

# Emoji: Lingua Franca or Passing Fancy?

**George F. Hurlburt**  
STEMCorp, Inc.

**Editor:**  
Jeffrey Voas, NIST;  
j.voas@ieee.org

Many express awe over the creative use of emoji. Others disdain 🤔 the perceived dissolution of proper English.

Yet others fully embrace 🙌 the free expression of emotions that emoji enable 😊, while some are infuriated 😡 by the wanton devolution of culture exemplified by

such primitive drawings. Many, however, remain indifferent. 🙄 🙋

Emoji are not new. The humble emoji, as a pictogram, a pictorial representation of an object, or an ideogram (a symbolic representation of a more abstract concept), enjoys a rather long heritage. One could argue that symbolic visualization extends back to prehistoric cave drawings.<sup>1</sup>

Legendary tribal norms, however, were mostly conveyed by aural means. Starting around 3200 BC, specially selected and educated scribes began etching Egyptian hieroglyphics into stone depicting nobility, conquests, and mysticism. Around the same time, the Sumerian cuneiform emerged as a pictographic script. It morphed over centuries to a more symbolic form of expression. Chinese calligraphy originated around 1200 BC as pictographic script. Around the same time, early pictograms predated the Aztec culture in Mesoamerica with its distinctive illustrative style of writing. In medieval times, educated monks scribed illuminated manuscripts, combining symbolic visual artistry with the written word to preserve religious history on paper. The hybrid visual rebus, also mixing emoji-like illustration with words, often in the form of visual puzzles, also enjoyed growing popularity.

Along the way, symbolic alphabets eventually enabled printing. Once in print, linear strings of symbols rapidly led to universal literacy-based education. After Gutenberg in 1450, knowledge became reproducible, portable, and essential. Printing, the very notion of linearity as reinforced by Newtonian physics, eventually led to production lines. Industrialized economies followed. Eventually, radio reopened aural space and television re-opened visual kinetics. In a short period of relative time, attention shifted from mass production to mass media starting around 1900 and culminating in the dynamic World Wide Web by 1999.

In response, post-modernism elevated consumerism to artfulness in the later 1900s. The smiley face pin became an overnight cultural icon in 1963. Around 1982, emoticons, the use of fonts to form facsimiles of human expression, became vogue. These font combinations conveyed emotion into otherwise dull texts. Influenced by Japanese graphics, Shigetaka Kurita first created the emoji in 1999. It burst quickly onto the Internet. Figure 1 loosely traces this long tail of visual language in human communications, leading to today's comic-inspired emoji.

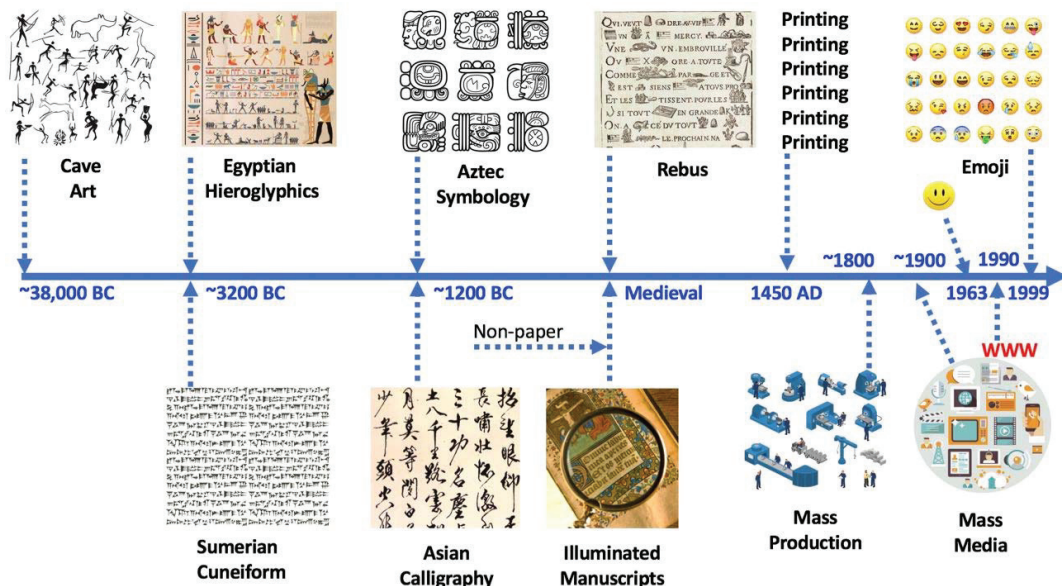


Figure 1. A conceptual timeline leading to the emergence of the emoji.

While emoji have a long heritage, they are also clearly a product of the digital age. They are now largely standardized into some 1,644 icons in the Unicode Emoji Version 11.0 released on June 5, 2018 (<https://unicode.org/emoji/charts/full-emoji-list.html>). Thus, they can be quickly produced via keystroke with no need for hand drawing. Using pictograms and ideograms, they frequently convey both thought and emotion. Emoji even follow loose syntax and grammatical rules. This suggests some degree of competence, perhaps even emoji literacy, to become a truly effective emoji communicator.<sup>2</sup> This gives rise to the question: Might emoji become the new lingua franca of the Internet?

## A NEW FORM OF EXPRESSION?

Some might agree that emoji is becoming the new universal language of Marshall McLuhan's "global village." For example, the Oxford Dictionary declared the "face with tears of joy" emoji 🤩 as its "Word of the Year" in 2015 (<https://en.oxforddictionaries.com/word-of-the-year/word-of-the-year-2015>). This is in recognition of the widespread global acceptance of the emoji as a popular means of expressing ideas and sentiment in an otherwise dry world of emotionless technocratic prose. The fact that a robust Unicode standard exists for emoji further reinforces a sense of universality. The need for maximum compression in Tweets, social media, text messages, and other digital media strongly encourages an economy of characters needed to express basic concepts. Whereas alphabets provide a finite set of characters to express any idea, many characters must be combined to do so. Emoji, at 144 pixels and 18 bytes, easily replace costly words with far greater economy.

RU 😊?

Advertisers, quick to pick up on trends, regularly target Internet users with hip emoji messages. The level of monetization even extends to the service economy where employees are encouraged to quite literally present a smiley face to their clients, much less to cope emotionally in an otherwise insensitive world.<sup>3</sup> Emoji appear to have "staying power" as an enduring visual code. Below is a list of a number of useful emoji-related websites.

- Unicode Emoji Standard V 11.0: <https://unicode.org/emoji/charts/full-emoji-list.html>
- Real-Time Twitter Emoji Usage Tracker: <http://emojitracker.com>

- Real-Time IOS Emoji Usage Tracker: <http://www.emojistats.org>
- Emoji Encyclopedia: <https://emojipedia.org>
- MIT NLP & AI-Based Sentiment Analysis: <https://deepmoji.mit.edu>
- Popular Emoji Grams: <https://emojisaurus.com>
- Personalized Emojis: <https://www.bitmoji.com>
- Moby Dick in Emoji: <http://www.emojidick.com>
- Worldwide Use of Emoji: <http://nlp.ffzg.hr/data/emoji-atlas>
- Emoji Statistics: <https://worldemojiday.com/statistics>

Intended meanings of many emoji, however, can too easily be misconstrued. While the Unicode standard defines “core emoji,” many more (less well-defined) emoji continue to emerge daily worldwide. Soon a set of scientific emoji are poised to appear. This leads to a bit of a tower of Babel situation, as emoji are often culturally or contextually dependent. In fact, cultures with varying economic descriptors as defined by the Hofstede Culture Index are liable to use emoji differently to describe their particular relationship to the world. For example, people from countries with high uncertainty-avoidance scores tend to disfavor emoji that express positive emotion.<sup>4</sup> Moreover, the same emoji might carry different meanings as determined by the culture where it is being used.

While the Unicode standard for emoji tends to reinforce meaning, there are at least 17 different proprietary platform-based fonts in place that significantly render the same Unicode emoji differently. The Unicode site (<https://unicode.org/emoji/charts/full-emoji-list.html>) shows 11 different platform-based renderings of standard emoji. Thus, a given standard Unicode emoji can appear differently on iOS than it does on an Android device. This leads to statistically different interpretations of both sentiment and meaning when specific standard emoji codes cross platforms. Nonetheless, variation in interpretation also occurs within the same platform, although to a lesser degree.<sup>5</sup>

The rather generalized lack of commonality in emoji interpretation suggests that emoji are actually less than a universal form of expression. As noted, cultural influences, context, and symbolic variation can potentially compromise intended meaning. Worse, it would appear that emoji are less than a complete form of expression.

Standardized emoji codes do not really exist for personal pronouns or most intransitive verbs. This limits the expressiveness of the language, while simultaneously opening the door for creativity in usage among various user cliques. It is the case, however, that volunteers using Amazon’s Mechanical Turk encoded the entire text of Melville’s *Moby Dick* into a book entitled *Emoji Dick*. *Moby Dick*’s iconic first sentence, “Call me Ishmael,” was emoji encoded as follows:



Figure 2. The first sentence of *Moby Dick* in emoji.

While clearly a period novel, the use of a telephone (a nonexistent item in the time of the novel) induces a form of contextual irony. Likewise, *Alice in Wonderland*, a rebus-friendly text by the intent of author Lewis Carroll, has also been translated fully into emoji. In both cases, however, the level of effort necessary to successfully navigate these annotated texts exceeds the ability of most readers. Emoji datasets, while highly creative, become highly subjective, induce repetition, and become exceedingly difficult to contextualize.<sup>6</sup> In other cases, multiple emoji must be creatively combined to suggest common items. For example, “sweetheart” might be written as a piece of candy next to a heart, hardly a literal translation.

The same emoji might carry different meanings as determined by the culture where it is being used.



Ultimately, emoji are technically oriented. As such, they are driven by advancing technology. Thus, as natural language processing (NLP) and artificial intelligence (AI) join forces to reinforce the effectiveness of vocal interaction, emoji might give way to vocalized inflections. Moreover, the number of bot-generated emoji could potentially overpower human users, much like spam often overwhelms the inbox. Both trends could signal a setback for emoji advocates.

The notion of emoji as an emergent universal language seems to be limited at best. The use of emoji as a hybrid form of expression to augment regular text, however, appears to be a strong and growing possibility in a world that increasingly demands symbolic economy and some level of personalization. Together with otherwise impersonal texts, selective use of emoji sets the tone for satisfying communication. Emoji tend to defuse what otherwise might be considered offensive messages with a friendly salutation, closing, or strategically placed emoji intended to add a more conciliatory tone.

## EMOJI IN A NETWORK AGE

Emoji represent a network phenomenon. An analysis of an early August 2018 snapshot of the frequency of emoji usage on Twitter using the website <http://emojitracker.com> reveals a clear power curve relationship. Figure 3 shows this plot in the form of a vertical bar graph.

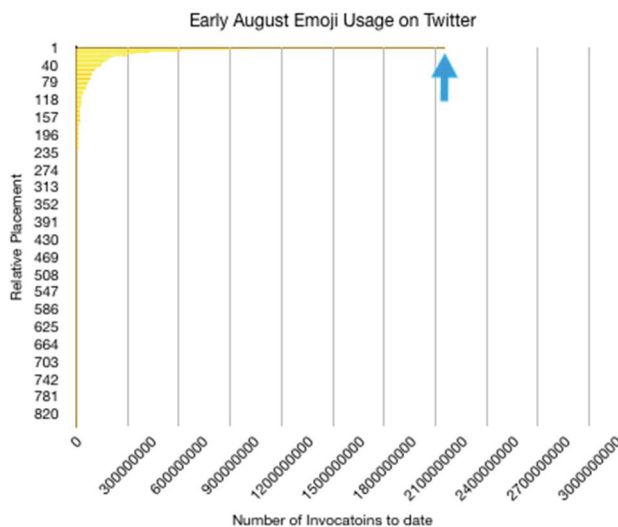


Figure 3. Distribution of 846 popular emoji on Twitter in early August 2018.

In this figure, the emoji occupying the top position was the familiar “face with tears of joy.” This emoji was invoked 2,145,510,490 times. The emoji at the last-used position, number 846, was called only 132,848 times. It was an emoji for uppercase Latin letters. The top 10 emoji were: face with tears of joy, a single heart, the recycling symbol, face with hearts for eyes, a slimmer single heart, a sad crying face, a simple happy face, a face with a furrowed brow and frown, a double heart, and a kissing face (see Figure 4).



Figure 4. Top 10 emoji on Twitter in August 2018.

It is interesting to note that the majority of the popular emoji are positive in nature, which is in keeping with most research on the use of emoji. Other research shows that applied network science techniques outperform state-of-the-art methods, including NLP for sentiment analysis.<sup>7</sup>

As noted, printing introduced a prevalent linear relationship that helped usher in an industrial age, enhancing the world's economy. The advent of mass media, especially the Internet, awoke other sensitivities. The rise of the emoji as a popular means of visual expression suggests a return to age-honored visual space. Moreover, despite distinct cultural differences in usage, the world-wide emoji acceptance is itself significant. It represents a broad-based trend toward the reality of networked global sharing. Steeped in older linear technology models, many people fail to appreciate or perhaps even fear such openness. To some, emoji represent nothing short of a tragic fallback to primitive behaviors. Further social research along these attitudinal lines might better help further delineate growing protectionist movements in many nations.

As industrialization engaged, literacy-focused education became indispensable. Now formal education increasingly seeks creative online outlets, and traditional literacy-based instruction seems somehow outdated. Yet computer literacy continually gains credence. Importantly, the growing cost of formal higher education leaves many indebted well beyond any entry-level thresholds. Perhaps it is time to acknowledge the shift from book-borne portable personal knowledge to online networked general knowledge. Such a shift likely has a profound effect on future educational strategies. Here, new forms of digital literacy become prerequisite for future opportunity. The increased and sustained use of emoji might suggest new innovative research initiatives to help identify new educational vectors, perhaps even extending to mathematics.<sup>8</sup>

Finally, visualization is endemic. For example, most nations regulate driving behavior by varying shapes and color cues. Emoji only represents one form of the resurgence of visualization in the digital world. As a case in point, augmented reality and virtual reality are opening new perceptual doors. Graphical representation of data is also increasingly pressing. Networks of all types frequently involve large sparse matrices. The ability to visualize these diverse datasets becomes an increasingly critical skill. Conceptualizing and constructing such graphs require new mathematical insights and new means of depicting their hidden realities accurately and convincingly. More importantly, the ability to evaluate and interpret such visual representations on their merit is equally important for an informed citizenry.

To this end, the ability to acquire visual literacy, including the use of emoji, becomes an increasingly important skill—not only for dedicated data scientists, but across virtually all the increasingly entwined domains of human knowledge.

The increased and sustained use of emoji might suggest new innovative research initiatives to help identify new educational vectors, perhaps even extending to mathematics.

## REFERENCES

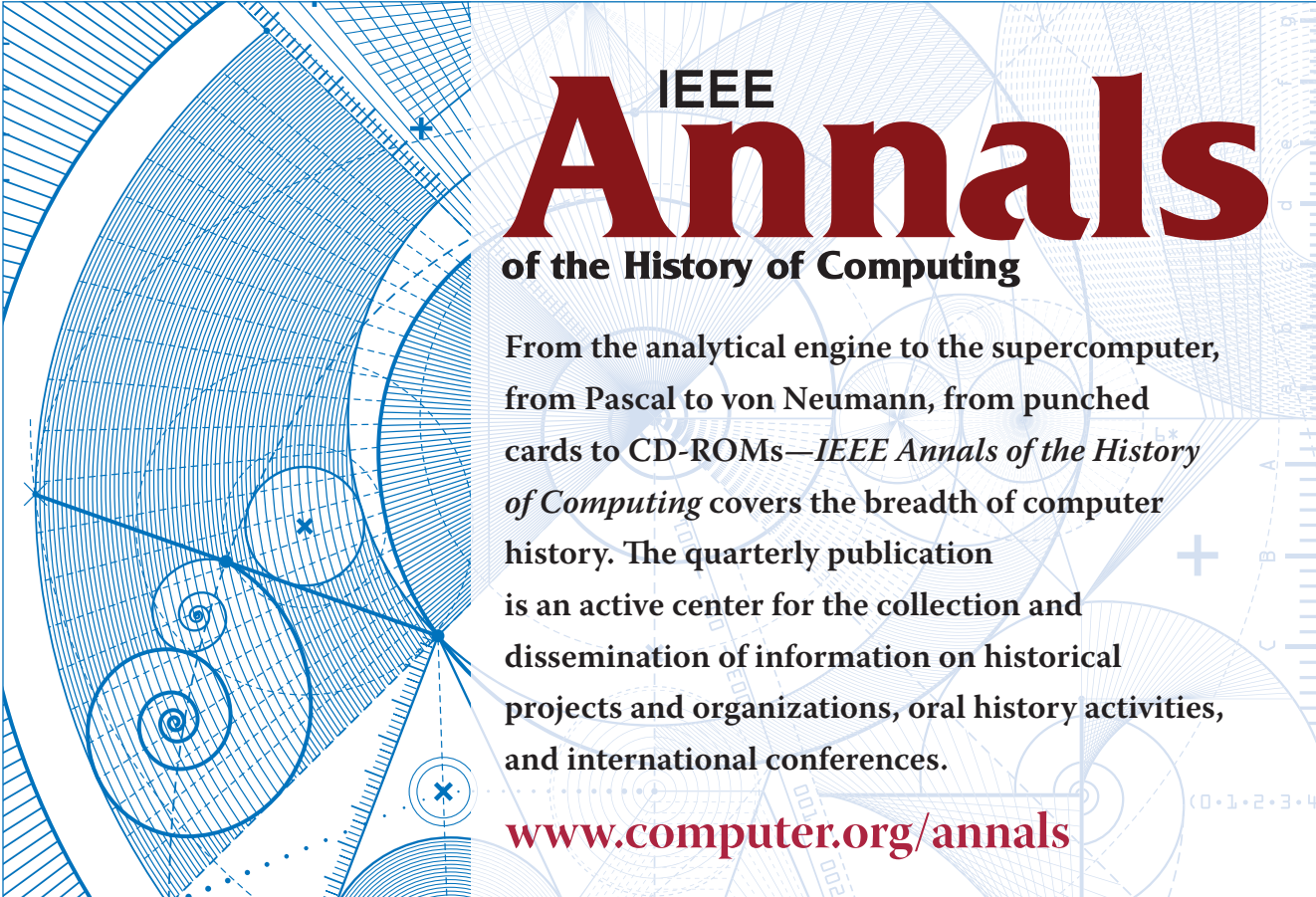
1. G. Hurlburt and J. Voas, "Storytelling: From Cave Art to Digital Media," *IT Professional*, vol. 13, no. 7, 2011, pp. 4–7.
2. M. Dansi, *The Semiotics of Emoji*, Bloomsbury Academic, 2017.
3. L. Stark and K. Crawford, "The Conservatism of Emoji: Work, Affect and Communication," *Social Media + Society*, vol. 1, no. 2, 2015; doi.org/10.1177/2056305115604853.
4. X. Lu et al., "Learning from the Ubiquitous Language: an Empirical Analysis of Emoji Usage of Smartphone Users," *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*, 2015, pp. 770–780.

5. H. Miller, J. Thebault, and I. Johnson, "Blissfully Happy' or 'Ready to Fight': Varying Interpretations of Emoji," *10th International Conference on Web and Social Media (ICWSM)*, 2016, pp. 259–268.
6. W. Radford et al., "Call me Ishmael!: How do you Translate Emoji?," *Proceedings of Australasian Language Technology Association Workshop*, 2016, pp. 150–154.
7. A. Illendula and R. Yedulla, "Learning Emoji Embedding using Emoji Co-occurrence Network Graph," *International Workshop on Emoji Understanding and Applications in Social Media*, pending publication, 2018; <https://arxiv.org/abs/1806.07785>.
8. T. McCafery and P.G. Mathews, "An Emoji Is Worth a Thousand Variables," *The Mathematics Teacher*, vol. 111, no. 2, October 2017, pp. 96–102.

## ABOUT THE AUTHOR

**George Hurlburt** is chief scientist at STEMCorp, a nonprofit that works to further economic development via adoption of network science and to advance autonomous technologies as useful tools for human use. He is engaged in dynamic, graph-based IoT architecture. Hurlburt is on the editorial board of *IT Professional* and is a member of the board of governors of the Southern Maryland Higher Education Center. Contact him at [ghurlburt@change-index.com](mailto:ghurlburt@change-index.com).

*This article originally appeared in  
IT Professional, vol. 20, no. 5, 2018.*



**IEEE**  
**Annals**  
**of the History of Computing**

From the analytical engine to the supercomputer, from Pascal to von Neumann, from punched cards to CD-ROMs—*IEEE Annals of the History of Computing* covers the breadth of computer history. The quarterly publication is an active center for the collection and dissemination of information on historical projects and organizations, oral history activities, and international conferences.

[www.computer.org/annals](http://www.computer.org/annals)