Critical Humanities

Volume 2
Issue 2 *Teaching in the Time of ChatGPT and other forms of AI*

Article 10

January 2024

"The Hard Work of Programming Germinates Soft Pleasures": Creating Synthetic Comics with AI Collaboration

Barbara Postema
University of Groningen

Ilan Manouach
Metalab, Harvard

Recommended Citation

Postema, Barbara, and Ilan Manouach. ""The Hard Work of Programming Germinates Soft Pleasures": Creating Synthetic Comics with AI Collaboration." *Critical Humanities* 2, 2 (2024). DOI: https://doi.org/10.33470/2836-3140.1053

"The Hard Work of Programming Germinates Soft Pleasures": Creating Synthetic Comics with AI Collaboration

A Discussion between Barbara Postema and Ilan Manouach

Barbara Postema
University of Groningen

Ilan Manouach
FNRS / Université de Liège, metaLAB@Harvard

Barbara Postema is a lecturer at the University of Groningen. She is known for publications on wordless comics, Canadian and American comics (including a roundtable in Seth's *Clyde Fans* for *The Comics Journal),* and on form and narratology. Her work has been recognized by associations in the US, Brazil, New Zealand, and Germany, leading to invitations to present keynote lectures and to the translation of her monograph, *Narrative Structure in Comics*, into Brazilian Portuguese. She was part of the research group "History in Comics" based out of Palacky University in Czechia in 2021-22. Barbara is a former president of the Canadian Society of the Study of Comics and a founding member of the Comics Studies Society. She is a co-editor of the Wilfrid Laurier UP book series *Crossing Lines* and regularly contributes peer reviews on manuscripts and journal articles for a variety of publishers and journals.

Ilan Manouach is a comics scholar and comics creator with an international career in creating and studying conceptual comics. In his introduction to a scholarly collection engaging Manouach's comics, Pedro Moura describes him as "very critical and discriminating towards the whole field at the same time that he shows an acute understanding of, and even heartfelt appreciation for, the comics medium and its history."[1] Manouach is an ambassador of sorts for reconceptions of what the comics form can be and do, perhaps best known for *Shapereader*, a tactile narrative system first developed in 2013 for people with visual impairments. In his processes of creating altered, redrawn, or collaged comics, Manouach has long since been using digital tools. However, in the last few years, he has been using machine learning to train computer programs to take on increasing shares in the work of creating comics, eventually launching the AI-generated *The Neural Yorker* on Twitter, posting new single-panel comics daily (since 2020), and publishing the "first synthetic comic book" *Fastwalkers* in 2021 (published by Echo Chamber). More information can be found here: https://ilanmanouach.com/.

[1] Moura, *Ilan Manouach*, 3

Barbara Postema: The description of *Fastwalkers* on your website says that for this book, "all text and images have been produced by machine learning," and that "*Fastwalkers* is an amalgamation of different community datasets, proprietary algorithms, indexing regimes, beta testing, and generative models, all trained on millions of data units and text bodies specifically to make this book." (From https://ilan-manouach.com/work/fastwalkers). Could you explain more about this process? For example, what kinds of datasets were used; and did AI only produce texts and images, or was it also involved in the other choices that go into making comics: panel breakdowns and page layouts, text balloons choices and placements and so on? In other words, where is the machine hand, and where is the human hand in this work?

Ilan Manouach: For the purposes of training the generative models for the production of *Fastwalkers,* we used Danbooru2020, a large-scale crowdsourced and tagged anime illustration dataset containing 4.23M media annotated by a large community of users with a total of 130M tags. The Danbooru dataset has been developed with a clear purpose for deep learning applications. Besides the fact that each image contains on average 25 tags, allowing users to form very fine-tuned collections of images belonging to the same category for different training purposes, the dataset, that is updated every year with new material and metadata, also comes in different sizes in order to accommodate various processing capabilities. Gwern, one of the main researchers and artists behind the Danbooru experiment has been a great inspiration for my work. I understood quickly that managing a dataset of such proportions can be daunting. The very high precision and granularity of the dataset, a result of the multiple annotations, features a long tail of image categories and tags that are only present in a few hundreds of images. This wouldn't necessarily be the most productive thing for training the generative models we needed for *Fastwalkers,* as each conditional model required a large amount of media. I decided along with my collaborator Yannis Siglidis to explore Danbooru's thousand most populated tags in order to directly address the largest distributions and to understand better what was the material in hand. This decision was certainly important as the capabilities for model conditionality based on the StyleGan architecture we were using, were limited and very difficult to program. By subsetting and regrouping a few categories from Danbooru, we were able to minimize the risks of model underfitting, train domain-specific models on the different elements we were interested in, and also conceptually define the visual building blocks for *Fastwalkers*. The model training was also extremely resourceful and we would have never been able to do it without benefit from a generous Nvidia grant consisting of computational credits. We were able to put them in good use by occupying an A100 cluster for almost 4 months. For generations, we used our in-house instances in order to avoid large file

transfers. The selection from the generated output was entirely human-curated and was a long process of parsing literally millions of image data with the help of my assistant, Luca Reverdit. The same goes for the text: I had the chance to be one of the early testers for OpenAI and decided to use the conclusion of my doctoral thesis as an initial prompt. This started a two-month co-creation with the GPT-3 model that led to a final text; a hallucination, a certain a manifesto on "metacomics" (a term that was not in my thesis), forecasting that the production of speculative comics will be sourced by the aggregation of data flows, cosmic dust and bodily fluids. Regarding the more formal elements of comics: While we have developed a solution that would automatically generate different layout compositions, we preferred to manually compose a set of basic and dense layout grids unfolding over 512 pages. The same goes for balloon generation: in that specific problem we created, using data augmentation, a vast synthetic dataset consisting of 100k balloons, that we decided to keep for a future application. The entire process was one of co-creation and not full automation, which is for instance the case for the *Neural Yorker,* our cartoon generation on Twitter.

BP: You contrast the co-creation of *Fastwalkers*, in the creative process of which numerous humans were still involved (from the crowdsourcing of Danbooru, to various collaborators, assistants, and technicians), to the "full automation" of *The Neural Yorker.* Could you explain the creative process for that project a bit more? Presumably there is no dataset for *New Yorker* cartoons, like the Danbooru dataset offered, though those cartoons do have a very recognizable style and format, of course. And does the automation extend to posting the cartoons to Twitter (now X), or do you check them before they go up?

IM: Despite the multiple technical components that need upgrading, maintenance and retraining, the production for *The Neural Yorker* can be more closely described as a fully automated content production. *The Neural Yorker* is a cartoon generator on Twitter, now X, in the tradition of the famous North American periodical. Its content is automatically generated by multiple AI conditional models that have been trained on hundreds of thousands of cartoons, and punchlines collected from a multitude of online repositories and databases with their own systems of classification and labeling. It is deployed by an on-premises server and consists of a sequence of different operations: The service "wakes up" every day at 9am and crawls a list of thousands online newspapers, periodicals and blogs that have been sorted in advance according to the media sector (social news, financial press, sports, etc.), their geographic focus and interest (regional, national and international), their territorial coverage (EU, South East Asia), and their general positioning in a simplified binary political spectrum, (progressive or conservative)- we hope in the future to be able to have a more granular classification of political affiliation, but

this task goes way beyond our existing capacities. Then, the next process performs a subtask of information extraction on the collected headlines called Named Entities Recognition (NER), and classifies the different objects into predefined categories such as celebrities, fashion items, etc. Here we would also like to be able to expand the range of our classifiers in order to allow for a more fine-tuned description of the different objects. A couple of hours later, using these newly acquired tokens (the Tories, cricket, Boris Johnson, or whatever happens to be in the news that day), the next process generates a large amount of cartoons, based on the collected headlines and our generative conditional models. A penultimate process is busy assigning a score to each one of them, according to a humor-detection algorithm entirely informed by heuristics- and whose actual performance is quite difficult to assess, and at the end a final deployment process it selects the most funny one and uploads it on twitter. The operation might seem to be of unprecedented proportions for the production of press cartoons, but to paraphrase Raymond Devos, "humor is a very serious matter and should never be entrusted to buffoons!" There are also a couple of other parallel operations involving different activations and user interactions on Twitter, but these are not entirely relevant to synthetic comics.

BP: Both of these synthetic comics display elements with uncanny qualities. For *Fastwalkers*, this is not just in the images and the texts, but also the placement of text balloons, which tend to overlap multiple panels and have tails that point in directions that are unintuitive, in terms of reading direction. Some even point to the gutter, rather than a figure in a panel. At other times, panels cross the middle gutter of the book, again a counter-intuitive placement for panels. Did the AI come up with such placements, or did you and the team play up these kinds of uncanny effects?

IM: *Fastwalkers* was not for me an opportunity to entirely relinquish control on a fully automated system of comics production. Whatever happens in the book is always the result of a decision I took with my collaborators based on a constant interaction with the production system we put in place. The book is certainly not a conventional comic book with a storyline, narrative arcs, an identifiable character and so on. These were never, in my sense, essential or structural elements of the comics medium, but just "sedimented habits". *Fastwalkers* is not a book that should be read from cover to cover. It is more an invitation for an alternative mode of access to the text, where readers trace their own path in a space of very high information density, not unlike machine learning models that explore a dataset's latent space. As for the more specific question related to the balloons, it was a necessary decision to point them to different places; the lack of constant and identifiable subject would make any requirements for interpellation unnecessary. The balloons were merely used as textual containers in the most bland way. While I have experimented in

the past with the deconstruction of comics, such as in the book *Compendium of Francobelgian Comics,* by 'freeing' the formal elements from the requirements of storytelling and trying to see what are the stories these elements might have to tell in themselves, there is a long history of experimentation in comics that *Fastwalkers* consciously draws from, such the Brazilian *poema processo* and its use of speech bubbles where the comics elements are used as floating signifiers, rather than fulfilling specific medial functions.

BP: You mention (the expectation of) storylines and identifiable characters as "sedimented habits" in comics. This applies to both comics creation and comics reading, and so a project like *Fastwalkers* will force readers to change their reading habits as well, browsing the text rather than following a set path. How do you see this working with *The Neural Yorker*? Single panel cartoons are always "browsed" in a sense. Besides the intervention in the process of creation for *The Neural Yorker*, what intervention do you feel you're making in the reading process?

IM: What I immensely enjoy with the *Neural Yorker*, is that it positions the reader in a specific suspicious mode of reading. In this context, the reader assumes a role reminiscent of that of a detective meticulously poring over textual and graphic evidence, all in pursuit of establishing deliberate causality within the material. This investigative process involves a constant effort to unveil the shortcomings of machine-based comprehension. *The Neural Yorker* cartoon is constantly under heavy scrutiny, standing accused of either faking the proceedings of multilayered irony or indulging in absurdist non sequitur. Its driving force is an unwavering determination to expose humor uniquely as a figure of dislocation, a machinic version of Lautreamont's "chance meeting on a dissecting-table of a sewing-machine and an umbrella". Within this domain, the doubting of authorial intentionality becomes a central theme of synthetic comics, underscoring the reader's immersion in a mindset characterized by hesitation, skepticism and distrust. Invariably, the reader of *The Neural Yorker* finds themselves ensnared in an endless loop of questioning, delving deeper and deeper into the layers of multimodal articulation, searching for the repressed implications laying beyond the lacquer of dada-like quid pro quo. Consequently, *The Neural Yorker*'s effect is to pathologize the very act of critical interpretation itself, portraying it as an unrelenting pursuit of concealed motives and hidden agendas. The reader becomes entrapped in a relentless quest for uncovering the subtext, forever caught in the intricate web of analysis and doubt.

BP: The @NeuralYorker Twitter account has over 2,600 followers, which I would guess is more than there were copies printed for *Fastwalkers*. Are these synthetic cartoons your most-viewed, most-engaged comics? And what do you hope those followers take from their encounters with these works?

IM: I won't casually monitor our Twitter account, although we have clear data about its interactions through our developer dashboard. Since X changed its internal algorithms, accounts with no clear interactions such as comments and retweets are dropping in visibility and we have to develop new strategies to enhance *The Neural Yorker*'s ranking. The project launched in 2020 and has since received some media attention and maintains a small but constant following. My guess is that its subscribers need to have in their daily twitter feed, along with other quirky meme generators and "shitposting bots", a healthy dose of absurdity and open-ended humor.

BP: One description of *Fastwalkers* mentions that the "models were trained on anime and hentai imagery". Why did you choose to use this overtly erotic, even pornographic genre of manga? Was this "shock value," or would this somehow pose a particular challenge for AI? Did you also choose specific anime genres? The cover image evokes *Akira*, and both text and imagery in the book use science fiction tropes. Was this genre chosen to link with the futurism of artificial intelligence?

IM: Many researchers might argue that pornography is the driving force behind an array of technological advancements with far ranging applications in all strands of our daily life, from Machine Learning, AR and smart wearables, to the development of search engine optimization and the refinement of UI design craft. As far as I am concerned, there is nothing shocking about pornography and any "shock value" deriving from sexual representations touches a very specific readership (hopefully not mine). The speculative nature of the text, which you describe as a science fiction trope, comes from the very nature of my initial prompting. I used the conclusion of my doctoral thesis, a speculation on the necessity to evolve comics beyond terrestrial attachments and earthly infrastructures, as a starting point of a two-month dialogue with GPT3, on the production of a metacomics manifesto that runs on the entire book.

BP: Some of the most standard qualities of anime and manga–identifiable characters and ongoing narratives–are absent in this book. The volume is also atypical for printed manga, which tends to come in easy-to-carry formats. Instead, this book is over 500 pages long and in a very large format. All these features make the book difficult to read, physically and mentally. Was this intentional, and if so, what was your thinking behind that?

IM: My goal was never to produce a genre-conforming object, neither in *Fastwalkers* nor elsewhere.

The format of the book needed to account for the mass of data that was used for training purposes as well as the data that was discarded from our selection. It is especially a product of media accumulation, and I thought that any given format, such as the very precious 48-page

Francobelgian BD or the pocket-size manga, would not rightly serve. Formats come with very specific expectations and reading modalities. The small format serialized manga publication has for instance a very precise temporality folded in it, one of *incompletion*: the reading experience is starting with a full realization that any sort of *denouement* will only come much later, and through many other books––an expectation that the best mangakas have been very good at deceiving. On the contrary, *Fastwalkers* needed to embrace a non-recognizable, a non-standardized form, in order to deflect any specific reading expectations about duration and temporalities. It needed to invent its own format, it needed to look like itself and to not point to any specific reference in terms of industrial fabrication that would run the risk of overdetermining its reception. As for the difficulty of the reading experience, I would assume that a linear approach to mainstream comics and graphic novels would definitely not work. *Fastwalkers* is a book to be skimmed, not necessarily to be read, at best to be read in a discontinuous way jumping from page to page, with a very loose idea of closure.

BP: Much of your work is a commentary on the comics industry. In your comics, you have been slowly stripping away the creative–or even the human–hand from comics: you have used existent comics or partial images for altered works and collage comics (such as *Noirs* and *Compendium of Franco-Belgian Comics*); you have had people (amateurs, non-comics creators) reproduce *Peanuts* strips (in *Peanuts minus Schulz*), you have even published a completely blank book (*Blanco*). Now you have been experimenting with AI, in *The Neural Yorker* and *Fastwalkers*. Is there an indictment of the comics industry in that progression?

IM: In retrospect, the publications you are referring to constitute a belated response to the unprecedented media accumulation and increasing easiness of access to comics media that I had to reckon with during my formative experience as a comics artist in the beginning of the 2010s. I should mention that the early networked affordances such as the online archives, digital libraries, torrents trackers and file sharing networks as rudimentary as they were, were nevertheless a catalyst for an aspiring comics artist growing up in Athens, with limited access to comics media and relatively far from comics communities, in no way comparable to my Belgian colleagues for whom comics has always been a well-documented form of vernacular expression and where access to comics books and prints was granted both by public libraries and specialized bookstores. The media philosopher Matthew Fuller writes that digital abundance, articulated through a variety of older and new distribution channels (and here the importance of online communities can never be overstated) is pushing against the limitations of our current conceptual frameworks; it challenges our preconceptions about what defines a work and a professional expertise by overwhelming the traditional media capacities. The largely unmediated access to entire

comics collections online, first and foremost challenged my own capacities as a reader, and then accustomed me from early on to understand information management as entirely convergent with my artistic projects. Navigating media collections, from *Compendium*'s Francobelgian comics, to *OnePiece* and Schulz's entire corpus was a way for me to respond, in my capacity as a reader, to the following question: how can comics account for the expansion of our relation to memory and the variety of attention regimes brought forward with the digitization of the comics infrastructure?

It's trivial to claim that the increasing democratization of network-, digital-, and now, generative tools, in comics is producing effects that have been nothing short of profound. The informatization of comics through the aggregation, the storage, and the retrieval of comics media is gradually laying the ground for an industry-wide adoption and integration of generative algorithms into the production of synthetic media content. As a consequence, the industry's scope of knowledge is currently expanding exponentially and challenges the traditional models of production and the professional capacities while setting the stage for a painfully slow but transformative institutional reckoning in academia, comics research and certainly comics craft. Notably, this tectonic shift in the comics media ecosystem extends its ramifications to the very core of comics professionals expressing unease and denial to the latest advances in generative AI—-sentiments that have clear corollaries with previous moments in the history of comics, such as the advent of digital tools for photo editing, which is now a staple of conventional comics craft. The increasing networking capabilities of machine learning tools are currently challenging what it means to be a comics artist and what counts as creativity in comics. My personal ambition is that these tools will help demonstrate to which extent comics craft was never about just creating images. In *my* comics history, the comics artists I care about are not image producers but designers and engineers of systems–-systems of production and systems of signification that are social, evolutionary and dynamic. Always attentive to the infrastructural dynamics at play and the contractual obligations implied by the comics industry, these artists are presented nowadays with the opportunity to leverage new methods and forms of knowledge that could inform a novel comics praxeology. While I would be wary to minimize the impact of generative AI as one more technological advancement which professionals sooner or later will have to come to terms with, I have the feeling that it's a multitude of different responses to these technologies, whether they come from commercial, institutional, technological and artistic bodies, that will simultaneously define the comics industries of the future.

BP: So the comics industries of the future would have the AI become part of the systems of production, that seems inevitable, and this makes sense in evolutionary and dynamic systems. What will this mean for the social systems of

comics? Do you feel your works *Fast-walkers* and *The Neural Yorker* already shed any light on that?

IM: Definitely. I think that generative AI will gradually be integrated in the comics toolbox. Generative AI might be an opportunity to broaden the creative influence of comics professionals. I perceive it as an opportunity to develop a deeper understanding of comics' inherent storytelling abilities, which, when combined with knowledge of cutting-edge data visualization techniques and generative models for multimedia content, can break down traditional boundaries within the medium and decompartmentalize comics from its relatively narrow range. By becoming proficient in handling vast amounts of data, comic artists can identify trends, patterns and anomalies (outliers), experiment with novel data visualization methods and diagrammatic representations, use their storytelling skills in visual data analysis, and even develop their own customized generative models for generating image, text, and multimedia content. These newfound abilities may redefine the role of a comics artist in a technology-driven landscape, redefine creativity, and potentially provide insights into addressing ethical concerns related to authorship, attribution, and shared ownership––a discussion that I believe hasn't yet matured. Synthetic comics, and other works that are programmatically generated, have the potential to articulate various interests within the realms of institutions, education, and commerce. Whether these interests can be incorporated into broader strategies that could significantly impact the comics industry, depends on our comprehension of these innovative tools and their role in generating fresh perspectives within comics. It also hinges on the futures we can envision and build based on these newfound insights, and whether they can find applications in existing artistic practices.

References

Manouach, Ilan. *Fastwalkers*. Brussels: Echo Chamber / 5c, 2021.

—. *Compendium of Francobelgian Comics*. Brussels: Echo Chamber / 5c, 2018. https://ilanmanouach.com/. Accessed 12 January 2024.

Moura, Pedro, ed. *Ilan Manouach in Review: Critical Approaches to His Conceptual Comics*. New York: Routledge, 2023.