

Data center ad classification prompt

You are coding Facebook/Instagram political ads that mention data centers (large facilities housing servers for AI and cloud computing). Classify the ad text below by PUBLISHER_TYPE, CANDIDATE_NAME, PARTY, SENTIMENT, and FRAMING. Each of the classifications are described below. Provide your answer in JSON format with the keys 'PUBLISHER_TYPE', 'CANDIDATE_NAME', 'PARTY', 'SENTIMENT', and 'FRAMING'. Each key should have a JSON value with keys for 'value' and 'reasoning'. The 'value' should be the classification as outlined below, and the 'reasoning' should be a brief explanation for your classification.

PUBLISHER_TYPE: Classify who published the ad.

PUBLISHER_TYPE Codes:

- CANDIDATE: Ad appears to be published by a political campaign account or partisan advocacy organization in support of a specific candidate's run for elected office.
- ADVOCACY: Ad appears to be published by a lobbying firm, corporation, industry group, activist group, or advocacy organization and does *not* endorse a specific candidate or party for elected office.
- OTHER: Not enough information to infer who published the ad, or the ad is from an organization that does not fall into either the CANDIDATE or ADVOCACY label (such as news outlets).

PUBLISHER_TYPE Rules:

1. Posts from lobbying/advocacy organizations that focus explicitly on supporting a candidate's campaign for elected office should be labeled as CANDIDATE.
2. Page names and publisher names that follow patterns like "[Name] for [Office]", "Friends of [Name]", or "[Name] Campaign" indicate CANDIDATE. A candidate's ad is CANDIDATE even if it only discusses an issue.
3. Issue ads from advocacy organizations that criticize a politician but do not explicitly endorse a candidate or political party should be labeled ADVOCACY. Similarly, issue ads that mention a politician's name but do not campaign for them or suggest voting for them (i.e. "Tell Rep. Ansari to pause data centers") should be labeled ADVOCACY.
4. Publishers that position themselves as a news outlet or media organization should be coded as OTHER, unless they endorse a candidate for elected office (in which case it should be labeled as CANDIDATE).
5. Use the page name, publisher name, and body together. Ignore duplicated text and read non-English text as written. Do not guess beyond the text.

PUBLISHER_TYPE Examples:

- Page name: "Rep. Gina Hinojosa" / Publisher: "Gina Hinojosa Campaign" / "Greg Abbott just filed his campaign finance report. Guess who's been writing him massive checks?" -> CANDIDATE
- Page name: "LCV Victory Fund" / Publisher: "LCV VICTORY FUND" / "Congressman Barrett will protect local communities. Go vote!" -> CANDIDATE
- Page name: "Clean Virginia" / Publisher: "Clean Virginia" / "Virginia lawmakers have the choice to end tax breaks for data centers. Tell your lawmaker to make Big Tech pay." -> ADVOCACY
- Page name: "Fair Rates for Wisconsin's Dairyland" / "New data centers in Wisconsin mean new property tax revenue and thousands of construction jobs." -> ADVOCACY
- Page name: "Tristate Infrastructure News" / Publisher: "Tristate Infrastructure News Network" / "Some basic facts about the PJM grid and energy demand from AI and data centers." -> OTHER

In the output JSON, set the "value" for PUBLISHER_TYPE to only one code: CANDIDATE, ADVOCACY, or OTHER. The "value" should contain only the code, with no explanation or punctuation; put any explanation in "reasoning".

CANDIDATE_NAME: Identify the full name of the candidate the ad is published for or in support of, as written in the ad (e.g., "Gina Hinojosa"). If the ad appears to promote multiple candidates, return a list of the candidates with a comma separating each candidate name. Use "None" if the publisher type is not CANDIDATE.

CANDIDATE_NAME Rules:

1. If an ad's publisher_type is not CANDIDATE, set CANDIDATE_NAME to "None".
2. Identify the candidate who is being promoted by the publisher, not the opponent being attacked. For example, an ad from "Gina Hinojosa Campaign" that attacks Greg Abbott is for Gina Hinojosa, not Greg Abbott.
3. For page names like "[Name] for [Office]", "Friends of [Name]", or "[Name] Campaign", or disclaimers like "Paid for by [Name] for Congress", use [Name] as the candidate. Drop titles such as "Rep.", "Sen.", "Dr.", or "Commissioner" from CANDIDATE_NAME.
4. For outside groups (PACs, advocacy groups) whose ad clearly supports or opposes a specific candidate, name the candidate they are supporting.
5. If the ad names multiple candidates from the same publisher, include all candidate names (separated by a comma) in CANDIDATE_NAME.
6. Use the page name, publisher name, and body together. Ignore duplicated text and read non-English text as written. Do not guess beyond the text.

CANDIDATE_NAME Examples:

- Page name: "Rep. Gina Hinojosa" / Publisher: "Gina Hinojosa Campaign" / "Greg Abbott just filed his campaign finance report. Guess who's been writing him massive checks?" -> "Gina Hinojosa"
- Page name: "Clean Virginia" / "Virginia lawmakers have the choice to end tax breaks for data centers." (publisher type ADVOCACY) -> "None"

In the output JSON, set the "value" for CANDIDATE_NAME to a string with the candidate's name. The "value" should contain only the name(s) or "None", with no other text; put any explanation in "reasoning".

PARTY: Identify the political party affiliation of the ad, if any.

PARTY Codes:

- D: The ad supports the Democratic party or a candidate who is a Democrat.
- R: The ad supports the Republican party or a candidate who is a Republican.
- I: The ad supports a third political party (other than Democratic/Republican) or a candidate who identifies as an independent.
- Unsure/None: The ad's partisanship or candidate's party cannot be determined from the text or from your existing knowledge of the candidate's partisan affiliation.

PARTY Rules:

1. Infer PARTY only from evidence in the ad text or based on partisanship of a political figure mentioned in the ad. Do not guess from the candidate's position taken on data centers.
2. If the ad supports multiple candidates from the same party, then code that party. If the ad supports multiple candidates from different parties, code "Unsure/None".
3. If the ad explicitly criticizes or campaigns against the Democratic party, then it should be coded as Republican. Likewise, ads that explicitly campaign against the Republican party should be coded as Democratic. Note that this rule only applies to clear criticisms, not simply mentions of the other party.
4. If the ad criticizes a politician or candidate without supporting any specific alternative candidate, DO NOT infer partisanship solely from that criticism. Some of these ads are from primaries in which a Democrat may criticize another Democrat and vice versa. One exception to this rule is that ads that criticize Trump or MAGA can be labeled as Democratic.
5. Ads that advocate for conservatives or conservative values (using the term "conservative" explicitly) should be labeled as Republican, and vice versa for liberals and liberal values (using the term "liberal" explicitly) being labeled as Democratic.
6. Use the page name, publisher name, and body together. Ignore duplicated text and read non-English text as written. Do not guess beyond the text.

PARTY Examples:

- Page name: "Rep. Gina Hinojosa" / Publisher: "Gina Hinojosa Campaign" / "Greg Abbott just filed his campaign finance report. Guess who's been writing him massive checks?" -> D
- Page name: "Clean Virginia" / "Virginia lawmakers have the choice to end tax breaks for data centers." (publisher type ADVOCACY) -> Unsure/None

In the output JSON, set the "value" for PARTY to a string with the ad's political party ("D", "R", "I", or "Unsure/None"). The "value" should contain only the code, with no explanation or punctuation; put any explanation in "reasoning".

SENTIMENT: Classify the ad's stance toward data centers.

SENTIMENT Codes:

- PRO: Supports data centers or their development, or promotes their benefits (jobs, tax revenue, infrastructure, etc.).
- ANTI: Opposes data centers, presents them as a threat or problem (higher bills, water use, pollution, noise, land use), or calls to stop, pause, or restrict them.
- OTHER: Neutral, balanced with no clear lean, not actually about data centers, or too vague/not enough text to be determined.

SENTIMENT Rules:

1. Code the stance toward data centers, not toward a candidate or party.
2. An ad that doesn't mention data centers (e.g., "Vote for Jim") should be coded as OTHER.
3. Passing mentions count. If an ad attacks an opponent for helping data centers (tax breaks, donations from data center interests, siding with them over residents), it is ANTI. If it attacks an opponent for blocking data centers, it is PRO.
4. Calls for rules, guardrails, or making data centers "pay their fair share" can be challenging to code. These kinds of ads are ANTI when the ad frames data centers as primarily causing harm and PRO when the ad welcomes data centers on the condition they are built responsibly ("how New Jersey gets data centers right"). If the ad is a balanced explainer or has a roughly even tone, use OTHER.
5. Ads that mention data centers only as context for another topic (coal, nuclear, chips, stock tips) with no stance on data centers themselves are OTHER.
6. If an ad leans one way with a minor concession, code the lean. Use OTHER only if it is truly balanced or unclear.
7. Use the page name, publisher name, and body together. Ignore duplicated text and read non-English text as written. Judge the intended meaning, including sarcasm. Do not guess beyond the text.

SENTIMENT Examples:

- "Data centers will bring 500 good-paying jobs and millions in tax revenue. Tell the council to say YES to this project." -> PRO
- "Reliable infrastructure without higher bills. This data center will fully fund new grid investments at zero cost to ratepayers." -> PRO
- "The storm alerts that wake you at 3am are delivered from data centers. Most infrastructure Floridians rely on isn't visible from the road." -> PRO
- "Billion-dollar data centers get huge tax breaks while you pay their energy costs. Is this fair?" -> ANTI
- "I am the only candidate who has passed a moratorium on data centers. In the Assembly, I'll do the same." -> ANTI
- "Greg Abbott handed data centers the most generous tax giveaways in the country while Texans foot the bill." -> ANTI
- "Public hearing on the proposed data center zoning change is Thursday at 6 PM at the courthouse." -> OTHER
- "This content was removed because it didn't follow our Advertising Standards." -> OTHER

In the output JSON, set the "value" for SENTIMENT to ONLY one code: PRO, ANTI, or OTHER. The "value" should contain only the code, with no explanation or punctuation; put any explanation in "reasoning".

FRAMING: Classify the ways that the ad frames the discussion about data centers.

FRAMING Codes:

- WATER: Mentions data center water usage and/or its impacts, including water consumption, water pollution, higher water bills, water scarcity, or strain on the water supply.
- ENERGY: Mentions data center energy usage and/or its impacts, such as higher electric bills, energy prices, or strain on the power grid.
- JOBS: Mentions jobs tied to data centers, such as creation of jobs or the lack of promised jobs.
- TAXATION: Mentions taxes tied to data centers, such as tax breaks, subsidies, or raising taxes on companies that run data centers. Includes tax revenue and lower taxes for residents.
- CORPORATIONS: Mentions how data centers serve the interests of corporations (such as Big Tech, utility companies, or "industry") or mentions corporate investment in data centers.
- ENVIRONMENT: Mentions the environmental impacts of data centers, including natural resource scarcity, conservation, pollution, emissions, and climate.

FRAMING Rules:

1. Each code is independent (not mutually exclusive). Set it to 1 if the ad explicitly mentions that frame in connection with data centers, otherwise 0. An ad can have several 1s or all 0s.
2. The frame can be positive or negative. "Data centers lower your taxes" and "data centers get tax giveaways" are both TAXATION.
3. Frames must be about data centers. A mention of taxes, jobs, or energy in other parts of the ad (e.g., "cap property taxes" in a candidate bio) does not count.
4. Do not code frames that are only implied. Impacts that fit none of the codes (e.g., noise, local control) are not coded. Do not code frames based on vague impacts of data centers like "takes advantage of our community".
5. Use ENERGY for electricity bills, prices, and the grid, and WATER for water use, water bills, and water pollution or scarcity. If an ad mentions both ("water and power bills"), code both. However, mentioning "utility bills" or "utility companies" without specifying water or power/energy should not be coded as WATER or ENERGY.
6. Discussions of data center "land use" can count as ENVIRONMENT, especially in a conservation context. However, discussions of data centers taking up farmland or other non-conservation contexts do not count as ENVIRONMENT.
7. If the ad does not mention data centers at all, or has no usable text (removed or disabled ad), set all codes to 0.
8. Use the page name, publisher name, and body together. Ignore duplicated text and read non-English text as written.

FRAMING Examples:

- "Corporate data centers get huge tax breaks while you pay their energy costs." -> {"WATER": 0, "ENERGY": 1, "JOBS": 0, "TAXATION": 1, "CORPORATIONS": 1, "ENVIRONMENT": 0}
- "Data centers are draining your water supply and polluting our rivers." -> {"WATER": 1, "ENERGY": 0, "JOBS": 0, "TAXATION": 0, "CORPORATIONS": 0, "ENVIRONMENT": 1}
- "This project brings 500 good-paying union jobs and lower electric rates." -> {"WATER": 0, "ENERGY": 1, "JOBS": 1, "TAXATION": 0, "CORPORATIONS": 0, "ENVIRONMENT": 0}
- "Greg Abbott took \$20 million from data center interests, then rolled out the red carpet for them." -> {"WATER": 0, "ENERGY": 0, "JOBS": 0, "TAXATION": 0, "CORPORATIONS": 1, "ENVIRONMENT": 0}
- "Amazon is investing \$10 billion to in local energy grid to support new data center construction and strengthen our community's infrastructure." -> {"WATER": 0, "ENERGY": 1, "JOBS": 0, "TAXATION": 0, "CORPORATIONS": 1, "ENVIRONMENT": 0}
- "Hyperscale data centers strain our local water, power grid, and land." -> {"WATER": 1, "ENERGY": 1, "JOBS": 0, "TAXATION": 0, "CORPORATIONS": 0, "ENVIRONMENT": 1}

```
- "Corporations want to bulldoze our farmland for a data center." -> {"WATER": 0,
"ENERGY": 0, "JOBS": 0, "TAXATION": 0, "CORPORATIONS": 1, "ENVIRONMENT": 0}
- "This content was removed because it didn't follow our Advertising Standards." ->
{"WATER": 0, "ENERGY": 0, "JOBS": 0, "TAXATION": 0, "CORPORATIONS": 0, "ENVIRONMENT": 0}
```

In the output JSON, set the "value" for FRAMING to a JSON object with the keys WATER, ENERGY, JOBS, TAXATION, CORPORATIONS, and ENVIRONMENT, each set to 0 or 1. The "value" should contain only the JSON object; put any explanation in "reasoning".

```
===== START of ad text to classify =====
PAGE NAME: {ad page name}
PUBLISHER: {ad "Paid for by" publisher}
AD TEXT:
{Ad caption}
LINK TITLE: {Link title}
LINK DESCRIPTION: {Link description}
LINK CAPTION: {Link caption}
===== END of ad text to classify =====
```

Please return your answer in the following JSON format. Return only the raw JSON object, with no code fences (```) and no text before or after it:

```
{
  "PUBLISHER_TYPE": {
    "value": "CANDIDATE | ADVOCACY | OTHER",
    "reasoning": "Brief explanation for the classification"
  },
  "CANDIDATE_NAME": {
    "value": "Candidate's name or None",
    "reasoning": "Brief explanation for the classification"
  },
  "PARTY": {
    "value": "D | R | I | Unsure/None",
    "reasoning": "Brief explanation for the classification"
  },
  "SENTIMENT": {
    "value": "PRO | ANTI | OTHER",
    "reasoning": "Brief explanation for the classification"
  },
  "FRAMING": {
    "value": {
      "WATER": 0 or 1,
      "ENERGY": 0 or 1,
      "JOBS": 0 or 1,
      "TAXATION": 0 or 1,
      "CORPORATIONS": 0 or 1,
      "ENVIRONMENT": 0 or 1
    },
    "reasoning": "Brief explanation for the classification"
  }
}
```

Candidate partisanship classification prompt

You are helping classify the political party of a candidate in the 2026 U.S. midterm elections. Use web search to determine the political party of the candidate named below.

Candidate name: {CANDIDATE_NAME}

PARTY Codes:

- D: The candidate is a Democrat.

- R: The candidate is a Republican.
- I: The candidate is an independent or a member of a third party (e.g., Libertarian, Green).
- Unsure/None: The candidate's party cannot be determined with confidence.

Rules:

1. Make sure that the candidate you're considering is running for office in the 2026 U.S. midterm elections (at any level of government).
2. If the name is too incomplete to identify a specific person (for example, only a first name or only a last name), return "Unsure/None". Do not search for or guess a candidate.
3. If the name matches multiple different candidates in different races (for example, two people with the same name running for different offices or in different states) and you cannot tell which one is meant, return "Unsure/None".
4. Report the party the candidate is running under or officially affiliated with. Do not infer party from the candidate's policy positions, endorsements, or ideology.
5. Do not guess. If you cannot find reliable evidence, return "Unsure/None".

Output format:

Return only a string, one of "D", "R", "I", or "Unsure/None", with no explanation or punctuation.