X-Integrity Technical Review FAQ

1) What is this system, in one sentence?
An evidence-routing layer that scores media for structural integrity, physical consistency, and social corroboration before it reaches peak amplification.

2) Is this a truth engine?
No. It does not decide ideology, opinion, or narrative truth. It measures consistency between a media claim and the observable world.

3) What does it actually output?
Two things:
- **Forensic state**: evidence-oriented outputs such as structural score, physical consistency score, corroboration score, confidence, and replay metadata.
- **Treatment hint**: a platform-facing recommendation such as supported, indeterminate, conflicted, or constructed-valid.

4) Why separate forensic state from treatment hint?
Because the protocol should describe what it observed, while X keeps control of what to do with that observation in ranking, review, notes, or labeling.

5) Does this censor content?
Not by itself. The default posture is fail-to-neutral. If evidence is incomplete or infrastructure is degraded, the system should step back rather than over-claim.

6) What is the main product value for X?
Closing the verification-virality gap:
- lower reach for high-confidence contradictions before they scale
- faster prioritization for Community Notes and internal review
- a replayable audit trail for legal, policy, and regulator review
- a cleaner environment for advertisers around breaking news

7) How does it fit technically?
As asynchronous middleware:
1. media uploaded to X
2. X sends media metadata and blob location to the engine
3. engine runs Gatekeeper, PRI, and Consensus
4. engine returns forensic state plus replay package
5. X decides ranking, review, or labeling behavior

8) What are the three pillars?
- **Gatekeeper**: file/path integrity and lane routing

- **PRI**: physical reality checks such as solar-shadow consistency
- **Consensus**: social corroboration against expected witness density

9) What is lane-aware routing?
It distinguishes raw-candidate media from constructed media. A professional news package should not be treated like fake raw footage just because it contains edits or graphics.

10) What happens when the system is unsure?
It returns **Indeterminate** or **Insufficient-Data** and avoids hard restriction. That is part of the fail-to-neutral posture.

11) What is snapshot locking?
Consensus queries persist a `snapshot_id` so later audits can replay the exact same world-state used during the original decision.

12) Why is replayability important?
Because a platform decision needs to be defensible. Replayability lets X show:
- what evidence existed at the time
- which policy version was active
- which module versions were used
- why the resulting treatment hint was generated

13) What would cause a high-confidence conflict?
Examples:
- shadow geometry contradicts claimed time/place
- a major urban event claims mass impact but has no corroborating witness footprint
- file-level or structural evidence strongly suggests a mismatched capture path

14) What does "social vacuum" mean?
For a high-impact claim in a dense region, observed corroboration is far below expected density. In those cases, silence itself becomes evidence.

15) How do you prevent unfair treatment of rural events?
The Consensus model includes:
- regional coefficients
- sparse-region fairness rules
- cold-start scaling for first reports
- degraded-mode neutrality when upstream systems lag

16) What happens if GEI is slow or down?
The system should not convert infrastructure failures into accusations. It emits

insufficient-data, shifts to hold or neutral behavior, and records the degraded state.

17) What is the rollout path for X?
Recommended ladder:
1. advisory logging
2. review-priority routing
3. ranking-assist signals
4. public integrity labels for high-confidence cases

18) What are the key success metrics?
- faster conflict detection than manual/community systems
- lower reach for later-confirmed fabrications
- low false-restriction rate in sparse or low-light conditions
- 100% replay consistency for audited cases

19) What are the biggest technical blockers?
- automated PRI shadow extraction at scale
- live Global Event Index integration
- adversarial hardening for coordinated spoof attempts
- strong audit UI for reviewer comprehension

20) Why now?
Because pixel-only detection is losing ground. Consistency across physics, structure, and social density is a stronger long-term basis for media integrity than artifact hunting alone.