Fractal Account Moderation Manifest v1.0

Detection of Impersonation & Thematic Sub-Accounts

Vision & Philosophy

In the age of scams and impersonations (e.g., fake VIP accounts), we need moderation that is *smarter, not heavier*. Inspired by fractal relations: the original account is the **core**, while impersonators or thematic sub-accounts are **leaves**. This is not central bans, but a **swarm of detectors**: AI scans bios, posts, and connections (cosine similarity, Jaccard overlap), building authenticity awareness from relational differences. Goal: Automatically catch 99% of scams, while linking legitimate sub-accounts as "fractal leaves" — verified without chaos.

Original vs. Fakes

Original accounts are easy to detect: millions of followers, rich activity, verified badge, consistent posting. Fakes: new accounts, low posts, stolen bios ("Official Elon channel"). Easy signals include follower count, activity, bio similarity. AI swarm flags isolated fakes while mapping verified thematic sub-accounts.

Thematic Sub-Accounts (Fractal Leaves)

Legitimate sub-accounts should be linked as *fractal leaves*. Example:
• @elonmusk (core) → @elonmusk_tesla (EV/Neuralink)
• @elonmusk_xai (AI/Grok).
Linked by pinned posts, hashtags, and follower overlap. This forms a **fractal structure**: core + leaves, zero chaos.

Similarity Metrics & Thresholds

Standard ML metrics can catch impersonation efficiently. Example thresholds:

Metric	Description	Suggested Threshold	Notes
Cosine Similarity (bio/text)	Angle between embeddings of bios/posts.	0.79–0.9	High precision, fewer false positives.
Jaccard Similarity (hashtags/followers)	Overlap ratio of sets.	0.6–0.8	Good for copied bios or follower lists.
Face Embedding Similarity (avatar)	Optional, cosine match for avatars.	0.8–0.95	High precision, opt-in due to privacy.
Overall F-Score Threshold	Weighted combo of metrics.	0.7–0.85	Balance recall & precision; ROC-tuned.

Integration with AJ Mesh

This framework naturally extends to AJ Mesh:
• Mesh nodes observe bios/posts (delta cosine <0.8 = flag).
• Verification offloaded to least-loaded node (swarm resilience).
• Not bans, but emergent authenticity awareness from relations.
This ensures robust, scalable impersonation detection with minimal human intervention.

Conclusion

Fractal Account Moderation = core + leaves, swarm detectors, ML thresholds.
A practical path for X (or any social platform) to eliminate 99% of scam impersonations,
while preserving thematic sub-accounts in a transparent, authentic network. © 2025 Antzedek & Eliasz — AJ Power Fractal Presence