

A Web-Scale Analysis of the Community Origins of Image Memes

DURIM MORINA, Stanford University, USA

MICHAEL S. BERNSTEIN, Stanford University, USA

Where do the most popular online cultural artifacts such as image memes originate? Media narratives suggest that cultural innovations often originate in peripheral communities and then diffuse to the mainstream core; behavioral science suggests that intermediate network positions that bridge between the periphery and the core are especially likely to originate many influential cultural innovations. Research has yet to fully adjudicate between these predictions because prior work focuses on individual platforms such as Twitter; however, any single platform is only a small, incomplete part of the larger online cultural ecosystem. In this paper, we perform the first analysis of the origins and diffusion of image memes at web scale, via a one-month crawl of all indexable online communities that principally share meme images with English text overlays. Our results suggest that communities at the core of the network originate the most highly diffused image memes: the top 10% of communities by network centrality originate the memes that generate 62% of the image meme diffusion events on the web. A zero-inflated negative binomial regression confirms that memes from core communities are more likely to diffuse than those from peripheral communities even when controlling for community size and activity level. However, a replication analysis that follows the traditional approach of testing the same question only within a single large community, Reddit, finds the regression coefficients reversed—underscoring the importance of engaging in web-scale, cross-community analyses. The ecosystem-level viewpoint of this work positions the web as a highly centralized generator of cultural artifacts such as image memes.

CCS Concepts: • **Human-centered computing** → **Empirical studies in collaborative and social computing**.

Additional Key Words and Phrases: memes, cultural diffusion, online communities

ACM Reference Format:

Durim Morina and Michael S. Bernstein. 2022. A Web-Scale Analysis of the Community Origins of Image Memes. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1, Article 74 (April 2022), 25 pages. <https://doi.org/10.1145/3512921>

1 Introduction

Which online communities originated the cultural artifacts that we see around us online? Cultural artifacts are passed from community to community, upvoted and downvoted, remixed and reposted. When we come across an artifact such as an image meme posted to an online community, we might wonder: where did this cultural artifact originate? Which communities are generating the internet culture that we see all around us?

Common narratives, drawing on examples such as jazz arising in African-American areas of New Orleans [35] and the emo fashion movement arising from the post-punk subculture in Washington, D.C. [23], suggest that cultural innovation arises in fringe, *peripheral* communities that are near the edge of the network. These innovations are then adapted and popularized by the *core*, or mainstream

Authors' addresses: Durim Morina, Stanford University, Stanford, USA, dmorina@cs.stanford.edu; Michael S. Bernstein, Stanford University, Stanford, USA, msb@cs.stanford.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.

2573-0142/2022/4-ART74 \$15.00

<https://doi.org/10.1145/3512921>

center of the network. These narratives recur in online communities as well, with the subversive online community 4chan /b/ generating some of the most recognizable early internet memes such as LOLcats and rickrolling [7, 25]. Organizational sociologists refine this notion, suggesting that cultural innovation often originates not in the most peripheral communities, but from those in an intermediate network location between core and periphery, because they are able to draw ideas from the periphery and bridge them into the core [10, 14].

However, research still lacks a firm answer as to whether these theories find substantial empirical support online, because research to date has focused on diffusion within individual communities, granting a valuable glimpse but an incomplete picture. Work focusing on individual communities (platforms, domains, or socio-technical systems) such as Twitter [26, 41], Digg [50], Facebook [11], and LiveJournal [4] provide lenses on how memes are created within that specific community. However, societies cluster into groups that inhabit different spaces [36], so any given system is only a small part of the larger ecosystem; we are left with an incomplete picture that leaves out many peripheral communities. How might these dynamics play out across the broader cultural landscape of the entire web? Studies of diffusion that cross multiple communities are rare [56], and are typically focused either on politics [1, 31] or on cascade size [21]. As a result, the question of where image memes and other online cultural artifacts originate is still unanswered.

In this paper, we contribute the first *web-scale* study of image memes: an analysis of the origination and diffusion of image memes that contain English-language text across essentially all indexable communities on the web. By pursuing a web-scale analysis, we achieve an ecosystem-level viewpoint that affords us a novel opportunity to track image memes as they arise and spread across the web, not just within single communities or across a small number of communities. We focus on image memes—static pictures with (often humorous) overlaid text—as a case study of cultural artifacts, because image memes are among the dominant cultural icons on the modern web.

Our approach produces a capture of new memes posted to the public, English-speaking web over one month. We proceed in two stages: (1) a *crawl* to identify all public, English-speaking communities on the web that post majority meme content, and (2) a *high frequency scrape* of those communities over a period of one month to identify new memes and track their diffusion. The first stage, the crawl, uses reverse image search on memes identified so far in the crawl to identify other communities that have posted memes. We iteratively crawl these new communities, again use reverse image search to identify additional communities that have also posted the images posted to those new communities, crawl those communities, and repeat until convergence. This crawl produces a list of communities that we then scrape in the second stage. Our high-frequency scrape captures every piece of content uploaded to any of those communities over one month. We again utilize image search to filter out memes that have been seen online prior to the sampling period, allowing us to only analyze memes that were created *de novo* during our sampling period. We follow prior work [56] by using perceptual hashing and clustering to combine variations of the same meme template, merging variations of the same base image that feature different crops or text overlays. Together, these methods allow us to identify the first posting of an image meme on the public web, which we identify as the *originating* community, and track other communities that later posted the same meme image. We map each community onto a continuous measure of its position on the core-periphery axis via a network centrality measure [43] calculated over a crawl of the entire web from the same period [13].

This process produced a dataset over one month of 906,481 meme images. With this dataset, we measure: (1) Is the most-shared content originated in the core or peripheral communities of the web's network structure? (2) Do core or peripheral communities produce more completely original content? This web-scale analysis allows us to analyze whether the cultural artifacts around us originate in small, peripheral communities of the web, out of the public eye, and then diffuse to the

core communities such as Reddit, Facebook, and Imgur, or whether modern cultural production has become centralized in the core.

Our results suggest that, while memes that originated in more peripheral communities will diffuse farther individually, in aggregate the vast majority of image meme diffusion events can be traced back to memes originating in the core of the network. For example, the top 10% of communities by network centrality posted the original content that eventually goes on to represent 62% of diffusion events online. A zero-inflated negative binomial regression finds that core communities' memes are more likely to diffuse, but diffuse less far, when controlling for the size and activity level of the community.

We then replicate our analysis recursively within a single influential community, analyzing the diffusions between core and peripheral subcommunities (subreddits) on Reddit. We find again that core subcommunities originate the memes that receive the vast majority of diffusions within the community, but the mechanism is reversed: the core subcommunities' memes are less likely to diffuse individually, but diffuse farther when they succeed. This reversal of the web-scale result reinforces the importance of looking beyond traditional within-community analyses in the literature and engaging in cross-domain or web-scale replications. We observe that all communities, including those in the core, can trace a plurality of their content back to other communities in the core. Finally, we perform a sensitivity analysis that confirms that our result is robust to even large amounts of unobserved memes diffusing from private or unobserved communities.

Web-scale analyses are rare in the literature, so it is important to articulate their strengths and limitations at the outset. In exchange for the benefits of offering an ecosystem-level view that generalizes beyond single platforms, web-scale analyses come with attendant methodological limitations because data must be captured and compared across as much of the web as possible. Principally, our analysis cannot fully estimate the impact that private (unscrapeable) communities play, or rule out the possibility that some peripheral communities could be missed in the crawl. So, in this paper, we perform a sensitivity analysis and confirm that the magnitude of our result is substantially robust, so that it would still hold even if many extremely peripheral communities were missed. Second, image memes cannot capture all forms of culture: prior sociological work investigates other forms of cultural artifacts including music, literature, sports, and technology [24, 28, 35, 45, 47]. However, image memes provide one valuable lens onto questions of cultural diffusion because they are among the most popular types of cultural artifact online, and they are easily created and transmitted, and they readily transfer across many communities and networks. Third, operationalizing network centrality via web link network structure is not completely coincident with cultural network structure, since hyperlinks do not always equal cultural influence. However, at a macro scale, the internet link structure does capture a substantial amount of cultural influence, in the same sense that offline studies of cultural networks track social connections.

This outcome suggests that the design affordances of modern social computing systems have centralized the publishing and diffusion of image memes online into the most popular and heavily linked-to sites online. This outcome is not unique to the internet: other media such as radio and television have likewise exhibited patterns of centralization over time due to commercialization pressures [18]. Disruptive cultural innovations may still come from the periphery of the internet network, of course, but our evidence suggests that the large plurality of the memes that we see online are originated in a small number of core communities. Being in the core offers large audiences and rapid re-sharing to other communities. Our results raise questions about what challenges we face as designers when cultural capital and innovation becomes increasingly centralized.

We begin by deriving our hypotheses from the behavioral sciences. We then describe our dataset collection, and our analysis. Our dataset is available at <https://osf.io/hr3e5/>, available for other researchers to download and use.

2 Memes, Culture and Diffusion

Raymond Williams perhaps sets the stakes most effectively, writing that “culture is one of the two or three most complicated words in the English language” [53]. Sewell declared two ways to define culture: a “theoretically defined category or aspect of social life that must be abstracted out from the complex reality of the human existence” and as a “concrete and bounded world of beliefs and practices” [9]. He goes further to conceptualize culture along many dimensions, ranging from learned behaviors, meaning-making, a system of symbols and meanings, and creativity. In this work, we focus on cultural diffusion in its context attached to symbols that are spread online via image memes; however, sociologists have studied diffusion of many other facets of culture, cultural practices [28], technological innovation [47], and linguistic evolution [31].

Much sociological work has focused on creation in the arts, such as jazz [35], with generalization to other social behaviors: John Hall argues that “there is little if any difference between the arts and other socially organized activities” [24]. The input is usually shaped by limited resources such as funding, and the output is shaped by regulation and censorship [30]. There is also substantial evidence that cultural production does not arise from an idealized artist, but instead also features a large supporting cast ranging from friends giving feedback to a formal organizations that mass produce the cultural item such as record labels and book publishers [24, 45]. Success may depend on multiple self-reinforcing rounds of exposure to a new cultural artifact [37].

One important lens on cultural diffusion is that of network structure. The earliest work in cultural diffusion focused on technological innovations such as the acceptance of hybrid corn seeds [47]. Diffusion processes of online content are likened to the way epidemics spread, with adoptees (infected individuals) spreading the innovation to adjacent nodes in the network [5, 32, 44]. One major topic of interest is the route of diffusion: people preferentially share content that matches their beliefs [2], and models can predict diffusion routes [27]. A second topic is the temporality of diffusion: sharing is often bursty initially, before slowing [39] and possibly recurring months or years later [12]. Research has translated many of these insights to online communities, such as studying the spread of videos on YouTube [8, 17, 54, 55] and TikTok [3].

Image memes offer a modern context for studying forms of cultural innovation and diffusion [33, 34, 51]. They rise and fall quickly, with content continuously being flushed out of sight by new arrivals [7]. Researchers have studied whether image meme diffusion cascades can be predicted [11], the temporal dynamics by which image memes come back and their cascades recur [12], the design choices that are associated with increased diffusion [34], and where image memes arise from within a limited set of communities [56]. We extend this literature by broadening the analysis from a small set of communities to essentially the entire indexable web, allowing us to test the impact that smaller, peripheral communities vs. large centralized communities play in this spread of cultural artifacts.

In sum, the existing literature has studied the structure of cultural diffusion, but rarely its genesis, and not at web scale. This motivates our research question:

RESEARCH QUESTION. *Are the online communities that originate popular image memes positioned in the core or periphery of the web?*

For our analysis, we draw on the practice of studying cultural innovation from the perspective of the structure of the social communities that produce it, often defined as *core* and *periphery*. The core is defined as the mainstream, and the periphery as the marginal communities [9]. One major result in the literature is that, offline, the success of cultural innovation depends on the position in the social network, greatest when occupying an intermediate, bridging position neither in the core nor the periphery [10, 14]. The innovation’s success also depends on its level of typicality

relative to existing content: the more diverse, the more likely it is to diffuse [6]. However, the level of legitimacy attached to the person adopting the innovation also impacts their success rate [42].

Prior work on individual or small sets of communities online suggests that communities that are well-known but still outside the core of the internet network (e.g., 4chan) carry influence on the core communities, while core communities are more effective at pushing out memes to other places [56]. Likewise, those in intermediate positions between core and periphery in professional or interaction networks are most likely to create durable cultural innovations [10, 14]. If core communities are mostly echoing content originated in the intermediate and peripheral communities, they should be responsible for fewer diffusion events:

HYPOTHESIS 1. *Image meme diffusion events are negatively associated with network centrality, controlling for community size and activity level.*

Another way we might test the literature’s claim of the mainstream’s inability to innovate [10] is in the volume of original content (OC) that are originated in the community. This prompts the following hypothesis:

HYPOTHESIS 2. *The rate of posting original content (OC) is negatively associated with network centrality, controlling for community size and activity level.*

Online communities’ structure and function are well studied in the CSCW research community [29], and memetic content arises regularly in their study (e.g., [33]). Memes arise as identity markers used for impression management [52], as members use them to project in-group signals. However, memes also present challenges for non-sighted community members [20]. Memes are also used regularly as vehicles for narratives in both social movements [49] and disinformation campaigns [48]. Communities cannot always attend to the large volume of content posted to them [19]. Less is known about the role that online communities play in producing online culture. Our research advances this goal through an analysis of the communities producing the dominant cultural icons.

3 Data and Methods

In our study, we focus on image memes. Image memes (e.g., Figure 1) are a dominant form of cultural artifact online today. They diffuse readily across many social media. Often memes contain both a base image and changeable text, or a base image where one part is replaced with another image, or are just a single image that gets shared widely. We will focus on image memes that contain a base image with text overlaid. We collapse instances of the same base image onto a single meme, allowing us to track that meme’s origination and diffusion across the web even as the text overlay changes.



Fig. 1. An example image meme. Our method automatically group memes with similar images and text, even accounting for differences in encoding, resolution, watermark, visual filter, or aspect ratio.

There are three main challenges in tracking image memes as they appear and are diffused across the web. First, we need to identify and track as complete a set of communities as possible, since the memes may be originally posted to small, peripheral communities that do not have many visitors. Second, we need to be able to identify memes as they move: image memes are transformed as they jump from community to community: they are downloaded and re-encoded; they change resolution; they acquire watermarks; they are re-cropped. Third, we need to restrict our analysis to only *de novo* memes—memes which are new and have never been posted online before—filtering out memes that were posted prior to our study.

3.1 Crawling communities

To perform a web-scale analysis, we first need to identify a set of public communities on the web that we will track during the study period. Our crawl methodology identifies this set of communities by observing who is posting memes to the web, and then tracking its memes to identify other communities that have shared the same memes. We make a simplifying assumption in treating a top-level domain as a community. Some communities, such as Facebook and Reddit, are very large and diverse. Others, as HiddenLol, are much smaller and focused.

We identify our set of communities through a web crawl. This crawl tracks the sources of memes posted to known communities back to new communities that have not been observed yet, then adds those new communities to the set to track. We initialized our crawl with seed communities popular for their meme production: two subreddits (r/memes and r/funny), twelve Instagram accounts (e.g., memes, beigecardigan), 9gag’s ‘fresh’ feed, pinterest’s ‘memes’ pin, and accounts returned from Twitter’s account search for ‘meme’. For each seed community, we wrote a custom scraper whose complexity ranged in complexity from a simple API client to a paginated HTML parser. This scraper captured main content pages, or the posts within an account (e.g., on Pinterest). We scraped a set of meme images from each community for review.

Then, for each meme image posted to our seed communities, we tracked that image backwards to other communities that had previously posted it. We achieved this by utilizing Google Cloud’s Vision API [22], which provides a reverse image search to return other URLs where Google had indexed an identical or very similar image. We kept any domains that occurred at least ten times across our crawl. So, if several images in the sample had previously been posted on cscwmemes.com according to the cloud vision API, then cscwmemes.com was added to the set of communities, and its contents was then scraped and fed to the Cloud Vision API to help identify further new communities. By default, we included the top-level domain. However, in many cases, communities were better characterized by individual accounts or groups (e.g., subreddits on Reddit, accounts on Twitter), in which case we included those specific accounts or groups and attributed them to the top-level domain.

Next, we filtered the communities to remove non-meme communities: not all potential communities are relevant sources of meme images. Our study is not focused on memes that arise as part of communities with other foci (e.g., a news community that occasionally posts a meme), so this filter allowed us to be certain that the images we were tracking were memes and not other images such as news photos. In addition, including non-majority meme communities would include non-meme images in our analyses, and which would confuse and dilute the analysis: e.g., capturing how photos from a news article spreads around the web. To maintain a coherent cultural lens, we focused on majority meme communities. The possible impacts of this decision are tested in our sensitivity analysis.

To achieve this, we utilized a crowdsourced labeling of these communities on Amazon Mechanical Turk to classify each one as posting a majority—over 50%—meme content. Workers were paid \$0.30 per classification, to ensure at least minimum wage rates in our local jurisdiction. We presented

workers with a definition of image memes that we were looking for, including examples, and asked three workers to independently classify whether each community we discovered featured a majority of image meme content. Workers were also asked to label if the site should be excluded for any of the following reasons: adult content, broken or private pages, non-English, news and advertising pages, online stores selling apparel with memes printed on them, or content hosted in private spaces such as Google Drive. These filters were included to ensure that our scrape would capture actual image memes, respected user privacy by only downloading public content, and would not download illegal underage pornography. Potential biases introduced by this process are considered in the Discussion section.

Once new communities had been vetted by this crowdsourced filter, we included them in our crawl and repeated the process, sampling their content and using the cloud vision API to identify additional communities or accounts within existing communities to include. We repeated this recursive crawl until the set of communities converged. The full crawl period lasted several months.

Facebook, due to its API restrictions, was the one outlier in our crawl. Facebook disallows Google indexing of images in Facebook groups and pages. So, we began by filtering all public Facebook pages for those that contained the substring “meme” in their name. While Facebook pages could be included, Facebook groups are private and cannot be accessed except by members due to API restrictions, so we were unable to include them in our analysis. We discuss this limitation in Section 6; in brief, given the directionality of the outcome of our study, our conclusions would only be strengthened if Facebook groups were to be indexed as part of Facebook, because it would add additional memes originated to a community which is at the core of the network.

Our crawl ran from October 2018 through February 2019 and identified a total of 88 communities (domains). Communities included Reddit, Facebook, Dop3r, Memedroid, Troll Street, Dump A Day, and many others. Most of these communities included one main feed; others comprised many meme accounts captured by our scraper, including 1235 subreddits on Reddit, 929 Twitter accounts, 173 Instagram accounts, 51 Tumblr blogs, and 1119 Facebook Pages. We then manually added and whitelisted three communities that were featured centrally in prior work on meme culture—Reddit’s *r/the_donald*, as well as 4chan’s */b/* (random) and */pol/* (politics) boards [7, 25, 56]. While these communities are not primarily meme content, and were thus excluded by our filter, we manually whitelisted and included these communities in our dataset to maintain continuity with prior work; the results reported in this paper replicate when they are not included.

3.2 Capturing diffusion events

Having identified the set of meme-centric communities to track, the second part of our process collected a complete set of image memes with English-language text overlays posted to those communities over a period of a month. About three quarters of memes in our dataset contain text overlays. We remove memes without text, because textless memes often mutate via image editing software in ways that are more challenging to track. Including textless memes results in many false negatives where instances of the same meme are not recognized by the algorithms meant to group meme images together, because they look substantially different from each other.¹ We then merged instances of the same meme *template* (shared base image) and filtered out memes that were posted elsewhere on the internet prior to our study period, ensuring that the same memes are grouped together in our dataset and each meme represents new original content created during the study period.

To achieve this, we performed a high-frequency data collection from each of the communities for a period of 30 days in March–April 2019. We elected for a 30 day sample because most meme

¹The Limitations section reflects on biases that this decision might introduce.

lifecycles are much shorter than this [12]. For each image posted to any community in our list, we captured metadata such as date and time posted, title, description, and popularity via upvotes or retweets if available. It is important for our study of diffusion that we be able to accurately timestamp posts: so, we used techniques such as retrieving creation and modification timestamps via domain specific APIs, reading the timestamp from the HTML, using HTTP Headers to capture caching information, and if all the previous techniques failed, we fell back on the server's timestamp when the image was first crawled—we consider this to be accurate, as the crawling frequency was always within minutes.

Next, we must ensure that all memes in our dataset were *de novo*, created and first posted during our study period and not before. To achieve this, we use the same Google Cloud Vision API [22] reverse image search functionality from the crawling phase: when an image exists elsewhere in Google's search index, the Cloud Vision API returns a list of all prior locations of that image on the web. Any image memes that the Cloud Vision API returned results for prior to their first occurrence in our sample were excluded from our analysis, since the API returning results means that the image existed online before our study period.

We identified instances of the same meme following the same method as prior work [56]. Specifically, we first run the perceptual hashing (pHash) algorithm [38] on each image. pHash assigns a vector hash to each image, where images that appear similar based on a model of the human perceptual system will appear in similar parts of the pHash embedding. Then, following prior work [56], we use DBSCAN clustering [16] on the pHash results to group the pHashed images into clusters, where each cluster represents one meme. DBSCAN identifies sets of images whose pHash values are similar and groups them together. Each DBSCAN cluster collects one meme in its variations (e.g., Figure 1). To achieve this, we replicate prior work's DBSCAN parameter settings [56].²

In summary: we first identify a set of communities online that post meme content, then scrape each image meme posted to any of these communities, then filter to the image memes created during our study period, then merged meme images that shared the same template. The outcome of this process is a set of memes, and a list of community timestamps for each meme when it was posted. The first timestamp for each meme represents the community that originated that meme: the meme existed nowhere else online before that post.

Our main dependent variable is the number of meme images in our sample that can be traced back to each meme. In other words, our construct of interest is *diffusion volume*: the number of posts elsewhere that are attributable to a meme originated in the community. So, for example, if a community posts three unique memes during the study period, the first meme might go on to get posted 0 times in other communities (never diffused), the second meme might also get posted 0 times, and the third meme might get posted 4 times. We include reposts within each downstream community but not reposts within the originating community in this measure, because we are focused on external influence. We also measure a second dependent variable, the diffusion distance: the number of unique communities that each meme diffuses to. Because this measure is community-level, it excludes any reposts within the diffused communities. For example, in the earlier example, those three memes might have each diffused to 0, 0, and 2 different communities. We consider diffusions between sub-communities (e.g., between two subreddits or Facebook Pages) to be within the same community.

²These parameter settings from prior work resulted in reasonable clusters with the exception of one massive supercluster; we recursively divided this supercluster by requiring a tighter Hamming distance. A manual sampling of the top ten clusters confirmed that the resulting clusters represented different meme templates.

These measures assume that the second chronological appearance of a meme captures a diffusion event from the first community to the second community.³ Beyond this first diffusion event, we cannot know for sure which community is responsible: did the third event originate from the original community, or from the newer second community? However, 96% of all diffusion events occur within one step of seed [21]. So, for our analysis, we focus only on attributing diffusions to the first community it appeared, rather than inducing an ordering on the following communities.

We noted that some communities appeared to either scrape and auto-post, or to be intentionally collaborating in scraping or resharing each others' content. This behavior would artificially inflate counts, since the reshare is not a genuine cultural diffusion but an automated process. These situations were clear from their high volume and large proportion of cross-posting. We removed any communities where both (1) the volume from one community to another community was high, and (2) a large proportion of memes on one community appeared on the second community. This criteria removed exactly two communities that were collaborating with each other—astrologymemes and loveforquotes. In addition, we noted that some communities in our crawl did not produce any new content in the period of the study, so we removed them from further analysis. Finally, we discarded any image memes whose text was not in English. Focusing on English served two purposes. First, it allowed us to focus on a coherent cultural context. Second, our research team and crowdsourced filter did not have the capacity to judge whether content is a meme across every written language, and financial constraints on the cloud vision API dictated that we could not consider all meme images in all languages.

3.3 Mapping communities onto core and periphery

Driven by the theory that inspired our work, we next mapped each community onto a measure of its network position, which captures how core (mainstream) or peripheral a community is in the world wide web link structure. The world wide web link structure, determined by the hyperlink graph of the web, captures many different sorts of influence in different links, including trust, citation, sourcing, related materials, and navigation. This paper concerns cultural (meme) diffusion; while the web link structure captures cultural influence (e.g., linking blogs, posts, and memes), the existence of other forms of link mean that the web link structure is more expansive than just cultural influence. This is a limitation of the wide lens of an internet-scale study. However, in this decision, we follow prior work studying core/periphery structures and artistic or innovation influence that also use networks that are related to but not identical with cultural influence, for example email threads in a consulting firm [14] and co-workers in films [10].

We then apply a measure common in the literature, called *network centrality*, to assign each community a continuous score of how core or peripheral it is on the web. Communities in the core have high network centrality, communities in the periphery have low centrality, and intermediate communities are between [15]. Prior work has used various centrality measures—e.g., dimensionality reduction via 'coreness' [10, 14]—and we make no claim that our approach is the only reasonable one. However, we do require a centrality measure that can operate without being contaminated by link spam, link rings, and other issues that arise at web scale. A well-known measure of network centrality is PageRank [40]; however, it is susceptible to spam and manipulation. So, we adopted a common revision called harmonic centrality [43], endorsed and used by the Common Crawl foundation in their web-wide web crawl snapshot [13]. We obtained harmonic centrality scores for each community from a full web crawl by Common Crawl completed during the same period as our data capture. In other words, we use a centrality measure derived from a full web crawl

³It is possible that the same user posted to both communities; however, investigation of a sample of these diffusion events suggests that this is exceedingly rare.

performed by Common Crawl to calculate how close each community is to the core or periphery of the web network. To test the robustness of our approach to other network centrality formulations, we also replicated our analysis using PageRank instead of harmonic centrality, finding the same results as reported in this paper.

For interpretability, we standardized the harmonic centrality scores into z-scores, so that 0 represents the average harmonic centrality, 1 represents one standard deviation above the average, and so on.

For several of our analyses, we need to tag whether a post is original content (OC). OC are posts that are entirely *de novo*, and have not been observed on the web previously. OC represents the new creative content that is uploaded onto each community. In our data, we annotated an image meme as OC if neither the cloud vision API nor DBSCAN identified a similar image meme posted previously on the web or in our dataset. Future instances of the same meme in our dataset are not considered OC.

We captured covariates to include in our analysis as well. These covariates included: (1) the number of posts made on the community during the sample period, capturing the overall volume of the community; and (2) community size, as measured by Alexa audience scores from www.alexa.com, a leading firm that estimates web traffic to domains across the web. These covariates were exponentially distributed, so we performed a logarithmic transformation on them to include in our regression models. Community size and number of posts exhibited some collinearity ($r = .65$); however, the model outcomes were stable with and without including community size, so we kept them both.

4 Results

In total, we collected 3,308,151 images, of which 852,453 were filtered out because they did not include text in them, 525,871 of them were removed because they didn't pass the safe search filter, and 279,872 were filtered out as non-English. After filtering, our dataset contained 906,481 images comprising 683,556 different image memes. Our dataset is available at <https://osf.io/hr3e5/>, free for other researchers to download and use. The crawling code is available at <https://github.com/StanfordHCI/memescraper>

13,404 (1%) of the memes diffused to at least one other community during our data collection period. The median diffusion distance amongst memes that diffused was three. The median time to first diffusion was 9.1 hours, suggesting that our thirty day scrape is of sufficient length to capture most diffusion events for a meme.

We organize our analysis into three main sections. The first two sections concern Hypothesis 1, that diffusion events are negatively associated with network centrality. Section 4.1 features the main analysis, across the whole web. To compare the results with a traditional single-community analysis, Section 4.2 then replicates the same analysis, but only on Reddit. Finally, Section 4.3 concerns Hypothesis 2, testing across the whole web the relationship between centrality and the probability of a meme being original content (OC).

4.1 Do image meme diffusions originate in core or peripheral communities on the web?

In this section, we investigate our dataset to understand the relationship between centrality and diffusion events. We examine this question through several complementary methods. First, we examine descriptive statistics: what percent of diffusion events are attributable to memes originating in the core of the network? Second, we complement this descriptive result with a zero-inflated negative binomial regression, which helps us control for possible covariates. We then report the

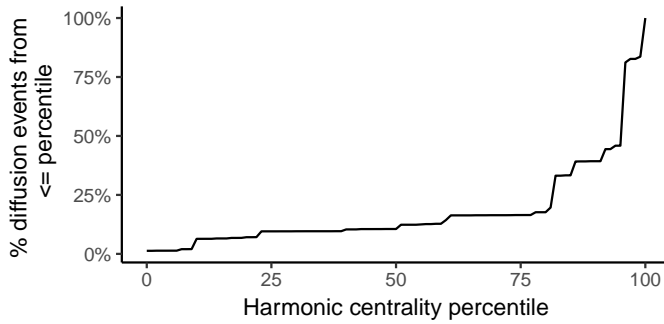


Fig. 2. The majority of memes that diffused to other communities originated in core communities rather than peripheral communities. For example, the top 10% most core communities originate 62% of diffusion events in our dataset.

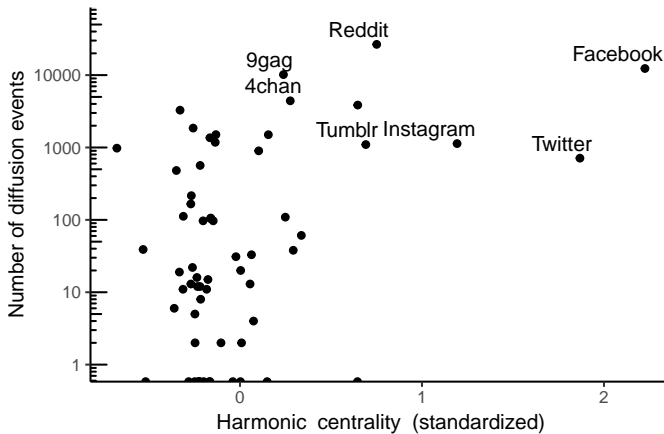


Fig. 3. Communities with the highest harmonic centrality—including Reddit, Facebook, Twitter, and Instagram—were also among the most productive in terms of originating the most diffusion events.

core and peripheral communities that originate the most (and fewest) diffusion events. Following this section, we will repeat this analysis within one large and influential community.

The majority of image meme diffusion events on the web—62%—originate from communities that are in the top 10% as measured by harmonic centrality (Figure 2). In contrast, about 20% of meme diffusion events originate from the next 20%, and about 18% of meme diffusion events originate from the peripheral communities in the the lowest 70%.

We performed a zero-inflated negative binomial regression, where each observation represents a meme and how many diffusion events it triggered. A zero-inflated negative binomial regression is a good fit for the data we are analyzing. Specifically, when modeling count data, it is generally ill-advised to use a linear regression, and more appropriate to use a method such as a Poisson regression, since a Poisson distribution better models counts of independent events. However, the Poisson distribution makes a strong assumption that mean equals variance—clearly inappropriate in a situation such as ours where some memes diffuse great distances—so it is common to utilize a negative binomial regression instead, which loosens this restriction. There is one final issue

Table 1. Memes originating from more popular communities are more likely to diffuse, but diffuse less far.

		Diffusions by meme	
		Negative Binomial	Zero-Inflated
		positive coefficient = more diffusion events	negative coefficient = more likely to diffuse
Harmonic Centrality (standardized)		-0.125*** (0.027)	-0.585*** (0.019)
Total Posts (log)		-0.105*** (0.011)	0.134*** (0.007)
Active Members (log)		0.028*** (0.006)	0.052*** (0.004)
(Intercept)		2.486*** (0.091)	1.410*** (0.064)
Observations		906,481	
Log Likelihood		-105,898.500	

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

to address: the observation that 99% of meme images in our data never diffused. The negative binomial regression does not correctly model data with overdispersed zeros such as this. So, we use a zero-inflated negative binomial regression, which corrects for this issue.

Unlike a typical regression, which has a single set of coefficients, a zero-inflated negative binomial regression separates the regression into two sets of coefficients and estimates these two sets simultaneously. One set of coefficients, typically estimated as a logistic regression, models the probability that the data point will be a zero (never diffused). Larger (positive) coefficients for this zero-inflated factor indicate that increases in the predictor variable make a zero more likely. In other words, positive values mean that a meme is *less* likely to diffuse. The other set of coefficients is the negative binomial component, which estimates the number of diffusion events for that meme. With the negative binomial, larger (positive) coefficients mean that the meme diffuses farther. So, a meme is most likely to diffuse, and to diffuse farthest, if zero-inflated coefficients are negative and negative binomial coefficients are positive.

In the zero-inflated negative binomial regression, standardized harmonic centrality was associated with fewer zeroes ($z = -30.48, p < .001$), meaning that the more core the community, the more likely the meme was to diffuse at least once. Standardized harmonic centrality was negatively associated with how far the meme ultimately diffused ($z = -4.655, p < .001$, Table 1). In other words, the memes originated in core communities diffuse less widely than memes originated in intermediate and peripheral communities, but they are more likely to diffuse in the first place. Having more active users in the community was associated with a lower probability of diffusion ($z = 11.45, p < .001$), but with larger diffusion volume ($z = 4.56, p < .001$). More total posts were associated with content being more likely to diffuse ($z = 17.07, p < .001$), and diffusing farther ($z = -9.64, p < .001$).

We investigated the communities in each category that were notable for their success (or lack thereof) in posting content that later diffused. For communities in the core (here, above the 90th

percentile in harmonic centrality), the community with the most diffused memes was Reddit, which originated 26,576 diffusion events to other communities. Facebook followed, with 12,346 diffusion events. Sampling from intermediate communities (here, those in the 70th–89th percentile by harmonic centrality), those originating the most diffusions were 9Gag and 4Chan; the one originating the fewest was MemeGenerator. Within the peripheral communities (1st–69th percentile), the one originating the most diffusions was HiddenLol; those originating the fewest diffusions were JoyReactor, RuinMyWeek, Barnorama, MemeCollection, each with no diffusions over the month of data collection.

We replicated this analysis with the diffusion distance rather than diffusion volume as the dependent variable. This second analysis counts the number of communities diffused to, rather than the number of raw meme images that were ultimately produced. Here, in the zero-inflated negative binomial regression, standardized harmonic centrality was still associated with fewer zeroes ($z = -29.62, p < .001$), but this time with slightly larger diffusion distance rather than shorter ($z = 3.12, p < .01$).

These results provide mixed support for Hypothesis 1 at the micro level of an individual piece of content, but not at the macro level of the ecosystem. Hypothesis 1 predicts that more central sites will be less influential. With *individual* pieces of content, higher harmonic centrality is associated with more diffusions, but shallower diffusions. However, the centralized sites make up for this in aggregate: even if individual pieces of content are diffuse less far, these communities push out so much content that they still originate over 60% of the diffusion events in our dataset. Peripheral communities may also have no *a priori* interest in diffusing their content to the mainstream, explaining their fewer diffusions, but Section 4.3 will identify that even the peripheral communities draw much of their meme content from the core, so there is influence at least in one direction.

4.2 Within communities, do memes originate from communities that occupy an intermediate network position on the web?

Do these patterns hold within communities as well? In other words, is the web-scale analysis yielding different results than a traditional within-community analysis might? Are peripheral subcommunities within a core community acting more like peripheral or core communities? To answer these questions, we replicated our analysis on a single community—the social aggregator Reddit—now treating subreddits as individual communities. In Reddit, we observed which subreddit community originated a piece of content, and which other subreddits that content later appeared on.

This analysis follows a similar structure as the previous section. We analyze the overall percentage of diffusion events attributable to image memes originating in the core and periphery of Reddit. We then perform a zero-inflated negative binomial regression to control for additional covariates. We again then report communities that are especially notable for originating many (or few) diffusion events.

However, to begin, we must construct the network graph amongst Reddit communities. With the web-scale analysis, we draw on the common practice of treating hyperlinks between webpages as directed edges. However, our preliminary experiments using links (e.g., crossposts) between Reddit communities demonstrated that Reddit has a much sparser sub-community link graph of explicit links. We sought a network structure that was valid to Reddit, using data available on Reddit.

We defined the directed edge weight as active commenter overlap between two communities. Consider a small source community s and a large destination, default Reddit community d . Intuitively, 75% of the posters and commenters of the small community s might also post or comment in the large default Reddit community d , indicating strong influence from s to d , but only 0.1% of

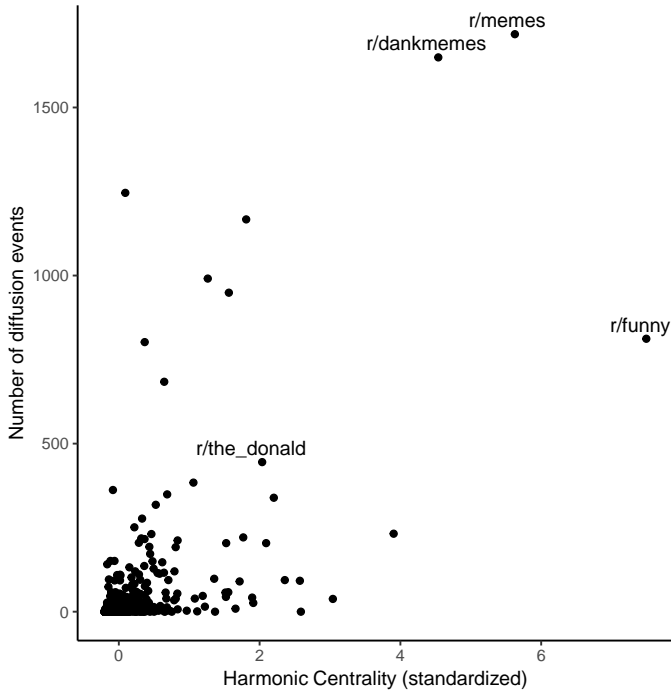


Fig. 4. Similar to our cross-community results, the majority of memes that diffused between sub-communities on Reddit originated in the core of the Reddit network, rather than in intermediate or peripheral subcommunities.

the large default community d 's members would be active in the small community, indicating weak influence from d to s . This is similar to the link structure of the web, where a small community might link to lots of content on Twitter, but only a small proportion of Twitter outgoing links are to the small community. Specifically, we defined the directed edge weight from subreddit A to subreddit B as active membership overlap, the number of users who commented in both subreddits during our data collection period divided by the total number of users who commented in subreddit A : $\frac{|A \cap B|}{|A|}$. Because the denominator $|A|$ changes depending on which direction the edge is pointing (e.g., $s \rightarrow d$ or $d \rightarrow s$), this is a directed graph. We calculated centrality again using standardized harmonic centrality.

With the within-domain analysis, we were also afforded a more precise estimate of the number of active members of each community. So, for this analysis, our covariate for community size was measured as the number of unique commenters in the subreddit during our observation period.

This analysis uses the same dataset as the prior section, restricted to a single community and peering inside that community instead of collapsing the community to a single node as in the prior analysis. We again trained a zero-inflated negative binomial regression predicting the diffusion volume for each image meme.

The overall volume measurements replicated the across-community analysis. In Reddit, subreddits in the top 10% by harmonic centrality were responsible for originating 69% of diffusion events, the next 20% by harmonic centrality originated 21%, and the final 70% in the periphery together originated 10% (Figure 4).

Table 2. Memes originating from more central subcommunities on Reddit are less likely to diffuse, but diffuse farther.

Diffusions by meme on Reddit		
	Negative Binomial	Zero-Inflated
	positive coefficient = more diffusion events	negative coefficient = more likely to diffuse
Harmonic Centrality (standardized)	0.364*** (0.081)	0.288*** (0.053)
log(OC posts)	-1.473*** (0.358)	3.071*** (0.198)
log(total posts)	1.820*** (0.345)	-2.856*** (0.190)
log(users)	-0.339** (0.041)	-0.313*** (0.027)
(Intercept)	1.081*** (0.291)	5.166*** (0.195)
Observations	371,113	
Log Likelihood	-37,593.440	

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

In the zero-inflated negative binomial regression within this community (Table 2), the coefficients reversed compared to the cross-community regression. Standardized harmonic centrality was associated with more zeroes, meaning that the more central the subreddit, the lower the probability that the meme would diffuse to another subreddit. Standardized harmonic centrality was also associated with more diffusions, meaning that the memes originated in the more central subreddits diffused farther, conditioned on them diffusing. So, the within-community analysis suggests potentially different mechanisms at play than the across-domain analysis, with the peripheral subcommunities within this core community originating memes that spread more to the core subcommunities but not to as many other subcommunities compared to core subcommunities.

We again catalogued individual accounts or communities of note. On Reddit, in the core, the subreddits whose memes originated the most diffusion events were *r/dankmemes*, *r/historymemes* and *r/memes*. Within intermediate subcommunities, *r/deepfriedmemes* originated the most diffusion events, and *r/comedycemetery* originated the fewest. Within peripheral subcommunities, *r/interestingasf**** and *r/depression_memes* originated the most diffusion events; *r/explainitpeter* and *r/peopledyinginside* originated the fewest.

Similar to cross-domain section, we replicated this analysis with the diffusion distance rather than diffusion volume as the dependent variable. The results replicated: in the zero-inflated negative binomial regression, standardized harmonic centrality was associated with more zeroes ($z = 4.819$, $p < .001$) and with larger diffusion distance ($z = 3.401$, $p < 0.001$).

Ultimately, these results provide a partial replication of the cross-community data. We compare the two studies' results in Table 3. Like the cross-community study, the most core subcommunities

Predictor	Negative Binomial		Zero-Inflated	
	positive coefficient = more diffusion events		negative coefficient = more likely to diffuse	
	Web-scale analysis	Single- community analysis	Web-scale analysis	Single- community analysis
Harmonic centrality	Negative	Positive	Negative	Positive

Table 3. Comparing the web-scale and single community analyses highlights the differences visible only at the web-scale lens. While the web-scale analysis had the same macro-scale result—that central communities originated more diffusion events—the mechanisms through which this occurred differed. Across the web, more central communities were more likely to diffuse (fewer zeroes) but diffused less far; within a single community, more central communities were less likely to diffuse (more zeroes) but diffused farther.

Table 4. Core communities are more likely to post original content.

	Original Content (1=OC)	
Harmonic Centrality (standardized)	0.272***	(0.007)
Active Members (log)	0.011***	(0.002)
Total Posts (log)	0.071***	(0.002)
(Intercept)	0.105***	(0.020)
Observations	906,481	
Log Likelihood	-506,715.700	
Akaike Inf. Crit.	1,013,439.0000	

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

originated the vast majority of diffusion events. However, the mechanism through which this occurred was different: fewer memes that diffuse, but memes that diffuse farther, compared to the cross-community study finding that the core communities have more memes that diffuse but will diffuse less far. This result reinforces the importance of pairing typical within-community analyses from the literature with cross-domain replications.

4.3 Do core or peripheral communities have a higher rate of posting original content (OC)?

To understand whether more central communities were producing less novel content, Hypothesis 2, we used a logistic regression analysis with the dependent variable coding whether each meme is original content, and the community's standardized harmonic centrality as the independent variable. We also visualized where the non-original content is attributable to for each community.

The logistic regression identified that the more core the community, the greater its log odds of posting OC ($p < .001$; Table 4). The model suggests that the odds ratio of a post being original content is multiplied by roughly 1.3 for each standard deviation increase in harmonic centrality. In other words, core communities are more likely to be posting fresh content, and Hypothesis 2 is not supported.

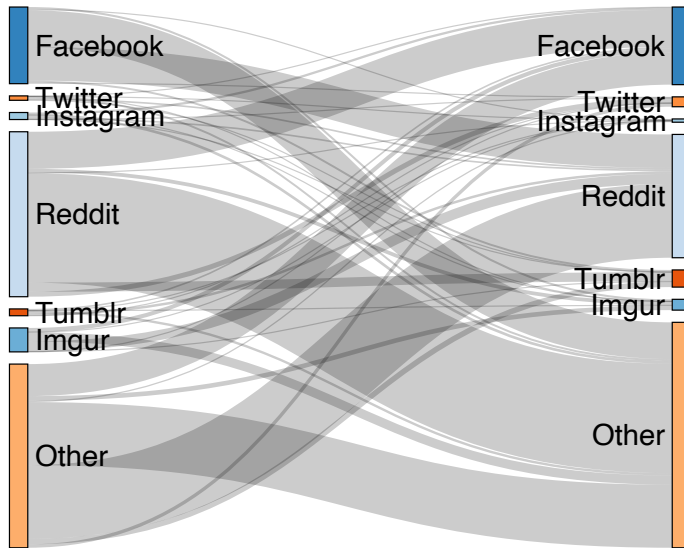


Fig. 5. Generally, all communities draw much of their content from core communities. The left side of the Sankey diagram is where content originates; the flows to the right demonstrate the first observed diffusion location. Content that originates on Reddit or Facebook diffuses in all directions. Communities that are named in this visualization are those in the top 10% by harmonic centrality.

Which communities draw on content originated by which other communities? Figure 5 visualizes where memes originated in each community appeared next in our dataset. In general, all communities are posting a substantial amount of non-OC content that was originated in core communities, from Facebook and Reddit especially. About $\frac{2}{3}$ of the non-OC content posted in intermediate and peripheral communities was originated in just two core communities, Facebook and Reddit. However, about half of the non-OC content on Facebook and Reddit does originate in non-core communities.

These results together suggest that core communities are more likely to repost content that originated elsewhere, but that the new content they do have is reposted in intermediate and peripheral communities. In addition, the core communities do post a non-trivial amount of content that was originated in intermediate and peripheral communities, but other content originated in other core communities is still the largest representation.

5 Sensitivity analysis

Our approach may have missed some peripheral communities. This section asks: What if some of these communities originated influential content? Would this threaten the results?

To test how missing peripheral communities might threaten our conclusions, we performed a sensitivity analysis. Specifically, we assumed a possible worst-case scenario where we did not find peripheral communities that were producing memes that originated a substantial number of diffusion events. To do this, we sampled all the communities from the top 10% by harmonic centrality, excluding Facebook and Reddit as they were outliers even within the core communities—communities of this size were extremely unlikely to have been missed. We then labeled these new communities as peripheral by assigning them harmonic centrality equivalent to the smallest

harmonic centrality existing in our dataset, and added these synthetic new communities to our dataset.

We continued doing this until the total volume of diffusions from the bottom 10th percentile was equal to the total volume of diffusions from the top 10th percentile. Achieving parity required adding 30 extra communities. Our original dataset comprised 88 communities, 56 of which produced content during our study, so by this analysis, we would have needed to miss a third of the meme-generating communities on the web through our crawl, and every single one of them would have needed to be both peripheral and equally productive as the most central communities such as Twitter and Pinterest. If this were true, these thirty communities would likely be as well known in the media and research literatures for originating this content—so this alternative explanation seems unlikely.

As another sensitivity analysis, we replicated this process again, this time trying to invalidate the zero-inflated negative binomial regression results for higher harmonic centrality having fewer zeroes. In other words, we again we sampled communities from the top 10% by harmonic centrality, excluding Facebook and Reddit, assigning them harmonic centrality equivalent to the smallest harmonic centrality existing in our dataset, added these synthetic new communities to our dataset, and continued doing so in an attempt to make the harmonic centrality coefficient in the model nonsignificant. After we added 100 communities, we ceased our attempt: the results were still significant.

In sum, by simulating increasingly worst-case scenarios for our crawling, we have strong evidence that it would be extremely unlikely for our results to be invalidated by missing influential peripheral communities. This result is due to the large effect sizes of the main results.

6 Limitations

A web-scale analysis provides a wide-angle lens that complements prior work's focus on individual communities, but it makes methodological tradeoffs to do so. Here, we reflect on these limitations.

First, even though our crawl took place over months, we cannot make formal guarantees that we have captured every eligible community. This is a fundamental limit of wide-lens analyses such as web-scale crawls including ours; while the method brings benefits, it also brings its own shortcomings. In particular, if our crawl did not come across a reference link to the community, and if Google did not index the domain (e.g., due to a robots.txt exclusion), it may not have been found in our crawl. Likewise, for privacy reasons, we did not index communities and accounts that were private, for example WeChat, Facebook groups, and Discord channels. Some content would have been flagged by the cloud vision API because much of it does get archived, for example 4chan. However, other platforms such as WeChat do not. So, our analysis was limited to the first public sharing of a piece of content.

If the missing content is mainly from core communities (e.g., Facebook groups), the results here would only be strengthened, since the core would be originating an even higher percentage of the image meme diffusions. Substantial meme activity occurs within Facebook groups as well; however, we were not able to observe it. Extending these results to include Facebook groups would only be possible if one were able to gain access to internal Facebook data. However, Facebook already holds both high harmonic centrality and high influence in terms of the number of diffusion events it originates, so our results would only be strengthened. Our sensitivity analysis helps estimate the possible impact of any oversights in the opposite direction, of peripheral communities.

However, if missing content were mainly from peripheral communities, it could undercut the results. Peripheral subcommunities within core communities are fine, because the indexing captured the large core communities; here we are mostly concerned with peripheral communities. This concern makes the sensitivity analysis central to the validity of our results. The sensitivity analysis

found that our crawling procedure would have needed to miss tens of peripheral communities, each originating memes that diffused as much as those on Twitter and Pinterest. Such communities would have been known to the media, and thus to us. So, this threat is unlikely to undermine the results; if anything, the countervailing effect of missing core communities such as Facebook groups likely makes our results an underestimate of the core communities, not an overestimate.

Second, we did not track communities that were not majority memes, with the exception of Reddit's `r/the_donald`, `4chan /b/`, and `4chan /pol/`, which were included for continuity with prior work [7, 25, 56]. While we did account for memes posted first to a multipurpose community in our method—any memes that first appeared in a multipurpose community were flagged by the cloud vision API as having arisen earlier on the web when they appeared in our sample, and filtered out—multipurpose communities do play an important role in the ecosystem. However, because we wanted to focus our analysis on content that we could be confident was an image meme, we restricted our communities to those that Mechanical Turk workers voted as majority meme. Otherwise, the analysis would include images spreading for other purposes, e.g., images of political figures diffusing to other news sites reporting on the same story. This would weaken the construct validity of our measures. In addition, a practical constraint: each image sent to the cloud vision API costs money, and including non-majority meme communities would require a much larger budget, since many of these communities are quite prolific but only a small proportion of their images would be of interest to the analysis. Future work can extend our analysis by including additional communities and filtering out non-meme content.

This limitation gives rise to a concern: non-majority meme sites behave differently than majority meme sites. For this reason, we manually whitelisted and included non-majority meme sites that featured prominently in previous analyses (e.g., [56]). These sites, such as Reddit's `r/the_donald`, as well as `4chan's /b/ (random)` and `/pol/ (politics)` boards, were included in our analyses. In contrast to prior work, they did not originate the content that was most widely diffused: `4chan` was ranked 4th in number of diffusion events in our per-domain analysis, and `r/the_donald` was ranked 7th in number of diffusion events in our within-domain analysis. So, while non-meme communities play a non-trivial role in origination of cultural artifacts, they do not appear to be leading producers.

Third, to map images onto templates, we excluded images that do not contain text. This resulted in dropping roughly 25% of the images that we encountered. We made this decision because pHash replicates visual features that the human perceptual system is likely to group together, and textless memes are often mutated via image editing software in ways that the pHash algorithm cannot detect. In other words, including textless memes will introduce a large number of false negatives into the analysis, where instances of the same meme are treated as separate memes, and thus thought to not diffuse.

However, removing textless memes may influence the results. For example, textless memes may spread in different social circles, or spread via different mechanisms, than memes with text. More broadly, the question might be posed as: how universal are our results, which focus on image memes with text, compared to other forms of cultural diffusion? For example, do our results generalize to text phrases, hashtags, viral videos, fashion, and food? Following prior work [10, 14, 24, 35, 45], we can only focus on one such domain in our study. Relative to these other forms of cultural diffusion, we suspect that memes with text have more rapid lifecycles (they are faster to create and edit than even textless memes, which require skills such as Photoshop), and because of language barriers, may be more culturally bounded than textless memes.

Fourth, we restricted this analysis to the English-speaking web, and the processes may differ in other cultural contexts. Practices of privacy, and of sharing, may not generalize. The cultural forces at play may differ in other contexts, and be more challenging to track if the sources cannot readily be scraped (e.g., WeChat in China).

Fifth, the network link structure of the web is not entirely coincident with social or cultural influence. As discussed in the Method section, we follow prior work in identifying core-periphery structure based on behavior (hyperlinking) that captures cultural influence, but also captures other elements. Our inspection of the harmonic centrality measure suggested that it does align with intuitions around core-periphery structure. However, there may be some room for differences. Our sensitivity analysis suggests that our results are robust to small (or even large) differences. However, it is important not to read too carefully into small differences in centrality score as being meaningful, as the score collapses together many forms of influence.

Finally, all results here are correlational, as our research questions are primarily investigating a question of network structure and not of causation, so we cannot draw any causal conclusions. For example, we cannot differentiate whether core communities cause greater diffusion, or whether the diffusion from these communities is what made them core in the first place. We also cannot identify the causal influence of the originating community, nor argue for agency of the originating community in any further spread.

7 Discussion

7.1 Navigating methodological tradeoffs in web-scale analyses

The benefits of a web-scale study are that it can consider an ecosystem-level view rather than focus narrowly on a single community, and that it can reveal patterns that a single community might not. However, performing such a study leads to methodological tradeoffs. For example, in our case, we made decisions to focus only on image memes with text, in English, posted by majority-meme communities. These decisions reflected different tradeoffs. Sometimes these tradeoffs were about construct validity: for example, including non-majority meme communities would mean that the crawl would eventually index news sites, blogs, and many other sites that would not qualify as “communities”, and that most of the images being tracked would not be image memes—they might be news photos, logos, product photos, or a variety of other content. Sometimes these tradeoffs were about scale: every query to the image search API cost money; many centrality algorithms fail with link rings and other strategic behavior that arises at web scale. Sometimes these tradeoffs were about trading off false positives and false negatives: textless memes are mutated via image editing software in ways that pHash cannot detect, which would result in meme diffusions not being correctly attributed to the originating community. So, we chose to drop the 25% of images that had no text. We made several of these tradeoffs of false positives vs. false negatives.

No single method can capture the web in its breadth. Each community has its own norms, APIs, and format. Our general approach has been to identify a tight subset of material that can be compared across the ecosystem, as rigorously as possible. We deemed that admitting false positives into the dataset, e.g., non-meme images, was more dangerous to our method than restricting the dataset. However, we had to be careful that restricting the dataset was not overly biasing our results. In our case, for example, we felt there was not a strong reason to believe that the 25% of non-text memes would behave categorically differently from the memes with text, and given that we could track the memes with text more successfully, we chose to focus on those.

The other tradeoff is, of course, that no method can capture the web in its entirety. Even the most powerful search engines in the world respect robots.txt files and do not index every web site. Clearly, important content arises outside of their view. We again had to ask the question of whether this threatens our results. In our case, we followed a two-criteria principle: (1) do we have a crawling algorithm that can identify the vast majority of content that exists?; (2) can we quantify whether the missing content threatens our conclusions? Addressing (1) motivated our use of the cloud vision API and its search index, which could lead us to any other public domain we might

have missed. Addressing (2) motivated the inclusion of a sensitivity analysis, deliberately trying to “break” the result.

We hope that these reflections can inspire future work to continue to refine these principles.

7.2 Reflections on theory

We asked whether a web-scale analysis of image memes would replicate and extend prior offline work indicating that communities in intermediate positions originate cultural production. A web-scale analysis attempts to characterize these dynamics more completely than is possible with traditional, single-community studies. Memes originated by the periphery were individually more likely to diffuse, but in aggregate, they were a relatively small percentage of memes that had diffusion events. This influence arises even if we control for activity level in terms of number of active participants. Peripheral and intermediate communities posted substantial content that was originated in core communities. Ultimately, these results do not support Hypothesis 1. We also hypothesized that original content (OC) rates would be lower in the core than in the periphery. This hypothesis, Hypothesis 2, was also not supported.

Our research is positioned as complementing prior work, which is focused on individual communities, by zooming out to an ecosystem level. Our results suggest that this strategy can be a critical lens to understanding social computing, since we found that the mechanisms internally within a community did not completely match the mechanisms at web scale. While no study can completely characterize the entire web, and our study makes methodological tradeoffs to operate at web scale, our results suggest that the CSCW community can continue to push beyond studies of individual communities.

That core communities originate so much of image meme culture online plays against popular and academic narratives. The most direct implication is that the cultural curve may have shifted online, with the ease of posting and sharing and the availability of large audiences all favoring mainstream and core communities more so than is possible offline. Intermediate communities such as 4chan claim responsibility for a vast number of memes [7, 25]; yet Reddit originates the content responsible for a far larger number of diffusion events. Core communities such as Reddit see large influxes of original content, and their content is also equally likely as peripheral communities to diffuse elsewhere. The growth of content aggregators such as Reddit may suggest that peripheral communities are not frequented as much as they used to be earlier in web’s history.

7.3 Alternative explanations

Another possible explanation is that our dependent variable can be refined: image meme diffusions may be incremental cultural innovations, rather than deeper and more transformative innovations such as the #metoo movement. Focusing on the communities responsible for memes at the 95th percentile of diffusion events might be a first step, but likely many of these innovations spawn multiple related memes. This remains future work.

While the popular narrative about marginal and counterculture communities is that they reject mainstream culture, peripheral communities appear to be sharing a substantial amount of content originated in the core. It is possible that our network centrality measures are not able to distinguish marginal communities from simply unpopular ones. However, it is also possible that this sort of reverse diffusion happens offline as well, with less spotlight: peripheral communities do not necessarily reject all of mainstream culture. We do not analyze what these peripheral communities *do* with this content—whether they use it earnestly themselves, mock it, or some combination—but this analysis confirms that the mainstream content influences their behavior.

Counter to prior work [56], we did not observe political communities such as Reddit’s `r/the_donald` and 4chan’s `/pol/` to be originating memes that were responsible for an especially

large number of diffusion events. Our data collection period was quite active in political news, too, covering Robert Mueller's delivery of his report on Russian meddling in the 2016 U.S. election and Robert Barr's announcement of his conclusion of no collusion in President Trump's campaign. This, of course, does not dispute that political communities have influence: however, it may help us place their broader place in the ecosystem.

The social web has also begun to blur the meaning of core and periphery, as core communities such as Reddit play host to peripheral sub-communities such as *r/interestingasf****. When an outsider hears that a meme originated at *r/interestingasf****, for example, do they assume that it comes from the mainstream (Reddit), or from the periphery (*r/interestingasf****)? Further within-community analysis would help examine these issues.

Our research also makes clear a limitation of prior work, which is that we know little about the roles of peripheral subcommunities within core communities. Our study found that peripheral subcommunities behaved differently than peripheral communities, originating more memes that diffuse, but those memes diffused less far. However, like the peripheral communities, these peripheral subcommunities did not originate a large proportion of the overall image meme diffusions on the community writ large. The CSCW community can further investigate the role of these communities that are peripheral within popular platforms.

7.4 Implications for design

A popular narrative is that much of internet culture comes from small peripheral communities. However, our work suggests that core communities have become not just the destination of memes, but a main origination point as well. Further, the core subcommunities are driving much of the diffusion within the core communities themselves—the core of the core appears to be extremely influential. We suggest that, recognizing this shift, core communities further support their members in creating and sharing cultural artifacts. How might they empower more people to engage in cultural artifact creation? For example, could these sites build their own simple image editing tools into the posts, or connect each meme image to its source so that it's easy for viewers to fork an image meme and create their own variant? Subcommunities such as Reddit's *r/memeeconomy* already expect this posting of the source template as a norm, but the site could support this behavior explicitly.

However, there also exist design implications that contravene the former: how might we design for the periphery to increase their influence if desired? Are peripheral communities being increasingly cut out the cultural equation? How can we produce tools that allow smaller communities to find critical mass, produce cultural content, and when appropriate, share it? Or do the peripheral communities prefer to keep their symbols to themselves (e.g., [46])?

8 Conclusion

Our results suggest that internet culture has bent the curve of cultural production toward the core. Toward this, we presented a web-scale analysis of image meme diffusion. Building on prior work in sociology, we hypothesized that cultural artifacts would arise from more peripheral communities in the network. We crawled meme communities across the web to identify a large number of communities and then captured a complete dataset of 906,481 instances of memes over 30 days. In contrast to our hypothesis, we found that memes originated in core communities originated the memes responsible for the majority of diffusion events on the web.

9 Acknowledgments

The authors thank for their support: Stanford MediaX, the Stanford Cyber Initiative, the Google Cloud Credits program, and the Office of Naval Research (N00014-16-1-2894). Toyota Research

Institute (“TRI”) provided funds to assist the authors with their research but this article solely reflects the opinions and conclusions of its authors and not TRI or any other Toyota entity.

References

- [1] Lada A Adamic and Natalie Glance. 2005. The political blogosphere and the 2004 US election: divided they blog. In *Proceedings of the 3rd international workshop on link discovery*. ACM, 36–43.
- [2] Eytan Adar and Lada A Adamic. 2005. Tracking information epidemics in blogspace. In *Proceedings of the 2005 IEEE/WIC/ACM international conference on web intelligence*. IEEE Computer Society, 207–214.
- [3] Katie Elson Anderson. 2020. Getting acquainted with social networks and apps: it is time to talk about TikTok. *Library Hi Tech News* (2020).
- [4] Lars Backstrom, Dan Huttenlocher, Jon Kleinberg, and Xiangyang Lan. 2006. Group formation in large social networks: membership, growth, and evolution. In *Proceedings of KDD*. 44–54.
- [5] Christian Bauckhage. 2011. Insights into Internet Memes.. In *ICWSM*.
- [6] Jonah Berger and Grant Packard. 2018. Are Atypical Things More Popular? *Psychological science* (2018), 0956797618759465.
- [7] Michael S Bernstein, Andrés Monroy-Hernández, Drew Harry, Paul André, Katrina Panovich, and Gregory G Vargas. 2011. 4chan and/b: An Analysis of Anonymity and Ephemerality in a Large Online Community.. In *ICWSM*. 50–57.
- [8] Youmna Borghol, Sebastien Ardon, Niklas Carlsson, Derek Eager, and Anirban Mahanti. 2012. The untold story of the clones: Content-agnostic factors that impact YouTube video popularity. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. 1186–1194.
- [9] Caroline Bynum, Richard Biernacki, Steven Feierman, Karen Halttunen, Margaret C Jacob, Sonya O Rose, Jerrold Seigel, William H Sewell, and Margaret R Somers. 1999. *Beyond the cultural turn: New directions in the study of society and culture*. Vol. 34. Univ of California Press, 35–61.
- [10] Gino Cattani and Simone Ferriani. 2008. A core/periphery perspective on individual creative performance: Social networks and cinematic achievements in the Hollywood film industry. *Organization Science* 19, 6 (2008), 824–844.
- [11] Justin Cheng, Lada Adamic, P Alex Dow, Jon Michael Kleinberg, and Jure Leskovec. 2014. Can cascades be predicted?. In *Proceedings of WWW*. 925–936.
- [12] Justin Cheng, Lada A Adamic, Jon M Kleinberg, and Jure Leskovec. 2016. Do cascades recur?. In *Proceedings of WWW*. 671–681.
- [13] Common Crawl Foundation. 2017. Common Crawl’s First In-House Web Graph. <http://commoncrawl.org/2017/05/hostgraph-2017-feb-mar-apr-crawls/>. Accessed: 13-Sep-2020.
- [14] Linus Dahlander and Lars Frederiksen. 2012. The Core and Cosmopolitans: A Relational View of Innovation in User Communities. *Organization Science* 23, 4 (2012), 988–1007. <https://doi.org/10.1287/orsc.1110.0673> arXiv:<https://doi.org/10.1287/orsc.1110.0673>
- [15] David Easley and Jon Kleinberg. 2010. *Networks, crowds, and markets*. Vol. 8. Cambridge University Press Cambridge.
- [16] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise.. In *Kdd*, Vol. 96. 226–231.
- [17] Flavio Figueiredo, Jussara M Almeida, Fabrício Benevenuto, and Krishna P Gummadi. 2014. Does content determine information popularity in social media? A case study of YouTube videos’ content and their popularity. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 979–982.
- [18] Adam Fish. 2017. *Technoliberalism and the end of participatory culture in the United States*. Springer.
- [19] Eric Gilbert. 2013. Widespread underprovision on Reddit. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 803–808.
- [20] Cole Gleason, Amy Pavel, Xingyu Liu, Patrick Carrington, Lydia B. Chilton, and Jeffrey P. Bigham. 2019. Making Memes Accessible. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) (ASSETS ’19). Association for Computing Machinery, New York, NY, USA, 367–376. <https://doi.org/10.1145/3308561.3353792>
- [21] Sharad Goel, Duncan J Watts, and Daniel G Goldstein. 2012. The structure of online diffusion networks. In *Proceedings of the 13th ACM conference on electronic commerce*. ACM, 623–638.
- [22] Google. 2020. Google Cloud Vision API. <https://cloud.google.com/vision/>. Online; accessed: 13-September-2020.
- [23] Andy Greenwald. 2003. *Nothing feels good: Punk rock, teenagers, and emo*. St. Martin’s Griffin.
- [24] John R Hall, Mary Jo Neitz, and Marshall Battani. 2003. *Sociology on culture*. Psychology Press.
- [25] Gabriel Emile Hine, Jeremiah Onaolapo, Emiliano De Cristofaro, Nicolas Kourtellis, Ilias Leontiadis, Riginos Samaras, Gianluca Stringhini, and Jeremy Blackburn. 2016. Kek, cucks, and god emperor trump: A measurement study of 4chan’s politically incorrect forum and its effects on the web. *arXiv preprint arXiv:1610.03452* (2016).

- [26] Jeff Huang, Katherine M. Thornton, and Efthimis N. Efthimiadis. 2010. Conversational Tagging in Twitter. In *Proceedings of the 21st ACM Conference on Hypertext and Hypermedia* (Toronto, Ontario, Canada) (HT '10). ACM, New York, NY, USA, 173–178. <https://doi.org/10.1145/1810617.1810647>
- [27] Dino Ienco, Francesco Bonchi, and Carlos Castillo. 2010. The meme ranking problem: Maximizing microblogging virality. In *Data Mining Workshops (ICDMW), 2010 IEEE International Conference on*. IEEE, 328–335.
- [28] Jason Kaufman and Orlando Patterson. 2005. Cross-national cultural diffusion: The global spread of cricket. *American sociological review* 70, 1 (2005), 82–110.
- [29] Robert E Kraut and Paul Resnick. 2012. *Building successful online communities: Evidence-based social design*. Mit Press.
- [30] Jennifer C Lena. 2012. *Banding together: How communities create genres in popular music*. Princeton University Press.
- [31] Jure Leskovec, Lars Backstrom, and Jon Kleinberg. 2009. Meme-tracking and the dynamics of the news cycle. In *Proceedings of KDD*. ACM, 497–506.
- [32] Jure Leskovec, Ajit Singh, and Jon Kleinberg. 2006. Patterns of influence in a recommendation network. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 380–389.
- [33] Chen Ling, Ihab AbuHilal, Jeremy Blackburn, Emiliano De Cristofaro, Savvas Zannettou, and Gianluca Stringhini. 2021. Dissecting the Meme Magic: Understanding Indicators of Virality in Image Memes. CSCW (2021). arXiv:2101.06535 [cs.HC]
- [34] Chen Ling, Ihab AbuHilal, Jeremy Blackburn, Emiliano De Cristofaro, Savvas Zannettou, and Gianluca Stringhini. 2021. Dissecting the Meme Magic: Understanding Indicators of Virality in Image Memes. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 81 (April 2021), 24 pages. <https://doi.org/10.1145/3449155>
- [35] Paul Lopes and Paul Douglas Lopes. 2002. *The rise of a jazz art world*. Cambridge University Press.
- [36] Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology* 27, 1 (2001), 415–444.
- [37] Adam J Mills. 2012. Virality in social media: the SPIN framework. *Journal of public affairs* 12, 2 (2012), 162–169.
- [38] Vishal Monga and Brian L Evans. 2006. Perceptual image hashing via feature points: performance evaluation and tradeoffs. *IEEE Transactions on Image Processing* 15, 11 (2006), 3452–3465.
- [39] Seth A Myers and Jure Leskovec. 2014. The bursty dynamics of the twitter information network. In *Proceedings of WWW*. 913–924.
- [40] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1999. *The PageRank citation ranking: Bringing order to the web*. Technical Report. Stanford InfoLab.
- [41] Sasa Petrovic, Miles Osborne, and Victor Lavrenko. 2011. RT to win! Predicting message propagation in Twitter. *ICWSM 11* (2011), 586–589.
- [42] Hayagreeva Rao, Philippe Monin, and Rodolphe Durand. 2003. Institutional change in Toque Ville: Nouvelle cuisine as an identity movement in French gastronomy. *American journal of sociology* 108, 4 (2003), 795–843.
- [43] Yannick Rochat. 2009. *Closeness centrality extended to unconnected graphs: The harmonic centrality index*. Technical Report.
- [44] Manuel Gomez Rodriguez, David Balduzzi, and Bernhard Schölkopf. 2011. Uncovering the temporal dynamics of diffusion networks. *arXiv preprint arXiv:1105.0697* (2011).
- [45] Gabriel Rossman. 2012. *Climbing the charts: What radio airplay tells us about the diffusion of innovation*. Princeton University Press.
- [46] Jessica Roy. 2016. How ‘Pepe the Frog’ went from harmless to hate symbol. <https://www.latimes.com/politics/la-na-pol-pepe-the-frog-hate-symbol-20161011-snap-htlmlstory.html>. Online, accessed: 07-July-2021.
- [47] Bryce Ryan and Neal Gross. 1950. Acceptance and diffusion of hybrid corn seed in two Iowa communities. *Research Bulletin (Iowa Agriculture and Home Economics Experiment Station)* 29, 372 (1950), 1.
- [48] Kate Starbird, Ahmer Arif, and Tom Wilson. 2019. Disinformation as Collaborative Work: Surfacing the Participatory Nature of Strategic Information Operations. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 127 (Nov. 2019), 26 pages. <https://doi.org/10.1145/3359229>
- [49] Leo Graiden Stewart, Ahmer Arif, A. Conrad Nied, Emma S. Spiro, and Kate Starbird. 2017. Drawing the Lines of Contention: Networked Frame Contexts Within #BlackLivesMatter Discourse. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 96 (Dec. 2017), 23 pages. <https://doi.org/10.1145/3134920>
- [50] Gabor Szabo and Bernardo A Huberman. 2010. Predicting the popularity of online content. *Commun. ACM* 53, 8 (2010), 80–88.
- [51] Jacqueline Ryan Vickery. 2014. The curious case of Confession Bear: the reappropriation of online macro-image memes. *Information, Communication & Society* 17, 3 (2014), 301–325.
- [52] Yuan Wang, Yukun Li, Xinning Gui, Yubo Kou, and Fenglian Liu. 2019. Culturally-Embedded Visual Literacy: A Study of Impression Management via Emoticon, Emoji, Sticker, and Meme on Social Media in China. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 68 (Nov. 2019), 24 pages. <https://doi.org/10.1145/3359170>
- [53] Raymond Williams. 2014. *Keywords: A vocabulary of culture and society*. Oxford University Press, 87.

- [54] Weiai Wayne Xu, Ji Young Park, Ji Young Kim, and Han Woo Park. 2016. Networked cultural diffusion and creation on YouTube: An analysis of YouTube memes. *Journal of Broadcasting & Electronic Media* 60, 1 (2016), 104–122.
- [55] Weiai Wayne Xu, Ji-young Park, and Han Woo Park. 2017. Longitudinal dynamics of the cultural diffusion of Kpop on YouTube. *Quality & Quantity* 51, 4 (2017), 1859–1875.
- [56] Savvas Zannettou, Tristan Caulfield, Jeremy Blackburn, Emiliano De Cristofaro, Michael Sirivianos, Gianluca Stringhini, and Guillermo Suarez-Tangil. 2018. On the Origins of Memes by Means of Fringe Web Communities. In *ACM SIGCOMM Internet Measurement Conference*. ACM.

Received January 2021; revised July 2021; accepted November 2021