

Web Scraping and Large Language Models for Geothermal Operations: Mapping Water Availability and Ordinances

Slater Podgorny, Whitney Trainor-Guitton, Faith Martinez Smith, & Pavlo Pinchuk

15013 Denver West Parkway, Golden, CO 80401

Slater.Podgorny@nlr.gov

Keywords: large language models (LLMs), enhanced geothermal systems, water availability, groundwater guidelines

ABSTRACT

Estimating the potential national build-out of emerging geothermal technologies such as enhanced geothermal systems will require understanding state and local regulatory policies and permitting around water use. Water use for enhanced geothermal systems is an outstanding issue, especially for the arid western United States where elevated subsurface temperatures offer promising geothermal resources. We developed a large language model (LLM) application to gather water rights, availability, and permitting data in 98 water districts in Texas. We present the challenges of validating the web-scraped and LLM-retrieved information, and we map this information with the collocated geothermal potential. In Texas, our results show that 58 districts do not have extraction limits for groundwater wells, but 63 districts require meters on the wells to track groundwater pumping. The higher geothermal potential of Texas runs along the Gulf Coast where qualitatively, half of the districts do not have annual extraction limitations nor costs for exporting the water out of the conservation district.

1. INTRODUCTION

Energy deployment in the United States is expected to advance rapidly in the coming years in response to the demands of data centers and other large load deployments. Achieving this growth will require a mix of generation resources and a comprehensive understanding of local siting ordinances as they have been shown to have a significant impact on nation-scale build-out estimates for wind and solar (Lopez et al. 2025). Manually collecting policy information on these ordinances and regulatory frameworks can be time-consuming and become out of date quickly (Levine et al. 2023; National Laboratory of the Rockies 2026). Previous efforts to build and maintain datasets related to energy policy have required considerable amount of manual labor. In an effort to improve the pace of data extraction in a way that is cost effective and efficient, the National Laboratory of the Rockies (NLR) developed the Energy Language Model (ELM) to automate this process using novel machine learning technologies to generate structure data from policy related to setbacks for wind infrastructure (Buster et al. 2024).

Based on previous U.S. Department of Energy (DOE) analysis, geothermal development has the potential to increase up to an estimated 60 gigawatts-electric of electricity generating capacity by 2050 (GeoVision 2019). Alternative scenario projections of growth for enhanced geothermal systems (EGS) could include 38 gigawatts-electricity of geothermal capacity by 2035 and 90 gigawatts-electricity by 2050 (Augustine et al. 2023). Higher EGS suitability is understood to be located west of the Mississippi river in the United States (Figure 1), but water resources are limited and, in most instances, are already allocated to other beneficial uses such as potable water end users (i.e., individual customers for drinking water), industrial, agricultural, etc. As such, policy and regulatory requirements for water use and availability will become more important for geothermal development (Fuchs 2026). Seligman (2025) reviewed previous published estimates for geothermal and next-generation water use, including Schroeder et al. (2014), which recognize the lack of public water-use data for geothermal operations. This is especially acute for enhanced geothermal systems (EGS), where there are only two active projects: the DOE-funded Frontier Observatory for Research in Geothermal Energy (FORGE) has reported 10% water loss (“Utah FORGE Circulation Tests” 2024)”, and Fervo has an aspirational water loss of 1% (Fervo Energy 2025). In the absence of numerous EGS operating fields and data, a possible analog for water use is unconventional oil and gas, where both injection and production also are performed. It has been estimated that fracturing an oil and gas well requires anywhere from half a million to 11 million gallons of water per well (Cath et al. 2021).

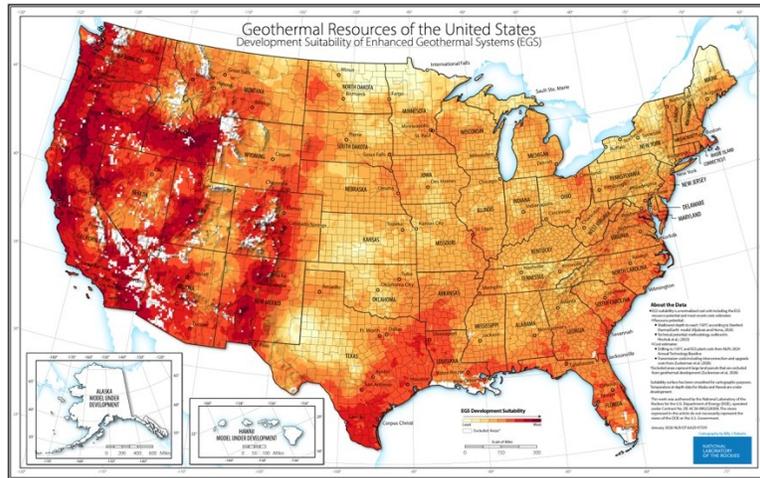


Figure 1: EGS suitability including estimates of the subsurface temperatures, costs and technical exclusions (cultural, ecological and structural) (Roberts 2025).

Based on increased developer interest in EGS as well as clear guidelines for surface water use in Texas via prior appropriation and “rule of capture” for groundwater use, straightforward water rights and regulations, the NLR team selected Texas as a starting point to further explore the potential impacts of water rights and availability for EGS development. Existing oil and gas companies have indicated interest in utilizing existing leases for EGS development, however, there are concerns associated with water availability. Because of this interest and understanding that most surface water use in Texas has been previously appropriated, the NLR team’s objective was to retrieve information that details groundwater well permitting in Texas using an LLM to extract unstructured data from text documents. Our approach included the use of retrieval-augmented generation (RAG), which allows for simultaneously considering multiple sources of information. This has the flexibility to answer more general questions using the most current information and to consider regulatory data that can come from several agencies and offices; however, this can come at the cost of accuracy, thus requiring manual verification.

2. METHODS

Building on work conducted by Buster et al. (2024) to extract wind setback restrictions from local ordinance documents, we used a similar approach designed to determine existing groundwater regulations and permitting requirements.

2.1 Document Processing

Our workflow, shown in Figure 2, begins with an automated document collection effort designed to retrieve relevant documents. In previous work, the LLM extraction process assumed a single source of information where all relevant ordinance information is in one document. We found that the data that we were targeting were typically spread across multiple documents. Using NLR’s ELM, five questions were submitted to an automated search engine targeting three key document types for groundwater conservation districts: groundwater management plans, rules and regulations, and permit application documents. The retrieved documents were then evaluated through a series of LLM prompts to determine whether they contain relevant information and refer to the location of interest. In previous workflows, a single document was evaluated, and only the relevant portions of text—those that referred to wind ordinances—were stored for further analysis.

To accommodate the consideration of multiple sources, we designed a small RAG pipeline with a vector database. RAG refers to the process of providing context to LLMs such that the generated responses use domain-specific information, improving the likelihood of a relevant and accurate response (Zhang et al. 2024). Creating the vector database requires breaking the documents into smaller portions that are easier for the model to process. We used a simple, fixed-size selection strategy where the document is broken into portions of approximately 500 tokens,¹ and then embeddings are generated for each section. Text embeddings are numerical representations of text that capture semantic meaning, and a vector database stores the embeddings for all the processed documents, effectively creating a knowledge base of relevant text to query when extracting information. Using a vector database allows us to simultaneously consider information from multiple documents.

¹ <https://help.openai.com/en/articles/4936856-what-are-tokens-and-how-to-count-them>

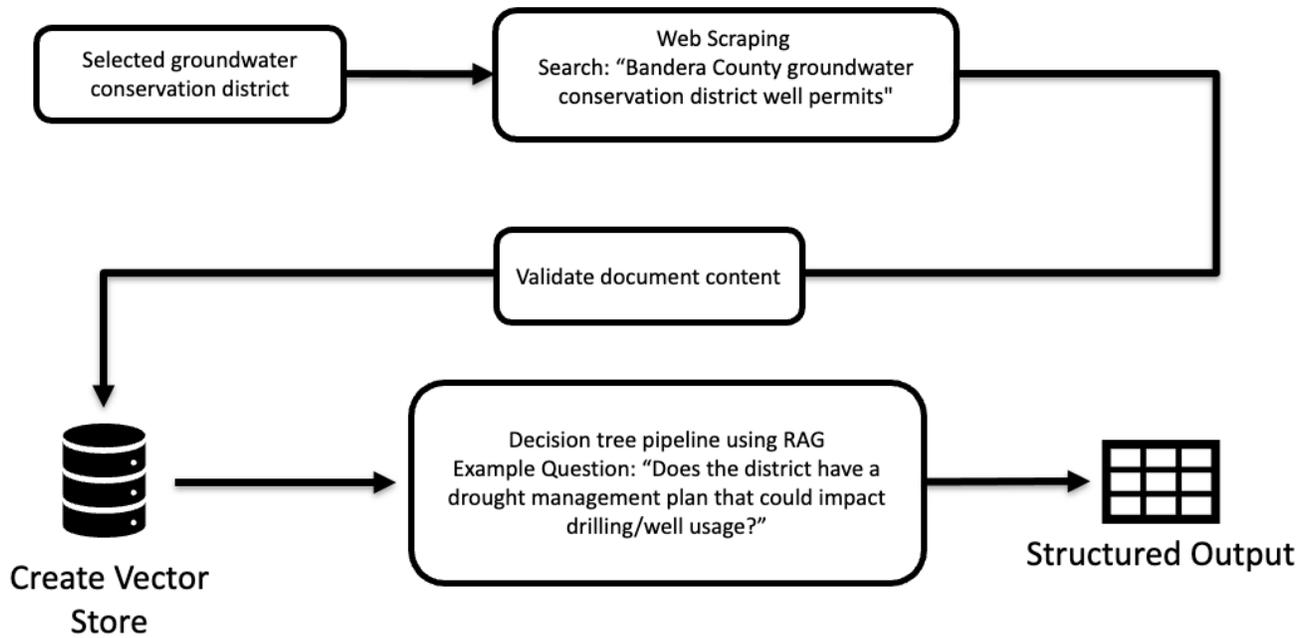


Figure 2: Water rights extraction workflow.

2.2 Water Guideline Extraction

We identified 13 key pieces of information to target for this analysis based on data that would inform ease of water access:

- General permit requirements
- Groundwater extraction limits
- Drilling window for wells
- Production costs
- Export Costs
- Well spacing requirements
- Metering device requirements
- District- and well-specific drought management plans
- Well plugging requirements
- Production reporting
- Existence of setback requirements
- Redrilling restrictions

We initiated the extraction process by first querying the vector database to retrieve feature/question-specific information. For example, when targeting information related to extraction limits in Bandera County, we queried the vector database with “Are there groundwater extraction limits in the Bandera County groundwater conservation district?” and returned the most relevant text. This method relies on cosine similarity, a measure of how similar two pieces of text are (Porter 2024). Cosine similarity scores fall between -1 and 1, with higher scores indicating greater similarity. We compared the query vector to each vector in the database, and we sorted the similarity scores to select only the 50 most relevant portions of text to be considered in the decision tree logic. In this way, the context that is provided to the LLM considers information from multiple documents but remains topic and location specific.

Decision tree execution follows processes developed by Buster et al. (2024): The retrieved text provides context for the LLM, and the extraction of a specific feature is broken down into smaller steps and transitions that create a “preprogrammed multi-prompt LLM conversation based on subject matter expertise” (Buster et al. 2024). This process is displayed in Figure 3. The decision tree prompt asks the LLM to respond with “Yes” or “No” based on the information provided, and depending on the answer, the next prompt is sent to the LLM. Each prompt serves as a logical check and asks a question that guides the model toward the desired information. The final prompt in the decision tree instructs the LLM to respond based on the entire conversation and to provide a structured output in JSON format.

User: Does the following text mention how far a new well must be from another well? Focus only on spacing requirements between wells and ignore spacing that is specific to other features such as property lines or septic systems. Begin your response with either ‘Yes’ or ‘No’ and explain your answer.

Assistant: Yes. The text explicitly mentions spacing requirements between wells. Specifically, it states that:

- Exempt wells must be spaced at least 100 feet from existing wells (Rule 3.3, section j).
- To reduce interference between wells during pumping, wells producing more than 25,000 gallons per day must be at least 200 feet from any existing permitted wells in the same producing intervals.
- Wells permitted to produce more than 72,000 gallons per day must be spaced a minimum of 1,000 feet from existing wells completed in the same producing formation.
- Variances to these spacing requirements may be granted by the Board if sufficient data shows insignificant impact on existing wells.

These provisions clearly set minimum distances between wells based on their production capacity and the producing formation, focusing on well-to-well spacing rather than setbacks from other features.

User: What is the spacing limit mentioned in the text? Include the units associated with the limit (example: feet or yards).

Assistant: The spacing limits mentioned in the text are as follows:

- Exempt wells must be spaced a minimum of 100 feet from existing wells.
- Wells producing more than 25,000 gallons per day must be spaced a minimum of 200 feet from any existing permitted wells in the same producing intervals.
- Wells permitted to produce more than 72,000 gallons per day must be spaced a minimum of 1,000 feet from existing wells completed in the same producing formation.

User: Do the spacing requirements mentioned apply specifically to the required distance between two water wells? Begin your response with either 'Yes' or 'No' and explain your answer.

Assistant: Yes. The spacing requirements mentioned specifically apply to the required distance between two water wells. The text refers to minimum distances that new wells must maintain from existing wells, such as exempt wells needing to be at least 100 feet from existing wells, and larger capacity wells requiring 200 feet or 1,000 feet spacing from other wells in the same producing intervals or formations. These distances are clearly about well-to-well spacing to reduce interference and protect groundwater resources.

User: Is the spacing limit dependent upon well characteristics such as depth or production capacity? If so, include that metric in the response (example: gallons per minute).

Assistant: Yes. The spacing limit is dependent upon well characteristics, specifically the production capacity and the producing formation. The text states:

- Wells producing more than 25,000 gallons per day must be spaced at least 200 feet from existing permitted wells in the same producing intervals.
- Wells permitted to produce more than 72,000 gallons per day (equivalent to 50 gallons per minute) must be spaced at least 1,000 feet from existing wells completed in the same producing formation.

Thus, both the production rate (in gallons per day or gallons per minute) and the producing formation influence the required spacing between wells.

User: Respond based on our entire conversation so far. Return your answer in JSON format (not markdown). Your JSON file must include exactly four keys. The keys are "spacing", "units", "qualifier", and "explanation".

Assistant:

```
{
  "spacing": [100, 200, 1000],
  "units": "feet",
  "qualifier": "production capacity in gallons per day and producing formation",
  "explanation": "The spacing requirements specify minimum distances between wells based on their production capacity and the producing formation. Exempt wells must be at least 100 feet from existing wells. Wells producing more than 25,000 gallons per day require 200 feet spacing, and wells producing more than 72,000 gallons per day (50 gallons per minute) require 1,000 feet spacing from wells in the same producing formation."
}
```

Figure 3: Example of Decision Tree Logic.

2.3 Decision Tree Development

Creating the decision trees requires subject matter expertise and multiple rounds of iteration to ensure that the output aligns with the expected values. For example, when extracting the required spacing between two wells, the model occasionally returned values that applied to the spacing requirements between wells and other features, such as property lines or septic systems (see results in Table 1). Our initial prompt left room for interpretation: “Does the following text mention restrictions related to well spacing? Such information typically dictates how far apart two wells must be and could prohibit an individual from drilling a well within a certain distance of another well.” As shown in Table 1, this prompt returned all values that related to well spacing, but it also returned a figure specifying that wells cannot be placed within 50 feet of a property line. Although this is helpful information, it is not what we were targeting.

Table 2 shows the source of this error and the difficulties of isolating the value of interest. The Bandera County River Authority and Groundwater Conservation District Rule 6.1j describes a restriction that lists a spacing value for both property lines and existing wells. The first prompt broadly asks for requirements related to “well spacing,” which allows the model to reasonably include both values in the response.

To resolve this issue, we designed a more specific prompt that adds instruction: “Does the following text mention how far apart a new well must be from another well? ... Focus only on spacing requirements between wells, and ignore spacing that is specific to other features, such as property lines or septic systems.” The resulting output captures all relevant values and excludes the property line requirement from the previous case.

Table 1: Results from (left) original well spacing prompt and (right) updated prompt specifying well-to-well distance.

	Original	Improved
Prompt	“Does the following text mention restrictions related to well spacing? Such information typically dictates how far apart two wells must be and could prohibit an individual from drilling a well within a certain distance of another well.”	“Does the following text mention how far apart a new well must be from another well? ... Focus only on spacing requirements between wells, and ignore spacing that is specific to other features, such as property lines or septic systems.”
Results	'spacing': [200, 1000, 100, 50], 'qualifier': ['production greater than 25,000 gallons per day in the same producing intervals',... 'exempt wells from existing wells', 'exempt wells from property line']	'spacing': [200, 1000, 100], 'units': 'feet', 'qualifier': ['wells producing more than 25,000 gallons per day', 'wells permitted to produce more than 72,000 gallons per day (50 gallons per minute)', 'exempt wells']

Table 2: Bandera County River Authority and groundwater conservation district rules.

Rule	Spacing Requirements
3.4p	To reduce the interference between wells during pumping, the following spacing is required between the wells: Wells that produce more than 25,000 gal / day must be a minimum of 200 ft. from any existing permitted wells in the same producing intervals. A well permitted to produce more than 72,000 gpd, (72,000 gpd = 50 gpm) must have a minimum of 1,000 feet from existing wells completed in the same producing formation. A variance by the Board may be made if the applicant submits sufficient acceptable data to show insignificant impact on existing wells.
6.1j	An exempt well must be spaced, at a minimum, 50 feet from the property line and 100 feet from existing wells . After a review of the application, the General Manager may grant a variance. A variance must be requested on forms provided by the District and the District may request additional information.

We present another example of when subject matter input is needed to verify the LLM results. Developing decision trees for the most accurate extraction requires input from subject matter experts (SMEs). SMEs can help inform what kind of data should be targeted, the language that should be used to extract that information, and the accuracy of the output. This dynamic was displayed when identifying extraction limits throughout Texas. The initial decision tree that was developed for this task returned a single district-wide limit. Although the value was accurate, an SME noted that there are often multiple limits within a district and that they can be specific to well permit type, application, or aquifer. An example is shown in Table 3, which is an excerpt from the Lost Pines Groundwater Conservation District Guide for Understanding Rules and Management Plan² showing six limits, each associated with a different aquifer. Based on this feedback,

² For additional information please see: <http://lostpineswater.org/191/District-Rules>

we added a node to the decision tree to determine if there are multiple extraction limits as well as instructions to include “multiple values based on permit, well or aquifer,” if applicable.

Table 3: Lost Pines Groundwater Conservation District extraction limits.

Aquifer	Maximum Annual Production per Surface Acre per Year (Acre-Feet)
Hooper	0.5
Simsboro	1.6
Calvert Bluff	0.5
Carrizo	0.8
Queen City	0.2
Sparta	0.3

3. TEXAS RESULTS

We now present some of the quantitative water district results. For reference, the computation time for the two stages is presented in Table 4. Document processing takes more compute time than value extraction as it involves two subprocesses: validating document content for relevance and generating embeddings for the document.

Table 4: Performance metric of extractions for Texas groundwater guidelines.

Extraction Step	Average Completion Time per District‡
Document processing§	4 minutes, 36 seconds
Value extraction	1 minute, 48 seconds
Total	6 minutes, 24 seconds

3.1 Existence of Metering Device Requirements and Annual Extraction Limits of Groundwater Wells

Water availability is limited in Texas, with additional stress on water resources in the western portion of the state. For instance, the volume of water use is higher than what is naturally recharged through precipitation and/or collection; however, Figure 4 shows that both metering devices (left) and extraction limits (right) exist throughout the state. White areas indicate the absence of a groundwater district (Texas Water Development Board 2019). Of the 98 total, 63 have a metering requirement, 30 do not, and no data was retrieved for 5 districts. Regarding extraction limits (Figure 4b), 58 districts have limits, and 35 do not. No information was found for 5 districts.

‡ There are 98 total groundwater conservation districts in Texas.

§ Document processing involves web scraping, document validation, and embedding.

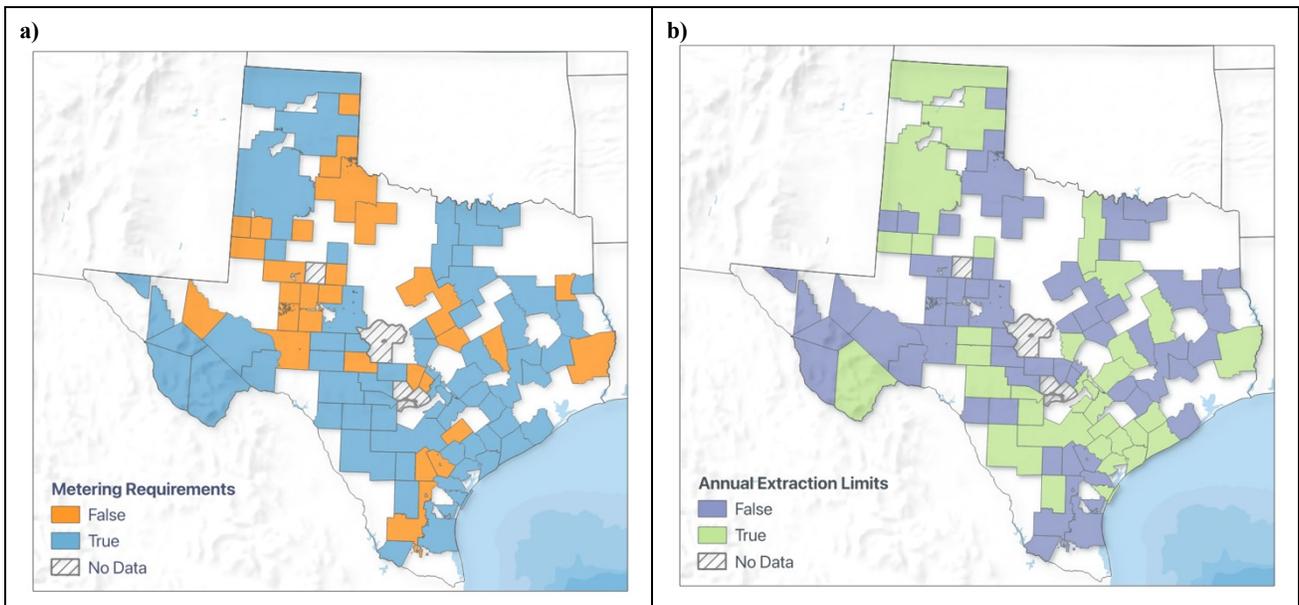


Figure 4: a) Metering device requirements, b) annual extraction limits. White areas do not have a water district. Defined boundaries within Figure 3a and Figure 3b are groundwater conservation district boundaries.

3.2 Annual Extraction Limits for Groundwater Wells

Of the 98 groundwater conservation districts, only 35 have some type of extraction limitation, with varying definitions and associated regulatory requirements. The first category defines the extraction by area or landownership acreage (e.g., 2 acre-feet per acre of land). The second category illustrates the total volume of groundwater that can be extracted over a year (e.g., 100 acre-feet per year). The last category is illustrated by permit where the limits are specific to each well application. These are shown in Figure 5.

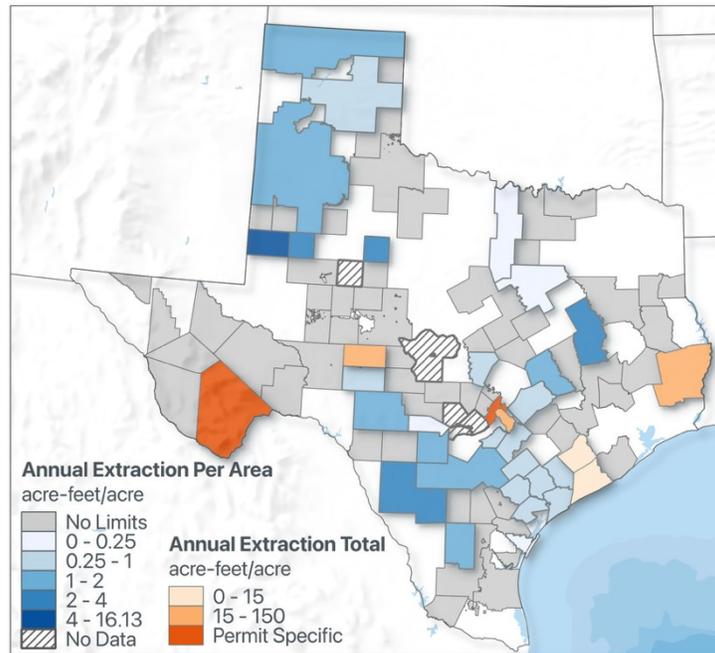


Figure 5: Annual extraction limits for groundwater wells based on extraction area or landownership acreage, total volumes, and permit type.

3.3 Groundwater Export and Potential Associated Usage Fees

Export fees are of particular interest because it is possible that an EGS development in a water-stressed district could import water from a nearby or neighboring district. Our analysis shows that export fees are common throughout the state and vary from \$0.02 per gallon to more than \$0.20 per gallon. Figure 6 shows that transporting water from one district to another would likely come with additional fees and application processes (i.e., blue).

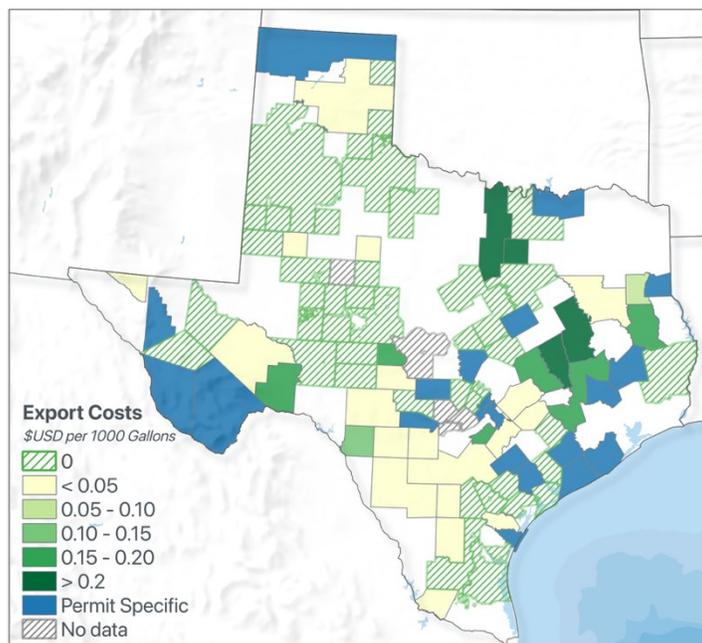


Figure 6: Export costs by groundwater district.

4. DISCUSSION

Validation efforts for this work involved evaluating the accuracy of extracted data for a subset of 15 randomly selected groundwater conservation districts. For these 15 districts, results were manually verified. The results were key indicators of the effectiveness of this approach as well as guidance for further development. When the target field was relatively straightforward to extract, such as metering device requirements, the model performed well. In our validation sample set where we specifically asked whether or not a metering device is required, this was extracted with 100% accuracy.

Limitations of our methodology, however, are exposed when the targeted metric involves complexities such as exemptions to rules for specific well types. This is displayed with the production cost field where accuracy dropped to 63% and errors were often the result of production fees being enforced but numerous exemptions applying to the rule including agricultural wells, residential wells, and wells extracting water below a given threshold**. Such scenarios can be resolved with further decision tree refinement to address cases when exemptions should be considered, but this must be done with careful consideration to ensure that the effectiveness of the decision tree is not diminished when applied to other districts or different states.

There were two conservation districts in our validation set where the web scraping process could not locate the rules and regulations document for a district and answers were generated with insufficient information. As expected, these answers were inaccurate. In the future, this issue could be addressed with continued web scraping development or additional validation steps to determine whether there is enough information to confidently generate data and if not, the location can be flagged for further evaluation.

During the validation process it was also discovered that “subsidence districts” and the Edwards Aquifer Authority overlap groundwater conservation districts, but are not groundwater conservation districts as defined under the Texas Water Code, but have the ability to regulate groundwater production. These geographically overlapping (shown in Figure 7) jurisdictions are another added complexity that requires clarification beyond the LLM results.

Our evaluation of groundwater ordinances in Texas provides a reminder of the complexities inherent to groundwater policy in the Western United States. Despite the challenges associated with understanding and extracting information from these documents, this analysis suggests that LLMs can be an effective method to automate data collection in a way that is cost effective and timely. As was the case with Texas, applying the process to additional states will be informed by subject matter expertise and improved with multiple rounds of iteration.

**Example “production” cost with exemptions – Calhoun County GCD Rule 10.4: <https://www.calhouncountygcd.org/files/03b09a3aa/CCGCD+-+Rules+of+the+District+-+Adopted+20241028.pdf>

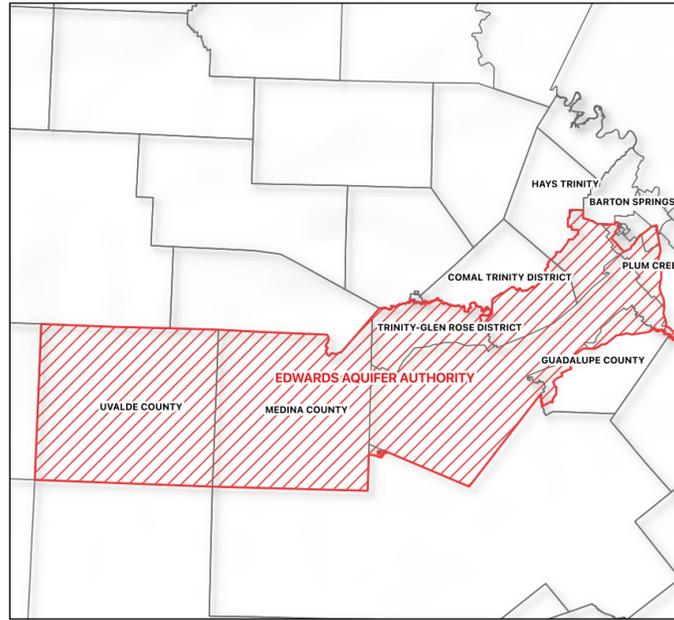


Figure 7: Geographical overlap of groundwater conservation districts with Edwards Aquifer Authority.

5. CONCLUSIONS & FUTURE/CURRENT WORK

Meeting the projected growth of geothermal systems in the United States in the coming decades will require a comprehensive and current understanding of water policy and availability. Our work evaluating potential groundwater availability and relevant regulatory frameworks or permitting requirements in Texas suggests that using LLMs that are guided by decision trees can provide an efficient method to create structured data from water policy documents. Validation suggests that minor errors in the results are likely, and we do not suggest using these data in work with little error tolerance.

When comparing the warmer colors of Texas in Figure 1 (the estimated higher geothermal potential) to the results from this study, we qualitatively conclude that about half of the districts along the Gulf Coast do not have annual extraction limitations nor costs for exporting the water out of the conservation district.

Future work includes applying this workflow to additional states in the western United States, including Nevada, Utah, California, and New Mexico. Expanding to new states will require developing state-specific methods for document collection and refining decision trees to ensure accuracy and flexibility.

After completing our initial efforts in Texas, Nevada was selected as the next state to analyze as it has the most operating geothermal power plants and a clearly defined regulatory framework for use of groundwater resources for geothermal utilization including a specific permitting structure. Texas does not have a statewide permitting structure; rather, the responsibility to determine how to conserve groundwater is given to the groundwater conservation districts. The Nevada water management structure provides the next opportunity to evaluate the effectiveness of our methodology; however, it maintains few similarities to the Texas groundwater conservation districts. Rather than permit processes being managed by individual districts, Nevada well permits are issued at the state level, and designated basins are defined to preserve water supplies in critical areas (Water in the West, n.d.; Nelson and Perrone 2016). With this configuration we intend to extract permit application information from state-level statutes and well-specific information from designated basin management plans.

Addressing state-level differences will primarily require adapting the document collection procedure because the search queries and document validation need to be tuned to locate and validate different documents of interest. Given that we will generally be attempting to extract similar data points for each state, the final decision trees developed for Texas will initially be applied to Nevada as well; however, it is expected that state-level differences will require some additional prompt development and validation of the results.

Output data from the extraction efforts will be published to the Renewable Energy Siting Lab on OpenEI.³

³ https://data.openei.org/siting_lab

ACKNOWLEDGEMENT

We thank Sid Ramavajjala at the National Laboratory of the Rockies for his mapping of our data, and Diana Acero-Allard and Stuart Cohen for their valuable feedback on the paper. We thank Andy Adams and Sean Porse for their early support of this research. This work was authored in part by the National Laboratory of the Rockies for the U.S. Department of Energy (DOE), operated under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Geothermal under Contract No. DE-EE0009962. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

REFERENCES

- Augustine, Chad, Sarah Fisher, Jonathan Ho, Ian Warren, and Erik Witter. 2023. *Enhanced Geothermal Shot Analysis for the Geothermal Technologies Office*. Technical Report NREL/TP-5700-84822. National Renewable Energy Laboratory. <https://www.osti.gov/servlets/purl/1922621/>.
- Buster, Grant, Pavlo Pinchuk, Jacob Barrons, Ryan McKeever, Aaron Levine, and Anthony Lopez. 2024. “Supporting Energy Policy Research with Large Language Models: A Case Study in Wind Energy Siting Ordinances.” *Energy and AI* 18 (December): 100431. <https://doi.org/10.1016/j.egyai.2024.100431>.
- Cath, Tzahi, Shankar Chellam, Lynn Katz, et al. 2021. *NAWI Technology Roadmap: Resource Extraction Sector*. https://docs.nrel.gov/docs/fy21osti/79895.pdf?utm_campaign=Newsletter%20-%20Enspired&utm_source=hs_email&utm_medium=email&_hsenc=p2ANqtz--255KUjYK1ZPQVuaPLxRY_2yq-AVCaOHPIDTS9QPESXHOBjvauV6BS_AJkcHPACxhzoPj.
- Fervo Energy. 2025. “Water Use and Impacts.” Insights. *Geothermal Mythbusting*, March 31. <https://fervoenergy.com/geothermal-mythbusting-water-use-and-impacts/>.
- Fuchs, Brian. 2026. “U.S. Drought Monitor.” <https://droughtmonitor.unl.edu/>.
- Levine, Aaron, Faith Martinez Smith, and Heather Buchanan. 2023. *Topics and Considerations for Developing State Geothermal Regulations*. Technical Report NREL/TP-6A20-86985. National Renewable Energy Laboratory. <https://docs.nrel.gov/docs/fy23osti/86985.pdf>.
- Lopez, Anthony, Gabriel Zuckerman, Pavlo Pinchuk, et al. 2025. *Renewable Energy Technical Potential and Supply Curves for the Contiguous United States: 2024 Edition*. NREL/TP-6A20-91900, 2500362, MainId:93678. <https://doi.org/10.2172/2500362>.
- National Laboratory of the Rockies. 2026. “RAPID Geothermal.” Regulatory and Permitting Information Desktop Toolkit. <https://openei.org/wiki/RAPID/Geothermal/Jurisdictions>.
- Nelson, Rebecca, and Debra Perrone. 2016. “The Role of Permitting Regimes in Western United States Groundwater Management.” *Groundwater* 54 (6): 761–64. <https://doi.org/10.1111/gwat.12467>.
- Porter, Spencer. 2024. “Understanding Cosine Similarity and Word Embeddings.” *Medium*, November 29. <https://spencerporter2.medium.com/understanding-cosine-similarity-and-word-embeddings-dbf19362a3c>.
- Roberts, Billy. 2025. *Geothermal Resources of the United States: Development Suitability of Enhanced Geothermal Systems (EGS)*. January. <https://www.nrel.gov/docs/gen/fy26/97591.jpg>.
- Schroeder, J.N., C.B. Harto, R.M. Horner, and C.E. Clark. 2014. *Geothermal Water Use: Life Cycle Water Consumption, Water Resource Assessment, and Water Policy Framework*. ANL/EVS-14/2. Argonne National Laboratory. <https://publications.anl.gov/anlpubs/2014/06/105260.pdf>.
- Seligman, Angela. 2025. “Water Requirements for Next Generation Geothermal Systems: The Need for Increased Reporting and Sharing of Lessons Learned in Early Projects.” *GRC Transactions* 49 (October): 1302–9.
- Texas Water Development Board. 2019. *Groundwater Conservation Districts of Texas*. https://www.twdb.texas.gov/mapping/doc/maps/GCDs_8x11.pdf.
- “Utah FORGE Successfully Completes Stimulation and Circulation Tests – Establishing Effective Communication.” 2024. *Utah FORGE*. <https://utahforge.com/press-release-stimcirc-tests/>.

Water in the West. n.d. "License to Pump: Groundwater Permitting in the West." Accessed January 23, 2026. <https://groundwater.stanford.edu/>.

Zhang, Tianjun, Shishir G. Patil, Naman Jain, et al. 2024. "RAFT: Adapting Language Model to Domain Specific RAG." arXiv:2403.10131. Preprint, arXiv, June 5. <https://doi.org/10.48550/arXiv.2403.10131>.