THE FEDERALISTS OF THE INTERNET? WHAT ONLINE PLATFORMS CAN LEARN FROM REDDIT'S DECENTRALIZED CONTENT MODERATION SCHEME

Content moderation didn't begin on the "front page of the internet," but the most anonymous site in the world has perfected the art of checks and balances.
-Courtney Linder[1]

Abstract

As online platforms become the primary carriers of news and information in society, their content moderation policies will have increasing power over users' daily lives. Centralized moderation is inappropriate and inefficient because it forces all online communities to adhere to the same (and often arbitrary) standards. Decentralized moderation, however, allows most moderation decisions to be made by users who are intimately aware of the contours of individual online communities. Such decisions are less likely to restrict content that is not violative of an online communities' standards.

This concept of subsidiarity mirrors the arguments made in favor of federalism in which the bulk of regulation takes place at lower levels of government and national regulation is reserved for particularly pervasive issues. In fact, as online platforms continue to develop, their moderation policies are beginning to resemble governance structures that are more akin to constitutions than they are to contracts. Reddit, for example, has not only embraced a decentralized content moderation scheme, but its policies mirror those of a federal system of government. Reddit's federalism has allowed it to identify pervasive issues that are appropriate for its centralized admins to moderate, leaving the day-to-day moderation of its platform to users. It has also increased Reddit's efficiency and reduced its power to censor users while simultaneously increasing the power users have in the governance of the platform. As online platforms continue to adapt their moderation policies, they should look to Reddit and its federal structure to reduce costs and empower users.

---

[1] Courtney Linder, *How Reddit Birthed a Community of Digital Vigilantes*, POPULAR MECHANICS (Nov. 1 2019), https://www.popularmechanics.com/technology/apps/a29620698/how-reddit-works/.

Healthy online communities must create "governance mechanisms that structure participation . . . to facilitate cooperation and prevent abuse."[2] In other words, online communities must moderate content. Moderation is necessary because online communities are akin to actual communities; just as physical communities cannot function without regulation, online communities cannot function without moderation.[3] The need for moderation of digital communities raises these questions: How much moderation is necessary and who should do the moderating? Attempts to answer these questions typically "track the usual legal debates about hierarchy and *federalism*."[4] Centralized moderation—like a strong national government—is consistent across the platform and relies on one set of policymakers to make decisions.[5] In contrast, decentralized moderation relies on local moderators who have "experience with their members' sense of what is off topic" in the creation and enforcement of policies.[6] It also promotes diversity and "permits those with ideological differences to agree and disagree."[7]

Reddit is often considered the prototypical model for a decentralized social media platform, while Facebook and Twitter are often considered to be the most centrally moderated. While it is true that Reddit is decentralized, this paper contends that Reddit's unique structure goes beyond mere decentralization and is, in fact, a governance model that is analogous to American federalism. Part I of this paper explains Reddit's content moderation scheme. Then, Part II juxtaposes Reddit's policies against American federalism to establish that Reddit's moderation scheme is a federal governance structure. Finally, Part III identifies a series of lessons that other platforms can learn from Reddit's federal model of digital governance.

I. REDDIT'S APPROACH TO CONTENT MODERATION

From its inception, Reddit has taken a libertarian approach to content moderation.[8] Reddit justifies its approach to content moderation on the belief that the internet was created to be a bastion for free expression; Reddit's role is to facilitate all manner of expression on the internet no

---

[2] James Grimmelmann, *The Virtues of Moderation*, 17 YALE J.L. & TECH. 42, 47 (2015).
[3] Linder, *supra* note 1 (asserting that "'[o]nline civilization is just another kind of civilization'").
[4] Grimmelmann, *supra* note 2, at 69 (emphasis added).
[5] *Id.* at 69–70.
[6] *Id.*
[7] *Id.*
[8] *See* Jonathan Peters, *Sexual Content and Social Media Moderation*, 59 WASHBURN L.J. 469, 484 (2020).

matter how distasteful, immoral, or objectionable it might be. Reddit's commitment to free expression was made explicit via a leaked memo sent by the company's then-CEO, Yishan Wong, in response to criticism the company faced for banning a subreddit (r/Jailbait) that posted nude photos of both adult and underage girls but allowed the user who posted these photos to remain on the platform.[9] The memo implied that the *only* reason r/Jailbait was banned was because it violated U.S. law.[10] Thus, only the posting of nude underage girls was problematic for Reddit; the posting of adult pornographic material, no matter how distasteful, was not a problem. The memo stated that Reddit "stand[s] for free speech . . . . We will not ban legal content even if we find it odious or if we personally condemn it."[11] The memo further explained that the company's goal is to "become a universal platform for human discourse . . . . [I]t would not do if . . . we decided to censor things simply because they were distasteful."[12]

Reddit's hands-off approach to content moderation is further emphasized in its content policy, the rules which govern all of the platform's users.[13] The policy contains only eight rules, each of which is comprised of no more than three sentences.[14] Furthermore, the entire policy is made up of less than 250 words and is flippant in tone.[15] Reddit's carefree approach to its content policy is indicative of the site's attitude toward rules of any kind; it imposes them because it has to, not because it wants to.[16]

---

[9] *See* Andrew Marantz, *Reddit and the Struggle to Detoxify the Internet*, THE NEW YORKER (12 Mar. 2018), https://www.newyorker.com/magazine/2018/03/19/reddit-and-the-struggle-to-detoxify-the-internet.
[10] *Id.*
[11] Dean Takahashi, *Reddit CEO Issues Internal Memo on Free Speech and Outing of Violentacrez*, VENTURE BEAT (17 Oct. 2021), https://venturebeat.com/2012/10/17/reddit-ceo-issues-internal-memo-on-free-speech-and-outing-of-violentacrez/.
[12] Marantz, *supra* note 8. In 2018, Reddit's current CEO, Steve Huffman, re-emphasized Reddit's commitment to free expression when he said that Reddit is "a place for open and honest conversations—'open and honest' meaning authentic, meaning messy, meaning the best and worst and realest and weirdest parts of humanity." *Id. See also* Grimmelmann, *supra* note 2, at 100 ("Reddit is passionate about creating strong communities but completely indifferent as to whether those communities collaborate for good or for ill.").
[13] REDDIT CONTENT POLICY, https://www.redditinc.com/policies/content-policy (last visited 2 May 2021).
[14] *Id.*
[15] For example, Rule Seven begins with the phrase "keep it legal" while Rule Eight whimsically states, in part, "Don't break the site." *Id.*
[16] When one compares the moderation policies of Reddit to a centrally moderated platform, such as Facebook, its libertarian approach becomes even clearer. Facebook's content policy is comprised of six "Parts" with a total of twenty-seven "Subparts." COMMUNITY STANDARDS, https://www.facebook.com/communitystandards/ (last visited 2 May 2021). Furthermore, not only are the community standards detailed and formal in nature, but they are comprised of nearly 4,000 words. *Id.* Each Subpart links to a separate webpage that explains the policy rationale for the rule and provides examples of rule violations. *See, e.g.*, VIOLENCE AND INCITEMENT POLICY RATIONALE, https://www.facebook.com/communitystandards/credible_violence/ (last visited 2 May 2021).

Reddit's reticence to moderate content is also illustrated through the formal mechanisms by which it polices content. There are three tiers of content moderation on Reddit which are organized hierarchically: admins (the first tier), mods (the second), and users (the third). Admins are Reddit employees responsible for enforcing Reddit's content policy; however, Reddit's decentralized content moderation scheme results in little moderation taking place at the admin level.[17] Rather, most of the moderation is carried out by mods—Reddit users who voluntarily supervise a subreddit.[18] Like admins, mods enforce Reddit's content policies, but they also create and enforce additional policies within their subreddit so long as they do not conflict with Reddit's content policy.[19] Furthermore, mods have significant editorial discretion as it relates to removing content or banning users from their subreddit.[20] As a result of Reddit admin's ambivalence to moderating content, and the wide latitude afforded to mods, mods removed nearly twice as much content as admins in 2020.[21]

Reddit's third tier of moderation, the users, also play a role in moderating content. Users moderate content by voting. Popular content (which gets "upvoted") is pushed to the top of the page while unpopular content (which gets "downvoted") is pushed to the bottom.[22] While unpopular content is never actually removed as a byproduct of being downvoted, it is less likely to be seen by users of the subreddit.[23] This not only moderates content by keeping only information relevant to the community displayed prominently, it also serves as a signal to users of the values and norms of the subreddit which are to be emulated by members who wish to contribute to the community.

---

[17] SPANDANA SINGH, NEW AMERICA, EVERYTHING IN MODERATION 26 (2019). Reddit employs about forty admins to enforce its content policy and much of their time is spent taking down illegal content or harassment as opposed to content that violates Reddit's other policies. *See* ROBYN CAPLAN, CONTENT OR CONTEXT MODERATION? 21 (2018).

[18] *See* SINGH, *supra* note 16, at 26.

[19] *Id.*

[20] Mods can even remove content merely for being "off-topic" even if it does not violate any specific rules. *Id.*

[21] TRANSPARENCY REPORT 2020, https://www.redditinc.com/policies/transparency-report-2020-1 (last visited 2 May 2021). The rate at which mods removed content as compared to admins does not consider the amount of content that was effectively buried due to the voting system. This indicates that the amount of content moderated (as opposed to removed) by users and mods is much greater than the data provided by the Transparency Report. However, since there are no quantifiable metrics to determine when content has been sufficiently downvoted so as to be effectively unseeable, there is no way to know exactly how much more users and mods moderated content compared to admins.

[22] Kevin Morris, *The Greatest Story Reddit Ever Told*, THE KERNEL (2 Nov. 2014), https://perma.cc/MH7A-KHAQ.

[23] *Id.*

II. ANALOGIES TO FEDERALISM

Ostensibly, it may seem inappropriate to compare a social media platform to a federal system of government. The proliferation of online platforms and the ubiquity of the internet more broadly, however, necessitates formalizing the way we classify and describe content moderation policies and how those policies interact with those who subject themselves to the jurisdiction of the company.[24] A comparison to government ends up being more fitting than one might initially think.[25]

In fact, online platforms are increasingly being compared to democracies,[26] constitutions,[27] and legal regimes.[28] Reddit itself classifies its approach to content moderation as a form of democracy,[29] and the company as a form of quasi-government.[30] That said—and while democracy and constitutions are all too familiar hallmarks of a federal government—the presence of democratizing factors in content moderation policies and analogizing oneself to a constitutional government is not sufficient to establish that Reddit's approach to content moderation is *federal* in nature. After all, many platforms describe themselves as being democratic or governmental when they are assuredly not federal.[31]

---

[24] *See* Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598, 1662 (2018) (arguing that comparing online platforms to a system of government is "both more descriptively accurate and more normatively useful in addressing the infrastructure of" online platforms).

[25] "It might seem outlandish to compare a website to one of the fifty states. But as technology continues to spread and more daily activities move from the physical world into cyberspace, it will become more natural to think of an online community as simply another type of territory in which people spend their time." James Smith, *Online Communities as Territorial Units*, 57 ST. LOUIS U. L.J. 839, 846 (2013); *see also* Kyle Langvardt, *Regulating Online Content Moderation*, 106 GEO. L.J. 1353, 1360 (2018) ("Online platforms are doing what we've only really expected constituted officials of sovereign powers to do.").

[26] *See* SINGH, *supra* note 16, at 5 (describing the democratization of online speech because of the "widespread adoption and penetration of online platforms").

[27] *See* TARLETON GILLESPIE, CUSTODIANS OF THE INTERNET 45 (2018) (opining that content moderation policies are more akin to a constitution than they are to instruction manuals).

[28] *See* Klonick, *supra* note 24, at 1644 (explaining that companies such as Reddit use "American law and legal reasoning as the foundation of their guiding principles").

[29] Reddit's Transparency Report 2020 stated that its approach to content moderation is "akin to a democracy, wherein everyone has the ability to vote and self-organize, follow a set of common rules, establish community-specific norms, and ultimately share some responsibility for how the platform works." TRANSPARENCY REPORT 2020, *supra* note 21.

[30] "[W]e consider ourselves not just a company running a website where one can post links and discuss them, but the *government* of a new type of community. The role and responsibility of a government differs from that of a private corporation, in that it exercises restraint in the usage of its powers." Yishan Wong, *Every Man is Responsible for his Own Soul*, REDDIT BLOG (6 Sept. 2014), https://redditblog.com/2014/09/06/every-man-is-responsible-for-his-own-soul/ (emphasis added).

[31] For example, Mark Zuckerberg is quoted as describing Facebook as "more like a government than a traditional company," but its centralized content moderation policies prevent it from being considered federal. *See* Langvardt, *supra* note 24, at 1357.

Thus, this Part will explain how Reddit's policies are not only democratic and governmental, but more specifically federal. Section A will establish an operative definition of federalism using American federalism as its model for a federalist state. Then, Section B will describe how Reddit's content policy structurally resembles American federalism. Section C will analogize Reddit's admins, mods, and users to the federal government, state governments, and American citizens respectively to demonstrate how Reddit's federal structure interacts with its constituent parts in practice.

A. Defining Digital Federalism: Lessons from the United States

First, it is necessary to point out that short of formally adopting the U.S. Constitution as its content policy, it is impossible for Reddit to exhibit all the components of American federalism. That is not, however, the purpose of this Section. Rather, the objective is to draw analogies between American federalism and Reddit's content policies to illustrate some examples which illustrate that Reddit is loosely federal in structure. In other words, from a functional perspective, Reddit operates in similar ways to American federalism even if it is impossible for its formal structure to exhibit all the components of American federalism. Thus, this Section will identify key tenets of federalism before discussing how those ideas manifest themselves in the American model of federalism. This will prime the later comparison between American federalism and Reddit's form of digital federalism.

The purpose of a federal system of government is to both reflect "foundational pluralism and to maintain the relationship between pluralism and union in the creation and reconciliation of different orders of government."[32] Inherent in this notion is the almost universally accepted assertion that federal governments are comprised of "both an overarching national or central government and a set of regional or subnational governments."[33] Maintaining union between the national and subnational governments, which is primarily accomplished through cooperation between both levels of government, while simultaneously respecting the sovereignty of subnational governments is of the utmost importance in a federal system. Furthermore, in promoting both union and sovereignty, federal governments develop and enhance the individual

---

[32] STEPHEN TIERNEY, FEDERALISM AND CONSTITUTIONAL THEORY 1 (working paper) (draft on file with author).
[33] THOMAS O. HUEGLIN & ALAN FENNA, COMPARATIVE FEDERALISM: A SYSTEMATIC INQUIRY 16 (2015).

rights of their citizens.[34] This is because federal systems are uniquely equipped to protect the enjoyment of individual rights.[35] As a result, a federal system of government is necessarily democratic in nature as democracy is a key component in ensuring individual rights are protected.[36]

Thus, the federal idea can be identified by five tenets. A federal state must: (1) be comprised of both a national government and set of subnational governments; (2) promote territorial sovereignty for subnational governments; (3) maintain some sense of a union that ties together the subnational governments notwithstanding their individual sovereignty; (4) enhance and protect individual rights; and (5) develop democratic processes in furtherance of the protection of individual rights. The first three tenets describe how governments interact with one another, while the last two describe how citizens interact with their government.

American federalism addresses the first three tenets of federalism primarily through the structure of the Constitution and voluntary arrangements between the two levels of government. Not much needs to be said by way of the first tenet of federalism; the Constitution clearly establishes a central government, and it is not disputed that the state governments existed before they were incorporated under the Constitution—and continue to exist today.[37]

As it relates to the promotion of territorial sovereignty, one way the U.S. accomplishes this is through the Tenth Amendment, which establishes that any powers not enumerated are reserved to the states or the people.[38] Thus, the laws of the national government are intended to be only a small portion of law in the U.S. while the states are empowered to actively legislate based on the needs of their communities.[39] That said, the national government does retain supremacy as it relates to its enumerated powers. Federal supremacy is partially maintained through the doctrine of preemption which stands for the proposition that federal law will displace state law when the two conflict.[40]

Although preemption limits the sovereignty of the states, it does so in the name of the third tenet of federalism: unity. The centralization and standardization of certain policies serves to create

---

[34] G. Patrick Lynch, *Protecting Individual Rights Through a Federal System: James Buchanan's View of Federalism*, 34 CONSERVATIVE PERSPECTIVE ON AMERICAN FEDERALISM 153, 160 (2004).

[35] JOSE WOEHRLING, NEGOTIATING DIVERSITY: IDENTITY, PLURALISM, AND DEMOCRACY 105 (2014).

[36] *See* PALMER EDMUNDS, LAW AND CIVILIZATION 324–343 (1959).

[37] *See generally* U.S. CONST.

[38] "The powers not delegated to the United States by the Constitution, nor prohibited by it to the States, are reserved to the States respectively, or the people." *Id.* at amend. X.

[39] *See* Antonin Scalia, interview by Charlie Rose, CHARLIE ROSE, (Jun. 20, 2008), https://charlierose.com/videos/15555.

[40] *See* U.S. CONST. art. VI., § 2.

an overarching standard that creates unity within a federal system. Furthermore, unity is promoted through voluntary interaction between the national and state governments. For example, the federal and state governments cooperate to respond to natural disasters, protect homeland security, and discuss how federal dollars can best be allocated to the states.[41] When the federal government interacts with state governments to either assist them in a time of need or to promote the national welfare, it breeds fidelity toward the nation as a whole that transcends one's individual conception of himself or herself as just a Texan or New Yorker; he or she is an American.

The U.S. also satisfies the fourth and fifth tenets of federalism which implicate how individuals interact with their government. The U.S. promotes individual rights through an entrenched bill of rights.[42] By defining a set of rights enjoyed by everyone, and preventing government from infringing upon those rights, the U.S. embodies the fourth tenet of federalism. The U.S. democratic process also protects individual rights. While the government has the means to subject citizens to its will—they do have tanks and bombs, after all—the American system works because government recognizes the value of a democratic system and chooses not to use the power that it so clearly has. Thus, in America, the people functionally maintain power through the democratic process. If enough citizens believe that the government is violating their rights, they have the power to vote elected officials out of office to effect change. Less drastically, citizens can publicly pressure or lobby their government to start or stop engaging in certain actions.

As a result of both formal and functional components of the American federal model, the U.S. serves as an exemplar of the five tenets of federalism and is a fitting prototype to analogize to Reddit's content moderation scheme.

B. Reddit's Content Moderation Scheme is Both Structurally and Functionally Federal

1. Hierarchy, Sovereignty, & Union: How Reddit/Admins (the National Government) & Subreddits/Mods (the State Governments) Interact

While Reddit does not have an entrenched constitution that guarantees its admins a prescribed set of enumerated powers, while leaving all remaining powers to the mods, the company has functionally created a system in which Reddit's powers are few and enumerated while the

---

[41] *See* COUNCIL OF GOVERNORS, https://www.nga.org/cog/ (last visited 2 May 2021).
[42] *See* U.S. CONST. amend. I–X.

mod's powers are many and unenumerated. As previously stated, Reddit has very few policies that apply to the entire platform, while subreddits are entrusted with the power to create and enforce rules specific to their communities.[43] When describing the structure of its content moderation policies, the head of Reddit's community team, Evan Hamilton, described Reddit as similar to the United States' governmental structure.[44] Hamilton stated that "[t]here are rules that everyone has to abide by . . . to ensure safety and consistency."[45] Beyond that, Reddit's goal is to "allow people to really build and curate the experience they want to have on the platform, and have some ownership" over their communities.[46]

For example, r/funny, one of Reddit's most populous subreddits, has its own rules that govern users.[47] Not only are r/funny's rules more numerous than Reddit's, but they are more specific. Compared to Reddit's eight rules comprising less than 250 words, r/funny imposes an additional ten rules comprising over 1,600 words on its users.[48] Every rule is explained in detail; examples of prohibited content are provided along with links to other subreddits that would allow the type of content that r/funny bans.[49]

Furthermore, r/funny's rules are specifically tailored to the values of the community in a similar way that state law is promulgated based on the specific needs and values of a state. r/funny's first rule states that "[a]ll posts must make an attempt at humor.[50] This rule makes sense as applied to "Reddit's largest humor repository."[51] However, such a rule would not be appropriate and would likely be detrimental to the community of a subreddit like r/alcoholism which is devoted to providing "[i]nformation and support for those affected by alcoholism."[52] Additionally, a rule this specific—if incorporated into Reddit's content policy—would not be conducive to Reddit's libertarian approach to content moderation. It would also interfere with Reddit's stated goals of being a place for "open and honest conversations" and "standing for free speech" if it banned all

---

[43] *See supra* notes 13–21 and accompanying text.
[44] David Pierce, *How a Screenshot Started a Fight That Took Over Reddit*, PROTOCOL (27 May 2020), https://www.protocol.com/reddit-powermods-war
[45] *Id.*
[46] *Id.*
[47] R/FUNNY POSTS, https://www.reddit.com/r/funny/ (last visited 2 May 2021).
[48] R/FUNNY COMPLETE RULES PAGE, https://www.reddit.com/r/funny/wiki/rules#wiki_rule_1.3A_all_posts_must_-make_an_attempt_at_humor. (last visited 2 May 2021).
[49] *Id.*
[50] R/FUNNY POSTS, *supra* note 47.
[51] *Id.*
[52] R/ALCOHOLISM, https://www.reddit.com/r/alcoholism/ (last visited 2 May 2021).

content that its admins decided did not make an attempt at humor. Subreddits, such as r/funny, also illustrate how Reddit's content policy preempts subreddit rules.[53]

Thus, Reddit's adherence to U.S. constitutional principles and legal doctrines—such as enumerated versus unenumerated powers and preemption—defines the contours of the relationship between the "first tier" of the Reddit government (admins) and the "second tier" (mods) while simultaneously affording mods sovereignty over their domains just like states have sovereignty over their territories in the U.S.[54]

Not only does Reddit's use of preemption foster unity on the site by creating an overarching standard to unite communities while balancing those standards against subreddit sovereignty, but voluntary associations between admins and mods foster unity too. The primary example of admins and mods interacting to promote unity is the Reddit Mod Council (RMC). The RMC "is a program that aims to increase collaboration between Reddit admins and moderators"[55] in recognition that "by building productive, healthy relationships [both admins and mods will] be able to achieve [the best for Reddit's communities.]"[56] The RMC is comprised of over fifty mods taken from a cross-section of subreddits and is continually growing.[57] Not only did the RMC contribute to recent revisions to Reddit's content policy, but it engages with admins to solve some of Reddit's most overarching and systemic problems.[58]

The RMC is not unlike the Council of Governors (COG) in the U.S. The COG is a mechanism for both state and federal officials to cooperate to address systemic issues implicating important matters to the U.S.; namely, national security.[59] Key areas of cooperation include "Catastrophic Disaster Response And Recovery, Cybersecurity, [and] Federal Budget Discussions."[60] Just as the COG brings both state and federal resources to bear on issues of both

---

[53] Reddit's moderator guidelines specifically state that mods "are obligated to comply with [Reddit's] Content Policy." MODERATOR GUIDELINES FOR HEALTHY COMMUNITIES, https://www.redditinc.com/policies/moderator-guidelines (last visited 2 May 2021).

[54] *See* Linder, *supra* note 1, for further discussion on how Reddit is "set up in layers" like a government ("Mods are 'tier two' of the government analogy. Mods have near complete freedom, so long as they are following the community guidelines Reddit has established overhead.").

[55] u/agoldenzebra, *An Update on the Reddit Mod Council and Adopt-an-Admin Programs*, R/MODNEWS (13 Jan. 2021), https://www.reddit.com/r/modnews/comments/kwr3yi/an_update_on_the_reddit_mod_council_and/.

[56] *Id.*

[57] *Id.*

[58] *Id.*

[59] COUNCIL OF GOVERNORS, *supra* note 41.

[60] *Id.*

local and national importance, the RMC brings together mods and admins to solve systemic issues important not only to individual subreddits but to the entire Reddit community.[61]

It is important to note, however, that there are some instances in which Reddit has fallen short in assisting mods in their ability to address systemic problems on their subreddits. Reddit is often criticized for its unwillingness to police hate speech on the site, even though mods are overwhelmed and lack the resources to respond effectively.[62] Furthermore, there is significant documentation of mods suffering severe mental health conditions as a result of their work.[63] Again, Reddit has not done much to address this issue. While Reddit has made some strides recently to address these systemic issues in collaboration with mods,[64] these efforts are in their infancy, and there is no way to predict their success or even if mods think Reddit is doing enough to help. However, even if it is not, Reddit's analogy to federalism is not necessarily undermined since there are also instances in which state governments seek help from the federal government in the U.S. and the federal government either refuses to respond or only responds after much controversy.[65]

Based on Reddit's ability to define a set of "national" policies that unite the platform under a single set of standards and its work to collaborate with its "subnational entities," Reddit satisfies the union tenet of federalism.

2. Individual Rights & Democracy: How Users Interact with Admins and Mods (and Each Other)

Reddit's respect for individual and vibrant communities—as reflected by its limited set of rules— promotes individual rights because it permits users to form any kind of community so long as the community engages in lawful conduct. For example, users are permitted to form objectively

---

[61] *See* Kim Renfro, *For Whom the Troll Tolls: A Day in the Life of a Reddit Moderator*, BUSINESS INSIDER (Jan. 13, 2016), https://www.businessinsider.com/what-is-a-reddit-moderator-2016-1 ("[S]ometimes [admins and mods] work in conjunction to troubleshoot site issues or come up with solutions to systemic problems.").

[62] Jason Abbruzzese, *Hate Speech is Drowning Reddit and No One Can Stop It*, MASHABLE (Oct. 26, 2014), https://mashable.com/2014/10/26/reddit-hate-speech-moderation/.

[63] Ben Plackett, *Unpaid and Abused: Moderators Speak Out Against Reddit*, ENGADGET (Aug. 31, 2018), https://www.engadget.com/2018-08-31-reddit-moderators-speak-out.html?guccounter=1.

[64] u/enthusiastic-potato, *Crowd Control and Other Safety Updates*, R/MODNEWS (Mar. 2, 2021), https://www.reddit.com/r/modnews/comments/lw9m6q/crowd_control_and_other_safety_updates/.

[65] For example, President Trump initially refused to provide federal assistance to California during its 2020 wildfires but eventually provided the relief after a personal appeal from Governor Gavin Newsom. *See* Scott Wilson & Tim Elfrink, *Trump Administration Rejects, Then Approves, Emergency Aid for California Fires, Including Biggest Blaze in State History*, WASH. POST. (Oct. 16, 2020), https://www.washingtonpost.com/nation/2020/10/16/trump-rejects-california-disaster-wildfires/.

benign communities such as r/corgis,[66] a subreddit devoted to endearing pictures of corgis. In contrast, subreddits can be devoted to disturbing and arguably immoral content such as r/LSD where people can post about their affinity for LSD and often imply that they are using it in excess, thus glorifying drug culture.[67] While many people do not want immoral content to be easily accessible, that does not deny those who want to have conversations about their immorality from being able to do so. Those who would rather engage with charming corgi photos as opposed to content related to drugs or sex can easily avoid subreddits devoted to such topics. Thus, Reddit protects the rights of both the corgi addict and the drug addict equally; a somewhat disturbing proposition that nonetheless is inherently tied to American ideals of freedom of choice and expression.

Reddit further enhances users' individual rights by providing them the necessary tools to curate content within subreddits. The site has long allowed users to toggle whether adult content will appear on their feed.[68] More recently, Reddit is developing tools that will allow users to further curate content in their subreddits. In an update regarding this developing curation system, Reddit said:

> Maybe you're cool with sexual content, but don't want the gore. Maybe you're ok seeing depictions of graphic medical surgeries or violence, but are recovering from addictions and don't want to see drugs or alcohol in your feed. As we evolve our classification system, we'll advance the tools that let redditors control their experience on the platform as well.[69]

As a result of Reddit's policies—both existing and developing—users can not only exercise their right to free expression by joining or creating a variety of different subreddits, but they can further curate the content that is visible to them within the communities they choose to join.

Reddit's democratic processes also enhance and protect individual rights on the platform by empowering users to cultivate norms in their subreddits. Cultivating norms in a subreddit is like how citizens cultivate norms in political discourse. For example, the First Amendment permits citizens to publicly criticize and condemn individuals with whom they disagree. If enough people

---

[66] R/CORGIS, https://www.reddit.com/r/corgi/ (last visited May 2, 2021).

[67] r/LSD, https://www.reddit.com/r/LSD/ (last visited May 2, 2021).

[68] u/ChiTownSando, *How to Turn Off NSFW Subreddit Search Filter and Allow NSFW Search Results to Appear in Feed*, R/HELP (Jun. 23, 2019), https://www.reddit.com/r/help/comments/c4g9tk/how_to_turn_off_nsfw_subreddit_-search_filter_and/.

[69] u/KeyserSosa, *Removing Sexually Explicit Content From r/all*, R/CHANGELOG (Feb. 11, 2021), https://www.reddit.-com/r/changelog/comments/lhnvok/removing_sexually_explicit_content_from_rall/.

speak out, the person espousing objectionable views may be cajoled into silence. At the very least, other members of the community become aware of what kind of behavior is unacceptable in the community.

Reddit accomplishes the same outcome with its upvote/downvote system. By pushing favorable content to the top of the feed and subjugating unpopular content to the bottom, users, at the very least, signal what kind of content is acceptable on the subreddit, but may also convince a troublesome member to stop posting objectionable content or leave the community altogether.[70] In fact, Reddit's decentralized approach to content moderation—which permits users to use social pressure to encourage good behavior—has resulted in "self-policing [which has] fostered an environment in which users call one another out for violating policies or posting objectionable content."[71]

Not only can users cultivate norms in their subreddits, they can also influence the actions of mods and admins through self-organization and public pressure. For example, when mods abuse their power, "those in the subreddit actually push back. They may push out a mod, create new rules, or split the subreddit population into two separate channels with slightly different topics."[72] One instance in which users pushed back against overzealous mods was on the subreddit r/politics.[73] In that case, users were concerned with the conservative slant of the moderators who banned domains from websites that produced "bad journalism."[74] After significant pushback and much debate, the moderators apologized and unbanned all of the previously banned domains.[75] Users, in cooperation with mods, can also pressure admins just like U.S. citizens and their state governments can pressure federal officials into adopting or avoiding certain policies.[76]

---

[70] Of course, there are instances where this system is abused. *See, e.g.*, Abbruzzese, *supra* note 62 (explaining how r/blackladies has been invaded by hateful users who are intent on posting racist content). However, the voting system in general promotes community norms on *most* subreddits. For those that are subject to abuse, some of Reddit's new tools and policies may be able to address them. *See* u/KeyserSosa, *supra* note 69.

[71] SINGH, *supra* note 18, at 26.

[72] Linder, *supra* note 1.

[73] R/POLITICS, https://www.reddit.com/r/politics/ (last visited May 2, 2021).

[74] *See* James Grimmelmann, *Anarchy, Status Updates, and Utopia*, 35 PACE L. REV. 135, 143 (2014).

[75] *See* Grimmelmann, *supra* note 74, at 152.

[76] For an example, *see* Adi Robertson, *Major Subreddits are Going Dark to Protest Reddit Allegedly Hiring a Controversial UK Politician*, THE VERGE (24 Mar. 2021), https://www.theverge.com/2021/3/24/22348255/reddit-moderator-blackout-protest-aimee-knight-uk-green-party (explaining how a number of subreddits set themselves to private—thus preventing any new members from joining—to protest Reddit's hiring of a former U.K. politician who arguably supported pedophilia. Reddit later fired the employee.)

When pressuring mods or admins does not work, users can "vote with their feet" and leave the subreddit to form their own community or join a similar community that is more in line with their values.[77] While these decisions come with unwanted costs to those who leave,[78] it cannot be denied that it is democratic and protects the individual rights of both content and discontent users of the subreddit. If enough users are unhappy with the policies of the original subreddit, then presumably those unhappy users will be yearning to join the new subreddit that is similar in topic to the original but does not contain the unwanted policies. Likewise, those who have no problems with the policies of the original subreddit are not forced to change their practices simply because some users want the community to look different.[79]

III. TAKEAWAYS

Analogizing Reddit to a federal system of government would not be particularly useful if the analysis stopped there. As such, this Part examines several takeaways from the Reddit federalism analogy. These takeaways would not only be beneficial to Reddit as it considers reforming some if its content policies, but they would also prove useful to other social media platforms with similar goals to Reddit as they reassess their content policies in a world that is increasingly demanding that online platforms exercise less power and control over their users.

A. A Federal Model of Digital Governance Requires a Strong Admin Presence . . . Sometimes

Over the years, it has become clear that in some instances, Reddit's fidelity to libertarianism has stifled free expression.[80] Any federal system of government—even online governance systems—must strike a balance between empowering subnational sovereignty while

---

[77] *See* Grimmelmann, *supra* note 74, at 143 (explaining that Reddit has "a gold-plated exit option to preserve user freedom. Any user (or 'redditor') can create a new section of the site (or 'subreddit'), automatically becoming its new moderator and establishing its rules.") (citations omitted).

[78] For example, leaving to "start another subreddit" is not always a fair response when a community is "taken over and locked down by amateur dictators" because "if you split the community, you hurt it." *See* Grimmelmann, *supra* note 74, at 144 (quoting PJ Vogt, *What it's Like When Redditors Ban Your Interview About Redditor's Content bans*, ON THE MEDIA (Nov. 1, 2013) (quoting Angela Motorman), https://www.wnyc.org/story/what-its-when-redditors-ban-your-interview-about-reddits-content-bans/.

[79] An analogy to the U.S. would be citizens moving from one state to another because the policies of the state they are moving to are more favorable to them. Furthermore, users boycotting a particular subreddit is no different than American citizens boycotting a company based upon the company's political donations.

[80] *See* Langvardt, *supra* note 24, at 1361 (explaining that some people avoid participating in certain communities "for fear of their personal and family safety").

promoting union and individual rights. Thus, when the subnational entities are unable or unwilling to protect the individual rights of their constituents, a strong federal response may be necessary.

Reddit has a history of trolls taking over subreddits by posting content that is contrary to the values of the actual members of the community, thus frustrating free expression on the site.[81] Oftentimes, mods are incapable of handling the influx of hateful content, and they report that they often do not receive a response from admins when they request help.[82] As a result, mods are forced to address hateful content that can also be mentally distressing with no training or assistance.[83] Experts report that the abuse mods face creates mental health problems which risks driving mods away from the platform.[84]

When mods face mental health issues or fear for their personal safety, Reddit must respond. A response from Reddit's admins to address such systemic issues would not be antithetical to its libertarian and federal ideals. For example, and as previously discussed, there are plenty of instances when the federal government steps in to assist state and local governments in addressing problems that they are unequipped to handle themselves. Additionally, the federal government has intervened in instances of a state government not protecting the individual rights of its citizens.[85] While Reddit has taken some measures to address hate speech—such as updating its content policy to explicitly ban hate speech—it has not provided mods with the resources or training that they need to manage these systemic issues and protect their mental health.[86] Some mods have recommended actions Reddit could take to better prepare mods for the abuse they may face, but Reddit has yet to adopt these suggestions.[87]

Any platform wishing to adopt a federal system of governance for its content moderation policies must know when a "federal" response is not only advantageous but necessary to preserve healthy communities on the platform. Reddit has made many mistakes as it tries to strike this

---

[81] *See supra* note 73.

[82] *See* Plackett, *supra* note 63.

[83] *Id.*

[84] *Id.*

[85] Exec. Order No. 10730 (Sept. 23, 1957) (sending federal troops to oversee the integration of Central High School in Little Rock, Arkansas).

[86] For more information regarding Reddit's revised content policy, see u/spez, *Update to our Content Policy*, R/ANNOUNCEMENTS (Jun. 29, 2020), https://www.reddit.com/r/announcements/comments/hi3oht/update_to_our_-content_policy/.

[87] *See* Plackett, *supra* note 63 (explaining that Reddit should: (1) inform new mods of the kind of abuse they might encounter and advise them on how to handle it; (2) provide clear policies for how it will respond when mods report abuse; (3) periodically check in on mods to make sure they are okay; and (4) offer resilience training to help mods cope with abuse).

balance, and despite some progress, it has yet to do so. Other platforms should learn from Reddit's mistakes—and their progress—to better understand when a strong admin presence is necessary.

B. A Federal Model of Digital Governance Empowers Users While Disempowering Moderators

As digital platforms continue to gain popularity as forums for the exchange of opinions in the public square, "a small number of politically-unaccountable technology oligarchs exercise state-like censorship powers" over users without being constrained by the rule of law.[88] Essentially, social media companies can ban content for any reason and without due process. If social media platforms are the most important places for the exchange of views in the modern era,[89] then the overwhelming and unconstrained power they have is concerning.[90] A federal model of digital governance redistributes the power from the platform and its moderators to the users themselves.

For example, the multiplicity of subreddits not only makes it easier for users to leave one subreddit for a similar subreddit, but it lends credibility to users who threaten to start their own communities on Reddit.[91] Such threats would not be taken seriously on a more centrally moderated platform like Twitter since the cost of leaving Twitter to create a similar online community would be too great for most users. Thus, the combination of Reddit's limited rules and the ease with which one can create a new subreddit creates an environment where user preferences must be adhered to more faithfully than a centrally moderated platform. In other words, while moderators on Reddit still have the technical power to behave poorly, it is much harder "to force users to hold still while they do."[92]

The increased power and mobility that users enjoy in a federal model of digital governance also benefits the diversity of the community. Ostensibly, it might appear that a platform like Reddit—which allows users to join communities that they agree with and leave those with which they disagree—would create echo chambers and prevent users from engaging with diverse

---

[88] Langvardt, *supra* note 25, at 1358.
[89] *See* Packingham v. North Carolina, 137 S. Ct. 1730, 1735 (2017) (explaining that social media is unequivocally the most important place for the exchange of views today).
[90] *See* Langvardt, *supra* note 25, at 1358 (asking how long society can accept this power differential before online platforms "fall into authoritarian hands, or before a sheepish CEO succumbs to governmental pressure in a time of national crisis").
[91] Grimmelmann, *supra* note 74, at 148.
[92] *Id.*

perspectives.[93] However, "Reddit's decentralized moderation scheme . . . actually results in a more diverse cross-section of opinions being shared . . . ."[94] This increase in diverse viewpoints is a direct byproduct of the power that a federal model of digital governance gives to users. When users moderate content in addition to admins and mods—such as through a voting system—content is prioritized based on what users "vote to be the most interesting or informative" as opposed to what admins believe should be featured.[95]

Centrally moderated platforms—which do not typically prioritize content based on user votes—traditionally promote content that has the most engagement regardless of whether it is positive or negative.[96] Thus, a centrally moderated platform that does not afford users the power to moderate is more likely to prioritize content that is divisive and partisan, thus muffling or suppressing moderate voices.[97] In contrast, a federal model of moderation empowers users to prioritize content that is in line with community values and gives power to the users to determine what content should be pushed to the top of the feed.[98]

C. A Federal Model of Digital Governance is More Efficient and Cost-Effective Than Central Moderation

Content moderation cannot prevent all forms of abuse without killing the proverbial commons.[99] Thus, moderation can only be expected to keep abuse within acceptable bounds— while facilitating open and productive conversations—without increasing the cost of moderation to unsustainable levels.[100] By having few central rules wherein admins only address systemic issues, locally created rules enforced by mods in subreddits, and values-based moderation by users,

---

[93] *See* Ethan Zuckerman & Chand Rajendra-Nicolucci, *What if Social Media Worked More Like Email*, KNIGHT FIRST AMENDMENT INSTITUTE BLOG (Nov. 3, 2020), https://knightcolumbia.org/blog/what-if-social-media-worked-more-like-email.

[94] Steven Johnson et al., Opinion, *Facebook Serves as an Echo Chamber, Especially for Conservatives. Blame its Algorithm*, WASH. POST (Oct. 26, 2020), https://www.washingtonpost.com/opinions/2020/10/26/facebook-algorithm-conservative-liberal-extremes/ (explaining a study that showed that a typical conservative Facebook user reads news that is thirty percent more conservative than the news they would typically read while a typical conservative Reddit user will read news that is fifty percent more moderate than the news they would read otherwise).

[95] *Id.*

[96] *Id.*

[97] *Id.*

[98] *See* Pierce, *supra* note 44 (opining that "outsourcing content moderation to the community empowers users to take ownership of their online communities, much like neighborhood watch groups in real life").

[99] *See* Grimmelmann, *supra* note 2, at 53.

[100] *Id.*

Reddit's federal model of digital governance is more efficient and cost-effective than centrally moderated platforms.

By delegating most of the moderation to mods, Reddit saves money and maintains the ability to focus on systemic issues.[101] For example, QAnon could not generate a following on Reddit before being banned, whereas it has proven nearly impossible to remove QAnon from centrally moderated platforms such as Facebook and Twitter.[102] While some of the credit for purging QAnon from Reddit is owed to its bifurcated structure—Reddit could simply find and ban all QAnon subreddits and their users while Twitter had to find and ban QAnon adherents individually—it is also true that admins had a greater ability to monitor QAnon's growth in real time than Twitter's or Facebook's admins.[103] While Twitter's or Facebook's admins must monitor every user at all times, Reddit's admins can trust that most users are being monitored by mods and save their time for greater threats. This flexibility likely contributed to Reddit's ability to excise QAnon from its community.

Reddit's federal digital governance model also minimizes the negative externalities caused by central and automated moderation. Moderators, whether they be human or automated, struggle with the nuanced judgments they must make when determining if content violates the platform's policies.[104] Human moderators are likely to struggle because any rule is likely to "contain gaps and conflicts" which create complexity and make moderation decisions difficult.[105] Likewise, automated moderation tools can be inaccurate and unable "to interpret context and understand the nuances of human speech."[106] Reddit mitigates these negative externalities in various ways.

First, its limited content policies create space for mods—who are more in tune with their community—to create and enforce rules. Mods are less likely to struggle when making moderation decisions compared to admins since they are actually members of the community. It is likely that mods had a role in creating the rule, which increases the odds that they know how it should operate—unlike the average Facebook moderator. Thus, mods are in a better position to make a

---

[101] Gicel Tomimbang, *Content Moderation and Removal Conference Recap*, HIGH TECH LAW INSTITUTE (Mar. 21, 2018), https://law.scu.edu/high-tech-law-institute/content-moderation-and-removal-conference-recap-by-gicel-tomimbang-3l-santa-clara-law/.
[102] Kaitlyn Tiffany, *Reddit Squashed QAnon by Accident*, THE ATLANTIC (Sept. 23, 2020), https://www.theatlantic.com/technology/archive/2020/09/reddit-qanon-ban-evasion-policy-moderation-facebook/616442/.
[103] *Id.*
[104] *See* Langvardt, *supra* note 25, at 1362.
[105] Klonick, *supra* note 24, at 1632.
[106] SINGH, *supra* note 17, at 6.

judgment call as to whether a post violates the community's policies than an admin, who may know nothing about the community and its values.[107]

Second, the use of many community mods allows Reddit to use automated moderation strategically while still ensuring humans are making most of the moderation decisions. With the increasing amount of content posted online, platforms will have no choice but to increase the amount of automated moderation they use to maintain a cost-effective moderation model.[108] However, in the absence of artificial intelligence (AI) that can be as discerning as humans, automated moderation is an imperfect solution to the moderation problem. Courtney Linder put it best when she asked: "Do we really want computers alone to tell us what to do?"[109] Until AI can be as adept at moderating as humans, Reddit has struck an appropriate balance. Its decentralized approach to moderation allows it to use a vast number of human mods to moderate content more accurately than a machine could.[110] However, given the human resources at Reddit's disposal, it can also strategically use AI to flag inappropriate content while allowing mods to make the final decision as to whether to remove the content or not.[111] Thus, Reddit's federal model of digital governance efficiently moderates content while strategically employing cost-saving mechanisms—such as AI—in ways that are not as feasible for centrally moderated platforms.

CONCLUSION

Federalism works. It works in the United States and several other democracies around the world. Federal governments—through the creation of multiple levels of government—promote the sovereignty of diverse communities while uniting them under a common identity. They champion individual rights and promote democracy such that people, not their governments, wield the power. Federalism works specifically for governments that prioritize the decentralization of power—thus affording most of the power to regional governments or its citizens. If decentralized online platforms like Reddit are operating like real-world governments, then why shouldn't federalism work for them too?

---

[107] *Id.* at 8 (explaining that "[d]ecentralized models . . . enable more localized, culture-specific, and context-specific moderation to take place").
[108] Platforms do not have unlimited capital and cannot afford to continue hiring mods to keep up with an exponential increase in content.
[109] Linder, *supra* note 1.
[110] *See supra* notes 106–109 and accompanying text (mods are superior to AI like they are superior to admins).
[111] *See* SINGH, *supra* note 17, at 27.