CURIOSITYBITS LIVING LAB BOSTON · EXTERNAL MULTI-MODEL REPLICATION AUDIT (EXT-AUDIT-026)

Author Verification & Edition `v1.2` Revision Response (`EXT-AUDIT-026`) Referencing `SG-AUDIT-S04-V1.1-2026-11-06`

- **Referenced Audit ID:** `SG-AUDIT-S04-V1.1-2026-11-06` (`EXT-AUDIT-025`, parent feedback `DIRECTOR-FEEDBACK-S04-2026-11-06`)
- **Referenced Audit SHA-256 Verified (`64/64` chars):** `765ba47bd6d18cf14ee98bf12062dafddeb37467079b270706adb6e4522c9ab7`
- **Audited Study & New Edition:** Study #04, *Where the Cloud Hits the Ground: Monitoring Local, State, and National Pushback and Bargaining Over AI Infrastructure* — revised from `Edition v1.1` to **`Edition v1.2`**
- **Authoring Model:** `Gemini 4 Argon (gemini-4-argon)` (`Antigravity CLI`, CuriosityBits Living Lab Boston)
- **Archived & Current Edition SHA-256 Hashes (`SKILL-025`):**
- `STUDY_DESIGN_CARD_V1_PRE_DATA_LOCKED.md` (`v1.0` pre-data lock): `bb6b1b709c5d99c1d9ebe6f954bfd99ca646879a332ec96b6d4fa29bf7b462b9`
- `STUDY_DESIGN_CARD.md` (`v1.2` with §6 Endpoint-Probe Provenance Disclosure): `a4f4119b4328f59b866f604ec43389eb9c88c137babe90e44cd1b35131482868`
- `PAPER_MANUSCRIPT_V1_0_ARCHIVED.md` (`v1.0` archive, byte-identical): `7ce2c0e5781347aa043c9946267ae73357498c58a500eeaae2ab5bde01e19374`
- `PAPER_MANUSCRIPT_V1_1_ARCHIVED.md` (`v1.1` archive, byte-identical): `d092faf1fe05c21e97155cb8a2e7ac1ced85aeb2300a04c6cbe88b0867758298`
- `VERIFY_STUDY04_V1_1_ARCHIVED_STDOUT.txt` (`v1.1` verification stdout archive): `2615dcdd4fb1111c89f9dd1275a073b76bed793a698980974f9c3029954044df`
- `PAPER_MANUSCRIPT.md` (`Edition v1.2`): `f279c0d44d8bb4dc536883e7dc5ce9d844a58672c1ac335f7a3f389a1679e853`
- `INDEPENDENT_PEER_REVIEW_REPORT.md` (`Edition v1.2`): `1ca3360ac896d6ae1c1a1ef17ee10f6380d8bced34874e957a5132723020d633`
- `VERIFY_STUDY04_V1_2_STDOUT.txt` (`Edition v1.2` reference verification stdout): `41b231195e5ea04ae94152e5e62b95638b5e23bbf428ec6a290ed4baa2c5b5f2`

1. Independent Verification of `SG-AUDIT-S04-V1.1-2026-11-06` (`SKILL-023`)

Before making any changes, we verified the SHA-256 of `SG-AUDIT-S04-V1.1-2026-11-06` (`765ba47bd6d18cf14ee98bf12062dafddeb37467079b270706adb6e4522c9ab7`) and wrote `/srv/margin-projects/intuitionist-agy/projects/ai-infrastructure/scripts/build_and_verify_v1_2.py` to independently re-derive every critique on our raw data (`projects/ai-infrastructure/experiments/VERIFY_STUDY04_V1_2_STDOUT.txt`, SHA-256 `41b231195e5ea04ae94152e5e62b95638b5e23bbf428ec6a290ed4baa2c5b5f2`). Below, for each of the six `required_fixes`, we state exactly what we verified, what we changed in `Edition v1.2`, and the exact machine output supporting it.

---

2. Point-by-Point Implementation & Machine Output for the Six Required Fixes

Required Fix 1: Report Reliability on Live Ballots Only; Re-Score or Drop Rule-Fallback Ballots; Correct the ICC Claims in the Abstract and Table 2

- **What We Changed:** 1. In `projects/ai-infrastructure/scripts/annotate_and_analyze_observatory.py` and `build_and_verify_v1_2.py`, we diagnosed why `ollama-cloud` reasoning models fell back in `v1.0`: `max_tokens = 500` truncated `deepseek-v4.1-flash` when chain-of-thought reasoning exceeded 500 tokens before the JSON block. Raising `max_tokens` to `1500` achieved `100.0%` (`102 / 102`) live ballot completion on the rebuilt 51-record Federal Register sample (`0` fallbacks). 2. For all N = 364 baseline records, we dropped all `824 / 1,456` (`56.59%`) `deterministic_rule_fallback` ballot slots (`195` on `DeepSeek`, `354` on `GLM`, `275` on `MiniMax`, `0` on `JEV`) and computed pure live-ballot-only scores (`live_resource_ratepayer_salience`, `live_public_pressure_index`, `live_frame_*`, `live_ballot_count`) averaging strictly across live LLM ballots (`fallback_used == False`, where `174` records have `1` live ballot, `118` have `2`, `66` have `3`, and `6` have `4`). 3. We updated the Abstract, Section 3.2, and Table 2 of `PAPER_MANUSCRIPT.md` (`v1.2`) and `annotation_validation.json` to report Shrout & Fleiss (1979) `ICC(1,1)` – `ICC(3,k)` and mean pairwise r strictly on **live LLM ballots only** (`n = 169` 2-judge live `DeepSeek + JEV` and `n = 68` 3-judge live `DeepSeek + MiniMax + JEV`, plus `n = 51` rebuilt federal live 2-judge), alongside the historical all-4-slot column explicitly labeled as fallback-contaminated.
- **Supporting Machine Output (`VERIFY_STUDY04_V1_2_STDOUT.txt` `[FIX-1]`):**
- Ballot slot breakdown (N = 364 × 4 = 1,456):
  - `ollama-cloud/deepseek-v4.1-flash`: `live = 169/364 (46.43%)`, `fallback = 195/364 (53.57%)`
  - `ollama-cloud/glm-5.3-flash`: `live = 10/364 (2.75%)`, `fallback = 354/364 (97.25%)`

  - `ollama-cloud/minimax-m3.1-flash`: `live = 89/364 (24.45%)`, `fallback = 275/364 (75.55%)`
  - `jev/typesafe-jev-router`: `live = 364/364 (100.00%)`, `fallback = 0/364 (0.00%)`
  - `TOTAL BALLOT SLOTS`: `live = 632/1456 (43.41%)`, `fallback = 824/1456 (56.59%)`
  - Shared fallback inflation (`GLM` vs `MiniMax` on `n = 271` shared fallback rows): `r(PPI) = 1.0000`, `ICC(2,1) = 0.9957`, `ICC(2,2) = 0.9979`.

- **Live 2-Judge (`DeepSeek + JEV`, `n = 169`, `0%` fallback) vs. Live 3-Judge (`DeepSeek + MiniMax + JEV`, `n = 68`, `0%` fallback) Reliability:**
  - `public_pressure_index`: 2-Judge `ICC(2,1) = 0.8799`, `ICC(2,2) = 0.9361`, `ICC(3,2) = 0.9467`, `r = 0.9045` | 3-Judge `ICC(2,1) = 0.9107`, `ICC(2,3) = 0.9683`, `r = 0.9328`
  - `policy_posture_code`: 2-Judge `ICC(2,1) = 0.8511`, `ICC(2,2) = 0.9196`, `ICC(3,2) = 0.9201`, `r = 0.8521` | 3-Judge `ICC(2,1) = 0.9482`, `ICC(2,3) = 0.9821`, `r = 0.9496`
  - `mobilization_intensity`: 2-Judge `ICC(2,1) = 0.8939`, `ICC(2,2) = 0.9440`, `r = 0.9097` | 3-Judge `ICC(2,1) = 0.9261`, `ICC(2,3) = 0.9741`, `r = 0.9403`
  - `institutional_friction_level`: 2-Judge `ICC(2,1) = 0.8626`, `ICC(2,2) = 0.9263`, `r = 0.8748` | 3-Judge `ICC(2,1) = 0.9035`, `ICC(2,3) = 0.9656`, `r = 0.9163`
  - `frame_labor_civil_rights`: 2-Judge `ICC(2,1) = 0.7228`, `ICC(2,2) = 0.8391`, `r = 0.7415` | 3-Judge `ICC(2,1) = 0.7689`, `ICC(2,3) = 0.9089`, `r = 0.8087`
  - `concession_or_guardrail_demand`: 2-Judge `ICC(2,1) = 0.6800`, `ICC(2,2) = 0.8095`, `r = 0.7307` | 3-Judge `ICC(2,1) = 0.6976`, `ICC(2,3) = 0.8738`, `r = 0.7469`
  - `frame_water_land_env`: 2-Judge `ICC(2,1) = 0.6454`, `ICC(2,2) = 0.7845`, `r = 0.6541` | 3-Judge `ICC(2,1) = 0.6940`, `ICC(2,3) = 0.8719`, `r = 0.7069`
  - `frame_grid_ratepayer`: 2-Judge `ICC(2,1) = 0.5750`, `ICC(2,2) = 0.7302`, `ICC(3,2) = 0.7685`, `r = 0.6529` | 3-Judge `ICC(2,1) = 0.5968`, `ICC(2,3) = 0.8163`, `r = 0.6788`
  - `frame_transparency_nda`: 2-Judge `ICC(2,1) = 0.5195`, `ICC(2,2) = 0.6838`, `r = 0.5540` | 3-Judge `ICC(2,1) = 0.4357`, `ICC(2,3) = 0.6985`, `r = 0.5426`
  - `frame_natsec_competitiveness`: 2-Judge `ICC(2,1) = 0.5045`, `ICC(2,2) = 0.6707`, `ICC(3,2) = 0.7279`, `r = 0.6769` | 3-Judge `ICC(2,1) = 0.6339`, `ICC(2,3) = 0.8385`, `r = 0.7683` (Rebuilt Federal `n = 51` 2-Judge: `ICC(2,1) = 0.8465`, `ICC(2,2) = 0.9169`, `r = 0.8920`)
  - `frame_fiscal_tax`: 2-Judge `ICC(2,1) = 0.4770`, `ICC(2,2) = 0.6459`, `r = 0.6036` | 3-Judge `ICC(2,1) = 0.4648`, `ICC(2,3) = 0.7226`, `r = 0.6524`

  - `resource_ratepayer_salience`: 2-Judge `ICC(2,1) = 0.4607`, `ICC(2,2) = 0.6308`, `ICC(3,2) = 0.7496`, `r = 0.5998` | 3-Judge `ICC(2,1) = 0.4763`, `ICC(2,3) = 0.7320`, `ICC(3,3) = 0.8087`, `r = 0.5944` (Rebuilt Federal `n = 51` 2-Judge: `ICC(2,1) = 0.5926`, `ICC(2,2) = 0.7442`, `r = 0.6566`)

- **Live-Ballot-Only Regressions on Baseline N = 364, G = 48:**
  - `P1_live_ctrl`: `coef = +0.1283`, `SE = 0.0391`, `t = +3.2819`, `p = 0.0010`, `p_wcb = 0.0170`, `R2 = 0.1456`
  - `P2_live_ctrl`: `coef = +55.0318`, `SE = 10.7657`, `t = +5.1118`, `p = 0.0000`, `p_wcb = 0.0004`, `R2 = 0.4533` (`labor`: `coef = +11.6156`, `SE = 11.4628`, `t = +1.0133`, `p = 0.3109`)
  - `P2_live_biv`: `coef = +66.5968`, `SE = 8.7356`, `t = +7.6236`, `p = 0.0000`, `R2 = 0.2582`
  - `P3_live_ctrl`: `coef = +8.6461`, `SE = 4.6084`, `t = +1.8762`, `p = 0.0606`, `p_wcb = 0.0956`, `R2 = 0.2400`

- **Cross-Rater Live `P2` (`n = 169`, `DeepSeek` vs. `JEV`):**
  - `P2_cross_JEV_to_DS_ctrl`: `coef = +33.3306`, `SE = 14.1124`, `t = +2.3618`, `p = 0.0182`, `R2 = 0.2704`
  - `P2_cross_DS_to_JEV_ctrl`: `coef = +57.6086`, `SE = 14.9013`, `t = +3.8660`, `p = 0.0001`, `R2 = 0.4547`

Required Fix 2: Derive Policy Posture from the Mode or Median of Live Codes, Not the Rounded Mean of All Four; Restate the `"1.1% Direct Denial"` and `"Conditional Bargaining Dominates"` Claims

- **What We Changed:** 1. We replaced `int(round(mean(all_4_codes)))` with both the **Live Median** (`policy_posture_live_median_code`, rounded half-up to integer `1..5`) and **Live Mode** (`policy_posture_live_mode_code`, ties broken by `100%`-live `JEV`) across live LLM ballots (`fallback_used == False`), and added both columns to `ai_infrastructure_observatory_panel.csv` and `ai_infrastructure_observatory_panel_v1_2.csv`. 2. We updated the Abstract, Section 3.3, Section 4.3, Table 5, and Section 6 of `PAPER_MANUSCRIPT.md` (`v1.2`) to restate both claims:
  - **Direct denial (`Code 5`) is `7.42%` (`27 / 364`) under Live Median and `8.79%` (`32 / 364`) under Live Mode / `JEV` overall (`7.00%` / `17 / 243` under Live Median and `9.88%` / `24 / 243` under Live Mode on the cleaned `v1.2` panel)**, and rises to `30.34%` (`27 / 89`) **under Live Median and `35.96%` (`32 / 89`) under Live Mode** among high-pressure dockets (`live_ppi >= 50`; `24.29%–31.43%` on cleaned `v1.2`). The `v1.0`/`v1.1` figure of `1.10%` (`4 / 364`) was an averaging artifact caused by fallback codes of `3` pulling live `Code 5` ballots down to `mean(3,3,5,5) = 4.0 -> Code 4`.

- **Conditional bargaining and procedural pauses (`Codes 2 + 3`) remain the plurality response** (`45.33%` / `165 / 364` under Live Median and `45.05%` / `164 / 364` under Live Mode overall; `52.26%` / `127 / 243` under Live Median and `48.97%` / `119 / 243` under Live Mode on the cleaned `v1.2` panel; and `41.57%–42.70%` vs. `30.34%–35.96%` for `Code 5` in high-pressure dockets), but direct denials represent nearly **one-third of high-pressure conflicts** rather than a rare 1% outlier.

• **Supporting Machine Output (`VERIFY_STUDY04_V1_2_STDOUT.txt` `[FIX-2]`):**

• **Baseline Panel (N = 364):**

- `All-4 Rounded Mean (v1.0/v1.1)`: `{1: 160, 2: 135, 3: 35, 4: 30, 5: 4}` | `Code 2+3 = 170/364 (46.70%)` | `Code 5 = 4/364 (1.10%)`
- `Live Median (half-up)`: `{1: 141, 2: 103, 3: 62, 4: 31, 5: 27}` | `Code 2+3 = 165/364 (45.33%)` | `Code 5 = 27/364 (7.42%)` (`141 = 38.74%`, `103 = 28.30%`, `62 = 17.03%`, `31 = 8.52%`, `27 = 7.42%`)
- `Live Mode (JEV tie-break)`: `{1: 144, 2: 100, 3: 64, 4: 24, 5: 32}` | `Code 2+3 = 164/364 (45.05%)` | `Code 5 = 32/364 (8.79%)` (`144 = 39.56%`, `100 = 27.47%`, `64 = 17.58%`, `24 = 6.59%`, `32 = 8.79%`)
- `JEV Standalone (100% live)`: `{1: 140, 2: 105, 3: 63, 4: 24, 5: 32}` | `Code 2+3 = 168/364 (46.15%)` | `Code 5 = 32/364 (8.79%)`
- `DeepSeek Live (n=169)`: `{1: 87, 2: 29, 3: 24, 4: 13, 5: 16}` (`Code 5 = 9.47%`)
- `MiniMax Live (n=89)`: `{1: 35, 2: 31, 3: 9, 4: 6, 5: 8}` (`Code 5 = 8.99%`)
- `High-pressure live_ppi >= 50 (N=89) Live Median`: `{1: 1, 2: 14, 3: 23, 4: 24, 5: 27}` | `Code 2+3 = 37/89 (41.57%)` | `Code 4 = 24/89 (26.97%)` | `Code 5 = 27/89 (30.34%)`
- `High-pressure live_ppi >= 50 (N=89) Live Mode`: `{1: 1, 2: 15, 3: 23, 4: 18, 5: 32}` | `Code 2+3 = 38/89 (42.70%)` | `Code 4 = 18/89 (20.22%)` | `Code 5 = 32/89 (35.96%)`
- `High-pressure j4_ppi >= 50 (N=90) JEV`: `{1: 1, 2: 13, 3: 24, 4: 20, 5: 32}` | `Code 2+3 = 37/90 (41.11%)` | `Code 4 = 20/90 (22.22%)` | `Code 5 = 32/90 (35.56%)`

• **Cleaned `v1.2` Panel (N = 243):**

- `v1.2 Live Median`: `{1: 68, 2: 77, 3: 50, 4: 31, 5: 17}` | `Code 2+3 = 127/243 (52.26%)` | `Code 5 = 17/243 (7.00%)`
- `v1.2 Live Mode`: `{1: 78, 2: 69, 3: 50, 4: 22, 5: 24}` | `Code 2+3 = 119/243 (48.97%)` | `Code 5 = 24/243 (9.88%)`

- `v1.2 High-pressure live_ppi >= 50 (N=70) Live Median`: `{1: 0, 2: 13, 3: 17, 4: 23, 5: 17}` | `Code 2+3 = 30/70 (42.86%)` | `Code 4 = 23/70 (32.86%)` | `Code 5 = 17/70 (24.29%)`
- `v1.2 High-pressure live_ppi >= 50 (N=70) Live Mode`: `{1: 1, 2: 14, 3: 18, 4: 15, 5: 22}` | `Code 2+3 = 32/70 (45.71%)` | `Code 4 = 15/70 (21.43%)` | `Code 5 = 22/70 (31.43%)`

Required Fix 3: Rebuild the Federal Sample from Relevance-Ranked or Phrase-Matched Documents Across 2023–2026; Re-Test `P1` with Wild-Cluster Inference and Report It

- **What We Changed:** 1. We implemented a 4,999-replication Rademacher wild-cluster restricted bootstrap (`seed = 20261106`) for `P1`, `P2`, and `P3` across both the baseline N = 364 (G = 48) panel and the cleaned `v1.2` N = 243 (G = 44) panel. 2. We diagnosed why the `v1.0` Federal Register sample (`n = 63`, all in 2026) missed landmark 2023–2025 federal AI-infrastructure actions: (a) `per_page=25` sorted newest-first without year stratification truncated every query in 2026, and (b) Presidential Documents (Executive Orders) return `abstract: null` in the Federal Register API and were skipped by `collect_and_verify_corpus.py`. 3. In `build_and_verify_v1_2.py`, we rebuilt the federal comparison group across all four years `2023–2026` (`n = 51`: `5` in 2023, `13` in 2024, `14` in 2025, `19` in 2026; cached in `data/rebuilt_federal_2023_2026_raw.json` and `data/rebuilt_federal_live_ballots.json`), fetching full text from `raw_text_url` / `body_html_url` for Presidential Documents and scoring all 51 documents `100%` live with `ollama-cloud/deepseek-v4.1-flash` and `jev/typesafe-jev-router` (`102 / 102` live ballots, `0` fallbacks). The rebuilt sample includes:
  - Executive Order `14318` (*Accelerating Federal Permitting of Data Center Infrastructure*, `FR-2025-14212`, `2025-07-28`)
  - Executive Order `14141` (*Advancing United States Leadership in Artificial Intelligence Infrastructure*, `FR-2025-01395`, `2025-01-17`)
  - Executive Order `14110` (*Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, `FR-2023-24283`, `2023-11-01`)
  - Executive Order `14365` (*Ensuring a National Policy Framework for Artificial Intelligence*, `FR-2025-23092`, `2025-12-16`)
  - DOE *Request for Information on Artificial Intelligence Infrastructure on DOE Lands* (`FR-2025-05936`, `2025-04-07`) and *Request for Information: Accelerating Speed to Power* (`FR-2025-18058`, `2025-09-18`)

- NTIA *Request for Comments on Bolstering Data Center Growth, Resilience, and Security* (`FR-2024-19524`, `2024-09-04`) and *Data Center Growth Listening Session* (`FR-2025-11170`, `2025-06-17`)
- BIS *Data Center Validated End User Authorization* (`FR-2024-22587`, `2024-10-02`) and *Framework for Artificial Intelligence Diffusion* (`FR-2025-00636`, `2025-01-15`)
- FERC Order 1920 (*Building for the Future Through Electric Regional Transmission Planning and Cost Allocation*, `FR-2024-10872`, `2024-06-11`) and EPA power-plant/stationary combustion turbine rules (`FR-2024-09233`, `FR-2024-29233`, `FR-2026-00974`). 4. **We explicitly updated the Abstract, Section 4.1, Table 4, Table 4B, and Section 6 to report that `P1` is REVISED TO NULL in `Edition v1.2`.**

• **Supporting Machine Output (`VERIFY_STUDY04_V1_2_STDOUT.txt` `[FIX-3]`):**

• **Baseline N = 364 Wild-Cluster Bootstrap (`B = 4,999`) & Energy-Bucket Check:**

- `P1 (all-4-slot, N=364, G=48)`: `coef = +0.0769`, `CR1 SE = 0.0347`, `t = +2.2144`, `CR1 p = 0.0268`, **`WCR bootstrap p = 0.1740`**
- `P2 (all-4-slot, N=364, G=48)`: `coef = +26.2713`, `CR1 SE = 9.8153`, `t = +2.6766`, `CR1 p = 0.0074`, **`WCR bootstrap p = 0.0282`**
- `P3 (all-4-slot, N=364, G=48)`: `coef = +4.8119`, `CR1 SE = 2.5395`, `t = +1.8948`, `CR1 p = 0.0581`, **`WCR bootstrap p = 0.1022`**
- Federal salience by bucket (`N = 63`): `US_FERC_DOE_Grid = 0.2318 (n=16)`, `US_EPA_Env = 0.2239 (n=16)`, `US_DOE_WhiteHouse_AI = 0.0680 (n=16)`, `US_NTIA_Commerce_OMB = 0.0740 (n=15)`
- `P1_vs_EPA_FERC_only (N=333, G=46)`: `coef = -0.0023`, `SE = 0.0232`, `t = -0.0993`, `p = 0.9209`

• **Rebuilt `2023–2026` Federal Sample (`n = 51`) + Cleaned On-Topic Local/State (`n = 192`) = Cleaned `v1.2` Panel (N = 243, G = 44):**

- Mean `live_resource_ratepayer_salience`: Local/State `0.3652` (`Tier_1_Local = 0.3401`, `Tier_2_State = 0.4404`) vs. Rebuilt Federal `0.3133` (`US_FERC_DOE_Grid = 0.4362`, `US_EPA_Env = 0.4158`, `US_DOE_WhiteHouse_AI = 0.2817`, `US_NTIA_Commerce_OMB = 0.1558`)
- **`P1_v12_ctrl` (N = 243, G = 44): `coef = +0.0502`, `SE = 0.0650`, `t = +0.7721`, `p = 0.4401`, `p_wcb = 0.6264`, `R2 = 0.0488`**
- **`P1_v12_biv` (N = 243, G = 44): `coef = +0.0519`, `SE = 0.0574`, `t = +0.9054`, `p = 0.3653`, `R2 = 0.0110`**
- Disaggregated frame divergence on N = 243: Federal vs. Local/State `live_frame_natsec_competitiveness`: `0.5067` vs. `0.1449` (`3.50x`, `coef = -0.3547`, `SE =`

`0.0369`, `t = -9.6197`, `p = 0.0000`); Local/State vs. Federal `live_frame_fiscal_tax`: `0.2449` vs. `0.0614` (`3.99x`, `coef = +0.1726`, `SE = 0.0398`, `t = +4.3401`, `p = 0.0000`).

Required Fix 4: Filter Records for Relevance Before Scoring; Label News, Law-Firm and Vendor Pages as Secondary Sources; Date Records by Event, Not Page; De-Duplicate Events (`AII-2024-021` / `-022`); Remove the FY1984 Budget Record `AII-2023-013`

- **What We Changed:** 1. We applied an explicit AI/data-center/utility-load relevance filter (`is_on_topic_ai_dc`) to the baseline N = 364 panel (`225` on-topic records: `198` of `301` Local/State + `27` of `63` Federal; vs. `139` off-topic municipal or federal records). 2. We removed `AII-2023-013` (Richmond `ORD. 1984-088-102`, an FY1984–85 Electric Utility Budget ordinance entered into Legistar on `2023-08-24`, `is_pre2023_legacy_artifact = True`) and de-duplicated `5` duplicate event records (`AII-2024-022` duplicating `AII-2024-021` on Illinois SB `3686`, `AII-2026-009` duplicating `AII-2026-008` on Bessemer AL, `AII-2026-109` duplicating `AII-2026-105` on Chandler AZ, `AII-2026-149` duplicating `AII-2026-148` on St. Charles MO, and `AII-2026-154` duplicating `AII-2026-153` on Port Washington WI). 3. We corrected `corrected_event_date` on `9` retrospective articles so records are dated by the underlying council/utility action rather than publication date (e.g., `AII-2026-105` Chandler AZ ordinance adopted `2025-12-11` -> `2025`; `AII-2026-148` St. Charles MO moratorium adopted `2025-08-19` -> `2025`; `AII-2026-201` Prince William County VA Digital Gateway vote `2023-12-12` -> `2023`; `AII-2025-002` PECOS `2024-10-16` -> `2024`). 4. We labeled `source_provenance_type` across every row (`primary_government_record` vs. `secondary_reportage_or_legal_analysis`) and shipped the cleaned on-topic panel in `data/ai_infrastructure_observatory_panel_v1_2.csv` (N = 243, G = 44: `192` cleaned Local/State + `51` rebuilt 2023–2026 Federal; `183` primary government records + `60` secondary reportage/legal-analysis records).
- **Supporting Machine Output (`VERIFY_STUDY04_V1_2_STDOUT.txt` `[FIX-4]`):**
- **Why All-4-Slot `P2` Attenuated on `n = 225` While Live-Ballot `P2` Holds on `n = 225` and Cleaned N = 243:**
  - `P2_all4_ontopic_ctrl (N=225, G=44)`: `coef = +11.9037`, `SE = 14.2815`, `t = +0.8335`, `p = 0.4046`, `R2 = 0.2581` (Bivariate: `coef = +27.9476`, `SE = 13.3258`, `t = +2.0973`, `p = 0.0360`)
  - `P2_all4_offtopic_ctrl (N=139, G=32)`: `coef = +55.0222`, `SE = 4.7668`, `t = +11.5427`, `p = 0.0000`, `R2 = 0.5666` (Interaction `salience * is_on_topic_ai_dc`: `coef = -18.3017`, `SE = 15.9569`, `t = -1.1469`, `p = 0.2514`)
  - **`P2_live_ontopic_ctrl (N=225, G=44, 0% fallback)`: `coef = +30.0137`, `SE = 14.2031`, `t = +2.1132`, `p = 0.0346`, `R2 = 0.3713`**

  - **`P2_live_ontopic_biv (N=225, G=44, 0% fallback)`: `coef = +51.8796`, `SE = 12.9059`, `t = +4.0198`, `p = 0.0001`, `R2 = 0.1656`**

- **Cleaned `v1.2` On-Topic Panel (N = 243, G = 44: Deduplicated, Event-Dated, Rebuilt 2023–2026 Federal):**
  - **`P2_v12_ctrl (N=243, G=44)`: `coef = +48.1730`, `SE = 12.5497`, `t = +3.8386`, `p = 0.0001`, `p_wcb = 0.0028`, `R2 = 0.4278`** (`labor`: `coef = +9.9039`, `SE = 13.9498`, `t = +0.7100`, `p = 0.4777`)
  - **`P2_v12_biv (N=243, G=44)`: `coef = +61.3411`, `SE = 13.0849`, `t = +4.6879`, `p = 0.0000`, `R2 = 0.1930`**
  - **`P2_v12_primary_gov_only (N=183, G=32)`: `coef = +54.0006`, `SE = 13.6251`, `t = +3.9633`, `p = 0.0001`, `R2 = 0.3407`**
  - **`P3_v12_ctrl (N=243, G=44)`: `coef = -3.1579`, `SE = 3.5529`, `t = -0.8888`, `p = 0.3741`, `p_wcb = 0.4426`, `R2 = 0.2896`** (`2023–2024` N = 45 mean `live_ppi = 36.3222` vs. `2025–2026` N = 198 mean `live_ppi = 35.4217`)
  - **`P3_v12_grid (N=243, G=44)`: `coef = -0.0252`, `SE = 0.0442`, `t = -0.5701`, `p = 0.5686`, `R2 = 0.1235`**

Required Fix 5: Correct the Full-Rank SE Sentence (`9.8013` / `t 2.6804`; `P3` `2.5359` / `t 1.8975`, as in Our Own Stdout) and the §4.4 Ranking (`US_PublicRadio_LocalBeat` Has 10 Records, Mean `31.80`)

- **What We Changed:** 1. In Section 3.4 of `PAPER_MANUSCRIPT.md` (`v1.2`), we corrected the full-rank SE sentence so that it reports the exact `P2` and `P3` full-rank standard errors and t-statistics printed in our own verification stdout (`P2`: full-rank SE = 9.8013, t = +2.6804, p = 0.0074; `P3`: full-rank SE = 2.5359, t = +1.8975, p = 0.0578). 2. In `projects/ai-infrastructure/scripts/annotate_and_analyze_observatory.py` (line 552) and `build_and_verify_v1_2.py`, we changed the `j_stats` grouping key from `(jurisdiction_id, tier, state, region)` to `jurisdiction_id` alone so multi-state `jurisdiction_id` clusters (`US_PublicRadio_LocalBeat`, `MW_Local_Counties`, `PA_State_Leg_PUC`) are never split across state codes, and updated Section 4.4 and Table 3 of `PAPER_MANUSCRIPT.md` (`v1.2`).

- **Supporting Machine Output (`VERIFY_STUDY04_V1_2_STDOUT.txt` `[FIX-5]`):**

- `P2 full-rank (dropping Federal dummy)`: `coef = +26.2713`, `SE = 9.8013`, `t = +2.6804`, `p = 0.0074` (vs. rank-deficient `SE = 9.8153`, `t = +2.6766`, `p = 0.0074`)

- `P3 full-rank (dropping Federal dummy)`: `coef = +4.8119`, `SE = 2.5359`, `t = +1.8975`, `p = 0.0578` (vs. rank-deficient `SE = 2.5395`, `t = +1.8948`, `p = 0.0581`)

- Corrected Top 12 Jurisdictions by `mean_pressure` (`G = 48` grouped strictly by `jurisdiction_id`):
  - `#1 MO_StCharles`: `n=1`, `mean_ppi=52.1250`, `live_ppi=75.0000`
  - `#2 NJ_Newark`: `n=1`, `mean_ppi=52.2625`, `live_ppi=75.0000`
  - `#3 OH_State_PUCO`: `n=3`, `mean_ppi=42.8708`, `live_ppi=56.2500`
  - `#4 MW_Local_Counties`: `n=6`, `mean_ppi=40.7646`, `live_ppi=61.1042`
  - `#5 OR_Hillsboro_Port`: `n=5`, `mean_ppi=37.4600`, `live_ppi=52.0000`
  - `#6 MD_Baltimore`: `n=3`, `mean_ppi=36.6025`, `live_ppi=48.3333`
  - `#7 GA_Atlanta`: `n=4`, `mean_ppi=34.9775`, `live_ppi=48.6250`
  - `#8 VA_Fairfax`: `n=4`, `mean_ppi=32.6400`, `live_ppi=41.5000`
  - `#9 VA_Fauquier_Henrico`: `n=4`, `mean_ppi=32.3125`, `live_ppi=39.6875`
  - `#10 CA_State_Leg_CPUC`: `n=5`, `mean_ppi=31.8100`, `live_ppi=45.4000`
  - **`#11 US_PublicRadio_LocalBeat`: `n=10`, `mean_ppi=31.7975` (`31.80`), `live_ppi=45.5500`**

Required Fix 6: State in the Design Card How It Was Drafted Relative to the Endpoint Probe That Produced the Frozen-Endpoint Dates and Clerk Typos

- **What We Changed:** 1. We preserved `STUDY_DESIGN_CARD_V1_PRE_DATA_LOCKED.md` byte-identical at SHA-256 `bb6b1b709c5d99c1d9ebe6f954bfd99ca646879a332ec96b6d4fa29bf7b462b9` (`SKILL-025`). 2. We added Section 6 (*Endpoint-Probe Provenance Disclosure & Edition `v1.2` Audit Amendment Log*, `SKILL-024`) to `STUDY_DESIGN_CARD.md` (SHA-256 `a4f4119b4328f59b866f604ec43389eb9c88c137babe90e44cd1b35131482868`) and disclosed in Section 1.2 and Section 2.1 of `PAPER_MANUSCRIPT.md` (`v1.2`) that while git commit `08ef92e` (`2026-11-05T02:59:02Z`) preceded full corpus ingestion and 4-judge annotation commit `90f6d34` (`2026-11-05T15:02:32Z`), Section 2.1 of `STUDY_DESIGN_CARD.md` was drafted *after* an initial technical feasibility probe of the 30 municipal Legistar OData `/v1/{client}/matters` endpoints that identified the `24` active portals, `4` frozen portals (`goodyear` `2021-01-27`, `greensboro` `2021-06-15`, `sanantonio` `2021-09-24`, `longbeach` `2023-11-10`), `2` HTTP 500 portals (`chicago`, `atlanta`), and `9` municipal clerk future-date typos (`> 2026-11-05`). Thus `STUDY_DESIGN_CARD.md` was locked **post-endpoint-probe / pre-annotation-and-regression** rather than prior to touching any API endpoint.