The World Wants Mangoes and Kangaroos: A Study of New Emoji Requests Based on Thirty Million Tweets

Yunhe Feng[1], Wenjun Zhou[1], Zheng Lu[1], Zhibo Wang[2], Qing Cao[1]

[1]University of Tennessee, Knoxville; [2]Wuhan University

{yfeng14,wzhou4,zlu12,cao}@utk.edu,zbwang@whu.edu.cn

ABSTRACT

As emojis become prevalent in personal communications, people are always looking for new, interesting emojis to express emotions, show attitudes, or simply visualize texts. In this study, we collected more than thirty million tweets mentioning the word "emoji" in a one-year period to study emoji requests on Twitter. First, we filtered out bot-generated tweets and extracted emoji requests from the raw tweets using a comprehensive list of linguistic patterns. Then, we examined patterns of new emoji requests by exploring their time, locations, and context. Finally, we summarized users' advocacy behaviors and identified expressions of equity, diversity, and fairness issues due to unreleased but expected emojis, and concluded the significance of new emojis on society. To the best of our knowledge, this paper is the first to conduct a systematic, large-scale study on new emoji requests.

CCS CONCEPTS

• **Information systems** → **Web mining**; **Social networks**; **Information retrieval**; • **Human-centered computing** → *Interactive systems and tools.*

KEYWORDS

emoji analysis; emoji mining; emoji petition; relatedness, fairness and equality in emojis; emoji categorization; emoji profiling

1 INTRODUCTION

The word emoji comes from the Japanese words e ("picture") and moji ("character") and has a history of nearly 30 years since it originated on Japanese mobile phones in the late 1990s. In 2009, a set of 722 emojis were first officially added into Unicode Standard 5.2 [28]. After Apple introduced the iOS emoji keyboard in 2011, the use of emojis grew rapidly [44]. By 2018, more than 2700 emojis had been added into Unicode Standard 11.0. According to a recent survey [44], almost everyone online (92% of the online population)

was using emojis. With the popularity of social networks, nowadays, emojis are used extensively on various social networking platforms, such as Twitter, Facebook, WhatsApp, and Instagram. In particular, nearly half of comments and captions on Instagram have emojis [9].

As the usage of emojis (and social media in general) evolves, new emojis are being continuously requested. The Unicode Consortium[1] updates the official list of Unicode emojis by judging and accepting proposals for new emojis annually. For each candidate emoji, its evidence of frequency from Google Search, Bing Search, Youtube Search and Google Trends must be submitted, and evidence from NGram Viewer and Wikipedia Search are optional. Besides substantial efforts to collect such evidence, this method has several additional drawbacks. First, not all objects with a higher frequency in search engines are more likely to be emojilized. For example, although the "mascot" is heavily searched, it is unlikely to be an emoji because there exists no specific image representing all mascots for different teams, events, and organizations. Second, it completely ignores emoji petitions directly generated by users who have first-hand information regarding the valuable usage context.

However, a systematic study on which new emojis are wanted, when, where and why these emojis are requested, and how to call for these emojis still remains unexplored. Few studies attempted to offer even partial answers to these questions. The emoji satisfaction survey [45] reported mobile message app users always desired more emoji choices, but provided no further detailed answers to the above specific emoji questions. Thomas Dimson [9] showed the emoji usage trend on Instagram from 2010 to 2015 during which a large number of new emojis were proposed, but those new emojis were not studied. Yonatan Zunger [52] analyzed the animal emoji requests of one day (August 3, 2017) using Twitter's Search APIs, and demonstrated the world wanted raccoon and lobster emojis. However, it only studied emojis in a single category on a small-scale in terms of both the number of tweets and the tweeting time span.

To more comprehensively study the emoji requests, this paper investigated more than thirty million tweets mentioning the word "emoji," and proposed a new framework to answer the above questions. Specifically, the framework consisted of the following analyses on the thirty million emoji-mentioned tweets. First, we extracted the requested emoji descriptions and calculated their corresponding frequencies. After filtering out emojis that already existed (i.e., extant emoji requests) and explaining why they were still being requested, we proposed a WordNet [31] based emoji classifier to cluster requested emojis. Then, we studied spatiotemporal patterns of these requests and explored possible reasons why new emojis were requested. Moreover, we presented the common characteristics of requested emojis and advocacy behaviors. Finally, we illustrated the existing relatedness, fairness, and equality problems reflected

---

[1]https://unicode.org/emoji/

through emojis, and discussed the positive impacts of new emojis on society.

As the first step to conduct a systematic, large-scale study on new emoji requests, our contributions can be summarized as follows:

- We revealed new and strong evidence of frequency, i.e., the explicit and accurate evidence like "why is there no *foo* emoji", for the Unicode emoji community to evaluate emoji petitions.
- Our study explained why extant emojis were still requested by users, and provided multiple suggestions to enable the timely availability of newly released emojis to users.
- A comprehensive understanding of new emoji requests was offered by profiling spatiotemporal distributions, summarizing advocacy behaviors, and exploring factors that inspire requests.
- We discussed the equality, fairness and diversity in emojis, and presented the potential significance of new emojis in many aspects like business promotion and violence control.

2 RELATED WORK

There is a great deal of work that investigates how people use emojis to facilitate communication and social interactions. Expressing and strengthening emotions [22, 43, 47], and conveying humor [8, 20] and sarcasm [14, 48] are major functions of emojis in interpersonal communication. Emojis are also used to manage conversations, such as maintaining a conversational connection, and ending a thread [7, 51]. In addition, emojis show more diverse functions in close relationships. Kelly et al. [25] reported that emojis encouraged playful interactions, and created shared and secret uniqueness between people in mediated close personal relationships. Similarly, Wiseman et al. [50] examined the repurposing emojis for personalized communication between close partners, friends and family members. An example is using the shared love of the pizza 🍕 to represent romantic love between partners.

Emojis are also having the significant influence on relevant life domains of business, politics, religion, entertainment and food, etc. Various companies used emojis to enrich their promotions, create awareness and attract attention from consumers [19, 26, 30]. Political leaders from the United States [2], Australia [42], and Argentina [18] used emojis in official speech, during interviews, or on social networks. For religions, some users embedded emojis like the folded hands (the prayer hands) emoji 🙏 into their usernames. Another example is the recycling symbol emoji ♻, which was taking over on Twitter due to its extensive usage by Arabic speakers to represent a shared Islamic *Dua* (supplication or invocation) [40]. In the entertainment field, *The Emoji Movie*, an animated film based on emoji graphics, was released in 2017. Recently, many researchers [16, 23, 24, 47] associated foods with emojis, and suggested emojis to be an easy and non-verbal way to measure food-related emotions, especially for children [16].

As emojis are impacting various aspects of real life, problems of equity, diversity, and fairness caused by emojis are drawing more attention. To promote gender equality through emojis, researchers from Google proposed a set of emojis reflecting a wide range of professions for women (as well as men) with a goal of highlighting the diversity of women's careers [36]. The *person* emoji 🧑, an adult with no gender specified, had become available as a gender-inclusive alternative to the man 👨 or the woman 👩 since Unicode Standard

10.0 in 2017. A recent study revealed that 0.13% of all emojis send by Americans were either a rainbow flag 🏳️‍🌈 (commonly known as the lesbian, gay, bisexual and transgender pride flag), men holding hands 👬 or women holding hands 👭 emojis [21]. Since *"people all over the world want to have emoji that reflect more human diversity, especially for skin tone"* [5], the Unicode Consortium released five different skin tone modifiers, which was based on the six tones of the Fitzpatrick scale [15, 49], to enrich human diversity in Unicode Standard 8.0 in 2015. When a human emoji is immediately followed by one skin tone modifier character, the person(s) or body part will be rendered using the specified skin tone. For example, 👦 + {🏻,🏼,🏽,🏾,🏿} → {👦🏻,👦🏼,👦🏽,👦🏾,👦🏿}. To depict diverse hair colors and styles, Unicode Standard 11.0 introduced hair components including red-haired, curly-haired, white-haired, and bald components in 2018 [6]. Besides, some emojis enable multi-person groupings, which enhances family-related emojis diversity significantly, such as single-parent families 👩‍👧 👨‍👦 , and homoparental families 👨‍👨‍👧 👩‍👩‍👦.

3 IDENTIFYING NEW EMOJI REQUESTS

3.1 Data Curation

We used Twitter's Streaming APIs, which enable developers to filter and collect real-time tweets, to crawl all the English tweets containing the word "emoji". In our study, more than thirty million tweets of interest were crawled in total from Oct. 2017 to Oct. 2018. The collected tweets were formatted in JavaScript Object Notation (JSON) files with named attributes and associated values [46].

To eliminate the side effects of bots on Twitter, we followed approaches proposed by Ljubesic et al. [27] to filter out those bot-generated tweets. More specifically, we removed eleven users who produced on average more than 10 "emoji" tagged tweets per day. For each of users having more than 100 collected postings, we calculated the time (in minutes) between her/his two successive tweets and removed those users whose three most frequent time spans between postings covered more than 90% of their overall production. This method deleted overall 131 users and 43,461 tweets.

We then extracted information of interest such as user profiles, tweet contents, timestamps, and geo tags, from JSON files, and built the *complete dataset* and the *unique dataset* for emoji analysis subtasks in Section 4 and Section 5 respectively. The *complete dataset* consisted of four types of tweets including general tweets, retweets, quoted tweets and replies. The *unique dataset* only contained the original tweets and excluded duplicated tweets like retweets. As tweets are usually composed of incomplete, noisy and poorly structured sentences due to the frequent presence of abbreviations, irregular expressions, ill-formed words and non-dictionary terms, a series of preprocessing steps were applied to reduce noise in tweets. For example, we removed URLs and non-ASCII characters except Unicode characters reserved for emojis.

3.2 Emoji Extraction Using Linguistic Patterns

Note that not all collected tweets were petitions for new emojis, e.g., the tweet like "I love this emoji!" was crawled as well since it contained the keyword "emoji". Therefore, we needed to identify emoji-requested tweets and extracted wanted emojis. However, Twitter users could choose different words and sentence patterns to express their expectations of new emojis, which made the emoji

extraction challenging. Yonatan Zunger [52] assumed that mentions of phrases like "*foo* emoji" were positive statements about desiring such an emoji. Although this hypothesis was claimed to hold true when validating with spot-checks of the matching tweets, it suffered from false positives, e.g., a tweet like "I hate a *foo* emoji!" was incorrectly recognized as desiring the *foo* emoji.

Inspired by [52], we proposed fine-tuned linguistic patterns to detect desired emojis more precisely. Based on our observations, we summarized 49 frequent linguistic patterns and their 2620 variations to match emoji-requested tweets and extract emojis. Ten linguistic patterns are illustrated as follows, and the whole linguistic pattern list and corresponding tweet screenshots are available through http://yunhefeng.me/linguistic_patterns.html.

- why is there no *foo* emoji
- where is the *foo* emoji
- need a *foo* emoji
- look for a *foo* emoji
- demand a *foo* emoji
- have no *foo* emoji
- invent a *foo* emoji
- a *foo* emoji is overdue
- still no *foo* emoji
- give us a *foo* emoji

To broaden the matching scopes, we first adopted natural language processing techniques, including part-of-speech (POS) tagging, stemming and lemmatization, before checking tweet contents. For example, "look for a *foo* emoji" would match "looked for an *foo* emoji", "looks for the *foo* emoji", etc. We also considered characteristics of casual English on social networks [4] and fixed common problems, such as the punctuation omission/error (e.g., theres → there's), the wordplay (e.g., neeeeeed → need), and the censor avoidance (e.g., shlt, fck, f***). In addition, we took possible variations on sentence structures of linguistic patterns into account. For example, our linguistic patterns covered not only "need a *foo* emoji" but also "need an emoji of/with/for *foo*".

3.3 Requesting Extant Emojis

When examining extracted emojis, to our surprise, we found hundreds of emojis that had already been released by the Unicode Consortium were still requested extensively. Figure 1(a) demonstrates the top 25 most requested extant emojis. Seven out of the top 10 extant emojis came from the recent Emoji Version 5.0, which was released in May 2017 [11], while we started to crawl the data in Oct. 2017. It is also interesting to note that the percentage of Twitter users on mobile was about 80% [33], but they contributed more than 91.8% extant emoji requests, as shown in Figure 1(b).

We compared Twitter's post-a-tweet interfaces on different platforms to explore why users could not find extant emojis, especially on mobile. On Twitter's desktop site, the post-a-tweet interface offers an emoji picker which contains all latest official emojis. By contrast, on mobile devices, post-a-tweet interfaces of both the mobile site and mobile apps have no such emoji pickers. Instead, users have to rely on on-screen keyboards to type emojis, which may cause potential poor user experiences. First, keyboards may not incorporate the latest emojis promptly so that new emojis are unavailable for users. Second, users may not update keyboards to the latest versions to access the recently added emojis, or their mobile operating systems are too out-of-date to be compatible with the latest versions of keyboards. Third, bad emoji keyboard layouts make it difficult for users to find and type intended emojis even if these emojis have been included.
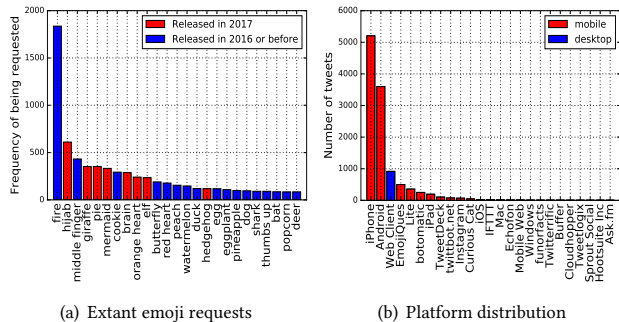

(a) Extant emoji requests   (b) Platform distribution

Figure 1: Requesting extant emojis

Accordingly, we have the following suggestions and recommendations to improve the user experience of inputting emojis. Mobile users should update installed keyboard software frequently and select the high-quality keyboard with proper emoji arrangements. It is keyboard developers' responsibility to merge newly released emojis into their products as soon as possible and highlight it in "What's New" descriptions of their keyboard apps to remind users that new emojis are available. For app developers, they can add the emoji picker or the search bar to enable users to input emojis without entirely relying on third-party keyboards.

3.4 Emojis Categorization

The emoji categorization plays a big role in facilitating emoji inputs for both mobile and desktop users. Almost all mobile emoji keyboards arrange emojis into categories to alleviate the problems of large lists. Most emoji pickers on social networks also group emojis to help users select wanted emojis quickly and effortlessly. When new emojis come, knowing the number of emojis in each category guides emoji input interface designers to adjust the emoji arrangement, such as increasing the number of emojis per screen. Especially when a large number of new emojis are requested, an automatic emoji classifier is necessary and helpful.

The Unicode Consortium officially categorizes emojis into eight groups, i.e., Smileys & People, Animals & Nature, Food & Drink, Activity, Travel & Places, Objects, Symbols, and Flags. The category of Flags is the easiest one to be detected, since each emoji belonging to this category contains the keyword of "flag". Therefore, we could simply search this keyword in descriptions of each requested emoji to determine whether it should be classified into the flag group.

However, for the rest categories, the method of searching keywords is obviously ineffective because of the difficulty in summarizing a set of keywords representing a particular category. We instead trained a semantic classifier based on WordNet [31], which is a widely-used lexical database for English. A critical concept in WordNet is the synset, namely a set of synonyms sharing a common meaning. In our study, we focused on noun synsets to which words in emoji descriptions belonged. For example, the similarity between two emojis e_1 and e_2 was calculated as the highest similarity score between the two noun synsets containing words in e_1 and e_2. For the similarity score between two synsets, we took the *path_similarity* score, a similarity metric based on the shortest

path that connected the senses in the is-a taxonomy, to denote how similar two word senses were. For a testing emoji e_{te}, we calculated its *path_similarity* with each training emoji e_{tr} in each category c. Then, for each category c, we sorted its similarity score list and summed up the top k similarity scores to represent the similarity of the unlabeled emoji e_{te} and the category c. Finally, we set the category with the largest summed similarity as the label for e_{te}.

We performed the 5-fold cross validation on the category dataset collected from Emojipedia[2] for 50 times and achieved an average accuracy of 71.1% with the top k set as 9. One may argue that the accuracy is too low. However, even the official category labels of some emojis are indeed ambiguous. For example, the monkey face (U+1F435) is categorized as animals & nature, but the see-no-evil monkey (U+1F648), hear-no-evil monkey (U+1F649), and speak-no-evil monkey (U+1F64A) are classified as smileys & people. In addition, prior research studies [32, 35] revealed that ambiguities in emoji categorization were common. So, we think the achieved accuracy is acceptable with the messy data.

4 PATTERNS OF NEW EMOJI REQUESTS

4.1 Requested Emojis by Category

We used the proposed keyword matching method and the WordNet-based classifier to categorize requested emojis. Considering requested emojis were too diverse, we only counted emoji requests greater than 10 times. As shown in Table 1, more than 31.8% of wanted emojis were from the Smileys & People category, which might indicate people's great passions for new emojis to express emotions. The public also desired many emojis, including kangaroos and mangoes, from the Animals & Nature and Food & Drink categories. Surprisingly, the number of tweets requesting symbol emojis was very large. After digging into related tweets, we found one tweet petitioning the anarchy symbol emoji, had been retweeted for over 6,000 times, which accounted for more than 27% of the total tweets in the Symbols category.

It is reasonable that categories of the Activity, Travel & Places, and Flags had relative fewer requests, since most emojis in these categories have been released. In addition, it takes a long time to evolve a new activity like a sports game, a new place like electric vehicle charging stations, or flags for new-born countries or influential social movements. The different demands of emojis per category may inspire emoji input interface designers to optimize emoji layouts, such as reserving spaces for new coming emojis, and displaying more emojis per screen. They can even regroup emojis, as suggested by Na'aman et al [32], to enhance user experience.

4.2 Temporal Distributions

We aggregated tweets petitioning the same emoji together by month. Figure 2 demonstrates emojis that were requested more than 1000 times throughout one year (Oct. 2017 - Oct. 2018). The circle diameter represents the number of requests made. Although the overall requested number was not very large, emojis of brooms, flamingos and kangaroos appeared consistently in all months. In contrast, heavily requested emojis like the lookout and the red carpet mainly appeared in one or two months. The fact that the broom, flamingo

[2]https://emojipedia.org/

Table 1: Emoji requests by category

Category	# emojis	# tweets	examples
Smileys & People	385	36,790	redhead, ass shaking
Animals & Nature	185	18,059	kangaroo, flamingo
Food & Drink	164	12,067	mango, waffle
Activity	42	2,421	slide, softball
Travel & Places	56	1,946	compass, brick
Objects	161	12,229	broom, red carpet
Symbols	170	21,443	anarchy, infinity
Flags	44	3,156	trans flag, Texas flag

and kangaroo emojis were selected as part of Unicode 11.0 in 2018 or Unicode 12.0 in 2019 implied those emojis that were requested continuously and by multiple users were more likely to be approved by the Unicode Consortium as they reflected the real needs of the majority of online users. We also found the extensive but relatively concentrated emoji requests were usually triggered by celebrities or their followers. For example, petitions of the red carpet emoji were retweeted more than one thousand times by fans of BTS, a South Korean boy band, within 24 hours. After that, there was nearly no petition for the red carpet emoji any more.
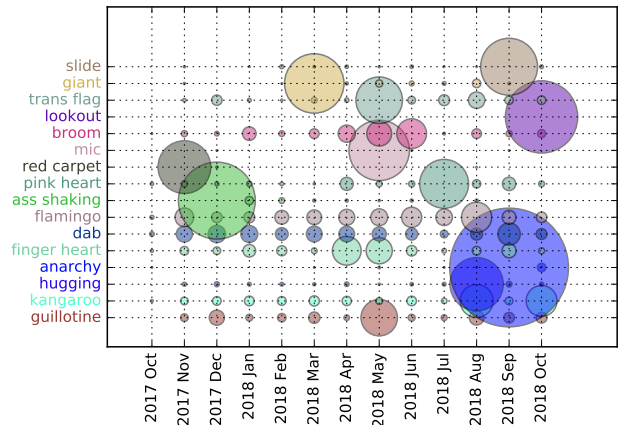

Figure 2: The number of requests per emoji per month. The circle diameter represents the number of requests.

4.3 Geographic Distributions

We utilized geotagged tweets (2.8% of tweets in the *complete dataset*) to profile geographic distributions of new emoji requests at both worldwide and national levels. We observed that people in as many as 110 different countries petitioned for new emojis. As we collected tweets written in English, English-speaking countries, such as United States (73.6%), United Kingdom (10.9%), and Canada (3.2%), contributed the most of emoji requests. It is interesting that non-native English-speaking countries, such as China, Japan, Brazil and Mexico, also expressed their desire for new emojis even in English, which might be one evidence of the world's passion for emojis.

Since most requests were made in the United States, we then focused on the United States to explore the geographic distribution

of emoji requests at the national level. As expected, states like California, Texas, and New York made a large number of requests, whereas those states lying at the heartland had low requesting percentages. We think this uneven distribution is caused mainly by the different populations in these regions. After normalizing by state population [39], the geographical distribution was relatively smooth and even across the country, which indicated that people in different states had a similar level of desire for new emojis.

5 BEHAVIORS OF NEW EMOJI REQUESTS

5.1 Context of Emoji Requests

5.1.1 Time-Related Events & Activities. During holidays and festivals, people requested time-sensitive and content-related emojis very frequently, such as the candy cane emoji on Candy Holidays, the carnation emoji on Mother's Day, and the waffle emoji on National Waffle Day. We also found emoji requests were related to popular entertainment products or events. Especially when trends of these popular elements emerged, related emojis were requested extensively by many participants such as movie audiences, music enthusiasts and game players. For example, shortly after Black Panther, a superhero film, was released in early 2018, hundreds of panther emoji requests occurred on Twitter. In addition, periodic reoccurring events like sports games might promote the expectation of new emojis. For example, during 2018 Winter Olympics, many users asked for the Olympic Rings. The yellow card (a serious warning sign in soccer) and red card (a sending-off sign) were petitioned widely in 2018 FIFA World Cup.

5.1.2 Place-Related Interests. Places of interests at different levels, like a single landmark, tourist attractions and even regions or countries, might encourage users to seek new place-related emojis. Many Twitter users visiting Paris claimed for an Eiffel Tower emoji, like "Paris first though!! why's there no Eiffel Tower emoji?!". Similarly, the Mickey and Minnie emoji was requested at Walt Disney World (WDW) Resort like "Guess where I am?!!! WDW (why is there no Mickey and Minnie emoji?!)". Note that the Unicode Consortium does not adopt emojis covered by trademarked logos or copyrighted designs [13, 17]. Residents in Hawaii and Texas looked for their state flag emojis respectively.

5.1.3 Twitter Influencer-Related Behaviors. Emoji requests made by prominent people on Twitter might trigger a widespread discussion of the requested emojis through a massive number of followers. In other words, people were more likely to interact with tweets created by Twitter influencers than those tweeted by unknown Twitter accounts. For example, Enya Umanzor, a famous YouTuber with over 800,000 subscribers to her makeup channel, tweeted "why is there no ass shaking emoji" and garnered 13,000 likes, 2,600 retweets and 34 replies. However, four non-prominent people tweeting for the same ass shaking emoji before Enya Umanzor only got three retweets, no like or reply in total.

5.2 Advocacy Behaviors

When wanted emojis were unavailable, nearly one in three Twitter users would use the symbol of '@' to mention some people or organizations for their attention. Figure 3(a) shows the top 10 Twitter accounts being mentioned by users, where eight of them

(except BTS_twt and realDonaldTrump) are specific apps or mobile operating system related Twitter accounts. It is reasonable for ordinary users to seek help from the apps or operating systems, because users thought it was these service providers' responsibility for the nonexistent emojis. The fact that people switched to Twitter to petition new emojis for other apps, e.g., WhatsApp and Discord, could be viewed as a justification for choosing tweet data to study the emoji request.
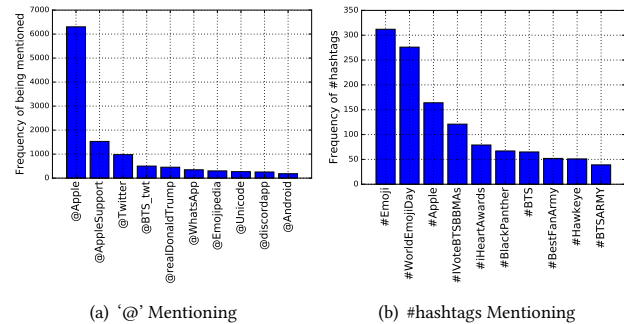
(a) '@' Mentioning   (b) #hashtags Mentioning

Figure 3: Frequency of '@' accounts and #hashtags

In addition to '@' other Twitter accounts, more than 12% of users inserted #hashtags in their tweets when wanted emojis were inaccessible. As we can see from Figure 3(b), the #Emoji is the most frequently created by users, which indicates that the primary concern of these tweets is about emojis. There are also some hashtagged words similar to those '@' mentioned words, such as the #Apple. According to Twitter, the #hashtags are mainly used for indexing keywords or topics, and popular hashtagged words are often trending topics. In our case, users attempted to advocate their desire for more emoji options with the aid of #hashtags.

5.3 Relatedness, Fairness, and Equality

An interesting scenario for the emoji request rose when users complained that there existed an emoji for A but no emoji for B in the same tweets. They thought it was unfair or unreasonable because A and B were usually very similar or related to each other. As emojis are ubiquitous in our lives, such concerns appear in diverse domains as shown in Table 2.

The gender, color, similar function and similar looking can cause a sense of unfairness and inequality. Gender equality and diversity in emojis are expected by both women and men. Women claimed for the female skier and the woman in tuxedo emojis, while men wanted male-holding-baby and pregnant man emojis. Also, the transgender flag was requested widely. The color is another factor leading to emoji inequalities. Since the blond hair, purple grape, and red ribbon emojis were released, people thought the red hair, green grape, and pink ribbon emojis should be available as well. The similar function can also be an excuse to request new emojis, like mobile phone emojis versus iPad/tablet emojis, guitar emojis versus ukulele emojis, alembic emojis versus test tube emojis, and trophy emojis versus Oscar emojis. Besides, the similarity in looks between two distinct objects can cause unfairness if one of them is

unavailable. For example, people reluctant to use the tortoise emoji to represent the turtle believed it was unfair to the turtle.

Emerging technologies, recent social movements, and the equality of political symbols motivate people to petition for new emojis. The bitcoin sign was approved in 2017 as a Unicode character, but not as an emoji. Twitter users wanted an emoji version of the bitcoin to be added. When #MeToo movement reached 1.7 million, Twitter gave it a custom emoji (three raising hands of different skin shades). However, this #MeToo emoji has not been officially supported by the Unicode Consortium and cannot be displayed across multiple platforms. In politics, the equality of both symbols and flags was considered. For example, as there existed the elephant emoji which could be used to represent GOP (the Republican Party), a donkey emoji representing the Democratic Party was requested.

Table 2: Requesting related emojis in the same tweets

Domain	Available Emoji (A)	Unavailable Emoji (B)
Human Diversity	breast-feeding 🤱	male-holding-baby
	man in tuxedo 🤵	woman in tuxedo
	blond hair 👱	red hair
Life	pancakes 🥞	waffle
	bed 🛏	pillow
	wine glass 🍷	white wine
Science & Tech	antenna bars 📶	Wi-Fi
	microscope 🔬	DNA
	mobile phone 📱	ipad/tablet
Nature	honeybee 🐝	fly
	tortoise 🐢	turtle
	crab 🦀	lobster
Business	Unicode (U+20BF) ₿	Bitcoin
	TOP arrow 🔝	bottom arrow
	bar chart 📊	pie chart
Society	#MeToo hashtag 🙌	#MeToo in Unicode
	Greenland flag 🇬🇱	transgender flag
	water pistol 🔫	real gun (AR15)
Politics	elephant for GOP 🐘	donkey for Dems.
	Guyana/Ghana flag 🇬🇾 🇬🇭	pan African flag
	United States flag 🇺🇸	Confederate flag
Entmt. & Arts	trophy 🏆	Oscar
	videocassette/DVD 📼 💿	cassette tape
	guitar 🎸	ukulele

6 SIGNIFICANCE OF NEW EMOJIS

New emojis contain both the unreleased emojis and the emojis needed to be re-designed by tech vendors, like Apple, Google, and Twitter. Identifying and introducing new emojis benefits the society a lot from many aspects, which explains why the Unicode Consortium and vendors update emojis continuously. The newly added hijab (woman with headscarf) emoji 🧕 through the Hijab Emoji Project campaign led by 15-year-old Saudi Rayouf Alhumedhi promotes inclusivity for about 550 million Muslim women on this earth [10, 37]. Researchers from the Johns Hopkins Bloomberg School of Public Health and the Bill & Melinda Gates Foundation proposed a mosquito emoji 🦟 to better explain mosquito-borne illnesses like malaria, Zika, dengue and yellow fever in 2017 [29]. Prior work also suggested creating a set of nursing emojis might facilitate health communications for patients and allow them to better understand their health data [41]. As branded emojis helped improve the amount of ads receive by almost 10% [38], brands like furniture company IKEA and fast food restaurant Tim Hortons have released app-specific branded emoji to iconize their products [1].

New appearances of emojis are always desired along with fixing design flaws, considering social influence, etc. When people found the original official lobster emoji 🦞 and the one 🦞 designed by Twitter were missing a set of legs, a new anatomically accurate lobster emoji was requested strongly and the four-legged lobster emoji 🦞 🦞 was available soon. The misplaced cheese in Google's burger emoji 🍔 sparked wide controversy online and was addressed soon by putting the cheese in its correct place 🍔 [34]. A more recent example is the Apple's bagel emoji 🥯 released in iOS 12.1 beta 2. Its lackluster appearance caused overwhelming complaints from bagel lovers and birthed the #SadBagel movement for a more appetizing design on Twitter. Apple added cream cheese to its forthcoming bagel emoji 🥯 after the social media outcry [3]. To curb visual representations of gun violence, all major vendors switched the realistic-looking pistol emoji to a toy water gun in 2018, e.g., Apple (🔫→🔫), Google (🔫→🔫), Microsoft (🔫→🔫), Facebook (🔫→🔫) and Twitter (🔫→🔫) [1, 12].

7 CONCLUDING REMARKS

In this paper, we proposed a framework for crawling and analyzing emoji requests on Twitter. We collected more than thirty million English tweets containing the keyword "emoji" throughout a year from Oct. 2017 to Oct. 2018. After filtering out bot-generated tweets, we extracted emoji descriptions using fine-tuned linguistic patterns. Surprisingly, some extant emojis were still frequently requested by many users, which were probably caused by out-of-date emoji keyboards or poor emoji keyboard layouts. For non-existing requested emojis, we categorized them into eight groups using a combination of keyword matching and WordNet-based classifiers. We then profiled temporal and geographic distributions of new emojis at different scales. Emojis requested consistently in every month and by multiple users were more likely to be approved by the Unicode Consortium. We next summarized three typical contexts of emoji requests, i.e., time-related events & activities, place-related interests, and Twitter influencer-related behaviors. Finally, we presented the equity, diversity, and fairness issues due to unreleased but expected emojis, and discussed the significance of new emojis on society. To the best of our knowledge, this paper is the first to conduct a systematic, large-scale study on new emoji requests.

REFERENCES

[1] Crystal Abidin and Joel Gn. 2018. Between art and application: Special issue on emoji epistemology. *First Monday* 23, 9 (2018).

[2] Matt Alt. 2015. *How Emoji Got to the White House.* https://www.newyorker.com/culture/culture-desk/how-emoji-got-to-the-white-house

[3] Jeremy Burge. 2018. *Apple Fixes Bagel Emoji.* https://blog.emojipedia.org/apple-fixes-bagel-emoji/

[4] Eleanor Clark and Kenji Araki. 2011. Text normalization in social media: progress, problems and applications for a pre-processing system of casual English. *Procedia-Social and Behavioral Sciences* 27 (2011), 2–11.

[5] The Unicode Consortium. 2018. *Unicode Technical Standard #51 - UNICODE EMOJI.* http://unicode.org/reports/tr51/#Emoji_Modifiers_Table

[6] The Unicode Consortium. 2018. *Unicode Technical Standard #51 - UNICODE EMOJI.* http://unicode.org/reports/tr51/#hair_components

[7] Henriette Cramer, Paloma de Juan, and Joel Tetreault. 2016. Sender-intended functions of emojis in US messaging. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services.* ACM, 504–509.

[8] Marcel Danesi. 2016. *The semiotics of emoji: The rise of visual language in the age of the internet.* Bloomsbury Publishing.

[9] Thomas Dimson. 2015. Emojineering Part 1: Machine Learning for Emoji Trends. Instagram Engineering Blog. https://instagram-engineering.com/emojineering-part-1-machine-learning-for-emoji-trendsmachine-learning-for-emoji-trends-7f5f9cb979ad.

[10] Melissa Eddy. 2016. *Muslim Teenager Proposes Emoji of Woman Wearing a Head Scarf.* https://www.nytimes.com/2016/09/15/world/europe/muslim-teenager-proposes-emoji-of-woman-wearing-a-head-scarf.html

[11] Emojipedia. 2017. *Emoji Version 5.0 List.* https://emojipedia.org/emoji-5.0/

[12] Emojipedia. 2018. *The Pistol Emoji.* https://emojipedia.org/pistol/

[13] Gabriella E. Ziccarelli Eric Goldman. 2018. *Emojis and intellectual property law.* https://www.wipo.int/wipo_magazine/en/2018/03/article_0006.html

[14] Bjarke Felbo, Alan Mislove, Anders Søgaard, Iyad Rahwan, and Sune Lehmann. 2017. Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. *arXiv preprint arXiv:1708.00524* (2017).

[15] Thomas B Fitzpatrick. 1975. Soleil et peau. *J Med Esthet* 2 (1975), 33–34.

[16] Katherine E Gallo, Marianne Swaney-Stueve, and Delores H Chambers. 2017. A focus group approach to understanding food-related emotions with children using words and emojis. *Journal of Sensory Studies* 32, 3 (2017), e12264.

[17] Eric Goldman. 2018. Emojis and the Law. *Wash. L. Rev.* 93 (2018), 1227.

[18] Kevin Gray. 2015. *Argentina's 'Facebook president' announces his cabinet with emojis.* https://splinternews.com/argentinas-facebook-president-announces-his-cabinet-wit-1793853317

[19] Robert D. Hof. 2016. *Picture This: Marketers Let Emojis Do the Talking.* http://www.nytimes.com/2016/03/07/business/media/picture-this-marketers-let-emojis-do-the-talking.html?_r=1#

[20] Tianran Hu, Han Guo, Hao Sun, Thuy-vy Thi Nguyen, and Jiebo Luo. 2017. Spice up your chat: The intentions and sentiment effects of using emoji. *arXiv preprint arXiv:1703.02860* (2017).

[21] SwiftKey Inc. 2015. *SwiftKey Emoji Report 2015.* https://www.aargauerzeitung.ch/asset_document/i/129067827/download

[22] Sara R Jaeger and Gastón Ares. 2017. Dominant meanings of facial emoji: Insights from Chinese consumers and comparison with meanings from internet resources. *Food Quality and Preference* 62 (2017), 275–283.

[23] Sara R Jaeger, Soh Min Lee, Kwang-Ok Kim, Sok L Chheang, David Jin, and Gastón Ares. 2017. Measurement of product emotions using emoji surveys: Case studies with tasted foods and beverages. *Food quality and preference* 62 (2017), 46–59.

[24] Sara R Jaeger, Leticia Vidal, Karrie Kam, and Gastón Ares. 2017. Can emoji be used as a direct method to measure emotional associations to food names? Preliminary investigations with consumers in USA and China. *Food quality and preference* 56 (2017), 38–48.

[25] Ryan Kelly and Leon Watts. 2015. Characterising the inventive appropriation of emoji as relationally meaningful in mediated close personal relationships. *Experiences of Technology Appropriation: Unanticipated Users, Usage, Circumstances, and Design* (2015).

[26] Chi Hong Leung and Winslet Ting Yan Chan. 2017. Using emoji effectively in marketing: An empirical study. *Journal of Digital & Social Media Marketing* 5, 1 (2017), 76–95.

[27] Nikola Ljubešić and Darja Fišer. 2016. A global analysis of emoji usage. In *Proceedings of the 10th Web as Corpus Workshop.* 82–89.

[28] Peter Edberg Mark Davis. 2017. *Unicode Emoji.*

[29] Jeff Chertack Marla Shaivitz. 2017. *Proposal for Mosquito Emoji.* http://www.unicode.org/L2/L2017/17268-mosquito-emoji.pdf

[30] Stanley Mathews and Seung-eun Lee. 2018. Use of Emoji as a Marketing Tool: An Exploratory Content Analysis. *Fashion, Industry and Education* 16, 1 (2018), 46–55.

[31] George A Miller. 1995. WordNet: a lexical database for English. *Commun. ACM* 38, 11 (1995), 39–41.

[32] Noa Na'aman, Hannah Provenza, and Orion Montoya. 2017. Varying linguistic purposes of emoji in (twitter) context. In *Proceedings of ACL 2017, Student Research Workshop.* 136–141.

[33] Christina Newberry. 2018. *28 Twitter Statistics All Marketers Need to Know in 2018.* https://blog.hootsuite.com/twitter-statistics/

[34] Sarah Perez. 2017. *WHEW, Google fixed the burger emoji in Android 8.1.* https://techcrunch.com/2017/11/28/whew-google-fixed-the-burger-emoji-in-android-8-1

[35] Henning Pohl, Christian Domin, and Michael Rohs. 2017. Beyond Just Text: Semantic Emoji Similarity Modeling to Support Expressive Communication. *ACM Transactions on Computer-Human Interaction (TOCHI)* 24, 1 (2017), 6.

[36] Agustin Fonts Mark Davis Rachel Been, Nicole Bleuel. 2016. *Expanding Emoji Professions: Reducing Gender Inequality.* http://unicode.org/L2/L2016/16160-emoji-professions.pdf

[37] Alexis Ohanian Jennifer 8. Lee Rayouf Alhumedhi, Aphee Messer. 2016. *The Hijab Emoji Project.* http://www.hijabemoji.org/

[38] EyeSee Research. 2016. *Branded emojis increase ads' attention by almost 10%!* http://eyesee-research.com/blog/branded-emojis-increase-ads-attention/

[39] World Population Review. 2018. *United States Population 2018.* http://worldpopulationreview.com/countries/united-states-population/

[40] Matthew Rothenberg. 2017. *Why the Emoji Recycling Symbol is Taking Over Twitter.* https://medium.com/@mroth/why-the-emoji-recycling-symbol-is-taking-over-twitter-65ad4b18b04b

[41] Diane J Skiba. 2016. Face with tears of joy is word of the year: are emoji a sign of things to come in health care? *Nursing education perspectives* 37, 1 (2016), 56–57.

[42] Mark Di Stefano. 2015. *Julie Bishop Describes Serious Diplomatic Relationships With Emoji.* https://www.buzzfeed.com/markdistefano/emoji-plomacy

[43] Channary Tauch and Eiman Kanjo. 2016. The roles of emojis in mobile phone notifications. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct.* ACM, 1560–1565.

[44] Emogi Research Team. 2015. *Emoji Report.*

[45] Emogi Research Team. 2016. *Emoji Report.*

[46] Twitter. 2018. *Introduction to Tweet JSON.* https://developer.twitter.com/en/docs/tweets/data-dictionary/overview/intro-to-tweet-json

[47] Leticia Vidal, Gastón Ares, and Sara R Jaeger. 2016. Use of emoticon and emoji in tweets for food-related emotional expression. *Food Quality and Preference* 49 (2016), 119–128.

[48] Benjamin Weissman and Darren Tanner. 2018. A strong wink between verbal and emoji-based irony: How the brain processes ironic emojis during language comprehension. *PloS one* 13, 8 (2018), e0201727.

[49] Wikipedia. 2018. *Fitzpatrick scale.* https://en.wikipedia.org/wiki/Fitzpatrick_scale

[50] Sarah Wiseman and Sandy JJ Gould. 2018. Repurposing Emoji for Personalised Communication: Why pizza slice emoji means "I love you". In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.* ACM, 152.

[51] Rui Zhou, Jasmine Hentschel, and Neha Kumar. 2017. Goodbye text, hello emoji: mobile communication on wechat in China. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.* ACM, 748–759.

[52] Yonatan Zunger. 2017. *Animal Emoji Requests on Twitter (The world wants raccoons and lobsters).* https://medium.com/@yonatanzunger/animal-emoji-requests-on-twitter-d514d5c729ec