

Regional Government Records Corpus

A Longitudinal Public-Meeting Archive for Local Governance Research in Virginia

Parent corpus overview · Version 3 · August 21, 2026

Tommy N. Turner · ORCID 0009-0005-7422-9881 · Williamsburg and Christiansburg, Virginia

What this is

The Regional Government Records Corpus is an independently assembled longitudinal archive of local-government public meeting materials spanning eleven Virginia jurisdictions across the Hampton Roads, New River Valley, and central Virginia regions.

The corpus supports reproducible research into local governance, public administration, land-use planning, budgeting, public participation, transcript analysis, and AI-assisted civic research methodology. Materials are normalized into a consistent naming and provenance framework so that researchers can trace policy decisions across agendas, minutes, transcripts, resolutions, ordinances, fiscal records, and related public documents.

Holdings total approximately **87,800 files**. Five jurisdictions are held at research depth. Two are held as comparison cases, and four as capability probes demonstrating that the construction method transfers to new localities at low marginal cost.

Coverage extends from the late nineteenth century through 2026 depending on document class. The strongest jurisdiction carries continuous minutes from 1978 and verbatim deliberation transcripts from 2012.

Holdings

Jurisdiction	Depth	Files	Anchor bodies
James City County	research	31,818	Board of Supervisors, Planning Commission, JCSA, six standing committees
City of Williamsburg	research	23,818	City Council and eighteen supporting bodies
Williamsburg–James City County Schools	research	16,006	Joint school division
York County	research	6,532	Board of Supervisors, Planning Commission, citizen boards, school division
City of Radford	research	6,074	City Council, EDA, regional authorities
Pulaski County	comparison	722	Board of Supervisors
Montgomery County	comparison	2,547	Board of Supervisors
Blacksburg, Christiansburg, Charlottesville,	probe	~200 combined	capability probes

Jurisdiction	Depth	Files	Anchor bodies
Harrisonburg			
Virginia planning-association materials	reference	93	statewide

Record classes

Meeting minutes; agenda packets and per-item attachments; verbatim deliberation transcripts; adopted resolutions and ordinances; operating budgets and capital improvement programs; Annual Comprehensive Financial Reports and audit correspondence; rezoning proffers; regional and joint authority records; state oversight material; and pre-portal records recovered through public records request.

Construction method

Records are retrieved from public channels: municipal meeting portals, government document repositories, official video channels, state financial-reporting systems, and public records requests where an online archive does not reach far enough back. Platform adapters exist for SharePoint, CivicEngage, CivicClerk, iCompass, BoardDocs, Granicus, Swagit, CivicMedia, and YouTube.

Every PDF carries a matching plain-text file with page boundaries preserved as explicit markers, so a quotation traces back to a page in the original. Digital PDFs are extracted directly; image-only scans are processed with optical character recognition. Meeting video is transcribed, with speech-model transcription applied where accuracy matters.

Each retrieval run writes a provenance manifest recording source URL, retrieval time, file size, and SHA-256 content hash. Failures are logged rather than passed over.

Scope of use and known limits

Vote tallies are taken from adopted minutes rather than from video transcripts. Automatically generated captions mishear roll calls, and they do so unevenly across speakers, which means a transcript-derived tally can be wrong in a non-random direction. Transcripts establish what was said; minutes establish what was decided.

Automatically generated captions carry no speaker attribution. Attributing a statement to a person requires the minutes or a timestamped review of the video.

Two recovery techniques are documented because they materially affect completeness. Adopted minutes are frequently reproduced inside the following meeting's agenda packet as a consent item, complete with roll-call tallies, which means an apparent gap in a minutes repository is often recoverable elsewhere. And minutes carry the date the body approved them rather than the date of the meeting they describe, typically one meeting cycle later; reading the approval stamp as the meeting date shifts events by weeks.

The decisive stage of most land-use matters occurs in pre-application consultation between applicants and staff, which generates no public record in any jurisdiction examined. Every document in this corpus is downstream of a process that was not written down.

Purpose

This corpus is a proof of concept for a statewide Virginia local-government records archive covering every county, city, and town. The construction methodology, platform adapters, naming conventions, and provenance framework transfer without

re-architecture. Three of the jurisdictions in the current holding were assembled in single-session builds from a documented seed brief.

The corpus serves as parent to derivative working papers, instructional case studies, governance analyses, methodological papers, and dataset extracts, each deposited separately and referencing this record.

AI-assisted construction

Anthropic Claude workflows were used extensively during corpus construction, retrieval scripting, normalization, and categorization. OpenAI systems including Whisper were used for transcription and transcript processing. Google Gemini and related systems were used selectively for external-source discovery and methodological review. AI-generated outputs were treated as provisional analytical aids and remained subordinate to verification against the underlying public records.

Rights

This deposit is the corpus overview document, licensed CC-BY-4.0. The underlying corpus compilation, indexing, canonical structure, metadata organization, original transcriptions, manifests, and derivative analytical materials remain © 2026 Tommy N. Turner.

Individual records within the corpus are public documents produced by Virginia local governments and are not subject to this claim.

Contact

Tommy N. Turner · tommy@gofast.pro · ORCID 0009-0005-7422-9881