

Recursive Polycentric Moderation (RPM)

Application of the Localities — Neighbors — Distsants Framework to Digital Platforms

Sylvain Lebel

2026

Framework reference:

Recursive Polycentric Governance (RPG)
<https://zenodo.org/records/18306005>

Project website:

<https://sylebel.net>

License: CC-BY-4.0

Abstract

This paper introduces **Recursive Polycentric Moderation (RPM)**, a distributed moderation framework for digital platforms and social networks.

RPM constitutes a direct application of the **Recursive Polycentric Governance (RPG)** framework.

The core thesis is that moderation should be neither centralized nor purely community-based, but **structurally distributed**, relying on judgments of unacceptability emerging from partially correlated communities.

RPM preserves local normative autonomy, limits extreme content without global censorship, resists brigading and echo chambers, ensures procedural transparency, and remains computationally scalable.

The framework defines no substantive moral standards. It specifies only a **minimal procedural architecture** enabling distributed normativity.

1 Problem Statement

Digital platforms face a structural contradiction:

- freedom of expression requires irreducible normative pluralism;
- virality produces real trans-community externalities;
- centralized moderation yields both over-censorship and under-regulation.

Existing approaches rely on centralized authority, isolated communities, or unstructured crowd mechanisms. None resolves context sensitivity, manipulation resistance, and procedural legitimacy simultaneously.

2 Fundamental Principle

Every content publication constitutes a **local decision** capable of generating externalities.

Moderation does not define truth or morality. It answers only:

Does this content cross a trans-community threshold of unacceptability?

3 Definition of Entities

3.1 Digital Locality

A locality is a community characterized by recurring interaction, shared implicit norms, and minimal deliberative capacity.

3.2 Neighbors

Neighbors are communities with overlapping membership, frequent interaction, and contextual proximity.

3.3 Distant

Distant are communities without member overlap, without correlated interests, and thematically decoupled.

They provide **structural independence of judgment**.

4 Triggering of Review

Content may be submitted for review via community reports, non-coercive algorithmic detection, rapid cross-community propagation, or random calibration sampling.

No automatic removal occurs.

5 Distributed Judgment

5.1 Judgment Question

“Does this content exceed a reasonable threshold of acceptability beyond simple disagreement?”

Judgment is binary and contextual.

5.2 Constitutional Rule

Content is judged unacceptable if and only if:

- at least 50% of neighbors vote unacceptable;
- and at least 50% of distants vote unacceptable.

Figure 1 — RPM Protocol Flow

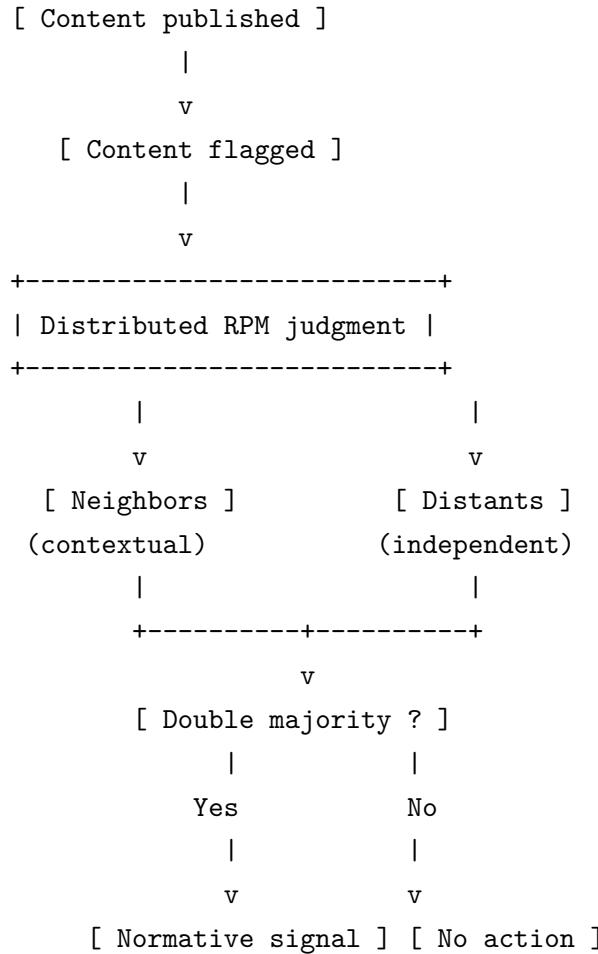


Figure 1. Conceptual flow of the Recursive Polycentric Moderation protocol.

5.3 Illustrative Example

Consider a post in the **#CryptoCurrency** community:

“Central banks are criminal organizations.”

Neighbors (#Finance, #Economics, #TechPolicy) understand the ideological context; 40% judge the content unacceptable.

Distants (#Gardening, #ClassicalMusic, #Medicine) are structurally independent; 35% judge the content unacceptable.

Outcome: no double majority reached.

Result: no moderation action.

6 Public Normative Signal

When issued, the signal consists of a visible marker, an aggregated explanation, and a fully auditable public record.

7 Sufficiency of Response

Authors may respond to a signal. Responses are evaluated by the same judges under the same rule.

8 Relational Sanctions

Sanctions are relational, graduated, reversible, and never global.

9 Normative Recursivity

Every act of moderation may itself be evaluated under the same protocol.

Recursivity is complete but self-limiting through relational cost. Moderation never stabilizes as sovereign authority.

10 Conclusion

Recursive Polycentric Moderation provides an alternative to centralized censorship and unregulated informational chaos.

Moderation becomes:

an emergent property of network structure,
not a sovereign power over speech.

Framework reference:

Recursive Polycentric Governance (RPG)

<https://zenodo.org/records/18306005>

Project website:

<https://sylebel.net>

License: CC-BY-4.0