

From Recommendations to Generation: Tools Without Interfaces

author: Lev Manovich

year: 2024

I realized finally why I love using GenAI tools for image making so much. Not because they are so fast, so efficient, so versatile. The reason is different.

For years already, my everyday culture and information intake was structured by AI systems. When I use YouTube, I let it recommend me videos. I almost never search for something myself---instead, I just refresh its home page (by clicking on the YouTube logo) to see what new videos it curates for me. I do this probably hundreds of times per day. The same goes for my news readings. Yes, probably once a week I may look at the New York Times, Financial Times, and Le Monde---but what I use everyday is the Google Chrome app on my phone and iPad, which automatically curates news for me.

Using AI-media generation tools (like Midjourney) fits perfectly into this paradigm. I tell the image generation tool what I like---and it “curates” a particular area of the infinite latent space for me, generating endless images. Just as with my culture and news consumption, I enjoy that I don’t have to make all choices, specify every detail, decide everything. And this is probably the key reason why I now prefer AI-image generation apps over my pens, brushes, drawing software like Procreate, or my professional photo camera. I no longer have to choose every line, every brushstroke, every time, every color, which objects to draw, what to photograph, what focus and exposure to set, and so on.

To summarize: based on my own experience, I feel that the generative AI-based creation process is similar to AI curating (e.g., “recommendation systems”) that many of us have been relying on for many years.

It’s the same paradigm. Moving from “deterministic culture” where you have complete control but have to make endless, and therefore often very tiring, choices to a new “probabilistic culture” (or “predictive culture”).

The emergence and growth of this “probabilistic culture” is logical. It’s a response to the information and media explosion of the 1990s and early 2000s---the phenomenal growth of available content of on the web (1993--). Algorithmic recommendation technologies emerged at the turn of the 21st century (Amazon was first already in 1996) as an alternative to search technology that made you specify what you wanted to find. With search, you have more control, actively declaring what you want to find, even if the results remain unpredictable.

Twenty years later, we arrived at the new stage of probabilistic culture---the emergence of “probabilistic creativity” tools for the masses. They don’t require learning math, programming languages, specialized art coding systems like Processing, or mastering complex interfaces like of professional software---Max/MSP/Jitter, PD, Avid, Ableton, After Effects, Premiere, Blender, and all systems that emerged in the PC era. Instead, you just direct GenAI tools using natural language. In other words---GenAI is probably the first media creation tool that does not have any dedicated interface! (Yes, you can also sing with your voice, make dance with your body, and press your hand to the cave wall to make original art from 100,000 years ago, but at least in the modern era, art required using specialized tools and years of training on how to use them properly.)

In the 1950s and 1960s, computers and algorithms gradually began to be adopted in the arts. First, computer music was composed on the Illiac computer in 1956; the first exhibitions of “computer graphics” and “computer-generated pictures” took place in 1965. Throughout the 1970s and 1990s, computer art remained largely confined to research labs and specialized studios equipped with expensive workstations. In the 2000s, new tools such as Processing language popularized algorithmic art and design, but creating this work still required specialized programming skills it still remained a specialized realm.

Only after 2022 with the new generation of “tools without interfaces” did “predictive creativity” become available to the masses. According to statistics by the SimilarWeb service that I checked today, a popular AI image generation app called leonardo.ai gets 12.7 monthly visits! And this is only one of many dozens of similar web applications for image generation---which is also available in all AI bots like ChatGPT, Google Gemini, Meta AI, claude.ai, and so on. (According to SimilarWeb, ChatGPT currently receives 500 million visits every month).

Obviously, the analogy between recommendations and generation does not apply in every respect. But I think on a fundamental level, the two are similar. Through different mechanisms, you tell (directly or indirectly) the system what you like. And it starts to either select from and recommend existing content---or generate new content. Another key similarity between these two cases is that the possibilities are practically endless. The number of images, songs, texts, etc. that you can generate with AI is infinite. And in the case of recommendations for online content (e.g., videos, songs, photos, and social media posts) it's practically infinite too. In both cases, we are dealing with new phenomenon at a super-human scale. Recommendation systems (along with search engines) emerge as one mechanism to deal with it. Generative AI applies the same logic to creation. For me, the new scale of digital culture became fully visible around 2003-2005, which led me to propose the concept of Cultural Analytics - analyzing culture through big data - in Fall 2005.

(For me, the new scale of media first became fully visible around 2003-2005. My Cultural Analytics idea which was provoked by it dates from Fall 2005.)

One difference between recommendation and generation is that recommendations stay within the same media type, the former tends to be one media to the same media type systems. You listen to songs and get recommended more songs; you explore book titles and get recommendations for more book titles. But at present, the dominant way of guiding AI generation systems, regardless of the media being generated, is to describe what you want with words.

But why does asking AI to generate an image, song, video, etc. from text work so well? For example, we know that in theory we can't describe every aspect and detail of a picture using language. It works in GenAI because... you are NOT describing every detail of the media you want generated! You give directions, and AI draws on a giant library of media content used to train it: compositions, color schemes, templates, real-world knowledge encoded in human photos, paintings, films, etc.

GenAI can be thought of as “post-modern” in this respect---or “post-production.” But it goes beyond this. A post-modern artist still had to assemble the new work herself, selecting and putting in all quotes and references manually. With GenAI, you are not making every detail yourself. You tell the AI model to select, remix, interpolate, and filter from its giant database of content, patterns, techniques, styles, and aesthetics. word knowledge (e.g., “in the evening, colors look this way” and “in the morning, they look that way”). And this is why commanding AI using language works so well. This is also why generative AI is only possible after the emergence of the web and its giant external digital culture archive. It's a new mechanism of accessing and remixing this vast cultural database. It's a new mechanism of using it.

In the 1990s, we already got 1) search, 2) recommendations, 3) remix. Later came republishing and commenting (via social media), and now---generation. In other words, while generative media can be considered a part of AI history that predates the web (1956--) it is perhaps more accurate to think of it in a very different way. It's actually just a new type of software for using the web (and its giant cultural archive)!

Perplexity.ai, an AI question-answering service, is a good example of how post-web 'AI' is different from classical AI. It compiles, reformats, and structures information from a number of web pages. These functions were not, of course, part of classical AI. It is also a mistake to focus on the idea of “AI creativity.” Today “AI” is not about being “creative” in the original Romantic meaning of this concept.

It's a new way to use the web, to take advantage of it. Generative AI is a new way to use the web, to take advantage of it.

And it is a completely logical step in digital culture evolution. Search engines that until now were our dominant mechanism of interacting with the web do not synthesize the most relevant information. They only try to select what already exists. This is why you often just click on the predictable result they typically feature on their first page---a link to a relevant Wikipedia article, where humans already synthesized many information sources. The behavior of search engines is logical because they were based on the ideas and information culture of the 1950s and 1960s, such as “information retrieval.” So, a search engine simply tries to “retrieve” the most relevant items from the endless web universe. Today, thirty years after the web started growing (1993--), it’s logical to expect new kind of software that starts synthesizing rather than simply retrieving. And this is what Generative AI is now doing.

In fact, we may wonder why it took so long.