

Faith in the present, memory in the martyr: Catholic education and the figure of Stepinac as sites of memory in Croatian digital media (2021–2025)

AUTOR

DigiKat projekt

Hrvatsko katoličko sveučilište

Abstract

Institutions transmit religious memory, but their public representation need not be historical. This article asks whether Catholic education in Croatian digital media connects present discourse to the past or appears mainly as a field of current institutional and value disputes. The analysis uses the rebuilt official DigiKat corpus and compares eight education-related institutional and biographical anchors through recurrence, temporal peaking, affective language, and an anchor-specific local link to past-related vocabulary. The 2021–2025 corpus window contains 367,146 posts, of which 117,180 mention at least one primary anchor. Educational institutions are recurrent but weakly linked to historical language: local past-linkage rates range from 5.6% to 9.0%. Alojzije Stepinac appears in 7,384 posts and has a markedly higher rate of 23.2%. He is also the most temporally concentrated anchor, with February as the annual peak in all four years for which February is observed, and his language has the highest anger, fear, sadness, and disgust shares and the lowest trust share. His classified source profile is somewhat more mixed than those of the educational institutions, although limited dictionary coverage makes that evidence conditional. The results support a division of mnemonic labour in public representation: educational institutions organize present-oriented debate, while twentieth-century rupture and commemoration concentrate around Stepinac. They also qualify the measurement claim. Of the strand, 31.0% contains an anchor and past vocabulary somewhere in the same document, but only 9.8% places them within 160 characters; the local rule removes 68.3% of apparent links. Digital memory research should therefore distinguish recurrence from historical activation and document-level overlap from local textual connection.

Keywords: collective memory; sites of memory; digital religion; Catholic education; Alojzije Stepinac; computational text analysis; Croatia

1. Introduction

Religious education is designed to connect new generations to inherited narratives, practices, and forms of belonging. Public media, however, do not observe that work directly. They encounter education through school calendars, curriculum disputes, parental rights, institutional announcements, moral controversies, and political argument. An institution may therefore transmit memory in practice while appearing in the public digital record almost entirely in the grammar of the present.

The distinction is important in Croatia. Catholicism remains closely connected to narratives of national identity, and the Church's twentieth-century history continues to organize public disagreement. Religious instruction in state schools, Catholic educational institutions, and the place of Christian values in public life recur as contemporary issues. Alojzije Stepinac belongs to the same religious field but enters it through a different temporal route. Archbishop of Zagreb during the Second World War, later cardinal, defendant in a post-war trial, prisoner of the communist authorities, and beatified Catholic figure, he condenses disputes over wartime conduct, repression, martyrdom, sainthood, nation, and historical responsibility (Biondich 2006; Schwabl 2024).

This combination creates an empirical puzzle. If Catholic education is a principal means by which a religious community hands on its past, should its mediated representation show a strong connection to memory? Or does public memory become concentrated in a symbolic figure whose contested biography is more readily commemorated, narrated, and disputed?

We examine that puzzle with the rebuilt official DigiKat corpus. Rather than beginning with an already accepted memorial event, the study compares eight primary *anchors*: four institutional or practice bundles and four named figures connected to Catholic education and public memory. It asks which anchors are simply recurrent and which also exhibit a local connection to historical language, a repeated commemorative calendar, and a distinctive affective register. It then examines the institutional origin of their circulation.

The new database matters analytically. The official corpus applies one religious-inclusion rule across both collection eras and removes a large amount of material retained in the earlier accumulator-based study. Re-estimation reduces the education strand from the previous 176,312 posts to 117,180. This is not a cosmetic update. Every denominator, source share, temporal profile, affect sample, figure, and conclusion reported below is derived anew from the official corpus.

The central contrast survives that stricter input. Teaching orders, discourse about upbringing and values, Catholic schools, and religious instruction are more frequent than Stepinac, but their local connections to past-related vocabulary remain below 10%. Stepinac's rate is 23.2%, February is his peak in every year in which that month is observed, and his affective profile is more conflict-oriented. The source result is real but more modest than in the earlier analysis: among the minority of posts covered by the source dictionary, Stepinac has the largest non-confessional share, yet the difference from values discourse and teaching orders is only several percentage points.

The article's contribution is therefore substantive and methodological. Substantively, it identifies a division of mnemonic labour in Croatian digital media: institutions charged with transmission are represented primarily through current concerns, while memory of twentieth-century rupture is concentrated in a contested person. Methodologically, it shows why neither prominence nor document-level co-occurrence should be treated as memory work. A local proximity rule removes more than two thirds of apparent anchor-past overlaps, while the convergence of independent signals distinguishes recurrent topics from a stronger candidate site of memory.

2. Sites of memory in a mediated religious field

Collective frameworks and sites of memory

Collective memory is a present social activity rather than a fixed inventory of historical facts. Halbwachs locates remembering within frameworks supplied by groups, relationships, and contemporary concerns (Halbwachs 1992). Media sources do more than carry memories formed elsewhere: by selecting speakers, genres, anniversaries, conflicts, and narrative cues, they help construct the conditions under which the past becomes meaningful.

Nora's concept of *lieux de mémoire* directs attention to places, persons, dates, rituals, texts, and institutions in which memory becomes concentrated (Nora 1989). A site of memory need not be a physical location, and it need not unite everyone who invokes it. A contested person may endure precisely because different actors repeatedly return to the same figure to assert incompatible accounts of the past.

The concept does not supply an automatic measurement rule. Recurrence can reflect routine administration; peaks can reflect one-off news; emotional vocabulary can come from journalistic genre; and historical words can be incidental to the focal anchor. We therefore treat four observable signals as comparative evidence rather than as a classifier. Recurrence indicates durability. Temporal concentration tests whether attention returns on a calendar. Affect describes the textual register surrounding an anchor. Local past linkage asks whether the anchor and a historical expression meet in the same passage. A candidate is interpreted as mnemonic only when the signals converge and plausible alternative explanations weaken.

Institutions, figures, and cultural memory

Jan Assmann distinguishes fluid communicative memory from cultural memory stabilized through texts, rites, figures, institutions, and repeated forms (Assmann 1995). Aleida Assmann similarly emphasizes selection, mediation, storage, and renewed activation (Assmann 2011). Education is central to such stabilization, but institutional function and mediated representation remain different objects. A school, catechetical program, or religious order may transmit inherited narratives internally while entering public media chiefly through contemporary policy, organizational life, or moral debate.

A named figure can condense time differently. Biography supplies a narrative sequence; anniversaries provide a calendar; suffering provides a moral register; and political conflict sustains reinterpretation. Stepinac can evoke the wartime Church, the post-war trial, communist repression, Catholic martyrdom, Croatian national identity, canonization disputes, and Croatian-Serbian conflicts over history. Those meanings are neither uniform nor consensual. Their concentration in one name is exactly what makes the figure analytically comparable with institutional anchors.

Strossmayer, Stadler, and Marija Petković provide useful contrasts. Each has historical importance, but the study's past-language register is intentionally centred on twentieth-century rupture. A low score for Strossmayer can therefore reflect register mismatch rather than absence of memory. Rare figures also pose a power problem for temporal and affective comparisons. The design retains them in the recurrence and proximity table rather than allowing frequency to decide candidacy in advance, but it does not force a complete four-signal profile where the evidence is too sparse.

Digital and connective memory

Digital media store, rank, recirculate, and recontextualize traces of the past (Dijck 2007). In Hoskins's connective turn, memory is continuously remade through networked circulation rather than retrieved unchanged from a stable archive (Hoskins 2011). Analysis must therefore extend beyond official Church communication to the wider mediasphere in which confessional institutions, mainstream outlets, community sources, forums, and individual accounts may repeat or contest the same anchor.

Recent computational work makes timing, actors, platforms, and language jointly observable. Ben-David, Meyers, and Neiger propose a framework for comparing mnemonic actors, temporal dynamics, and vernaculars across social media while warning that commemorative and non-commemorative discourse blur online (Ben-David et al. 2024). Jumle and Makhortykh show how a known digital site of memory changes in relevance and discursive composition over time (Jumle and Makhortykh 2026). The present study asks an earlier comparative question: before explaining how a site changes, which of several plausible religious anchors exhibits the clearest mnemonic pattern?

Religion, media, and the Croatian case

Mediatization research argues that media do not simply reproduce religious messages. Their institutional routines and genres shape the form in which religion enters public life (Hjarvard 2008; Hoover 2006). Croatian research likewise describes the internet as a setting in which religious authority, belonging, community, and practice are reorganized (Novak and Valković 2016). Pavić, Kurbanović, and Levak found that Croatian Catholic websites and Facebook pages remained closely connected to offline structures and tended to reinforce established authority (Pavić et al. 2017). Peran's distinction between Church media and media that cover religion is equally important: public Catholicism cannot be reduced to communication produced by the Church itself (Peran 2018).

Stepinac makes the boundary especially visible. Historical scholarship documents disputes over his wartime position, post-war trial, and subsequent commemoration (Biondich 2006). Work on Croatian memory culture situates his rehabilitation at the intersection of Church, state, national identity, and revisionist conflict (Schwabl 2024). Survey research indicates that many Croatian students encounter Stepinac through media and view him positively within the national image (Trbušić and Rimac 2020). This literature makes him a plausible site of memory, but it does not establish how strongly he differs from Catholic educational institutions within the same digital field.

3. Research questions

The study is descriptive and interpretive. It does not estimate causal effects of platforms, sources, or Church communication.

RQ1. Which Catholic educational and biographical anchors combine recurrence, temporal peaking, affective charge, and local past linkage, and which are represented primarily as present-day topics?

The pre-analysis expectation was comparative rather than absolute: named figures associated with historical rupture should show stronger local historical linkage than broad institutional bundles. The design nevertheless allowed Catholic schools, religious instruction, or teaching orders to emerge as mnemonic if their mentions repeatedly activated communism, war, post-socialist change, victimhood, or national history.

RQ2. Which sources carry each anchor, and through what affective register?

Source origin identifies where an anchor is produced, not what authors or audiences believe. A mixed profile can show circulation across institutional boundaries without implying interpretive agreement. Affect likewise describes lexically marked language, not felt emotion.

4. Data and methods

The rebuilt official corpus

The official DigiKat corpus contains 413,985 Croatian- and Bosnian-language posts dated from 1 January 2021 to 11 June 2026. It includes web portals, YouTube, Facebook, Twitter/X, Reddit, forums, comment sections, Instagram, and TikTok. The corpus applies the same inclusion rule to both collection eras: within the first 3,000 characters a post must contain at least one decisive and at least two total matches from the 119-term religious list, after which a second-pass ensemble is applied at a threshold of 0.70. The tracked corpus manifest records SHA-256 [15473a615bf301c02b5d4149d662a4db282927d3b3e98308c5ff54cbe1de520a](#).

This study retains posts dated from 1 January 2021 through 31 December 2025. That window contains 367,146 corpus posts. The capture method changed around July 2024, so one inclusion rule makes the eras rule-comparable

but not capture-comparable. We do not interpret changes in raw annual volume as changes in attention. February through May 2024 contain no text in the vendor feed. Temporal analysis therefore uses 56 observed months, excludes those four months rather than coding them as zero attention, and checks concentration separately within the two collection streams.

Anchors and the education strand

Four primary bundles cover institutions and practices: religious instruction (*vjeronauk* and religion-teacher forms), Catholic schools and higher education, upbringing/values/curriculum, and teaching orders. Four named anchors cover Alojzije Stepinac, Josip Juraj Strossmayer, Josip Stadler, and Blessed Marija Petković. A secondary ritual bundle includes pilgrimages, processions, the Croatian Catholic Youth Meeting, and similar gatherings; it serves as a calendar comparison but does not define the education strand.

Patterns are applied to lower-cased text using Croatian locale handling and common inflectional or orthographic variants. A post enters the strand when it mentions at least one primary anchor. The resulting strand has 117,180 posts. Web sources account for 84.1%, with the remainder distributed across the other eight platform types.

The probe favours recall, but known ambiguities are constrained where possible. Marija Petković requires a name or religious-order context. The Jesuit pattern uses explicit order forms and “Society of Jesus” rather than the bare possessive *isusov*, which frequently means “of Jesus.” The war token is restricted to Croatian inflections of *rat*, *ratov*- and *ratn*- so that agricultural and other unrelated *rat*- words are excluded. Broad bundles are decomposed before interpretation.

Four comparative signals

Recurrence is the number of strand posts in which an anchor appears. It measures sustained visibility, not mnemonic meaning by itself.

Local past linkage is an anchor-specific proximity measure. The past-language register covers communism, Yugoslavia, secularization, 1991, the Homeland War, war, victimhood, and pre-war or post-war periods. A link is recorded when a register expression begins within 160 characters of the focal anchor. Distance is calculated separately for each anchor, so one entity cannot borrow a historical expression located near another entity in the same post. The deliberately limited label reflects that proximity does not establish syntax, negation, quotation, or interpretation. It is a stricter proxy than document-level co-occurrence, not a semantic gold standard.

Temporal peaking is calculated on observed months only. The pooled coefficient of variation describes unevenness across the 2021–2025 series. Month-of-year profiles use mean posts per observed instance of each month before normalization, preventing the 2024 text gap from depressing February through May. Annual peaks identify whether the same month returns. Within-stream coefficients provide a check against the collection-method break.

Affective register is estimated on a deterministic union sample. Up to 800 posts are drawn for each of seven sufficiently frequent anchors, and overlapping selections are retained once, producing 5,526 unique posts. Because a sampled post can mention several anchors, anchor-specific membership counts in the union can exceed 800. The first 1,500 characters are lemmatized with the project’s Croatian UDPipe model. CroSentilex supplies signed polarity and lilaHR supplies an eight-emotion profile. CroSentilex covers 54.3% of lemmatized tokens and at least one polarity item in essentially every sampled document. Emotion profiles are conditional on documents with an emotion match. The measures cannot recover irony, narrative voice, or felt emotion and therefore remain supporting evidence.

Source composition and validation logic

Source composition is described through platform shares, leading named sources, and a web-actor typology based on reach and interaction. The broad actor types add little differentiation, so interpretation focuses on source origin. A project dictionary classifies 109 named sources or accounts as confessional, secular, other non-confessional, or unclassified. The comparison combines secular and other labelled sources as non-confessional and calculates shares only among classified posts. Coverage ranges from 34.5% to 47.2% for the four institutional bundles and Stepinac and is 23.5% for Strossmayer. These conditional shares cannot be generalized to all posts if unclassified sources differ systematically.

Stepinac functions as a positive control for the twentieth-century rupture register. The upbringing-and-values bundle is the predeclared present-oriented foil. If the local-linkage measure did not separate the two, the four-signal interpretation would fail. The analysis also retains document-level overlap to quantify how much an unrestricted co-occurrence rule would inflate the result.

All study stages read the official corpus without modifying it. The input resolver verifies its hash, dimensions, and date span; the slice cache is keyed to both the corpus and the probe; aggregate tables and figures are then generated from the same restricted slice.

5. Results

Recurrence and local historical linkage point in different directions

The most recurrent primary anchor is teaching orders, with 65,468 posts, followed by upbringing and values with 40,073, Catholic schools with 19,671, and religious instruction with 12,598. Stepinac appears in 7,384 posts. Frequency alone would therefore place institutional life at the centre of the strand.

Local past linkage reverses that ranking. Historical vocabulary occurs within 160 characters of Stepinac in 23.2% of his posts. The institutional rates are 9.0% for upbringing and values, 8.0% for teaching orders, 6.2% for Catholic schools, and 5.6% for religious instruction. Stepinac's rate is between 2.6 and 4.1 times those rates. The positive-control comparison is clear: Stepinac's 23.2% is 2.6 times the 9.0% foil rate.

The other named figures do not show the same pattern under this register. Strossmayer has a 4.9% local-linkage rate, Stadler 5.2%, and Petković 3.9%. These figures should not be treated identically. Strossmayer is frequent enough for a temporal profile, but a vocabulary centred on communism, Yugoslavia, and twentieth-century war is a poor instrument for nineteenth-century memory. Stadler and Petković are too rare for the same temporal, source, and affect comparisons. Their low rates therefore constrain the present instrument rather than settle their status in Croatian Catholic memory.

Figure [Figure 1](#) shows the main contrast. Institutional anchors occupy the high-frequency, low-linkage region. Stepinac is less frequent but stands well above them in local historical connection.

Table 1. Comparative profile of the eight primary anchors in the official corpus

Anchor	Posts	Local past linkage	Temporal rhythm	Source boundary	Affective profile
Upbringing and values	40,073	9.0%	CV 0.32; December peak in 4 of 5 eligible years	26.6% non-confessional (40.5% classified)	Trust-leaning present-oriented register
Teaching orders	65,468	8.0%	CV 0.26; December peak in 2 of 5 eligible years	27.3% non-confessional (38.9% classified)	Trust-leaning institutional register
Catholic schools	19,671	6.2%	CV 0.25; April peak in 1 of 4 eligible years	17.2% non-confessional (47.2% classified)	Highest trust; low conflict-language shares
Religious instruction	12,598	5.6%	CV 0.30; March peak in 2 of 4 eligible years	22.3% non-confessional (34.5% classified)	Trust-leaning; lower conflict-language shares
Alojzije Stepinac	7,384	23.2%	CV 0.75; February peak in 4 of 4 eligible years	30.2% non-confessional (37.4% classified)	Highest anger, fear, sadness and disgust; lowest trust
Josip Juraj Strossmayer	2,460	4.9%	CV 0.40; March peak in 1 of 4 eligible years	22.1% non-confessional (23.5% classified)	Trust-leaning; no distinct conflict profile
Josip Stadler	673	5.2%	Not estimated	Not estimated	Not estimated
Marija Petković	254	3.9%	Not estimated	Not estimated	Not estimated

Note. Local past linkage is an anchor-specific ± 160 -character proximity measure, not a semantic classification. Temporal rates exclude the four unobserved months from February through May 2024. Source shares use only classified posts. Affect is supporting evidence from the deterministic union sample.

The foil’s composition explains why frequency is not enough. Within the broad upbringing-and-values bundle, *value* appears in 25,878 posts, *upbringing* in 17,458, and *curriculum* in 775; the explicit phrase “Christian roots” appears once. These overlapping counts show a large field of contemporary moral and educational language, not a recurrence pattern driven by an explicit civilizational-memory phrase.

The unrestricted overlap comparison reinforces that interpretation. Past-related vocabulary appears somewhere in the same document as an education anchor in 31.0% of strand posts. Only 9.8% places a past expression within the local window of any primary anchor. Requiring local proximity removes 68.3% of apparent links. Figure [Figure 2](#) visualizes the reduction. The result does not prove that every retained link performs memory work, but it shows that most document-level overlaps are too distant to support even a local textual connection.

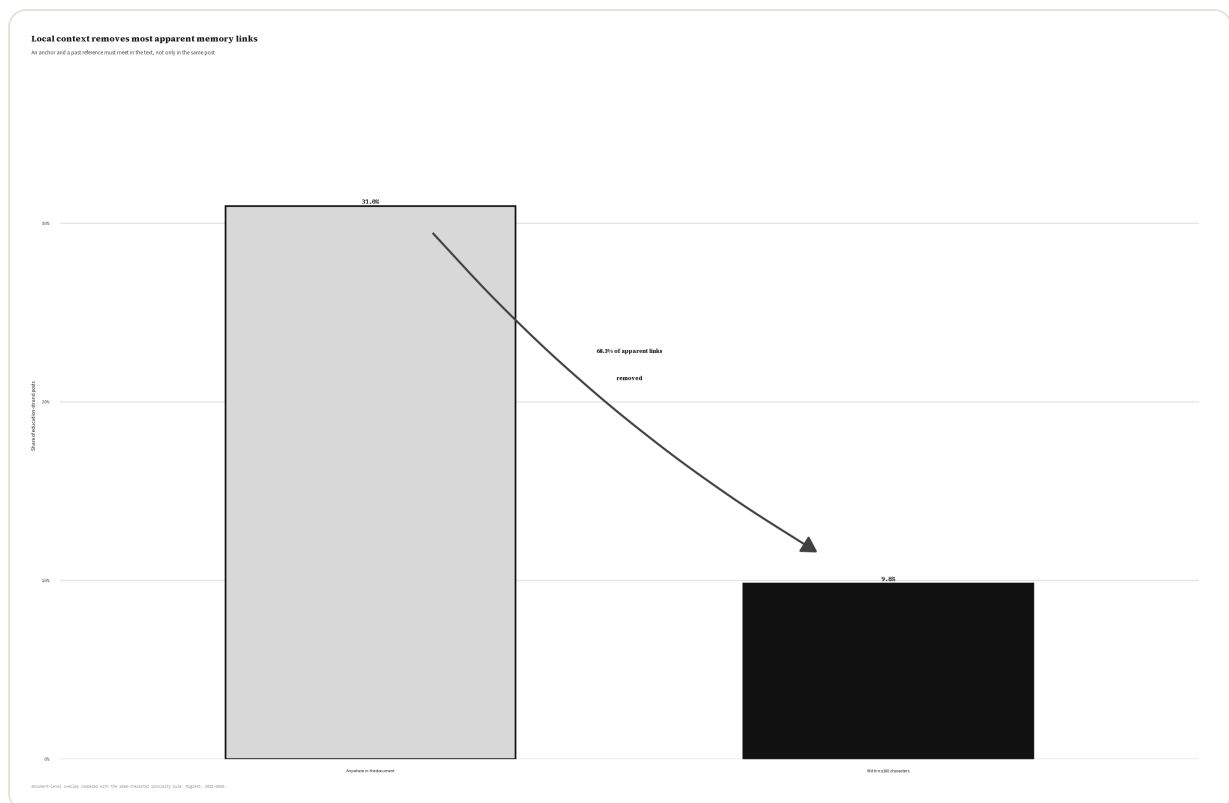


Figure 2: Most apparent education-and-past links disappear when the anchor and past reference must occur within 160 characters.

Stepinac has a commemorative calendar

Stepinac is the most temporally concentrated anchor. His pooled monthly coefficient of variation is 0.75. The institutional values range from 0.25 for Catholic schools to 0.32 for upbringing and values. Within the older collection stream Stepinac's coefficient is 0.75, and within the newer stream it is 0.81; the separation is therefore not produced only by joining the two collection eras.

February is Stepinac's annual maximum in 2021, 2022, 2023, and 2025. February 2024 is absent from the vendor feed, so the defensible denominator is four eligible years rather than five calendar years. Stepinac peaks in February in all four. The timing corresponds to *Stepinčevo* around the anniversary of his death on 10 February 1960. October, the month of his 1998 beatification, forms a smaller secondary rise in the exposure-adjusted profile.

Educational anchors have broader calendars. Upbringing and values peak in December in four of five eligible years, but their low temporal concentration and low local past linkage point to a year-end institutional rhythm rather than commemoration of a historical rupture. Religious instruction and Catholic schools have no equally stable peak. The secondary ritual bundle peaks in August in four of five years, providing a useful non-biographical example of recurring religious seasonality.

Figure [Figure 3](#) uses mean posts per observed instance of each month. The adjustment is visible in the February through May panels: the missing 2024 months do not enter as zero attention.

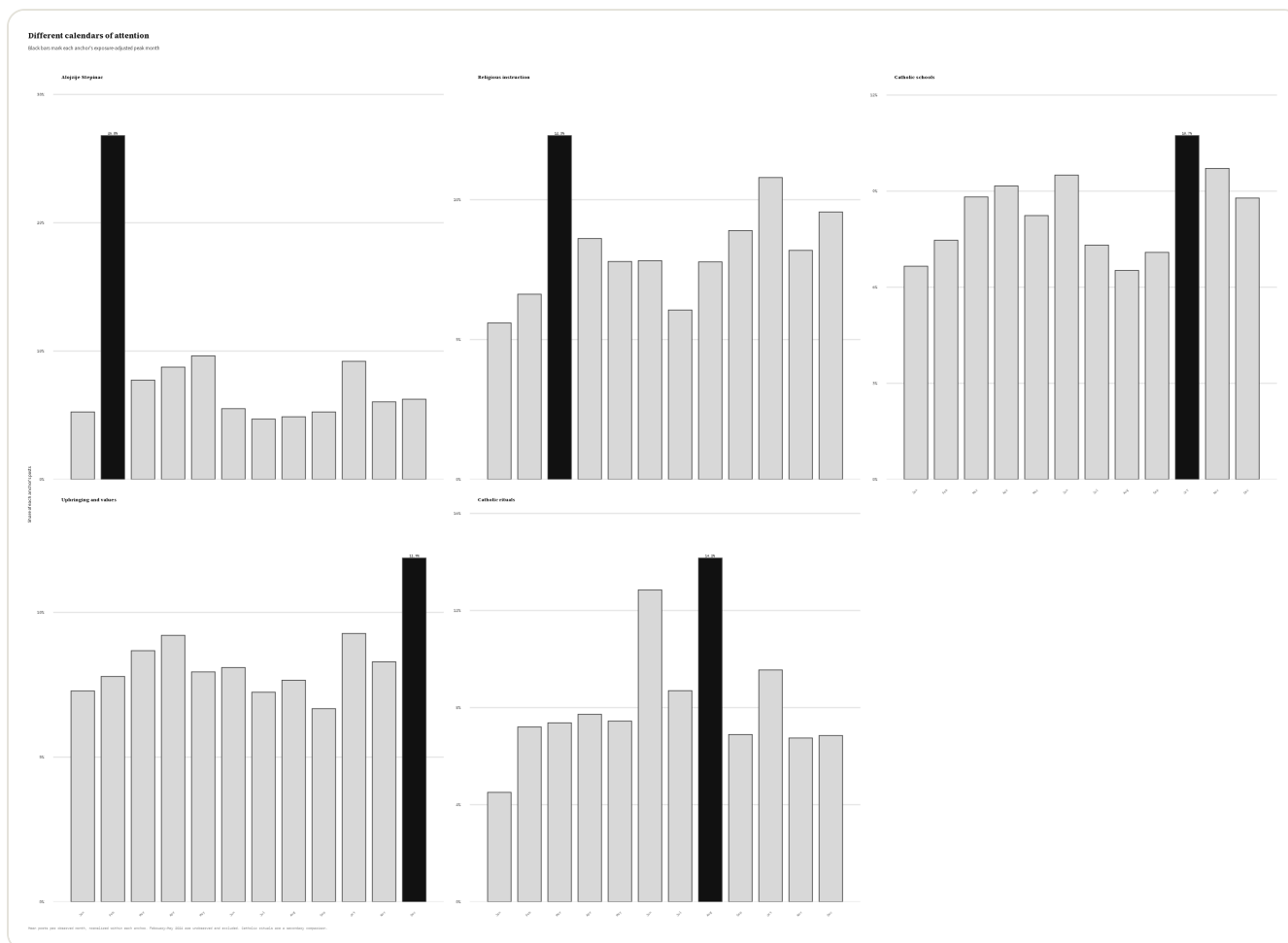


Figure 3: Exposure-adjusted monthly profiles distinguish Stepinac's February commemoration from the broader calendars of educational discourse.

Source circulation extends beyond confessional media, but the difference is conditional

Stepinac has the most mixed classified source profile among the five main anchors. Of his posts, 37.4% can be assigned a source-origin category. Within that classified subset, 69.8% are confessional and 30.2% non-confessional. The non-confessional shares are 26.6% for upbringing and values, 27.3% for teaching orders, 22.3% for religious instruction, and 17.2% for Catholic schools.

The result supports cross-boundary circulation, but only modestly. Stepinac differs sharply from Catholic schools and religious instruction, yet by only 3.6 percentage points from values discourse and 2.9 points from teaching orders. Moreover, 62.6% of Stepinac posts remain unclassified. If their composition differs from classified posts, the total-corpus share could change substantially. Figure [Figure 4](#) therefore reports only the conditional comparison and labels its denominator explicitly.

Named sources provide qualitative context. Stepinac's leading sources include Catholic and diocesan outlets alongside nationally oriented and mainstream portals. Religious instruction also appears prominently in Reddit's Croatia community and forum.hr, consistent with a present-day controversy frame. The web-actor reach/interaction typology does not separate the anchors clearly and is treated as a null descriptive result.

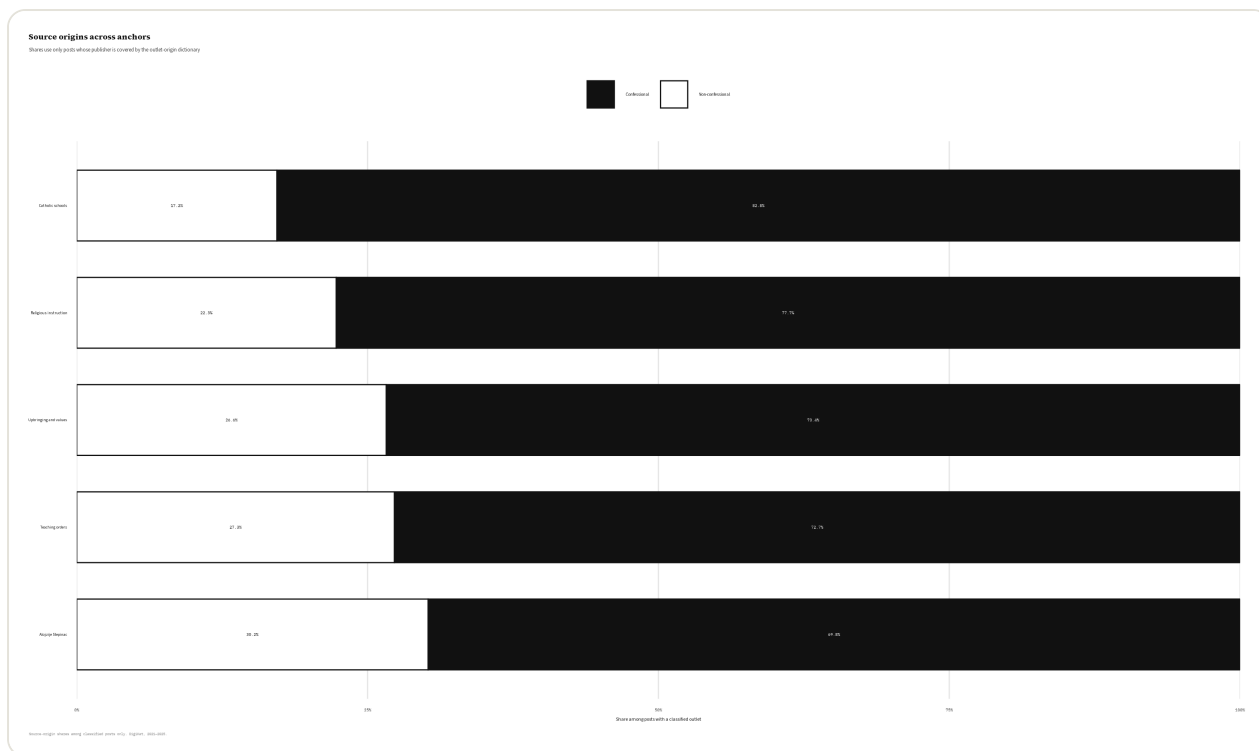


Figure 4: Source-origin composition among posts covered by the outlet dictionary.

Affect supports the contrast but does not establish it

CroSentilex polarity does not differentiate the anchors. Mean scores range narrowly from -0.432 for Stepinac to -0.405 for Strossmayer, and the near-universal non-neutral match rate reflects lexicon structure and shared news language more than substantive opposition. Polarity is therefore excluded from the main interpretation.

The lilaHR emotion profile provides a consistent but small contrast. Stepinac has the highest average shares of anger, fear, sadness, and disgust and the lowest trust share. Catholic schools have the highest trust share and low conflict-language values; religious instruction, values discourse, and teaching orders also lean more strongly toward trust. The ritual comparison has the highest joy share. These patterns fit the distinction between persecution and conflict around Stepinac, institutional formation around education, and celebration around rituals.

The differences are descriptive and lexicon-dependent. They cannot show what authors intended or readers felt, and they do not justify the mnemonic conclusion on their own. Their value lies in convergence with the much larger separation in local past linkage and temporal concentration.

Convergence

No single signal defines a site of memory. Taken together, however, the official-corpus results identify one strong pattern. Stepinac is recurrent without being the most frequent anchor; he is locally linked to the past at more than twice the rate of the strongest institutional foil; attention returns in every observed February; the surrounding language is relatively conflict-oriented; and circulation reaches beyond confessional sources within the classified subset. The institutional anchors are more frequent but have weaker historical linkage, flatter temporal profiles, and more trust-oriented registers.

The conclusion is comparative and register-specific. Under a twentieth-century rupture vocabulary, Stepinac is the clearest candidate site of memory in this education-related field. The analysis does not establish that Catholic institutions fail to transmit memory in their own practices, nor that historically important nineteenth-century figures lack mnemonic life.

6. Discussion

A division of mnemonic labour in public representation

The new analysis supports a division of mnemonic labour, but locates it carefully. Catholic education can transmit tradition in classrooms, families, parishes, rituals, and curricula. The corpus does not observe those practices directly. It observes how institutions and figures appear in a heterogeneous digital mediasphere. Within that public representation, educational anchors organize current discussion about schooling, values, formation, organizational life, and the role of religion. Stepinac more consistently opens a passage to a remembered twentieth-century past.

Assmann's distinction between communicative and cultural memory helps interpret the pattern ([Assmann 1995](#)). The institutional bundles circulate in a fluid present-oriented field, while Stepinac combines symbolic concentration, repeated calendar activation, and a conflict-laden historical register. The division is not ontological: institutions can carry cultural memory, and Stepinac is also invoked in current politics. It describes the relative work performed by these anchors in the observed media record.

The foil is especially revealing. Upbringing and values are more than five times as frequent as Stepinac, yet the bundle is driven by generic present-oriented language. The explicit "Christian roots" phrase appears once. Public moral vocabulary can be religiously important without functioning as an active historical bridge. Conversely, a less frequent person can become a more efficient mnemonic condensation point.

The martyr as contested public property

Nora's sites of memory do not require consensus. They can endure because different groups revisit and dispute them. Stepinac's classified source profile fits that model, although the new estimate calls for restraint. His non-confessional share is the largest among the main anchors, but it is not overwhelmingly different from values discourse or teaching orders, and most sources remain unclassified.

The stronger evidence for contested public property comes from the combination of source diversity, commemoration, and affect. Catholic, diocesan, mainstream, nationally oriented, and community sources sustain the same figure, while conflict and persecution vocabulary distinguishes his textual environment. This is shared production of an anchor, not shared belief about its meaning. The pattern aligns with scholarship that places Stepinac at the intersection of Church, state, wartime history, socialism, national identity, and rehabilitation ([Biondich 2006](#); [Schwabl 2024](#)).

The martyr label itself belongs to a Catholic mnemonic tradition. The analysis does not adjudicate Stepinac's wartime conduct or infer that every source accepts the martyr narrative. It shows that the public figure around whom such a narrative is organized has a distinctive historical and calendrical structure.

What the four-signal comparison adds

Digital memory research often begins with an established memorial date or accepted site and asks how actors, language, and attention change ([Ben-David et al. 2024](#); [Jumle and Makhortykh 2026](#)). The present design compares several plausible candidates before granting that status. The contrast between recurrence and local linkage prevents high-volume topics from winning by default. The temporal signal distinguishes recurring commemoration from isolated news. Affect supplies a weak independent register. Source composition asks whether circulation remains institutionally enclosed.

The framework should not be converted into an automatic composite score. Signals differ in validity and scale. The local rule is more directly connected to historical activation than lexicon affect; a stable anniversary has a different

meaning from a high coefficient of variation; and source coverage can limit boundary claims. The correct use is comparative triangulation with explicit controls and known register limits.

Strossmayer demonstrates that limit. His low score does not falsify his historical importance. It shows that an instrument calibrated to communism, Yugoslavia, 1991, war, and victimhood cannot fairly test every layer of Croatian Catholic memory. A future nineteenth-century register should be predeclared and validated separately rather than added after inspecting his score.

Co-occurrence is not memory work

The proximity result is the most transferable methodological finding. Long digital documents contain several themes, quoted passages, sidebars, and historical references. In this strand, 31.0% of posts contain both an education anchor and past vocabulary somewhere, yet only 9.8% bring them within 160 characters. Whole-document co-occurrence would multiply plausible links by more than three and make institutional discourse appear much more historical than it is locally.

Proximity remains imperfect. A short window can contain an unrelated list, negation, quotation, or syntactic separation. It can also miss a genuine relation expressed across paragraphs or through anaphora. The value of the rule is not that it proves memory work, but that it rejects a large class of clearly weak overlaps using a transparent, reproducible criterion. Future work should hand-code a stratified sample to estimate the precision of both the document-level and local measures.

7. Scope and limitations

The analysis concerns public digital representation, not memory held by citizens or practiced inside classrooms, families, schools, and parishes. Web sources constitute 84.1% of the strand, so the results should not be generalized to an evenly balanced platform ecology.

The database is rule-comparable across collection eras but not capture-comparable. The July 2024 method change prevents interpretation of raw cross-era volume, and February through May 2024 are unobserved. Coverage-aware temporal estimates and within-stream checks address the most direct artifacts, but they cannot reconstruct missing content.

The anchor patterns favour recall and remain heterogeneous. The teaching-order bundle includes monasteries and order names that are not always about formal education. The upbringing-and-values bundle is intentionally broad. The local past-language register is weighted toward twentieth-century rupture and under-measures nineteenth-century memory. The proximity rule has not yet been validated against a two-coder semantic gold standard.

Source-origin labels are proposed and cover only part of the corpus. Conditional differences may not hold among unclassified sources, and source origin says nothing about audience composition or interpretive agreement. Affect scores depend on Croatian lexical resources, truncate text at 1,500 characters, and cannot recover irony, quotation, or narrative voice. They should be read only as supporting textual profiles.

Finally, the study identifies patterned association rather than causation. It cannot determine whether Church communication, platform ranking, journalistic routines, anniversaries, or political events cause an anchor's observed profile.

8. Conclusion

The rebuilt official DigiKat corpus changes the scale of the evidence but not its central direction. The education strand contains 117,180 rather than 176,312 posts, and every reported rate, calendar, source profile, affect estimate, table, and figure has been recomputed. Under the official inclusion rule, Catholic educational institutions remain highly visible yet weakly connected to historical language. Their public representation is dominated by the present: schools, values, formation, curriculum, organizational life, and disputes over religion's public role.

Stepinac is different. His local past-linkage rate is 23.2%, compared with 5.6% to 9.0% for the institutional anchors. His attention is more temporally concentrated, and February is the peak in every year for which February is observed. His language carries relatively more anger, fear, sadness, and disgust and less trust. Within a partially classified source subset, his circulation is also somewhat more cross-boundary.

These results support a specific claim: in Croatian digital media from 2021 to 2025, the public work of Catholic memory is more concentrated in the contested figure of the martyr than in institutions formally charged with education. This is a claim about mediated representation, not about what schools, parishes, or families actually remember.

The re-analysis also narrows the methodological language. A 160-character relation is a local-linkage proxy, not proof of genuine memorial activation. Even so, it is far more informative than document-level overlap: 68.3% of apparent education–past links disappear when the terms must meet in the text. Comparative memory research should therefore separate frequency from memory, missing data from zero attention, and co-presence from local historical connection.

Data availability and declarations

The official corpus is derived from platform and vendor data that cannot be redistributed at row level. Its tracked manifest records provenance, dimensions, date coverage, inclusion rule, and SHA-256. Aggregate study tables, figures, input contracts, and analysis scripts are available in the DigiKat repository. The full analysis can be reproduced on an authorized machine with the official corpus by running the study pipeline from the repository root.

OpenAI Codex was used to audit the analysis pipeline, execute reproducible code, check numerical consistency, and assist with language revision. The authors retain responsibility for the research design, interpretation, source classification, and final text.

The authors declare no competing interests.

References

- Assmann, Aleida. 2011. *Cultural Memory and Western Civilization: Functions, Media, Archives*. Cambridge University Press.
- Assmann, Jan. 1995. "Collective Memory and Cultural Identity." Translated by John Czaplicka. *New German Critique*, no. 65: 125–33. <https://doi.org/10.2307/488538>.
- Ben-David, Anat, Oren Meyers, and Motti Neiger. 2024. "How Social Memory Works on Social Media: A Methodological Framework." *Memory, Mind & Media* 3: e20. <https://doi.org/10.1017/mem.2024.18>.
- Biondich, Mark. 2006. "Controversies Surrounding the Catholic Church in Wartime Croatia, 1941–45." *Totalitarian Movements and Political Religions* 7 (4): 429–57. <https://doi.org/10.1080/14690760600963222>.

- Dijck, José van. 2007. *Mediated Memories in the Digital Age*. Stanford University Press.
- Halbwachs, Maurice. 1992. *On Collective Memory*. Edited by Lewis A. Coser. University of Chicago Press.
- Hjarvard, Stig. 2008. "The Mediatization of Religion: A Theory of the Media as Agents of Religious Change." *Northern Lights* 6: 9–26. https://doi.org/10.1386/nl.6.1.9_1.
- Hoover, Stewart M. 2006. *Religion in the Media Age*. Routledge.
- Hoskins, Andrew. 2011. "Media, Memory, Metaphor: Remembering and the Connective Turn." *Parallax* 17 (4): 19–31. <https://doi.org/10.1080/13534645.2011.605573>.
- Jumle, Vihang, and Mykola Makhortykh. 2026. "Protean Memorialisation: International Holocaust Remembrance Day as a Digital Site of Memory on Twitter." *Politics and Governance* 14: 11830. <https://doi.org/10.17645/pag.11830>.
- Nora, Pierre. 1989. "Between Memory and History: Les Lieux de Mémoire." *Representations* 26: 7–24. <https://doi.org/10.2307/2928520>.
- Novak, Krunoslav, and Jerko Valković. 2016. "Religija i Internet—Novi Izazovi Življenja Vjere." *Obnovljeni Život* 71 (3): 389–402. <https://doi.org/10.31337/oz.71.3.8>.
- Pavić, Željko, Filip Kurbanović, and Tomislav Levak. 2017. "Mediatisation of Catholicism in Croatia: A Networked Religion?" *Revija Za Sociologiju* 47 (3): 241–70. <https://doi.org/10.5613/rzs.47.3.1>.
- Peran, Suzana. 2018. "Religious Media in the Service of Dialogue Within the Church: Media Communication of the Church in Croatia According to the Guidelines of the Croatian Bishops Conference, Crkva i Mediji (the Church and Media)." *Bogoslovska Smotra* 88 (3): 853–71. <https://hrcak.srce.hr/207880>.
- Schwabl, Natalie. 2024. "Between State and Church: Memory Culture and Revisionism in Croatia. The Role and Rehabilitation of Alojzije Stepinac, Archbishop of Zagreb." *Qualestoria* 52 (2): 67–87. <https://doi.org/10.13137/0393-6082/36702>.
- Trbušić, Davor, and Stipan Rimac. 2020. "Cardinal Stepinac as an Element in Shaping the Image of the National Brand of Croatia." *Communication Management Review* 5 (1): 84–103. <https://doi.org/10.22522/cmr20200156>.