

Regional Government Records Corpus

A longitudinal public-records archive of eleven Virginia localities, five held at research depth, with an institutional overview prepared for university partnership review

Volume I, Corpus Overview

Companion to Volume II, AI-Assisted Civic Research Methodology

Prepared by: Tommy N. Turner

ORCID: 0009-0005-7422-9881

Jurisdictions covered: Research depth: James City County • City of Williamsburg • WJCC School Board • York County • City of Radford. Comparison: Pulaski County • Montgomery County. Probes: Blacksburg • Christiansburg • Charlottesville • Harrisonburg

DOI (concept): 10.5281/zenodo.20088873

Document date: September 2026 (revision of August 2026 edition)

Status: For institutional review, university research-partnership discussion

Contact: tommy@gofast.pro

Confidentiality. *Prepared for review by prospective university partners and affiliated faculty. The underlying records are public; this overview document and its compilation, organization, methodology, and analytical framing are the original work of the author.*

Contents

1. Executive Summary.....	4
2. The Strategic Opportunity.....	5
3. Corpus at a Glance.....	6
4. Research Domains and Use Cases.....	7
5. Architecture and Methodological Anchor.....	10
6. Jurisdictional Detail.....	12
6.1 James City County.....	12
6.2 City of Williamsburg.....	14
6.3 Williamsburg–James City County School Board.....	16
6.4 York County, Virginia.....	17
6.5 City of Radford, Virginia.....	19
7. Cross-Jurisdictional Crosswalks.....	21
8. Demonstrated Analytical Yield.....	22
9. Known Limitations and Gap Inventory.....	24
10. Planned Expansion.....	25
11. Rights, Citation, and Contact.....	26

1. Executive Summary

The Regional Government Records Corpus is an independently assembled, source-grounded archive of local-government public-meeting and policy records across eleven Virginia localities. Five are held at research depth: James City County, the City of Williamsburg, the Williamsburg–James City County joint school division, York County, and the City of Radford. Pulaski and Montgomery counties are held as comparison cases, and four further localities as capability probes. The corpus currently contains roughly 87,800 files spanning meeting minutes, agenda packets, verbatim deliberation transcripts, adopted resolutions and ordinances, rezoning proffers, financial reports, operating budgets, capital improvement programs, public authority records, and adjacent regulatory filings. Coverage extends from the late 1970s in the strongest jurisdictions through the present, with selected captures as far back as 1887.

The corpus is not a document dump. Each file class is anchored to a canonical public source platform, normalized to a consistent filename and folder convention, deduplicated by SHA-256 content hash, paired with a JSONL provenance manifest, and crosswalked at the meeting-event level so minutes, packets, transcripts, votes, and follow-on actions can be retrieved together. The aggregate compilation, indexing, transcription, and methodological framework represent two years of independent effort built on the underlying public record.

The corpus has already produced peer-quality analytical work: two governance working papers on the James City County Board of Supervisors, a documentary monograph drafted from the October-2012-forward verbatim record, a longform investigative piece on the 2023 school-system breakup attempt and the 2025 reconciliation, and a five-deliverable package on Radford's March-2026 state fiscal-distress designation including an op-ed for The Roanoke Times, a senior-analyst forensic report, a quantitative analysis, and a normalized 23-tab longitudinal financial workbook covering FY2002 through FY2024.

Local government records are public. They are rarely usable as a research corpus without normalization, crosswalking, transcription, indexing, and provenance tracking. The Regional Government Records Corpus is the infrastructure that makes them usable.

This Volume I document describes the corpus at the level of institutional review. It opens with the strategic opportunity that motivated the build, summarizes scope, then walks jurisdiction by jurisdiction through the record classes, coverage windows, source platforms, known gaps, and demonstrated analytical applications. A companion Volume II document treats the AI-assisted research methodology in equivalent depth. The two documents are designed to be read separately or together by faculty, deans, research-center directors, journalism editors, and policy-program leadership evaluating a potential research partnership.

2. The Strategic Opportunity

2.1 Why local-government corpora matter, and why they are rare

American local government produces a continuous, structured, public stream of decisions that shape land use, school capacity, fiscal trajectory, public-safety policy, infrastructure investment, and the daily texture of civic life. Almost none of it is usable for longitudinal research without significant pre-processing. Minutes live on one platform, agenda packets on another, verbatim video on a third, financial reports on a fourth, and FOIA-only material on none of them. Filenames are inconsistent within a single body across years. Pre-2018 video is often archived in proprietary players that do not produce downloadable captions. Multi-source duplication is common but rarely deduplicated. A graduate student or a working journalist who wants to ask longitudinal questions of a single jurisdiction discovers within a week that the data layer does not exist.

The result is a research literature on local government that is overwhelmingly cross-sectional, case-survey, or interview-driven, and that struggles to deliver the kind of within-jurisdiction longitudinal analysis that would let practitioners and scholars trace decisions from staff recommendation through deliberation, public testimony, board action, fiscal consequence, and institutional memory. The corpus described here is a working answer to that gap, built on a methodology that scales replicably across additional jurisdictions.

2.2 Why now

Three conditions intersect at this moment. First, the public platforms hosting local-government records have matured. Most Virginia jurisdictions now operate on a small set of vendor stacks, CivicEngage, CivicClerk, Granicus, Swagit, BoardDocs, iCompass, and SharePoint, each of which exposes structured retrieval surfaces that allow scripted capture. Second, the cost of high-quality machine transcription has collapsed since the release of the OpenAI Whisper API, making verbatim coverage of pre-YouTube meeting eras feasible for the first time. Third, large-language-model assistants have become capable enough to act as research operations consoles, accelerating the corpus-construction, classification, indexing, and red-teaming workloads that previously required institutional staff. The combination makes a single independent researcher capable, with discipline, of producing infrastructure that twenty years ago would have required a funded center.

2.3 Why this corpus

The Regional Government Records Corpus is distinctive on four dimensions. It is multi-jurisdictional, covering eleven Virginia localities of varied scale, vendor stack, and fiscal posture, five of them at research depth. It is longitudinal, with continuous coverage windows of more than four decades in the strongest jurisdictions. It is verbatim where the public record allows, with audio-derived

transcripts back to October 2012 in the lead jurisdiction. It is meeting-anchored, with crosswalks that link minutes, packets, transcripts, votes, and follow-on actions at the event level. And it is methodologically auditable, with JSONL manifests, SHA-256 hashes, dated retrieval logs, and a transparent verification hierarchy that distinguishes official record from research aid.

3. Corpus at a Glance

The corpus currently spans roughly 87,800 files across eleven Virginia localities, five of them held at research depth. Headline numbers vary by record class, jurisdiction, and the underlying retention practice of the producing body. The table below presents the corpus by jurisdiction at the highest level; Section 6 disaggregates each jurisdiction by body and record class.

Jurisdiction	Anchor Body	Coverage Window	Files	Verbatim Era
James City County	Board of Supervisors	1978–2026 continuous	31,818	Oct 2012–present
City of Williamsburg	City Council + 11 bodies	2011–2026 (varies by body)	23,818	YouTube + Whisper
WJCC School Board	Joint School Division	2006–2026 (FOIA + BoardDocs)	16,006	Jul 2015–present
York County, Virginia	Board of Supervisors + 16 bodies	2010–2026 (FOIA + portals)	6,532	2015–present
City of Radford	City Council + regional auth.	2012–2026 (Wayback fills earlier)	6,074	2020–present
Cross-jurisdictional joins	Joint authorities, school funding	2018–2026	integrated	integrated

Counts reflect the current build pass and are version-specific. Some classes, agenda packets in particular, contain meeting-level rollup PDFs alongside per-item attachment PDFs, which inflates raw file counts above the number of distinct meetings. The figures above are intended for orienting institutional readers; the precise jurisdiction-by-jurisdiction breakdown follows in Section 6.

3.1 Record classes present

The corpus organizes records into a stable set of classes that recur across jurisdictions. Not every jurisdiction produces every class; the table below shows which classes are present where, at material scale.

Record Class	JCC	WMBG	WJCC	York	Radford
--------------	-----	------	------	------	---------

Meeting minutes	✓ continuous	✓ continuous	✓ FOIA + BoardDocs	✓ 2020+ portal, FOIA pre-2020	✓ AgendaCenter
Agenda packets	✓ continuous 2000+	✓ varies by body	✓ BoardDocs era	✓ 2015+ multi-platform	✓ AgendaCenter
Verbatim transcripts	✓ Oct 2012+	✓ selective bodies	✓ Jul 2015+	✓ 2015+ Swagit + YouTube	✓ 2020+ YouTube
Adopted resolutions	✓ 1993+ (6,756)	embedded in minutes	embedded in minutes	embedded in minutes	embedded
Adopted ordinances	✓ full (1,461)	embedded	n/a	embedded	embedded
ACFRs / audits	✓ FY2006, 2014+	partial (City)	embedded in school finance	✓ FY2010+	✓ FY2011+ continuous
Operating budgets	✓ FY2018+	partial	school transfer cycle	✓ FY2010+ continuous	✓ FY2014+ continuous
Capital improvement programs	✓ embedded	embedded	embedded	✓ standalone CIP series	embedded in budget
Rezoning proffers	✓ 1987–2025 (163)	minimal	n/a	present in packets	n/a
State oversight records	minimal	minimal	minimal	minimal	✓ APA + FAC integrated
Regional authority records	✓ JCSA	✓ HTRFA + transit	✓ entirely joint	✓ regional commissions	✓ NRVJRJ + 6 others
Pre-portal FOIA recoveries	✓ multiple	✓ City Clerk channel	✓ NextRequest 26-47	✓ pending Va. FOIA	✓ ArchiveCenter + Wayback

4. Research Domains and Use Cases

The corpus supports a wide variety of analytical work, organized below by domain and reader. The categorization is illustrative rather than restrictive; many of the most productive applications combine domains.

4.1 Public administration and governance research

Local-government boards are unusually generative research subjects because their behavior is recorded in dense, structured public deliberation. The corpus supports longitudinal study of board cohort dynamics across membership change, dissent and majority-formation patterns, staff-board interaction across decision sequences, and the institutional translation of public testimony into procedural action. The James City County Board of Supervisors record from January 2010 through February 2026, roughly 2,290 recorded votes, 940 named-commenter appearances by 468 distinct residents, and 388 coded minute files, is already a publishable analytic universe and has produced two working papers.

Representative research questions

- How do governing-body cohorts evolve their working majority across natural turnover and contested elections?
- Which procedural moves, work sessions, deferrals, motions to amend, postponement, predict vote-coalition change?
- Does public testimony alter outcomes, alter language, or alter neither, and how does that vary across decision type?
- When boards override staff recommendations, what predicts the override, and what are the fiscal or procedural consequences?

4.2 Public finance and fiscal-distress research

The Radford slice of the corpus is a fully assembled fiscal-distress case. It pairs the March 2026 Auditor of Public Accounts Recommendation Letter for State Assistance, the formal designation that triggered the Virginia distress framework, with twenty-three years of audited financial values normalized into a 23-tab longitudinal workbook covering FY2002 through FY2024. The financial layer integrates state-level monitoring documents, Federal Audit Clearinghouse filings, peer-city benchmarking (Hopewell selected as the closest analog), and regional-authority audit trails. The corpus permits one of the few open-source longitudinal datasets capable of supporting early-warning-signal research grounded in pre-distress public records.

Representative research questions

- How early do fiscal-distress signals become visible in public records before formal state intervention?
- Which audit findings, budget revisions, or fund-balance trajectories carry the strongest warning signal?
- How do governing bodies hear, translate, absorb, or proceduralize cumulative fiscal-burden testimony from residents?
- How does state intervention propagate through subsequent budget cycles, audit narratives, and council deliberation?

4.3 Planning, land use, and zoning research

Land-use decisions are unusually visible because they concentrate staff analysis, citizen testimony, and deliberation into structured public hearings. The corpus includes the JCC rezoning proffer set from 1987 forward, full Planning Commission minutes from 1976 in JCC, packet-level detail across multiple Hampton Roads jurisdictions, and the Special Use Permit and conditional-use record across the five research-depth jurisdictions. Verbatim transcripts make it possible to trace the migration of language from staff report through commission deliberation, board hearing, and follow-on litigation or institutional response.

Representative research questions

- Can meeting records trace the migration from zoning approvals to later school-capacity, infrastructure, traffic, or fiscal pressures?
- How do short-term-rental SUP cycles propagate decision rules across adjacent jurisdictions?
- What is the staff-board contradiction rate on contested rezoning, and does it cluster by property type or applicant class?
- Where deliberation language and minutes language diverge, what is being compressed out?

4.4 Public participation and citizen-voice research

Public-comment records are the connective tissue between democratic theory and observed institutional behavior. The corpus indexes named-commenter appearances by meeting and date in the JCC slice, where coding has been completed to the level of 940 named-commenter appearances across 468 distinct residents from January 2010 through February 2026. Verbatim capture is available for the post-October-2012 era; minutes-derived summary is available for the full window. The corpus permits analysis of testimony density, returning-speaker dynamics, topic concentration across decision sequences, and the procedural fate of repeated public objection.

Representative research questions

- Do governing bodies override sustained public opposition, and under what conditions?

- Does the same set of residents speak across multiple decision sequences, and what does that imply about civic capacity?
- How does named-coalition organizing, Save Rural James City County, Change.org Government Center, interact with formal participation channels?
- Does the compression from verbatim deliberation to formal minutes systematically remove certain language?

4.5 Joint governance and school-division research

The Williamsburg–James City County joint school division is governed by a school board appointed jointly by two distinct local governments under a shared funding agreement. The 2023 breakup attempt, the 2024 feasibility study, the prolonged 2025 contract negotiation, and the October 2025 reconciliation signing produced a dense record of inter-body interaction across three governing bodies. The corpus integrates all three rooms, Williamsburg City Council, James City County Board of Supervisors, and the WJCC School Board — through a longform piece already drafted from the verbatim record.

Representative research questions

- How does a joint school system distribute fiscal responsibility and political accountability across two parent localities?
- What procedural mechanisms hold a joint funding arrangement together when one parent threatens exit?
- How does the school division's deliberation reflect, ignore, or reframe parent-jurisdiction budget pressure?
- What does inter-body discourse look like during institutional rupture, and how does it recover?

4.6 Civic technology and AI-assisted research methodology

The corpus is also a secondary object of study. The construction workflow — discovery, scripted retrieval, normalization, transcription, indexing, AI-assisted extraction, hostile review, and human verification — is documented in Volume II to a level that supports replication. The methodology question itself is publishable: how can a documented, human-verified, AI-assisted workflow make local-government corpus construction replicable across counties and cities without sacrificing source provenance? Three of the five research-depth jurisdictions in this corpus were assembled in single-session builds from a documented methodology, demonstrating that the approach scales.

4.7 Pedagogy, teaching cases, and practitioner training

Local-government decisions translate well into MPA, planning-school, and law-school teaching cases because they concentrate institutional dilemma, public testimony, fiscal trade-off, and procedural choice into bounded decision sequences. Cases drawn from this corpus would use real Virginia

records rather than synthetic scenarios, would carry source-annotated exhibits drawn from the underlying record, and would support both discussion-leader notes and student-facing decision dilemmas. The JCC slice alone supports a dozen viable teaching cases on contested rezoning, capital-project financing, public-facility siting, and budget-versus-service trade-offs.

Representative deliverables

- Teaching cases with paired student handout and instructor notes
- Discussion exhibits: vote chronologies, public-comment thematic timelines, staff-board contradiction tables
- Replicable lab exercises for graduate methods courses on text analysis, transcript coding, and verification discipline
- Practitioner-facing workshops for clerks, planners, finance officers, and city managers on transparency and record retention

4.8 Journalism and civic investigation

The corpus has already been used to produce published or submission-ready investigative material. The Radford fiscal-distress reporting package combines the senior-analyst forensic report, a quantitative analysis, the longitudinal financial workbook, and a published op-ed in The Roanoke Times reaching the regional audience most directly responsible for oversight. The Three Rooms piece on the Williamsburg school-system breakup attempt draws on the verbatim record across three governing bodies. These products demonstrate that the corpus is not a research instrument only; it supports real journalism on real public-interest controversies, grounded in the underlying public record rather than interview reconstruction.

5. Architecture and Methodological Anchor

The corpus follows a source-first construction principle. Each document class has a canonical public source platform. Capture methods vary by platform, but the governing principle is consistent across all eleven localities and all record classes: pull from the official record, preserve the original filename and platform identifier where useful, normalize a canonical destination path, and record retrieval details in a manifest. Full methodological treatment is in the companion Volume II document; this section presents the anchors a reader needs to evaluate the corpus on its own terms.

5.1 Source platforms in active use

The corpus draws from a small set of vendor stacks that together cover the production of nearly all Virginia local-government records. The same stacks recur across all eleven localities, which means a build pipeline developed for one jurisdiction transfers to the next with modest adaptation.

Platform	Materials drawn from it	Where used
SharePoint document library	Minutes, agenda packets, resolutions, ordinances, proffers, committee materials	JCC
CivicEngage Archive Center	ACFRs, adopted budgets, narrative annual reports, department annual reports	JCC, Radford, York
CivicEngage AgendaCenter	Current-era agendas and minutes across 16+ bodies per locality	York, Radford
CivicClerk events portal	Post-migration agendas and supporting documents (2024 onward)	JCC
iCompass MeetingTypeList	City Council, ARB, BZA, EDA, and 17 other body records	Williamsburg
BoardDocs API	School-board agendas, minutes, attachments, video links	WJCC, York YCSD
Granicus video archive	Pre-2018 frozen video archives with paired agenda PDFs	York
Swagit voice-to-text	2017+ verbatim transcripts and paired agendas	York
earthchannel.com publisher feed	Pre-March-2018 BoS audio (Oct 2012–Jan 2018) for Whisper transcription	JCC
YouTube auto-caption track	BoS/Council/School Board	JCC, WMBG, WJCC, York

	transcripts, March 2018 to present	
Virginia Auditor of Public Accounts	APA Comparative Reports, Recommendation Letters, distress framework records	Radford
Federal Audit Clearinghouse API	Single Audit filings, auditee status, finding history	Radford peer benchmarking
Wayback Machine	Historical site captures, pre-portal recovery, legacy ArchiveCenter snapshots	all jurisdictions
Virginia FOIA / NextRequest	Records not retained on live portals (pre-BoardDocs, pre-2020, archived)	all jurisdictions

5.2 Filename convention and event-level crosswalk

Meeting-anchored files are renamed to a consistent ISO-date prefix with a meeting-type suffix in the form YYYY-MM-DD_{meeting_type}.{pdf|txt|srt|vtt}. The regular Board of Supervisors meeting minutes for July 14, 2015 land at minutes/bos/2015/2015-07-14_regular.pdf with a paired transcript at transcripts/2015-07-14_regular_eID314.{txt,srt} when verbatim deliberation was captured. The eID suffix on Whisper-derived transcripts preserves the source-publisher identifier, allowing back-trace to the original audio source. Resolution and ordinance filenames preserve the locality's original numbering scheme to permit traceability back to the public record.

This convention is the foundation of the meeting-event crosswalk. Minutes, packets, transcripts, votes, and follow-on actions for any given meeting can be retrieved together with a single date-keyed query. The vote register, which contains 1,542 roll-call votes parsed from the JCC minutes, joins back to the source minute file by date and motion sequence. This is the structural enabling condition for everything else the corpus supports.

5.3 Multi-source aggregation and deduplication

Several document classes appear on more than one source platform. JCC Board of Supervisors agenda packets, for example, exist both in the SharePoint document library for earlier years and in the CivicClerk events portal after the 2024 migration. The corpus retrieves both, computes a SHA-256 content hash on each file, and uses the hash to identify exact-content duplicates across sources. A keep-policy script collapses duplicates into a single canonical copy while preserving the alternate-source filename in a sidecar provenance record. This both reduces storage and preserves the audit trail back to either source platform.

5.4 Verbatim deliberation transcription

Verbatim deliberation capture uses two pipelines depending on era. From March 2018 onward, YouTube auto-captions are downloaded as VTT subtitle files alongside the matching meeting record. For the pre-2018 window, audio is downloaded directly from the publisher feed for each meeting and transcribed with the OpenAI Whisper API. In York County, a third pipeline harvests Swagit voice-to-text transcripts. In all cases the transcripts are research aids — not official records. The official record for vote tallies, motions, and procedural action remains the minutes.

5.5 Provenance and reproducibility

Every retrieval pass produces a JSONL manifest with one row per file: source URL, date of retrieval, file size, SHA-256 hash, and canonical destination path. Manifests are dated and stored alongside the data so that any future re-run can be reconciled against the original capture. Build scripts and intermediate working files are kept in a working bin separate from the canonical corpus tree, with citable provenance copies of scripts and manifests filed under a corpus-internal `_provenance` directory. Researchers using the corpus can reconstruct what was pulled, when, from where, and at what content hash.

6. Jurisdictional Detail

The following five subsections present each jurisdiction at the level of bodies covered, record classes captured, coverage windows, source platforms, known gaps, and demonstrated analytical applications. The jurisdictions are presented in order of build maturity, beginning with the lead jurisdiction.

6.1 James City County

James City County is the lead jurisdiction and the most fully developed slice of the corpus. The County operates a five-member Board of Supervisors elected from geographic districts, a Planning Commission, a Board of Zoning Appeals, the James City Service Authority (a publicly chartered water and wastewater utility), the Economic Development Authority, and a body of supporting departments and authorities. The County is governed under a council-manager form of government and is one of two parent localities of the Williamsburg–James City County joint school division (see Section 6.3).

Board of Supervisors

Minutes: 3,260 PDF files, scattered records back to 1887, continuous coverage 1978 through May 2026

Agenda packets: 1,961 PDFs, continuous 2000 through May 2026; post-2018 packets paired with text-extracted sidecars carrying per-page markers

Verbatim transcripts: 195 YouTube auto-caption VTTs covering March 2018–May 2026; 189 Whisper-transcribed audio files covering October 2012–March 2018

Vote register: 1,542 roll-call votes parsed from minutes (parser-coverage gap in 2015–2016; researchers cross-reference minutes for those years)

Public-comment index: 940 named-commenter appearances by 468 distinct residents, January 2010 through February 2026 (coded analytic universe for working-paper layer)

Planning Commission

Minutes: 501 PDFs covering 1976 through 2025 continuous

Agenda packets: 143 PDFs

Annual reports: 24 PDFs

Verbatim transcripts: Selected YouTube captions where available

County administrative and authority bodies

Board of Zoning Appeals: 158 agenda PDFs

James City Service Authority: 132 minutes PDFs

Economic Development Authority: Records present; EDA-specific ACFR coverage FY2020–FY2025

Fiscal and statutory record

Adopted resolutions: 6,756 PDFs covering 1993 through May 2026, indexed by resolution number

Adopted ordinances: 1,461 PDFs covering the same period

Rezoning proffers: 163 PDFs covering 1987–2025

ACFRs: 40 PDFs spanning County, EDA, and JCSA; FY2014 through FY2026 continuous, plus FY2006 recovered from unlisted archive ID

Adopted operating budgets: 21 PDFs covering FY2018 through FY2027 (proposed). Pre-FY2018 budgets confirmed by FOIA officer as not retained under Virginia Public Records Act schedule

Department annual reports: 47 PDFs: Police 2012–2024; General Services FY2010–FY2025; Planning Commission 2010–2021; Parks & Rec FY2012–FY2014; EDA FY2020–FY2025

Narrative annual reports: 28 PDFs covering historical 1974–2014 (with documented gaps) and modern FY2020–FY2025; 2015–2019 absent from County archive

Adjacent productions in the JCC channel

The corpus also includes the JCC podcast, *This Week in James City County*, inventoried at 234 episodes 2017–2026 and retained as audio with show-note metadata. Transcripts are on-demand rather than pre-built. FOIA correspondence is preserved in jcc/foia/. The petitions channel inventories citizen petitions filed with the County, including the named-coalition organizing artifacts — Save Rural James City County and Change.org Government Center — that figure in the override paper case set.

Demonstrated analytical applications, James City County

- The Working Majority: Public Dissent and Decision-Making on the JCC Board of Supervisors, 2010–2026 (working paper)
- A Board That Fought and Did Not Override: A Longitudinal Cohort Analysis of the JCC Board of Supervisors, 2010–2026 (working paper)
- Representative Judgment After Public Opposition: A James City County Companion Case on Capital Projects, Institutional Continuity, and Changing One's Mind (companion case)
- Mounts Bay: A James City County Story, 2012–2026 — completed monograph drawn from the verbatim deliberation record
- Documented analytic universe: 388 minute files, 295 agenda packets, ~2,290 votes, 940 commenter appearances

6.2 City of Williamsburg

The City of Williamsburg is the lead corpus's tourism-economy partner jurisdiction, sharing planning, school, service-authority, and recreation entanglements with James City County. The City operates a council-manager form of government with a five-member City Council and a particularly broad set of advisory boards, commissions, and authorities — twenty-one distinct bodies are represented in the corpus, the largest body-count of any jurisdiction in the build. The City record runs on the iCompass platform, supplemented by FOIA recoveries for sparse early periods.

Council and supporting bodies

City Council: 220 minutes PDFs covering 2011 through 2024 continuous

Council Work Sessions, Budget Retreats, Council Retreats, Special Called Meetings:

Approximately 150 additional minutes PDFs across these formats

Joint meetings: Captured where they occurred, including the 2018 Legacy Hall budget meeting and the October 10, 2025 joint contract-signing meeting at the JCC Recreation Center

Planning and land use

Planning Commission: 205 minutes PDFs

Planning Commission Work Sessions: Captured separately

Architectural Review Board: 516 minutes PDFs covering ARB and ARB special meetings — the largest single-body record in the City corpus

Board of Zoning Appeals and Board of Equalization: Captured under their respective channels

Economic development and tourism

Economic Development Authority: 264 minutes PDFs

EDA Business Roundtable: 28 audio recordings transcribed via Whisper covering February 2020 through October 2022 — pandemic-era small-business voice in a tourism-dependent City, a voice-led primary source not available in any other channel

Tourism Development Grant Review Committee: Captured under the tourism funding channel

Williamsburg Farmers' Market Advisory Board: Captured

Historic Triangle Recreational Facilities Authority: Captured (joint with JCC and York)

Arts, culture, and housing

Williamsburg-Area Arts Commission: Captured

Williamsburg Public Art Council: Captured

Williamsburg Redevelopment and Housing Authority: Captured

Agenda packets and supporting documents

Agenda packets across all bodies: 1,684 PDFs

Aggregate file count: Approximately 11,500 PDFs across the entire City build, organized by body

Coverage notes for the Williamsburg channel

The CIO confirmed in 2026 that the iCompass MeetingTypeList.aspx is the authoritative inventory surface for City records. All twenty bodies listed in that inventory are present in the corpus. Six bodies — ARB, BZA, WPAC, WAAC, BoE, TGR — operate as text-only deliberation channels by institutional choice, meaning the City does not produce or post video records for these bodies. Coverage gaps in older periods reflect the City's own retention practice; the City Clerk channel has been engaged via the iCompass portal and where necessary via Virginia FOIA to fill institutionally-documented gaps.

Demonstrated analytical applications, Williamsburg

- Three Rooms: Williamsburg and James City County and the Joint School System, 1955–2026 (longform investigative piece, drafted)
- EDA Business Roundtable pandemic-era small-business voice study (analytical seed, voice register validated)
- Cross-jurisdictional school-funding deliberation analysis through the 2023 breakup attempt and 2025 reconciliation
- ARB historic-district deliberation record — largest single-body record permits long-period style and norm analysis

6.3 Williamsburg–James City County School Board

The WJCC School Board is the joint school division governing the public schools across the City of Williamsburg and James City County. It operates under a joint funding contract between the two parent localities, with appointed members from each. The 2023 breakup attempt, the 2024 feasibility study, and the October 2025 contract signing produced an unusually rich record of inter-body interaction. Current WJCC records live on BoardDocs; the pre-BoardDocs era was recovered via Virginia FOIA.

BoardDocs era

Meeting index: 554 meetings covering 2015 through May 2026

Agenda HTML: 643 BoardDocs-rendered agenda files

Minutes: 12 PDFs plus 166 HTML minutes rendered through BoardDocs; HTML record is canonical for BoardDocs-era meetings

Attachments inventoried: 6,963 URLs available for on-demand retrieval rather than pulled en masse; raw URL list preserved in `wjcc/joins/`

Video captions: 205 YouTube VTT files covering July 2015 through May 2026

Pre-BoardDocs era (recovered via Virginia FOIA)

Coverage: Agendas and minutes for January 2006 through October 2008

Provenance: Virginia FOIA via NextRequest reference 26-47

Significance: Populates the channel that BoardDocs does not cover; the only public-records source for school-board deliberation in this window

Build infrastructure notes

The BoardDocs API was used to retrieve agenda and minute HTML and to inventory attachment URLs. The on-demand retrieval pattern for the 6,963 inventoried attachments avoids replicating thousands of files that are already available at stable URLs while preserving the ability to pull any specific attachment when needed for a particular analysis. Scripts and manifest are stored under `wjcc/joins/`.

Demonstrated analytical applications, WJCC

- Three Rooms longform: WJCC-side deliberation record across the 2023 breakup, 2024 feasibility study, and 2025 reconciliation
- Joint-funding-contract negotiation timeline, 2022 through October 2025 — verbatim deliberation across three governing bodies
- Pre-BoardDocs FOIA recovery as a methodological demonstration of how to fill a vendor-platform gap

- School-funding-cycle interactions between County and City on the annual transfer — case material on joint-governance fiscal mechanics

6.4 York County, Virginia

York County is adjacent to James City County on the east, shares the Bruton Heights school history, and sits in the Williamsburg–Yorktown employment corridor. York is the third Historic Triangle jurisdiction in the corpus. The build was completed in a single session in May 2026 from a mixed multi-platform stack — CivicEngage AgendaCenter for the current era, Granicus 2015–2017 archive, Swagit for the 2015–2026 BoS and PC archive, YouTube for streaming captions, and BoardDocs for the York County School Division. A Virginia FOIA request was filed May 14, 2026 to recover pre-2020 records; that request was substantially fulfilled on May 21, 2026, with 262 BoS minutes covering 2010–2019 delivered for a scope-narrowed fee.

Channels assembled

CivicEngage AgendaCenter: 1,266 PDFs across 16 county bodies including BoS, Planning Commission, Wetlands Board, Chesapeake Bay Board, Historic York District Committee, Library Board, EDA, Senior Center, Arts Commission, Board of Equalization, and others. Coverage concentrated 2020–2026 with scattered earlier entries.

Granicus 2015–2017 archive (frozen): 74 video clips with 71 paired agenda PDFs covering Board of Supervisors and Planning Commission

Swagit 2015–2026 canonical video archive: 356 video records with 285 agenda PDFs and 178 voice-to-text transcripts covering BoS and PC

YouTube @YorkCountyVirginiaUS: 49 auto-caption VTT files, sparse 2020–2024, dense 2025–2026

York County School Division BoardDocs: 607 meetings 2011–2026; 558 agenda HTML files; 304 minutes HTML; 6,546 attachment URLs inventoried

Financial records

Coverage: 53 PDFs covering FY2010 through FY2026: Adopted Budget, Capital Improvements Program, and ACFR series, plus the Auditor of Public Accounts distress-letter framework documentation

FOIA recovery (closed May 21, 2026)

Scope: 262 BoS minutes 2010–2019

Mechanism: Virginia FOIA via NextRequest #26-271; fee narrowed from full citizen-boards to BoS-only by negotiation

Result: Total York BoS minutes after ingestion: 415 files spanning 2010–2026 unbroken

Outstanding: PC and 14 citizen-board minutes 2010–2019 still pending; second FOIA request filed for that scope

Unified index

York meeting index v4: 2,623 rows joining all five channels by body and date at `york/joins/york_unified_meeting_index_2026-05-13.{jsonl,csv}`

Aggregate file count: Approximately 5,900 files in the York slice as of the May 21, 2026 build

Demonstrated analytical applications, York

- Comparative Historic Triangle governance — three adjacent localities under different vendor stacks and different transparency postures
- Coastal and environmental land-use record via Wetlands Board and Chesapeake Bay Board channels not present in JCC
- Tidewater-economy planning case material via Historic York District Committee and the working harbor record
- School-funding parallel-case material via YCSD BoardDocs for cross-jurisdictional joint-governance comparison

6.5 City of Radford, Virginia

The City of Radford is the corpus's fiscal-distress case and the most analytically deep slice on financial and state-oversight material. Radford received the Auditor of Public Accounts Recommendation Letter for State Assistance in March 2026 — the formal designation that triggers the Virginia fiscal-distress framework — following the November 2025 Preliminary Notification. The Radford slice was built between May 13 and May 21, 2026 from a mix of city portals, the Auditor of Public Accounts, the Federal Audit Clearinghouse, contemporaneous news reporting, and Wayback Machine recovery.

City Council and city-level bodies

City Council: Coverage from July 2, 2012 (earliest meeting on AgendaCenter) through spring 2026

Planning Commission: AgendaCenter coverage; multiple revisions captured

Economic Development Authority: AgendaCenter coverage

Archived / revised versions: 483 archived-or-revised files preserved with manifest at radford/agendas/council/manifest_2026-05-13.jsonl

Pre-2012 council records: Not on AgendaCenter; ArchiveCenter weekly summaries supply administrative continuity; Virginia FOIA request filed May 21, 2026 to recover pre-2012 council minutes

Financial records and state oversight (the analytical core)

ACFR coverage: FY2011 through FY2024 (audits) plus partial FY2025; PDFs and text extractions in radford/financial-reports/ and radford/analysis/

Budget coverage: FY2014 through FY2026 adopted budgets plus the FY2027 proposed budget

Statistical Section parsing: May 21, 2026 parse produced normalized longitudinal data FY2002–FY2024 covering all 14 GASB-recommended Statistical Section tables

APA Report ID 23684: March 2026 Recommendation Letter for State Assistance — formal distress designation, city-hosted as DC6317

APA Report ID 23683: November 2025 Preliminary Notification — precedes the formal designation

APA Comparative Reports: FY2000–FY2005 recovered from Wayback Machine and parsed into the workbook's Peer_Benchmarks tab; FY2006–FY2024 require direct APA staff outreach

Federal Audit Clearinghouse integration: EIN 546001548; FAC API used for auditee-status verification and peer-city benchmarking (FY2016 forward in the FAC system)

Hopewell peer comparison: Only Virginia small city in FAC with continuous FY2020+ ACFR coverage matching Radford's profile; selected for the workbook's Peer_Hopewell_Comparison tab

Regional authorities and adjacent records

Seven regional bodies tracked: Radford is member or major stakeholder; NRVJR audits extend back to FY2008

Radford City Public Schools (RCPS): Coverage back to 2020

Document Center: City-hosted documents pulled from radfordva.gov DocumentCenter: budgets, audits, policies, financial schedules, regulatory filings; includes DC6317, DC1203, and a sweep of administrative documents

ArchiveCenter weekly summaries: Predate AgendaCenter; provide pre-2012 administrative continuity

Wayback Machine snapshots: Historical city-website captures plus pre-2011 Radford audit material

Video and transcripts

Coverage: YouTube and Granicus meeting video, with auto-captions and Whisper-generated transcripts

Total transcripts: 293 transcript files across the Radford slice

News coverage

Coverage: Local and regional news pieces indexed for the fiscal-distress storyline; Roanoke Times beat reporter relationship confirmed for op-ed placement

Demonstrated analytical applications, Radford

- Op-ed in The Roanoke Times on Radford fiscal distress — published May 2026, first-of-two placement decision
- Documentary long-form on the fiscal-distress trajectory and the foundry-economy decline
- Senior-analyst forensic report on financial trajectory FY2002–FY2024
- Quantitative analysis report with normalized longitudinal Statistical Section data
- 23-tab Radford Financial Analysis workbook including Peer_Hopewell_Comparison tab
- Grede Foundry research memo — quantifying the \$1.3M annual electric-and-water revenue loss from the 2024 foundry closure on a city that is one of twelve Virginia municipal-electric localities

7. Cross-Jurisdictional Crosswalks

Because the corpus normalizes record classes across jurisdictions, several specific cross-jurisdictional joins have already been built and are available for analytical work. These joins are the strongest demonstration of why a multi-jurisdictional corpus produces analytical capability that no single-locality archive can match.

7.1 Joint funding contract negotiation, 2022–October 2025

The WJCC joint school division operates under a funding contract jointly executed by the City of Williamsburg and James City County. Negotiation of the renewed contract spanned 2022 through October 2025 and produced deliberation in all three rooms — City Council, Board of Supervisors, and the School Board itself. The corpus integrates the verbatim record across the three bodies via the meeting-event crosswalk, allowing researchers to follow the same argument as it travels between rooms. The October 10, 2025 joint signing at the JCC Recreation Center is captured as a discrete event in all three channels.

7.2 The 2023 breakup attempt vote (JCC Board of Supervisors)

On July 25, 2023, the JCC Board of Supervisors voted on a motion that would have initiated the dissolution of the joint funding arrangement. The motion failed, but the vote is the closest the joint arrangement came to formal severance. The City of Williamsburg's June 8, 2023 feasibility-study authorization is captured as the City-side preparatory action. Both meetings are present with minutes and verbatim transcript and join via the joint-funding crosswalk.

7.3 Budget-cycle interactions on the annual school transfer

The annual school transfer is a recurring inter-body negotiation between the County, the City, and the School Board over the parent-locality contribution to the joint division. The corpus indexes budget-cycle deliberation in all three bodies for each fiscal year from FY2018 forward, allowing comparison of how the same fiscal question is framed in each room and how the procedural mechanism translates the negotiation into the adopted budget on each side.

7.4 Three-jurisdiction Historic Triangle governance comparison

JCC, Williamsburg, and York are three adjacent localities operating on three different vendor stacks under three different transparency postures. The corpus captures all three, permitting direct comparison of deliberation language, procedural design, packet structure, and minutes-versus-transcript compression across the three rooms. Several specific land-use decisions — Bruton Heights, the working harbor record, short-term-rental cycles — span the boundary between localities and benefit from joint analysis.

7.5 Radford-to-Hopewell peer benchmarking

The Radford fiscal-distress case is paired with Hopewell as the closest Virginia small-city peer in the Federal Audit Clearinghouse system. The Peer_Hopewell_Comparison tab in the Radford Financial Analysis workbook normalizes ten-year audited financial series across the two cities. The cross-city analytical pattern extends to other Virginia distress-designated localities if their FAC coverage warrants it.

8. Demonstrated Analytical Yield

The corpus has already produced peer-quality analytical work. The list below records what has been built and where it stands, as evidence that the corpus is not a hypothetical research instrument. It supports real deliverables now.

8.1 Working papers and academic-format products

- The Working Majority: Public Dissent and Decision-Making on the James City County Board of Supervisors, 2010–2026
- A Board That Fought and Did Not Override: A Longitudinal Cohort Analysis of the James City County Board of Supervisors, 2010–2026
- Representative Judgment After Public Opposition: A James City County Companion Case on Capital Projects, Institutional Continuity, and Changing One's Mind

8.2 Monographs and longform investigative pieces

- Mounts Bay: A James City County Story, 2012–2026 — completed monograph, voice register validated against the underlying transcript record
- Three Rooms: Williamsburg and James City County and the Joint School System, 1955–2026 — longform piece on the 2023 breakup attempt and the 2025 reconciliation

8.3 Radford fiscal-distress reporting package (May 2026)

- Op-ed published in The Roanoke Times
- Senior-analyst forensic report on Radford's financial trajectory FY2002–FY2024
- Quantitative analysis report on the same window with normalized longitudinal data
- Radford Financial Analysis workbook — 23 tabs of normalized Statistical Section data, peer benchmarking, and APA framework documentation
- Grede Foundry research memo on the \$1.3M annual revenue loss from the 2024 closure and its compounding effect on a city already designated for distress

8.4 Methodological and infrastructure products

- Regional Government Records Corpus parent overview (Zenodo, concept DOI 10.5281/zenodo.20088873)
- Per-jurisdiction corpus briefs and inventories
- Build scripts, manifests, and provenance records for all eleven localities
- AI-assisted civic research workflow report (Volume II of this institutional review package)

8.5 Analytic universe summary, JCC

The James City County working-paper layer draws on a narrower coded analytic universe extracted from the broader JCC slice. The universe includes 388 minute files, 295 agenda packets, approximately 2,290 recorded votes, and 940 named-commenter appearances by 468 distinct residents, all for the January 2010 through February 2026 Board of Supervisors window. Coding rules, inclusion thresholds, exclusions, and known limitations are documented in the working-paper appendices.

9. Known Limitations and Gap Inventory

Every public-records corpus has limitations. The most important institutional discipline is to name them rather than to silently fill them. The list below catalogs the gaps and limitations a researcher will encounter when working with the corpus. Several are jurisdictional retention practices the corpus cannot alter; others are transcription or parsing constraints documented for transparency.

9.1 Transcription and verbatim capture

YouTube auto-caption is the verbatim layer from March 2018 forward in the lead jurisdiction. It misrecognizes clipped pronunciation. One supervisor's roll-call responses are systematically miscaptured because the speech recognition cannot parse that clipped delivery; verbatim verification of those votes defaults to the official minutes. Whisper-derived transcripts for the October 2012 through March 2018 window are more accurate but still imperfect for this speaker. The same caution applies to any speaker whose delivery produces low caption confidence.

9.2 Vote-register coverage

The JCC vote register contains 1,542 roll-call votes parsed from the minutes. The parser's coverage in the 2015–2016 window is known to be incomplete relative to the actual roll calls in those years. Researchers using the vote register for those years should cross-reference the underlying minutes. This is a parser limitation, not a missing-document limitation; the minutes themselves are present and complete.

9.3 Retention-driven gaps

Pre-FY2014 ACFRs and pre-FY2018 adopted budgets are not retained on the JCC live archive. The single FY2006 ACFR currently in the corpus was recovered from an unlisted-but-still-served archive identifier. FY2007–FY2013 ACFRs were not recoverable from any public source identified in this build. Pre-FY2022 proposed budgets are not retained under the applicable Virginia Public Records Act retention schedule; the JCC FOIA officer confirmed this in writing on May 6, 2026, and the response letter is preserved in the FOIA provenance folder. These are retention-driven gaps that no further recovery effort will close.

9.4 Pre-portal video and transcript ceiling

Board of Supervisors verbatim deliberation prior to October 2012 is not available from any public source identified in this build. The earthchannel.com archive for the JCC publisher does not extend earlier than that date. The pre-October-2012 record is therefore minute-only — researchers can know what was decided and roughly how, but cannot recover the verbatim language. Similar pre-platform ceilings apply in several of the other jurisdictions and are noted in their respective subsections.

9.5 Image-only PDFs requiring OCR

Two adopted budget PDFs in the FY2019–FY2020 and FY2026 cycles in JCC are unusually large — approximately 98 MB and 105 MB — and their page characteristics suggest scanned image content rather than born-digital text. Researchers running text-extraction pipelines should flag these for OCR before treating them as searchable. The corpus identifies and lists such files in its analytical readiness notes.

9.6 Vendor-platform completeness

Several jurisdictions inherit vendor-platform limitations. The City of Williamsburg's iCompass coverage is sparse for older periods, and a City Clerk channel has been opened to fill gaps. WJCC's BoardDocs era begins in late 2008; the 2006–October-2008 window was recovered via Virginia FOIA. York's CivicEngage AgendaCenter returns no pre-2019 BoS minutes, and the May 2026 FOIA recovered 2010–2019 but pre-2010 remains outside the scope of any public source identified. Radford's pre-2012 council records are limited to ArchiveCenter weekly summaries and pre-2011 Wayback Machine recoveries; a Virginia FOIA request was filed May 21, 2026 to recover full pre-2012 minutes.

9.7 Where minutes and transcript diverge

Minutes are summary; verbatim transcripts are not. Where the two diverge — and they do, routinely — the minutes are the official record for votes, motions, and procedural action. Transcripts are

research aids for deliberation language, tone, sequence, and quotation discovery. The compression from verbatim to minutes is itself an object of study; the corpus permits comparison but does not collapse the two into a single record. Researchers using verbatim transcripts for publication should verify against the original audio or video, not against the transcript alone.

10. Planned Expansion

Eleven localities are the current build: five held at research depth, two as comparison cases, and four as capability probes. The corpus is designed to extend without re-architecture. The methodology, naming conventions, source-platform adapters, and crosswalk patterns transfer directly to additional Virginia localities.

The objective is a statewide Virginia local-government records archive covering every county, city, and town. The current holding is a proof of concept for that objective rather than an endpoint.

10.1 What the current build already demonstrates

Expansion since the five-jurisdiction build has tested the method rather than merely enlarged the corpus. Pulaski and Montgomery counties were added as New River Valley comparison cases for the Radford fiscal work, and four further localities, Blacksburg, Christiansburg, Charlottesville, and Harrisonburg, were built as deliberately small capability probes.

The probes matter more than their file counts suggest. Each confirmed that a new locality can be brought to a usable baseline in a single session from a documented seed brief, across a different portal vendor, without modification to the naming or provenance framework. That is the claim on which statewide expansion rests, and it is now evidenced in six localities rather than asserted from three.

10.2 Expansion priorities

Priority is set by analytical value rather than by geography. Localities adjacent to an existing research-depth holding extend cross-jurisdictional questions already in progress. Localities facing a documented fiscal or land-use stress event supply comparison cases. Localities on a portal vendor not yet adapted extend the method's proven range.

On the first criterion, the Lower Peninsula and Hampton Roads neighbours of the Historic Triangle holding, Newport News in particular as the dominant regional employment center, remain the natural extension for regional-scale questions. On the second, the New River Valley cluster is already established and can be deepened. On the third, each new vendor stack encountered has so far required adaptation measured in hours rather than in redesign.

10.3 Higher-order and regional bodies

Planning district commissions, transportation planning organizations, transit authorities, and regional jails and service authorities produce records that bear directly on local-government decisions in the existing holdings. These bodies operate on their own platforms and require dedicated adapters. Bringing them into the same naming and provenance framework would allow regional decisions to be traced against the local decisions they constrain.

10.4 State-level integration

State oversight material is currently integrated for one jurisdiction, through the Auditor of Public Accounts and the Federal Audit Clearinghouse. Extending state-oversight integration across the other jurisdictions, particularly through the APA Comparative Reports series, which extends back twenty-plus years and is recoverable for some jurisdictions via Wayback Machine and direct APA outreach, would produce a corpus capable of state-level fiscal-trajectory analysis spanning multiple Virginia localities.

11. Rights, Citation, and Contact

11.1 Underlying records and aggregate compilation

The underlying government documents are public records and are reproduced at the level of the original public release. The aggregate compilation, indexing, canonical structure, metadata, manifests, transcription work, multi-source crosswalks, methodology notes, and derivative analytical materials represent original effort by the project author.

11.2 Copyright and reuse

Copyright © 2026 Tommy N. Turner. All rights reserved in the aggregate compilation, indexing, canonical structure, original transcriptions, manifests, metadata, methodology notes, and derivative analytical materials. Materials are shared for scholarly, evaluative, civic, and instructional review. Redistribution, republication, commercial use, or derivative instructional reuse requires written permission unless a specific derivative record states otherwise.

11.3 Researcher citation practice

Researchers using the corpus should cite both the underlying public record and the corpus release that delivered it. Whisper-derived transcripts and YouTube caption-derived files are research aids and should not be treated as official meeting records. Where verbatim accuracy matters for publication, researchers should verify against the original audio or video and against the official minutes.

11.4 Recommended citation

Turner, Tommy N. *Regional Government Records Corpus: A Longitudinal Public-Meeting Archive for Local Governance Research in Virginia*. Zenodo, 2026.
<https://doi.org/10.5281/zenodo.20088873>

11.5 Contact and review status

This document is prepared as a parent-project institutional overview for university partnership discussion. It is intended for scholarly review, classroom-use evaluation, methodological comment, and possible collaboration with public administration, public finance, planning, governance, civic technology, education governance, and AI-assisted research-methods scholars. Volume II — the AI-Assisted Civic Research Methodology document — accompanies this Volume I overview and treats the workflow in equivalent depth.

Author: Tommy N. Turner

ORCID: <https://orcid.org/0009-0005-7422-9881>

Email: tommy@gofast.pro

Website: gofast.pro

Zenodo concept DOI: [10.5281/zenodo.20088873](https://doi.org/10.5281/zenodo.20088873)