Local community opposition to data centers is clear and growing

Adam Rauh
Department of Political Science
University of Michigan

Mason Reece
Department of Political Science
Massachusetts Institute of Technology

Tyler Simko*
Department of Political Science
University of Michigan

April 20, 2026

Abstract

Rapid growth in computing and generative artificial intelligence is creating unprecedented demand for data center construction. Anecdotal media reports suggest residents often resist local governments approving data centers in their communities, citing concerns over issues like water consumption and expected increases in utility costs. However, studying the political reactions of communities across the United States to data centers remains difficult due to a lack of centralized data on local policymaking. We offer the first large-scale evidence of how U.S. communities discuss and respond to data center expansion by collecting and transcribing videos of local government meetings between 2007 and 2026. We show that the primary tension in local policymaking is between the economic development potential of data centers and expected local costs, such as noise, utility rate increases, and environmental impacts. We also find that residents are systematically less supportive of data centers than officials when discussing the same proposals.

Keywords data centers • local politics • land use • local government meetings

*To whom correspondence should be addressed. Email: tsimko@umich.edu. Website: https://www.tylersimko.com. We are grateful for helpful comments from seminar audiences at the University of Michigan and MIT.

1 Introduction

Data center investment has increased fivefold in the last decade, with the largest investments taking place in the United States (International Energy Agency, 2025). This pattern shows no sign of slowing—current projections anticipate more than an additional $1 trillion will be spent on data center investment before 2030 (JLL, 2026). Data centers already account for around 1.5% of global electricity consumption, and that rate is expected to more than double by 2030 (International Energy Agency, 2025). Political and economic pressures to meet this demand are high, as the transformative potential of artificial intelligence depends largely on access to computational power offered by data centers (CBRE, 2025).

Despite their global impacts, data centers are infrastructure projects that must be sited in local communities. Recent media accounts have documented how debates over new data center construction can become contentious and polarizing (Ambrose and Chen, 2026; Murphy and Feng, 2025; Green, 2026). Most often, existing accounts suggest these discussions feature tensions between the potential for economic development and local costs that residents may have to face, such as rising utility costs, noise, and traffic (Sharp, 2025; Wittenberg, 2025). Yet, beyond these individual accounts from specific places and projects, we know little about how local communities are systematically grappling with the increase in data center construction.

We offer the first large-scale evidence on how local communities across the United States are discussing data centers in their local governments. Although data center construction is growing rapidly, the decentralized nature of American government means that there is no easily accessible data on how local communities are discussing data centers (Barari and Simko, 2025; Holman et al., 2026). We address this lack of data by extending *LocalView*—the largest publicly available database of local government meetings—to systematically collect nearly 150,000 transcripts of local government meetings in 48 states over 2007–2026 (Barari and Simko, 2023). We use these transcripts to (1) characterize the extent of discussion about data centers in local governments over time, (2) identify the most commonly discussed issues, (3) the tone by which people discuss those issues, and (4) distinguish speakers as members of the public or local officials.

2 Results

First, we find that local government attention to data centers has dramatically increased in recent years. The left plot of Figure 1 shows the proportion of all local meetings in three month bins that discuss data centers. Since the beginning of 2023, we find local government attention to data centers (shown in maize) has increased by more than a factor of 10. Nearly 10% of local government meetings discussed data centers since March 2026, up from less than 1% of meetings between January and March 2023. We also compare the rise in attention to data centers to other examples of projects commonly referred to as "locally unwanted land use"—casinos, drug treatment clinics, and sports arenas. While attention to these other issues has remained relatively flat since 2023 (other than a slight uptick in attention to sports arenas, likely driven by the upcoming World Cup and Olympics), attention to data centers has skyrocketed with no signs of slowing.

We identify 1,707 meetings that discuss data centers and use text analysis to assign whether each data center discussion mentions topics related to community costs (e.g. safety, noise, traffic), economic development (e.g. jobs and tax base impacts), land use and infrastructure (e.g. comparing data center land use to alternatives), utilities and the environment (e.g. utility rate increases, water cooling), and data center projects in other cities.

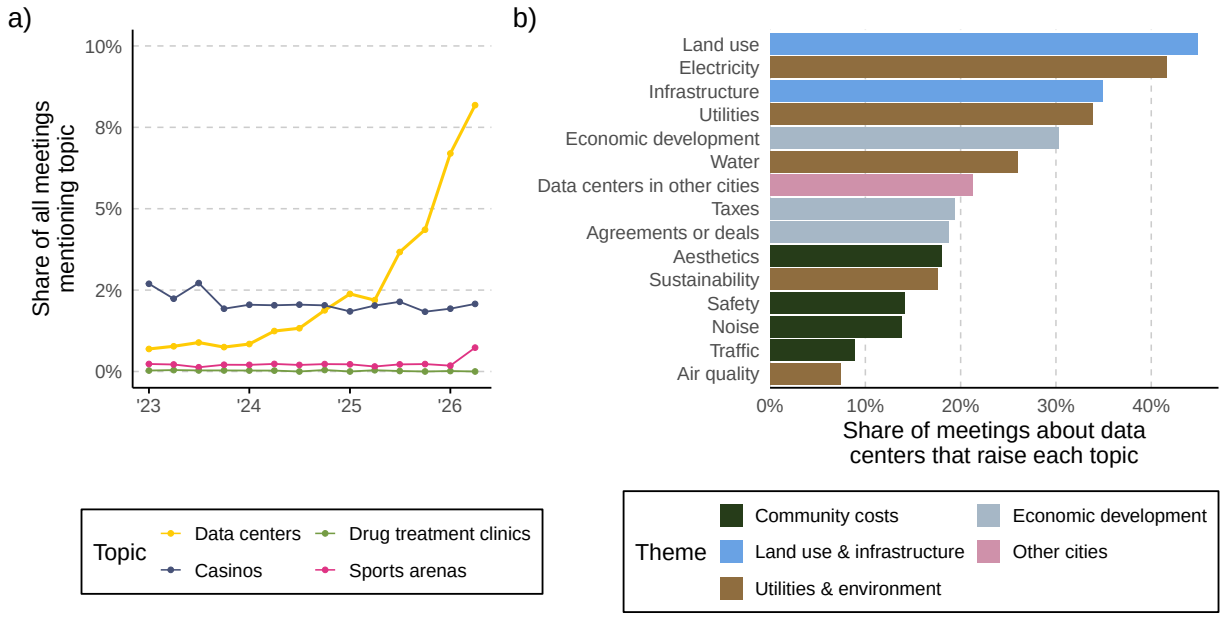

Figure 1: Mentions of data centers in local government meetings are growing, are most often about land use, infrastructure, utilities, and the environment. The left plot shows that attention to data centers in local government meetings is increasing over time. The x-axis shows meetings grouped into three month bins since the beginning of 2023 (we filter our meeting sample to start in 2023 for visual clarity), and the y-axis shows the proportion of meetings in that period that mention each issue. We compare discussions of data centers to several other types of local infrastructure projects. In the right plot we show the share of meetings about data centers dedicated to each topic. The x-axis shows the share of meetings in which the topic was mentioned. We group the topics into five themes for clarity.

The right plot of Figure 1 shows that land use and infrastructure, electricity, water, and other utilities are among the most common topics discussed about data centers (appearing in 25-40% of data center discussions), while other tangible concerns about noise and traffic are less common but still appear often (10-15%). Nearly all discussions related to economic development—most often related to jobs and tax benefits—appear around the middle of our distribution, and are present in around 20-30% of data center discussions.

Next, we use tone classifications described in *Materials and Methods* to show whether speakers who discuss each topic tend to support or oppose additional data centers in their communities. For each topic, we calculate the difference between the proportion of speakers who invoked it in positive and negative ways (we drop neutral passing references for this analysis). The left plot of Figure 2 shows how speakers tend to be more positive when discussing the potential economic impacts of data centers, such as the prospect of new jobs. Economic development is the most positive theme associated with new data center construction. Most other concerns—utilities, traffic, noise, sustainability, and especially air quality—are much more negative. We find this support pattern is largely asymmetric—speakers invoke more topics in a primarily negative way (i.e., to the left of the dashed line) than a primarily positive way (i.e., to the right).

Finally, we show that opposition to data centers is driven by members of the public, rather than local officials. In the right plot of Figure 2, we calculate the tone (positive or negative, dropping neutral) of our discussions separately by speaker as described in *Materials and Methods*, and show that

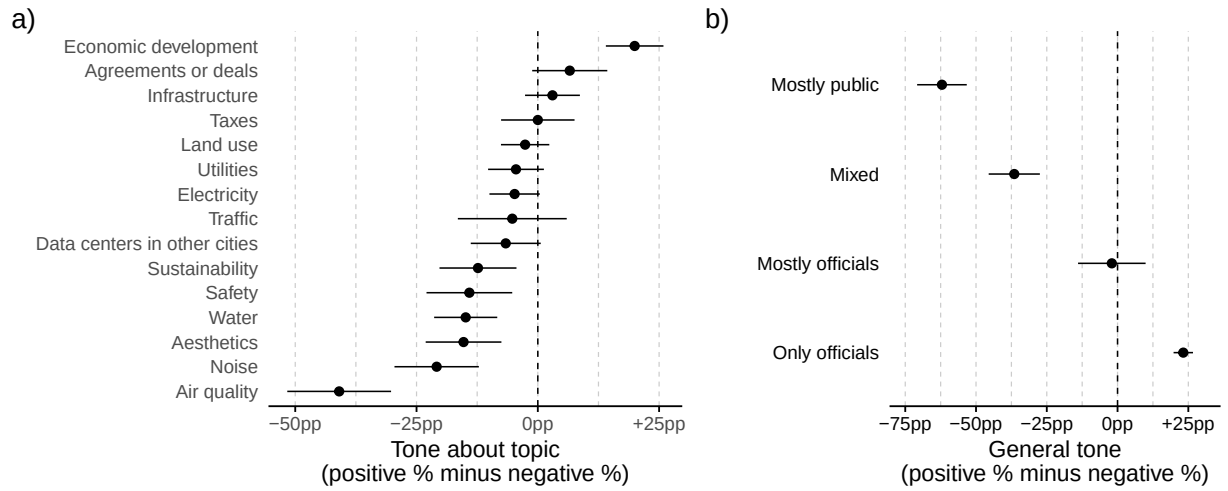

Figure 2: Communities discuss most topics negatively and members of the public are in general more negative than local officials. We use our tone classifications described in *Materials and Methods* to calculate, for each meeting-topic, the proportion of positive and negative discussions. The left plot shows the average difference between positive and negative proportions across all meetings (bars show 95% CIs). In the right plot we show the average tone across all topics for members of the public versus local officials. Each point shows the difference between overall positive and negative tone for four binned meeting types. We categorize meetings in which data center discussions were made by only officials (0% public share), mostly officials (less than 25% public share), mixed (between 25% and 75%), and mostly public (greater than 75%).

speaking members of the public are much less supportive of data centers than officials. This pattern is clear—as more members of the public speak in a given meeting, overall discussions about data centers become much more negative. In contrast, meeting discussions led only by local officials without community input discuss data centers in a more positive way. As above, we find evidence that this tone differential is asymmetric—while discussions led by local officials are slightly more positive overall (nearly 25pp net positive tone), our meetings with the highest proportion of public speakers are overwhelmingly negative (approximately 60pp net negative tone).

3 Discussion

We offer a first look into how local governments are considering a critical issue in America's adaptation to generative artificial intelligence. Although our analysis is preliminary—as the growth of the data center sector is ongoing and expected to further balloon in coming years—we see three contributions from our results. First, we validate emerging evidence about public opposition to data centers in communities across the country. In general, we find that local residents who participate in local policymaking debates about data centers are cautiously optimistic about the potential for future economic development, but are overall more concerned about how data centers will negatively impact their communities.

Second, our results underscore how American federalism complicates the study of policymaking for artificial intelligence. Despite increasing national attention to computational infrastructure, America's decentralized government means that many of the most influential actors for the growth of data centers are and will continue to be the subnational officials and local residents deciding

whether and how to construct data centers in their own communities. We contribute to a growing literature that attempts to overcome these decentralization challenges by studying local governments at scale (Barari and Simko, 2023).

Finally, we find evidence of a sharp divide between government officials and members of the public regarding data centers. While data centers may offer local communities the opportunity to host critical infrastructure for an evolving economy, current residents are the ones who will pay the most direct costs of these projects. Our results offer the most comprehensive evidence yet available about how communities across the country are discussing these trade-offs in their own neighborhoods, and highlight the need for robust public deliberation in data center siting.

4 Materials and Methods

We begin with the local government meetings collected and transcribed by the *LocalView* project (Barari and Simko, 2023). Then, we use keyword search to identify all meetings where data centers are discussed. For those 1,707 meetings, we download an audio file of the meeting and follow current best practices to produce a new, higher quality, transcript using OpenAI's Whisper model accessed through the `whisperx` Python package (Radford et al., 2022; Bain et al., 2023). Because local meetings often cover many different topics, we subset to only the part(s) of the meeting where data centers are discussed (~10,000). We call these individual units "discussions." This process is described in detail in the *Supplementary Information* (SI). For each discussion, we identify the topics mentioned, whether the speaker in that discussion was a member of the public or a local official, and whether the discussant was speaking about the topic in a positive, negative, or neutral way. We classify 1% of discussions manually, then build an LLM prompt with the manual classifications as examples. We pass the prompt, and all of the discussions, to Claude Haiku 4.5 (2025-10-01) for classification. We validate the classifications in the SI. Finally, we aggregate our data in several plots to the meeting level, so each variable is reported as a proportion of the data center part of the meeting.

A Supplementary Information

A.1 LLM Prompts and Validation

As discussed in the main paper, we identify all meetings in our dataset that feature a discussion of data centers using keyword searches. Meeting transcripts were preprocessed by collapsing extra whitespace, then segmented into sentences using a punctuation-based split that accounts for common abbreviations (e.g., Dr., vs., U.S.) to avoid false sentence boundaries. Sentences were then grouped into fixed-size blocks of ten consecutive sentences, which served as the unit of analysis. In the main paper, we refer to these chunked sections of meetings as "discussions." Then, we manually coded 128 of these chunks that discuss data centers for theme, tone, and speaker information. These were then provided, along with detailed prompts, as training information for an LLM to code the roughly 10,000 total chunks. The LLM coded each chunk for thematic content, tone, and identified information about the speaker. In this section, we provide some validation statistics, showing the level of agreement between our manually coded observations and the results of the same discussions passed to the LLM. We also include the prompts provided to the LLM for each. We used Claude Haiku 4.5 (2025-10-01) for all LLM tasks.

A.1.1 LLM Theme Coding Prompt

The following prompt was given to Claude for coding themes:

For each topic below, code 1 if it was discussed **at any point** in the section of the meeting, and 0 if it was not mentioned at all.

- *Environment*
  - `any_water`: Water use, wastewater, cooling, sewer, runoff, or water cost.
  - `any_electricity`: Energy demand of the data center, power grid references, or electricity cost.
  - `any_air_quality`: Air quality concerns or pollution.
  - `any_sustainability`: Sustainability or climate concerns.
  - `any_utility`: Generic public utility references (other than water or electricity), or utility cost.
- *Community Costs*
  - `any_noise`: Noise from the data center project (noise pollution, construction, server hum).
  - `any_traffic`: Traffic during construction or operation.
  - `any_safety`: Safety concerns: hazmat, waste runoff, fire risk, electromagnetic concerns, local emergency response capacity.
  - `any_infrastructure`: General infrastructure concerns related to data centers.
  - `any_aesthetics`: Neighborhood character, visual or aesthetic impacts.
  - `any_equity_concerns`: Equity in the siting process, or distribution of benefits and costs across residents.
- *Economic Impacts*
  - `any_economic_development`: Jobs, local hiring, pay, unionization, or number of positions.
  - `any_taxes`: Tax or fiscal benefits/burdens of the data center on the community.

  - `any_land_use`: Productivity or appropriateness of land use for a data center vs. alternatives.
  - `any_deal`: The agreement with the developer: cost sharing, tax incentives, infrastructure responsibilities.

A.1.2 LLM Speaker Coding Prompt

The following prompt was given to Claude for coding speaker information:

For each section, identify the primary category of speaker. Code 1 if the section is primarily spoken by that speaker, and 0 otherwise. Use your best judgment based on the transcript. Speaker categories are **mutually exclusive**:

- **Developer** (`developer`): Speech by a representative of the company proposing or building the data center—a developer, consultant, or lawyer speaking on the applicant's behalf.
  - *Example*: "I am here tonight to speak on behalf of Crytalis Developers…" or "As the project engineer for the applicant…"
- **Public** (`public`): Speech by a member of the public—not an elected/appointed official or a developer representative.
  - *Example*: "My name is Tyler Simko, and I strongly oppose this data center project on 5th…"
- **Official** (`official`): Speech by a formal local official—elected or appointed government staff.
  - *Example*: "I have always said I am a no on this project. I am going to vote no because of the extensive feedback I have heard…"
- **Other** (`other`): Speech by someone who fits none of the above—for example, an invited expert witness, a state agency representative, or an unidentifiable speaker.
  - *Example*: "As the state environmental liaison, I'm here to summarize the permitting requirements…"

A.1.3 LLM Tone Coding Prompt

The following prompt was given to Claude for coding tone:

Code which tones are present at any point in the section of the meeting. Multiple tone flags can be 1 simultaneously—a meeting may contain both supportive and critical speech. Each flag is independent.

- `any_positive_tone` (1 vs. 0): Was any portion of the section clearly supportive of the data center project, or speaking positively about data centers in general? The dominant thrust of that portion must be favorable.
- `any_negative_tone` (1 vs. 0): Was any portion of the section clearly opposed to the data center project, or speaking critically about data centers in general? The dominant thrust of that portion must be critical or opposed.
- `any_neutral_tone` (1 vs. 0): Was any portion of the section neutral—factual, procedural, or not taking a clear stance? Purely administrative segments (vote tallies, scheduling, introductions) always count as neutral.

Default rule: When a section contains mixed signals with no clear dominant thrust, count it as neutral, not positive or negative.

A.1.4 Validation of LLM Theme Coding

Manual and LLM codings are considered to be in agreement if values for the same section match. Using these matches, we then aggregate over all 128 sections and calculate the average number of same classifications.

Agreement between manual and LLM-assigned codes was generally high across thematic variables, though with some variation. Among environmental indicators, agreement was 95.31% for water, 85.16% for electricity, 96.09% for air quality, 86.72% for sustainability, and 84.38% for utility. Community indicators showed similarly strong agreement, with noise at 96.88%, traffic at 96.88%, safety at 91.41%, infrastructure at 75.00%, aesthetics at 85.94%, and equity concerns at 84.38%. Agreement on economic impact variables was 89.06% for economic development, 95.31% for taxes, 81.25% for land use, and 83.59% for deal.

A.1.5 Validation of LLM Speaker Categorization

We performed a similar validation exercise for speaker categorization. Speakers were either categorized as an elected or formal local official, a member of the public, a representative from a data center developer, or someone else.

Agreement on speaker categorization was more mixed across the four categories. Agreement was 78.13% for "developer", 81.25% for "public", 74.22% for "official", and 87.50% for "other". In our main text results, we use only the public and official categories.

A.1.6 Validation of LLM Tone Categorization

Sections were coded as either Positive, Negative, or Neutral. Agreement between manual and LLM classification on tone was 75.78% overall. However, this overstates the level of disagreement between the manual and LLM codings because many of these errors were driven by discrepancies in the "neutral" category, which we do not use in our analyses. Directional agreement between positive and negative codings was much stronger. Of the 50 observations coded as negative manually, only 3 were classified as positive by the LLM. Of the 17 observations manually coded as positive, only 1 was coded as negative by the LLM.

A.2 Strings Used in Figure 1

We search transcripts to identify meetings about data centers, drug treatment clinics, casinos, and sports arenas. We search for the following strings to identify meetings falling into each of the categories: **data center:** `data center, data centre, server farm`, **Drug treatment centers:** `drug treatment clinic, drug treatment center, opioid treatment clinic, opioid treatment center`, **Casinos:** `casino`, **Sports arenas:** `sports stadium, sports arena`.

References

Ambrose, A. and Chen, P. (2026). Fool's gold: The hidden costs of AI data centers for New Jersey. Technical report, New Jersey Policy Perspective, Trenton, NJ.

Bain, M., Huh, J., Han, T., and Zisserman, A. (2023). Whisperx: Time-accurate speech transcription of long-form audio. *INTERSPEECH 2023*.

Barari, S. and Simko, T. (2023). Localview, a database of public meetings for the study of local politics and policy-making in the united states. *Scientific Data*, 10(1):135.

Barari, S. and Simko, T. (2025). The promise of text, audio, and video data for the study of us local politics and federalism. *Publius: The Journal of Federalism*, 55(2):223–252.

CBRE (2025). U.S. data center trends report, year-end 2025. Technical report, CBRE Research, Dallas, TX.

Green, B. (2026). The public is getting fed up with data centers. politicians need to take notice. Tech Policy Press.

Holman, M., Johnson, R., and Simko, T. (2026). Measuring conflict in local politics. *Urban Affairs Review*, 62(2):303–338.

International Energy Agency (2025). Energy and AI. Technical report, IEA, Paris, France.

JLL (2026). 2026 global data center outlook. Technical report, JLL Research.

Murphy, R. and Feng, E. (2025). Why more residents are saying 'no' to AI data centers in their backyard. NPR.

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., and Sutskever, I. (2022). Robust speech recognition via large-scale weak supervision.

Sharp, J. (2025). Southern GOP governors embrace AI data centers, push back on clean energy subsidies. AL.com / Alabama Media Group.

Wittenberg, A. (2025). 'How Come I Can't Breathe?': Musk's Data Company Draws a Backlash in Memphis. POLITICO / E&E News.