Reporting artifacts in crisis event streams, measured — and a declaration standard for them

Paper B of the Fenocrisis founding program (PROMPTS.md P-2.2; Arc 2). The Q3 metrology companion to Paper A. Aimed at a methods/measurement audience, the descendant of a sibling programme's surrogate-audit methods result.

Author. Evan Tabak Atlas (ORCID 0009-0007-7374-2338).

Status. Complete draft, gated on owner review before submission (Owner queue). Every recommendation traces to a measured sensitivity — either measured on that stream in the tournament (`out/v1/x3/`, `out/v3/`), or the measured general rule of §3 (coupling magnitude is baseline-fragile and attention-absorbable; mechanism structure is baseline-robust) applied to the stream's declared artifact channel. The declaration schema is published as Appendix A and as machine-readable JSON (`papers/02-artifact-schema.json`).

Abstract

Crisis event data has a photon channel, and it is the newswire. Detection is plausibly self-exciting (crises attract journalists), reporting capacity ramps over decades (EM-DAT's own maintainers caution against pre-2000 trends), and several corpora define events partly by institutional *response* — a declared state of emergency, an international appeal, a policy intervention. A constant-baseline point-process fit attributes all of that to *mechanism*. No one in the crisis literature has corrected for it, and no machine-readable standard for declaring it exists. This paper does three things. First, it proposes such a standard: a thirteen-field **artifact-channel declaration** every stream ships before it may back a claim — detection mechanism, dated structural breaks, threshold criteria, response-endogeneity flag, license class, and vintage policy — published here as a JSON Schema and instantiated for the charter's eleven crisis streams. Second, it measures how much of a cross-domain coupling signal survives the channel, by recomputing every headline of a founding tournament under the whole declared reporting-baseline family and a shared-attention covariate. The cross-domain excess is **fragile to the reporting-baseline choice** — significant under some baselines, absorbed by the flexible spline — and is **absorbed below significance by a single shared contemporaneous attention covariate the per-stream baseline family structurally cannot represent**; the mechanism decomposition, by contrast, is baseline-robust. Third, it states the underlying identification problem formally — *a self-exciting detection filter over a Poisson occurrence process is observationally near-identical to a Hawkes process* — and names the infrastructure gap that makes it hard to close: there is no ALFRED for crisis data, no archival vintage record for event corpora that are silently re-edited with hindsight. Every number is a retrospective measurement on public, aggregate data; nothing forecasts,

warns, or scores. **Scope: the audited signal is cross-domain coupling in crisis onset *timing*; severity/magnitude channels are a separate deliverable.**

---

1. The problem: crisis data's photon channel

Every measurement instrument has a channel between the event and the record, and its imperfections masquerade as signal. In astronomy it is the photon path and the detector; in crisis data it is **detection and definition** — who noticed the event, why they wrote it down, and what institutional act the record is partly counting. Three properties of that channel corrupt a naïve point-process fit:

1. **Detection is plausibly self-exciting.** A crisis draws journalists and coders; coverage of one event raises the probability the next is recorded. A self-exciting *detection* filter over an underlying Poisson *occurrence* process is observationally near-identical to a Hawkes *occurrence* process (§4). A branching ratio fit to such data is not evidence of contagion.
2. **Reporting capacity ramps.** EM-DAT's maintainers explicitly recommend discarding pre-2000 trend interpretation; WHO's Disease Outbreak News begins in 1996 and shifts at IHR-2005 and COVID; UCDP GED coverage begins in 1989 and is re-released annually. A constant baseline reads the ramp as a trend in the process.
3. **Events are partly response-defined.** EM-DAT admits disasters on a declared state of emergency or an international appeal; Laeven–Valencia dates a banking crisis partly by the policy intervention that answered it; WHO weighs "need for assistance." When the record counts the response, the intensity is entangled with endogenous institutional behaviour.

The field has never corrected for any of this. The correction is a metrology contribution independent of whether any coupling is real.

---

2. The declaration standard (Appendix A; JSON in `02-artifact-schema.json`)

The prerequisite for auditing the channel is *declaring* it, machine-readably, so the audit can be mechanical rather than editorial. We propose a thirteen-field **artifact-channel declaration** — a YAML/JSON mapping every stream ships beside its loader, validated on load, so a malformed or incomplete declaration can never silently back a claim. The full field list, types, and rationale are Appendix A; the JSON Schema is `papers/02-artifact-schema.json`. The load-bearing fields:

- **`detection_mechanism`** — the photon channel in prose (newswire coding, relief/reinsurance reporting, administrative dating, market observation, editorial selection).

- **`structural_breaks`** — a list of `{date, description}`, *always dated*, because the reporting-baseline family μ(t) dummies exactly these dates and the audit (§3) recomputes every number with them dummied.
- **`response_endogeneity`** — the single most important flag: `flag: true` iff events are defined partly by institutional response. A response-endogenous stream may not support a self-excitation claim without a detection/response control (§4). Where a loader removes the endogenous subset (EM-DAT's declaration/appeal rows), the flag records that the *raw* source is endogenous even though the *loaded* subset is not.
- **`license.class`** — `open | registered | restricted`, read by a serialization gate that makes the charter's teeth impossible-by-construction: a registered or restricted row-level frame *raises* if a write targets a committed path; only a deliberately-constructed derived summary passes. Licensing discipline becomes a type error, not a matter of care.
- **`vintage_policy.type`** — `retrospective | real-time | revised | frozen | mixed | static`; the F5 hook (§5) that forbids calling a retrospectively-edited corpus "out-of-sample."

The charter's eleven crisis streams (SPECS §6) are declared against this schema and shipped with the program (Appendix B); UCDP GED's real-time Candidate vintage carries its own declaration file (`ucdp_candidate.yaml`), so twelve `*.yaml` declarations ship in all. The claim of the standard is not that these are the right thirteen fields forever — it is that *some* such declaration must precede *any* coupling or excitation number on crisis data, and that a validated, machine- readable one lets the audit run as a gate rather than a footnote.

---

3. The measured decomposition

We recompute every headline of a founding tournament (Paper A; record v1) under the whole declared reporting-baseline family and report the *range*, not a single specification — a claim survives the audit only if it holds across the family (record `out/v1/x3/`, seed `20260903`; and the attention control `out/v3/`, P-5.2). Two corpora: Corpus-A (Delannoy six-domain country-year, 173 countries, 1970–2019) and Corpus-B (UCDP GED conflict + WHO-DON epidemic country-month, 192 countries, 1990–2024).

3.1 The μ(t) reporting-baseline sweep

The reporting baseline μ_d(t) is swept across a declared family — **constant, linear, spline, and changepoint** (the last with the declared structural break dates dummied). The cross-domain coupling excess (Paper A) is **fragile** to the choice:

corpus	constant	linear	spline	changepoint	range (ΔBIC)	significant in	survives declared family
A	+40,995 (p=.012)	+11,530 (p=.160)	+21,555 (p=.012)	+6,371 (p=.086)	[6,371, 40,995]	2/4	**no**
B	+5,344 (p=.012)	+25,208 (p=.012)	+6,289 (p=.346)	+23,747 (p=.025)	[5,344, 25,208]	3/4	**no**

On Corpus-A the excess is significant under only two of four baselines; on Corpus-B, three of four, but the flexible spline absorbs it. Under neither corpus does the cross-domain coupling excess survive the declared baseline family as a robust claim. **The audit weakens the headline — stated plainly (F2/F4): the audited number is the headline number from here on.**

By contrast, the **mechanism decomposition is baseline-robust.** The excitation- given-ladder term of the nested hybrid (how much a Hawkes excitation is worth *after* conditioning on the latent regime; Paper C) is worth its parameters under *every* baseline on both corpora — Corpus-A ΔBIC {constant +17,459, linear +4,591, spline +1,256, changepoint +710}; Corpus-B {+20,577, +20,919, +12,428, +11,780}. The multi-horizon separation (msm−null Δ at the longest horizon) is baseline-robust on Corpus-B (spline +46,046 the extreme) but **not** on Corpus-A (constant +145, linear −29, spline +10,511, changepoint −1,021 — sign-unstable). The cross-country synchronisation statistic, audited on its own axis (per-country detrend degree ∈ {1,2,3} rather than μ(t)), is robust in five of six Corpus-A domains (all but GEOPHYSICAL) and in both Corpus-B domains. **What is fragile is the cross-domain coupling magnitude; what is robust is the within-domain mechanism structure** — a clean separation the audit earns.

3.2 Detection re-weighting

A constant under-detection cannot manufacture the excess: it is absorbed by the baseline intercept, and the excess is invariant to it (max relative deviation 0.005 on Corpus-A, 0.015 on Corpus-B). A *smoothly time-varying* detection is a baseline effect, already swept by §3.1. What remains is a *self-exciting* detection filter (§4), whose ACLED-calibrated capture–recapture bound is registered-access and pending — named, not assumed away.

3.3 The attention control (record v3)

The per-stream, smooth μ(t) family structurally cannot represent a **shared, spiky, contemporaneous** attention shock — the very structure a media-attention event imprints across domains at once. So the μ(t) sweep is *not* a substitute for an attention control (that is the point of declaring the channel separately). We add a shared contemporaneous attention covariate — a *deterministic maximal-shared-modulation* stress construct, the variance-maximal single shared modulation the residuals admit, a tight heuristic upper bound on any real attention shock's reach (not a live GDELT pull; the per-country GDELT DOC-API loader is built as code, the real run gated). Measured:

- **Corpus-B:** the covariate is **0.986 unspanned** by the μ(t) family (cross-domain residual correlation 0.201). Added to the model, it **absorbs the entire remaining excess, driving it to or below zero, below significance.** A placebo — the same covariate with its cross-domain timing destroyed — leaves the excess standing: the absorption is specifically the covariate's cross-domain contemporaneous alignment, not a generic regressor effect. Across the baseline family it absorbs the excess where it was significant (linear, constant) and leaves it standing under changepoint — downside-only, as pre-committed.
- **Corpus-A:** 0.837 unspanned (residual correlation 0.398); there was no significant excess on the primary baseline for the control to test (the μ(t) audit already carried the verdict), and the placebo does not cleanly confirm specificity here — read cautiously.

All three pending controls (attention, ACLED capture–recapture, real-time vintage) are **downside-only**: none can revive the claim, each can only shrink it, and each was pre-committed by name before its data was in hand.

---

4. The identification problem, stated formally

Let N(t) be the recorded event count, D(t) the underlying occurrence process, and suppose detection is a filter that raises the recording probability after a recorded event. Two data-generating stories then produce near-identical records:

- **(a) True self-excitation.** D is a Hawkes process: each occurrence raises the conditional intensity of the next, $\lambda(t) = \mu + \Sigma_{t_i<t} \alpha e^{-\beta(t-t_i)}$.
- **(b) Self-exciting detection over Poisson occurrence.** D is homogeneous Poisson (no contagion), but the *detection* probability π(t) is itself self-exciting — each *recorded* event raises π for a window — so the *recorded* process N inherits a self-exciting conditional intensity indistinguishable at second order from (a).

Hawkes (1971) proved that a self-exciting and a doubly-stochastic process can share identical second-order properties; Filimonov & Sornette (2015) showed a Hawkes fit to non-stationary/regime-switching Poisson data manufactures spurious near-critical branching ratios; Wheatley, Filimonov & Sornette (2016) showed misspecified immigration inflates branching-ratio estimates — **the bias always runs toward false self-excitation.** The detection-filter story (b) is the same degeneracy with the confound located in the instrument rather than the process. The consequence is a hard rule (F4): **a media-derived or editorially-selected stream cannot support a self-excitation claim without a detection-model control.** This is why the `response_endogeneity` flag and the attention control (§3.3) exist, and why the programme's own audited verdict on Corpus-B rests on the attention control, not the raw likelihood. The programme's discrimination atlas quantifies the degeneracy on known truth: regime-switching Poisson truth ($\eta = 0$) is fit by a Hawkes arm as $\hat{\eta} = 0.80$ (Paper A §S). The identification problem is not a philosophical caveat; it is a measured, near-critical bias with a named control.

---

5. The vintage-archive gap: no ALFRED for crisis data

Macroeconomics has ALFRED — the archival companion to FRED — which preserves every released *vintage* of a revised series, so a forecaster can fit on exactly the data available at a past date and test out-of-sample honestly. **Crisis data has no ALFRED.** UCDP GED is re-released annually as a *full* revision of prior years, not an append-only log; EM-DAT is continuously re-edited as reports arrive; the harmonized corpora are assembled retrospectively. A "held-out" split over such a corpus is edited with post-split knowledge — it is not out-of-sample, and calling it so is a category error (F5). Exactly one stream in the charter carries a genuine real-time vintage: **UCDP Candidate Events**, the monthly pre-revision release, which is the single legitimate out-of-sample check in the programme (wired, pending the loader). The infrastructure need is concrete and nameable: a crisis-data vintage archive — dated snapshots of every event corpus at release — without which no crisis point-process model can be validated out-of-sample, and every "forecast" in the literature is retrospective. The programme's discipline in the interim is to label every retrospective split as such and never call it out-of-sample, and to report the one real-time slice honestly as pending.

---

6. Recommendations, per stream

Each recommendation is tagged **[measured]** where it follows from a sensitivity measured on that stream in this tournament, or **[declared]** where it applies the measured *general* sensitivities of §3 (coupling magnitude is baseline-fragile and attention-absorbable; mechanism structure is baseline-robust; the excess is invariant to a constant under-detection) to that stream's declared artifact channel — the stream is charter-admitted but not yet fit, so it inherits the measured rule rather than a stream-specific measurement. Both tags therefore trace to a measured sensitivity; neither is an unmeasured assertion. All the declarations are in Appendix B.

- **Delannoy corpus (Corpus-A; open).** [measured] The cross-domain coupling excess is baseline-fragile (significant in 2/4, range [6,371, 40,995] ΔBIC); do not report a coupling number from this corpus without the full μ(t) family and the range. [declared] It inherits response-endogenous EM-DAT and expert-dated economic components and a GTD-1993 gap; the audit must probe them.
- **UCDP GED (Corpus-B conflict; open).** [measured] The conflict-domain synchronisation is detrend-robust and the excess is invariant to constant under-detection — but the coupling excess is absorbed by a shared attention covariate (§3.3). Report conflict coupling only after the attention control. [declared] Newswire detection is self-exciting even though events are fatality-defined (not response-defined); instrument it with GDELT. Pin the annual release; cross-version comparisons are vintage-sensitive.
- **WHO-DON (Corpus-B epidemic; open).** [measured] Epidemic-domain synchronisation is detrend-robust; the excess absorbs into the attention covariate as with conflict. [declared] `response_endogeneity: true` — WHO's "need for assistance"

criterion makes a DON item partly a response signal, so no self-excitation claim on this stream without a detection/response control; dummy the IHR-2005 and COVID breaks.

- **EM-DAT (disasters; registered).** [declared] Admit only the impact-threshold subset (≥10 deaths / ≥100 affected) and drop declaration/appeal rows at load — the `response_endogeneity` flag records that the raw source is endogenous even where the loaded subset is not. Dummy the 2000 reporting-capacity ramp; use post-2000 as the honest window. Fit-only; never commit rows.
- **Laeven–Valencia / financial chronologies (economic; open, sparse).** [declared] Dated partly by the policy intervention that answered the crisis — an outcome, never a covariate. Country-year / Arc-3 pairs only, below the minimum-N for regime claims (F3).
- **Powell–Thyne coups (open).** [declared] Press-reported overt acts with asymmetric detection of *failed* attempts; sparse — expect wide intervals and cite the atlas cell before any coupling number.
- **FAO Food Price Index (covariate; open).** [declared] The one detection channel independent of journalism (market price observation) — its value is precisely as an exogenous instrument for food-price→unrest pairs; index revisions are the only break.
- **GDELT (attention covariate only; open).** [measured] A shared contemporaneous attention covariate absorbs the Corpus-B cross-domain excess below significance and is 0.986 unspanned by the μ(t) family — so GDELT is the load-bearing detection control, not an event stream. [declared] Volume only, never events (date-index break 2013; ~55% key-field accuracy). Instrument the channel; do not count it.
- **ACLED (unrest; registered).** [declared] Its value here is the dual-coded capture–recapture detection bound on *self-exciting* detection — the one pending control §3.2 names. Derived statistics only; honest window 2018–; dummy the source-list expansions.
- **GTD (terrorism; restricted).** [declared] Fit-only, never leaves the local cache; four declared collection-regime breaks and a missing 1993 must be dummied before any within-stream excitation number.
- **V-Dem ERT (validator; open).** [declared] A latent-state external validator only, never an event stream (annual, thresholded, retroactively re-estimated) — use to validate a recovered regime path, not to fit.

The through-line: the two measured sensitivities — coupling magnitude is baseline-fragile and attention-absorbable; within-domain mechanism structure is baseline-robust — generalize to a rule the declaration standard operationalizes. *Declare the channel, sweep the baseline family, instrument the attention shock, and never call a retrospective split out-of-sample* — before any crisis coupling or excitation number is called real.

---

7. Limitations

- The attention control uses a maximal-shared-modulation upper bound, not a live per-country GDELT series over the full window (unavailable to a determinism-gated run; DOC-API coverage begins ~2017). It is downside-only; the real run can only push the audited claim further down.
- The ACLED capture–recapture detection bound is registered-access and pending; the vintage check awaits the UCDP Candidate loader. Both are named, not assumed.
- The audited signal is onset-timing coupling; severity/magnitude channels are a separate deliverable and are not audited here.
- Nothing here forecasts, warns, or scores; every number is a retrospective, dataset-level measurement on public, aggregate data (F7).

---

8. Data and code availability

```
python3 -m venv venv && venv/bin/pip install -r requirements.txt
venv/bin/python x0_verify.py            # the F1 anchor gate — must print X0 PASS
./fetch_data.sh                         # open-tier data (Delannoy, UCDP GED, WHO-DON)
venv/bin/python x3_audit.py --profile default --seed 20260903   # → out/v1/x3/ (the μ(t) sweep)
```

Records: `out/v1/x3/` (the μ(t) baseline sweep and detection invariance; `x3.json`, `x3_sensitivity_table.png`, `x3_decomposition.png`) and `out/v3/` (the attention control; `controls.json`, `v3_attention_absorption.png`, `v3_covariate_structure.png`). The declaration schema is `fenocrisis/data/ schema.py` (validator), `docs/artifact-schema.md` (prose), the `fenocrisis/data/artifacts/*.yaml` instances (the eleven charter streams as twelve declaration files — UCDP GED and its real-time Candidate vintage each carry their own), and `papers/02-artifact-schema.json` (this paper's JSON appendix). Environment: Python 3.11.15, NumPy 2.4.6, SciPy 1.17.1, Matplotlib 3.11.1. License: CC BY 4.0.

Deposit DOIs (published 2026-08-31). This paper is deposited with its declaration schema beside it.

- This paper (with `papers/02-artifact-schema.json`) — `10.5281/zenodo.22180507`.
- Record `out/v0` (the discrimination atlas) — `10.5281/zenodo.22180515`.

Each DOI above is **published**: the deposits went live on Zenodo on **2026-08-31** and every identifier resolves. The source repository is private and will remain so, so each deposit is standalone and carries its own reproduction materials; deposit metadata deliberately omits repository links. *(Corrected 2026-08-31 — publication day: until today this paragraph read "reserved on Zenodo and its deposit is a private draft … the record is not yet published", with "(Zenodo DOI reserved; publication pending)" wherever a DOI appeared. That staging state ended when the owner pressed Publish on 2026-08-31.)*

Appendix A — the F4 artifact-channel declaration schema

Every declaration is a mapping with all thirteen top-level fields; the validator reports every missing or mistyped field at once. Types and meanings:

field	type	meaning
`stream`	string	Stable machine id (matches the loader key).
`name`	string	Source database name + reference.
`serves`	string | list	Which question(s)/role the stream serves.
`resolution`	string	Native unit of observation.
`detection_mechanism`	string	**The photon channel**: how an event is detected/recorded.
`threshold_criteria`	string | dict | list	Inclusion thresholds defining an admitted event.
`response_endogeneity`	dict `{flag: bool, notes: string}`	**The F4 heart**: are events defined partly by institutional response?
`structural_breaks`	list of `{date, description}`	Known breaks **with dates** (ISO date, year, or recurring keyword `annual` / `rolling` / `monthly` / `version`).
`license`	dict `{class, name, url, redistribution}`	`class ∈ {open, registered, restricted}`; the gate reads `class`.
`vintage_policy`	dict `{type, notes}`	`type ∈ {retrospective, real-time, revised, frozen, mixed, static}` (F5).
`source`	dict `{url, release_version, checksum, checksum_policy}`	Retrieval + pinning; `checksum_policy ∈ {pinned, recorded-on-fetch, snapshot}`.
`coverage`	dict `{temporal, spatial}`	Temporal span and spatial scope.
`citation`	string	Canonical reference, verified before release (Standing Rule 11).

The machine-readable form is `papers/02-artifact-schema.json` (JSON Schema 2020-12). The gate reads `license.class`: **open** rows may be committed; **registered** rows are fit-only (derived statistics only); **restricted** rows are fit-only and never leave the local cache. A `registered` / `restricted` row-level frame raises if a write targets a committed path — the invariant is enforced by construction, not by discipline.

Worked example — WHO Disease Outbreak News (`who_don`, response-endogenous)

```
stream: who_don
detection_mechanism: >-
  WHO editorial selection of which outbreaks warrant a public Disease Outbreak
  News item — an institutional editorial filter, explicitly including a "need for
  international assistance" consideration.
response_endogeneity:
  flag: true
  notes: >-
    WHO's inclusion criterion incorporates "need for assistance," so a DON item is
    partly a response signal — media-derived/editorial detection cannot support a
    self-excitation claim without a detection-model control (F4).
structural_breaks:
  - {date: 1996, description: WHO DON series begins}
  - {date: 2007, description: IHR-2005 enters force, changing notification obligations}
  - {date: 2020, description: COVID-19 distorts DON publishing volume}
license: {class: open, name: CC BY 4.0 (Carlson et al. 2023 PLOS Global Public Health
derivative), redistribution: true}
vintage_policy: {type: retrospective, notes: never called out-of-sample (F5)}
```

The `flag: true` is what forbids a self-excitation claim on this stream without the attention control of §3.3 — the declaration is not documentation, it is a gate.

Dated note (2026-08-25) — a declaration standard audited against itself. This paper's reference list and the live `who_don.yaml` declaration disagreed on which paper describes the `cghss/dons` derivative: the reference list here had Carlson et al. 2023, *PLOS Global Public Health* 3(1):e0001083; the declaration file, the SPECS data charter and `docs/corpus-choices.md` had "Carlson et al. 2022, *Scientific Data* 9:709, doi:10.1038/s41597-022-01797-2". Both DOIs were checked against Crossref on 2026-08-25. The *Scientific Data* DOI resolves to Torres Munguía, Badarau, Díaz Pavez, Martínez-Zarzoso & Wacker (2022), "A global dataset of pandemic- and epidemic-prone disease outbreaks", *Scientific Data* **9**, 683 — a different team, a different artifact, and article 683 rather than the cited 709. The derivative's own repository names the retrospective-database paper by title in its README. **The reference list below was right and the declaration surfaces were wrong; they now agree with it**, and the wrong DOI is carried as an explicit do-not-confuse cross-reference in `who_don.yaml`. That the disagreement lived inside a provenance standard for two weeks is the honest illustration of why F4 declarations need a mechanical cross-surface check, which this repo does not yet have — recorded here rather than quietly fixed.

Appendix B — the eleven charter streams

key	class	serves	detection channel (short)	response-endog.
`delannoy`	open	Q1 substrate (Corpus-A)	inherits all six component channels	true (inherited)
`ucdp_ged`	open	anchor conflict; the one real-time vintage	newswire + fatality thresholds	false
`emdat`	registered	disasters (impact-threshold subset)	relief/reinsurance reporting; pre-2000 ramp	true (raw)
`who_don`	open	epidemic stream	WHO editorial selection incl. "need for assistance"	true
`fao_fpi`	open	food-price coupling covariate	market price observation (journalism-independent)	false
`laeven_valencia`	open	financial-crisis chronology	expert dating; policy-intervention by definition	true
`powell_thyne`	open	coup-contagion pair	press-reported overt acts; asymmetric failed-attempt detection	false
`gdelt`	open	media-attention covariate **only**	it *is* the artifact channel, instrumented (volume only)	n/a
`acled`	registered	dense unrest stream (2018– honest window)	newswire + source-list expansions	false
`gtd`	restricted	long dense terrorism stream (fit-only)	newswire; four collection-regime breaks; 1993 missing	false
`vdem_ert`	open	latent-state external validator **only**	expert-coder measurement model	n/a

The `ucdp_ged` row's real-time UCDP Candidate vintage carries its own twelfth declaration file (`ucdp_candidate.yaml`, `vintage_policy.type: real-time`), which SPECS §6 counts within the anchor conflict stream. Full declarations: `fenocrisis/data/artifacts/*.yaml` (twelve files; eleven charter streams).

References

(Verified against the actual papers before release — Standing Rule 11.)

- Carlson, C. J., Boyce, M. R., Dunne, M., Graeden, E., Lin, J., et al. 2023. The World Health Organization's Disease Outbreak News: a retrospective database. *PLOS Global Public Health* 3(1), e0001083. doi:10.1371/journal.pgph.0001083. (The `cghss/dons` machine-readable derivative. Title, venue, volume/issue, article id and author list verified against Crossref 2026-08-25.)
- Delannoy, L., et al. 2025. Dynamics of the polycrisis: temporal trends, spatial distribution, and co-occurrences of national shocks (1970–2019). *Global Sustainability*

8, e24. doi:10.1017/sus.2025.10008.

- Filimonov, V., & Sornette, D. 2015. Apparent criticality and calibration issues in the Hawkes self-excited point process model. *Quantitative Finance* 15(8), 1293–1314.
- Hawkes, A. G. 1971. Spectra of some self-exciting and mutually exciting point processes. *Biometrika* 58(1), 83–90.
- Sundberg, R., & Melander, E. 2013. Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research* 50(4), 523–532.
- Wheatley, S., Filimonov, V., & Sornette, D. 2016. The Hawkes process with renewal immigration & its estimation with an EM algorithm. *Computational Statistics & Data Analysis* 94, 120–135.

Full program reference list: SPECS.md §10.