Understand & Build Custom LLM Applications to Accelerate Governance Research

By: Umang Chaudhry, Mark Williams, Ethan I. Thorpe, J.B. Ruhl, and Dr. Mariah D. Caballero

VANDERBILT®
PRIVATE CLIMATE GOVERNANCE LAB

vaill
VANDERBILT AI LAW LAB

About the Authors

Umang Chaudhry M.S. is a Senior Data Scientist at the Vanderbilt Data Science Institute; **Mark Williams J.D.** is a Professor of Practice at Vanderbilt Law School and the Co-Director of the Vanderbilt AI Law Lab; **Ethan I. Thorpe** is the Climate Governance Fellow at Vanderbilt Law School's Private Climate Governance Lab; **J.B. Ruhl** is the David Daniels Allen Distinguished Chair in Law, Director of the Program on Law and Innovation, and Co-Director of the Energy, Environment, and Land Use Program at Vanderbilt Law School; and **Dr. Mariah D. Caballero** is an Assistant Professor of Environmental Studies at Wellesley College.

Abstract:
Artificial intelligence (AI) has rapidly become an undeniable part of the governance ecosystem. Of course, there are critical questions about the governance *of* AI, but for many practitioners, the bigger question is how to leverage AI to accelerate progress towards effective outcomes. Researchers at Vanderbilt University have developed a flexible, customizable, and easily maneuverable AI framework that can be applied in dozens of use cases. This piece will outline how we have used it, other potential applications, and a step-by-step guide to tailor the tool to your needs.

Table of Contents

How AI Works 4
Two Examples of AI in Practice...5
Similarities and Characteristics for Use 6
How Does this Tool Work? 6
Data Collection and Organization 6
Advanced Semantic Search with AI 7
Automated Analysis Using Retrieval Augmented Generation (RAG) 7
User-Friendly Interface...8
Building Your Own Tool 8
Define your Case 9
Replacing Underlying Data 9
Adjust Prompts and Analytical Questions 9
Customize the Interface 10
Validate Search and Retrieval Behavior 10
Conduct Real-World Testing 10
Requirements and Roles 11
Use Cases 11

How AI Works

Experts in every field are actively seeking ways to harness the power of AI to perform functions ranging from drafting emails to predicting energy supply and demand across thousands of square miles. At the time of this writing, ChatGPT has amassed over 400 million weekly users in less than two-and-a-half years. A recent survey of business professionals found that 78% of them use AI in at least one business function, up from 50% in 2022. AI is quickly being integrated into many facets of personal and professional life. It has exceptional potential to accelerate productivity, but it is important to understand how it functions to identify its limitations and where it may or may not fit your particular needs.[1]

At its core, AI involves training algorithms on vast amounts of data, enabling them to learn how to perform tasks such as recognizing patterns and making predictions. A specific, prominent application of AI involves Large Language Models (LLMs)—the type of AI driving much recent public attention—which generate text by predicting the most likely sequence of words.[2] This is the technology behind ChatGPT, Google Gemini, and Anthropic's Claude. It is also the kind of AI technology utilized by the tools described in this paperDeveloping and using an AI model involves two primary phases: **training** and **inference**

Training is the learning phase, where the AI is fed massive amounts of information. Using complex algorithms, it processes this data to identify patterns and learn connections, much like teaching someone to recognize different animals by showing them countless pictures. Through this process, the model effectively learns these patterns. Subsequently, **inference** is the application phase. Here, the trained model uses its learned patterns to analyze new, unseen data and make predictions, generate outputs, or perform the tasks it was designed for. This is

Umang Chaudhry M.S. is a Senior Data Scientist at the Vanderbilt Data Science Institute; Mark Williams J.D. is a Professor of Practice at Vanderbilt Law School and the Co-Director of the Vanderbilt AI Law Lab; Ethan I. Thorpe is the Climate Governance Fellow at Vanderbilt Law School's Private Climate Governance Lab; J.B. Ruhl is the David Daniels Allen Distinguished Chair in Law, Director of the Program on Law and Innovation, and Co-Director of the Energy, Environment, and Land Use Program at Vanderbilt Law School; and Dr. Mariah D. Caballero is an Assistant Professor of Environmental Studies at Wellesley College.

[1] You can see a list of some of AI's current limitations here: https://medium.com/@marklevisebook/understanding-the-limitations-of-ai-artificial-intelligence-a264c1e0b8ab.

[2] For a brief overview of generative AI and LLMs, see: https://ig.ft.com/generative-ai/ and https://www.3blue1brown.com/lessons/mini-llm. For more depth, see: https://poloclub.github.io/transformer-explainer/ for a visualization; https://www.youtube.com/watch?v=zjkBMFhNj_g&ab_channel=AndrejKarpathy for a video; or https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/ for a written explanation.

what happens when you ask an AI tool a question or ask it to perform a task; it looks through its memory for similar data and predicts an output based on those patterns.[3]

Two Examples of AI in Practice

In practice, AI allows users to interact with volumes of data that would otherwise take years to analyze, if possible at all. ChatGPT 3.5 was trained on roughly 300 billion words-worth of text, for example. More niche models can operate with less and achieve similar results, but a well-trained AI model is designed to make working with millions or billions of data points accessible.

We have used AI to build two innovative tools that accelerate research into pressing topics. The first, a climate adaptation policy tool, allows users to query, analyze, synthesize, and compare local climate adaptation plans.[4] For example, a local official might ask how comparable cities handle extreme heat to figure out a new adaptation strategy; a researcher might summarize all the plans in a particular region to accelerate work on regional adaptation; or a resident might explore how their city deals with flood risk to determine whether they should consider moving inland. Cities and counties around the nation are preparing for climate change in very different ways to meet their diverse needs. However, the extent to which these plans identify and meet those needs is unclear. We developed this tool to accelerate research and city planning to understand and improve climate adaptation initiatives.

The second tracks the development of AI regulations across jurisdictions. States are rapidly adopting legislation to regulate AI, but these bills vary drastically. This patchwork of regulations can lead to inefficiencies. The tool allows researchers and regulators to align their goals and create a more unified system of AI governance even without federal mandates. Consider a policy analyst working for a tech industry association who needs to advise members operating across multiple states on compliance obligations. Currently, they would have to manually search for dozens of state laws, track down relevant bills, and compare definitions and

---

[3] New developments of reasoning models include additional steps where the model breaks problems down into smaller steps and "checks its work." Read more about reasoning models here: https://techcommunity.microsoft.com/blog/azuredevcommunityblog/how-reasoning-models-are-transforming-logical-ai-thinking/4373194.

[4] These documents tend to be anywhere from 50 to 250 pages long, meaning that reading the ~100 of them that are currently uploaded would take weeks. The tool has all their summaries in its memory, meaning that when you ask it a question about a given plan or set of plans, it can read through the others to draw connections. Even this is not nearly enough data to fully train an AI model, so the underlying text generation comes from other models (ChatGPT or Anthropic's Claude). The model uses a context document to provide instructions for how it should respond (e.g. "only use information from the plans" and "cite the source of information").

requirements across dense legislative documents. With this tool, that analyst could simply search for a component of compliance, automatically uncover the right bills, and compare definitions without having to read through hundreds of pages of legislation.

Similarities and Characteristics for Use

On their face, these tools are very different. One handles report-style documents, the other legislation; one is about climate, the other tech. However, both of them use the same underlying code to generate responses. What similarities exist and how can those help guide users in applying this code?

1. Time Sensitive Issues: Climate change and AI governance are both rapidly developing, time sensitive, and confusing. Complexity and urgency mean that governance may not be well informed. These tools help simplify information sharing to respond more effectively.
2. Governance Patchwork: Regulations and standards, understandably, vary widely across the country. This can lead to inefficiencies in collective action. The tools adapt to this variety by allowing users to quickly compare across the landscape.
3. Access Problems: The underlying documents are often long, tedious, and difficult to find in one place. Despite this, seeing and understanding the variety can improve responses. It is much easier to ask a question to multiple documents than read through and analyze each one individually.

Most problems that meet these criteria can benefit from this style of AI integration. We have compiled a set of possible use cases on Page 11.

How Does this Tool Work?

These tools work like specialized search engines that help you navigate through complex fields without technical expertise. Here's an overview of how it works:

Data Collection and Organization

The tool gathers information from a defined source (uploaded PDF documents for the climate adaptation tool and the Legiscan API for the AI regulation tool) and organizes them in a **database**. This centralizes information that would otherwise be scattered across different websites and repositories.

Advanced Semantic Search with AI

The heart of this AI tool lies in its ability to understand *meaning*, not just keywords. Traditional keyword searches can miss relevant results if the exact phrase doesn't appear in a document. By contrast, this tool uses a powerful AI technique called **vector embeddings** to overcome that limitation.

1. **What are Embeddings?** When a user enters a question or phrase—like "regulations on AI in healthcare"—the tool transforms that input into a *vector*, which is a series of numbers that represent the meaning of the phrase in a high-dimensional mathematical space (e.g. the vectors for "book" and "reading" will have closer values than "book" and "running"). The same transformation is applied to documents like bills or plans.[5]

These vectors are designed so that content with similar meanings (**semantically similar content)** is located "closer" in this space. So even if the document says "clinical decision support tools" instead of "healthcare AI," the model can still identify it as relevant. This process is known as **semantic search** and allows users to uncover insights across complex, jargon-heavy documents—even if they don't know the exact terms to use.

Automated Analysis Using Retrieval Augmented Generation (RAG)

Unlike traditional Natural Language Processing (NLP) workflows that rely on hand-crafted rules or statistical models to extract meaning, this tool takes a modern approach centered around **Large Language Models (LLMs)** and **Retrieval-Augmented Generation (RAG)**.

1. **Preprocessing and Structuring**

Before any semantic analysis begins, the tool performs **basic data wrangling and cleanup**—such as converting PDFs to text, removing formatting artifacts, and standardizing metadata fields like jurisdiction or publication date.6 These steps ensure the content is clean and consistent enough to be queried effectively.

However, this tool doesn't use the usual NLP pipeline with steps (like named entity recognition or syntactic parsing).7 Instead of being explicitly told how to analyze the data, the LLM figures out the best ways to organize and search based on the massive amount of text it has already learned from.

---

[5] See a brief explanation of embeddings here: https://www.cloudflare.com/en-gb/learning/ai/what-are-embeddings/.

[6] Metadata is "data about data." Read more about what it is and why it is important here: https://www.ibm.com/think/topics/metadata.

[7] Learn more about NLP and popular pipelines here: https://www.oracle.com/ba/artificial-intelligence/what-is-natural-language-processing/.

2. **RAG Analysis**[8]

    - **Retrieval**: When a user submits a question, the system retrieves relevant document snippets using *vector embeddings*—a way of comparing meaning rather than keywords.
    - **Augmentation**: These snippets are passed into a LLM to provide context for its response.
    - **Generation**: The LLM generates an answer grounded in the retrieved text, ensuring factual accuracy and contextual awareness.

3. **With this tool you can:**

    - **Summarize long documents** in plain language
    - **Compare similarities** between plans or bills without hard-coding any topic models
    - **Generate responses to queries** even when the phrasing differs from the source documents

User-Friendly Interface

Users interact with the tool online, like you would any other website that allows you to search. The benefits of this include:

- Easy searching, even if you have no coding experience
- Customizable filters
- Side-by-side comparisons
- Plain-language summaries of complex text

The underlying technology is **modular**, meaning the same framework could be adapted for all sorts of applications simply by changing the data sources, any data pre-processing and search parameters.

Building Your Own Tool

This framework was originally built to support climate adaptation policy analysis. However, its modular architecture allows it to be repurposed for a wide range of domains—

[8] You can read more about how RAG works here: https://www.ibm.com/think/topics/retrieval-augmented-generation.

including but not limited to data privacy, supply chain risk, AI regulation, education policy, and environmental justice.

While the core components are reusable, **adaptation does require a meaningful investment of time and light coding expertise**. This section outlines the key steps involved in tailoring the tool to a new domain.[9]

Define your Case

Begin by answering the following questions:

- **What kinds of documents** will you analyze? (e.g., legislative texts, research reports, standards)
- **Who are your intended users?** (e.g., analysts, community groups, agency staff)
- **What will they want to do?** (e.g., compare documents, ask questions, visualize data)

Your answers will shape every part of the adaptation process.

Replacing Underlying Data

To repurpose the existing tool:

- **Swap in your own documents**: Replace the existing PDFs, metadata tables, and geospatial layers with files relevant to your domain.
- **Update metadata**: Modify supporting files such as CSVs or pickled dataframes to match your document collection.
- **Adjust or remove maps**: If your topic is not geographically focused, you may simplify the interface by removing geospatial features entirely.

Estimated effort: Moderate — will require someone comfortable with file formats and basic data preprocessing.

Adjust Prompts and Analytical Questions

The AI's usefulness is heavily driven by the instructions it receives. You will need to:

- **Update the prompt templates** to reflect your new subject matter.
- **Revise the structured questions** used in report generation and plan comparison tools.[10]

---

[9] It can be very useful to put the source code into a popular public model like Claude, Gemini, or ChatGPT. It will be able to answer questions you might have throughout this process and identify file names or places that you should alter to make specific changes.
[10] For tips on prompt engineering, see: https://arxiv.org/abs/2302.11382.

- **Test and refine** your questions to ensure outputs are accurate and helpful.

These instructions are stored in editable markdown files and can be tailored without changing the core code.

Estimated effort: Moderate — best done in collaboration with a subject-matter expert.

Customize the Interface[11]

This platform uses Streamlit, a Python-based UI framework designed for ease of use. You should:

- **Update section names** and instructional text across tabs.
- **Remove tabs or features** that don't apply to your use case.
- **Add new input options** if necessary to support your users.

Streamlit makes it possible to modify these elements with basic Python knowledge.

Estimated effort: Low to moderate — simple interface changes are easy; larger redesigns require developer time.

Validate Search and Retrieval Behavior

At its core, the tool uses **vector embeddings and Retrieval-Augmented Generation (RAG)** to retrieve and synthesize information from documents.

- **Confirm that your content is loading properly** into the vector database.
- **Test retrieval accuracy** using example queries from your use case.
- **Adjust pre-processing** (e.g., text chunking) if needed to improve response quality.

Estimated effort: Low — unless your documents require complex handling or customization.

Conduct Real-World Testing

Once adapted:

- Upload sample documents.
- Ask real-world questions users might pose.
- Evaluate whether answers are relevant, accurate, and grounded in the text.

This testing phase will often surface usability issues, edge cases, or prompt wording that need refinement.

---

[11] See note 6.

Estimated effort: Moderate — requires structured user testing and feedback.

Requirements and Roles

To successfully adapt the framework, most teams will need:

Domain expert	To define key questions and validate outputs
Technical lead	To implement code and data modifications
Test users	To provide feedback on usability and relevance

Most adaptations can be accomplished in **1–2 weeks of focused development time**, assuming clean input data and modest scope. Projects involving large-scale data ingestion, heavy customization, or new features will take longer.

Use Cases

1. **Environmental conservation**: Use land use data to identify where conservation is most important. Compare with conservation regulations to highlight discrepancies between conservation value and protection (this is a huge problem in the SE where we have huge biodiversity but almost none of the protected land).
2. **Critical supply chain nodes**: Map suppliers of critical goods, highlight the climate-related risks they might face, and predict how shocks could be avoided for the consumer.
3. **Data centers and clean energy**: Upload geographic data of data centers and energy supplier documents to show how data centers are likely to be powered, and any discrepancies in potential local and regional environmental impacts.
4. **Data privacy and AI**: Compare data privacy laws across states and localities to understand where gaps might be.
5. **Education policy**: Upload state or local educational standards and socioeconomic data to compare with academic outcomes. This could help local schools to adopt more effective pedagogies by learning from more successful, socioeconomically similar places.
6. **Cross-Border Compliance Monitoring**: The tool could analyze jurisdictional differences in regulatory approaches to create comprehensive compliance matrices for organizations operating across multiple states. As the tracker shows, states are taking dramatically different approaches to categories like risk assessments, labeling requirements, and third-party reviews. This application would help multi-state enterprises identify the most

stringent requirements they need to meet while highlighting state-specific obligations that might require targeted compliance efforts.

7. **AI governance taxonomy standardization**: Compare how different jurisdictions define and categorize AI systems (e.g., foundation models vs. automated decision systems vs. generative AI) and the varying obligations applied to each. The tracker reveals significant inconsistency in how states scope AI governance requirements, creating definitional challenges for companies trying to determine which regulations apply to their products. By identifying emerging classification patterns, this application would support efforts to develop more consistent taxonomies for AI governance that could benefit both industry and regulators.

8. **Individual rights protection landscape**: track the evolution of AI-related individual rights provisions (like opt-out requirements and non-discrimination protections) across jurisdictions to identify emerging standards and gaps. As shown in the tracker, some states are creating robust individual rights frameworks while others focus primarily on organizational obligations. This application would help consumer advocates identify model provisions worth promoting in states with minimal protections while helping organizations anticipate the likely direction of future regulation.

9. **AI governance transparency evolution**: Analyze how transparency requirements (like labeling, notification, and explanation obligations) are evolving across different states and system types. The tracker shows varying approaches to disclosure requirements, with some states focusing on up-front notification and others emphasizing post-decision explanations. This application would help identify best practices for meaningful AI transparency that balances consumer protection with practical implementation considerations.

Though this tool can add significant value, it remains an imperfect technology. Users should define key terms like "success" and "importance" in consultation with affected stakeholders. They should also be careful to identify research plans that comply with other regulations including data security because information may be routed to public servers. Human validation is still critical to ensuring accurate information. It can accelerate research but is not a substitute for normal review processes.

VANDERBILT®
PRIVATE CLIMATE GOVERNANCE LAB

vaill
VANDERBILT AI LAW LAB