into the NM-LLM need lite-admin access to the NM-LLM. The proposed request route:



Alternatively, requests may route through your Vice President for Research (VPR) or similar representative.

Where to Find



Git

<https://git.cybersecurity.nmt.edu/ecore-nm-rio-llm/llm.git>



Quick Setup Guide

<https://drive.google.com/file/d/1tDbPjjIX5vW7KIE9CWrAOESTZUW7dyD/view>



User Tutorial

IHLLM: <https://drive.google.com/file/d/1j1i1lqDrjI5R1njH7PFtiLKO3DqNymDD/view>
State-wide: <https://drive.google.com/file/d/189UmAGbFIT-3nhxIx3mpJGtmOi6CZvyb/view>



Admin User Guide

IHLLM: <https://drive.google.com/file/d/14FXzrWGw7HjggK4MgOyklepYbgWmiqbg>
State-wide: <https://drive.google.com/file/d/17hRVZQeAlGTiug1Jht-3FZ9z8n-OOpnR/view>



NMT Research LLM

<https://research-llm.local.cybersecurity.nmt.edu>



Youtube

<https://www.youtube.com/@NMResearchCollaboration>



SUPPORT

NMCCoE is here to help

- Technical support for setup, collection, and import into your institutional IHLLM
- Onboarding new institutions onto the NM-LLM
- Coordinating statewide research-collaboration infrastructure