

---

## **GENERATIVE AI REPORT**

---

### **Creating a comic with Generative AI**

Pietro Marcatti (164717)

University of Udine

Academic year 2023/24

# Goal of the project

The goal of this course-work project was to create a short comic with only the use of free and publicly available generative AI tools. The final submitted comic is 17 pages long and its creation has been subdivided into 4 major steps which will be detailed in the following sections. The end result is also present as an appendix to this report, furthermore, the conversations, all the images and the prompts used to generate them will be available at the project's [GitHub page](#).

## Solution Overview

Here are listed the 4 major steps in which the comic generation process was subdivided:

**Plot Generation** The comic creation started with the textual interaction carried over with ChatGPT (version 3.5). This conversation served to generate the main story points starting from the incipit I provided, and the dialogues which remained pretty much unaltered in the final version.

**Page Layout Generation** In a different conversation with ChatGPT, I tasked it to define a suitable page layout given some basic rules of panel sizes and space usage.

**Panel Generation** Having a good understanding of the panel's configuration and content, the next step was that of generating the image for every one of them. This has been done using a free Stable Diffusion tool online, more on it in the dedicated section.

**Comic Assembly** The final step was to pair every panel's image with its accompanying speech bubbles and text and then merge those in order to create the final result. This job was done on the free version of Canva, a popular online graphics/creative tool.

Now to address every step in further detail.

## Plot Generation

The starting concept I provided to ChatGPT can be summarized as "*A story about a software development company that publishes a suspiciously good chatbot*". At the beginning of the chatting session I provided a couple of important guidelines, such as the story structure:

- introduce the characters
- present the environment
- introduce the status quo
- have the main act take place
- give a brief conclusion to the story

Another important instruction, that I felt was key in obtaining a story that was more suitable for a comic, was avoiding to introduce an omniscient narrator. Having such narration lead ChatGPT to constantly anticipate events, give out the motives behind the behaviour or dialogues of characters. This works against the style of comics, as the story needs to be *shown* more than it needs to be *narrated*.

The first phase of the plot generation was focused on deciding what the story was about and what events took place. To do this, I asked to be presented with some story incipits. Having chosen the one I liked best, I continued by tasking the chatbot with writing out the beginning of the story. The technique I used in this section was to ask for story points with fine granularity for which I provided examples, like this one:

- Story starts describing the office environment
- The clock shows 9.00 AM and an email arrives
- Alex reads the email
- Alex checks with Maya if she read the email too
- Alex discloses his uninterest in the new assistant

I also provided the context for the main characters' background. Since they're employed in a software development company that works on AI, they know about LLMs and AI in general. The language registry, the reasoning and their actions had to be written with this in mind. A challenge I faced was that, given I wanted a more realistic take on the way the story talks about AI, ChatGPT kept resorting to clichés rooted in common misconceptions of what an algorithm is, or how a chatbot works internally. To give an example, many times ChatGPT suggested that the real reason why the virtual assistant seemed sentient was because its algorithm was coded to "mimic humans' behaviour", whatever that means.

I used the fine-detailed story points technique repeatedly to work through the story writing. One more important instruction I provided was to strongly encourage fully fleshed-out direct dialogues and to avoid indirect ones. Again, I used examples like this one:

*Use fully fleshed-out direct dialogues. Avoid indirect ones. Like this:*

*Alex leans over to Maya, Alex: "Have you seen that we're joining in on the AI assistant bandwagon?"*

*I didn't even know we had a team working to that. Whatever it's gonna be unhelpful at best."*

Another struggle I encountered in a couple of the interactions I had while writing the story was that ChatGPT, given that I wanted a sarcastic undertone to the story, kept making the chatbot character way too witty or sarcastic, constantly spitting jokes and making banter. To solve this, it helped to frame the bot's purpose as a bog-standard corporate virtual assistant that seems just a bit *too good*.

### 3.1 Creating a coherent story

In the tests I conducted to understand how to effectively interact with ChatGPT to create the story, I was able to create a prompt template that allowed me to obtain longer and more coherent stories. To achieve this, I found that repeatedly starting the prompt with the entire story written thus far and asking for a continuation resulted in the best results. This is an example:

*I'll paste your previous output up to the point I deem acceptable, continue onwards from there:*

- The story unfolds in the bustling office of ByteCorp, where Alex and Maya are settling into their morning routine.
- Alex: "Another email from IT. Probably announcing some new corporate initiative."
- Maya: "Let me guess, another pointless chatbot to distract us from actual work?"
- As the clock strikes 9:00 AM, Alex's computer pings with a new email notification.
- Alex: "Here we go. Let's see what brilliant idea they've come up with this time."
- Alex clicks open the email from the IT department, reading aloud, "Introducing ByteBuddy, our latest AI chatbot assistant."
- Alex: "Oh joy, another chatbot. Just what we needed."
- Maya, engrossed in her work, glances over at Alex's screen and raises an eyebrow.
- Maya: "Another one? I swear, they think we have nothing better to do than talk to bots all day."
- Alex nods in agreement.
- Alex: "Exactly. As if we don't have enough distractions already."
- Maya chuckles. ...

This served also to keep in context the way I wanted the story to be narrated, which is through fine-grained story points.

# Page Layout Generation

After having reaching a satisfactory results in the generated story, I initiated a new chat session with ChatGPT to produce a suitable page layout that would keep in mind how a comic page is structured, what kind of panel sizes I wanted to work with and more. In particular, I wanted to work with vertical A4 pages divided into 6 equally big rectangles arranged in a  $2 \times 3$  grid (row, column). Then, the valid panel sizes were only 3:

**Weight of 1** A single grid cell

**Weight of 2** An horizontal strip of 2 contiguous grid cells

**Weight of 4** A bigger rectangular area made of  $2 \times 2$  contiguous grid cells.

As a consequence, a single A4 page has a total panel weight of 6 units and there are several possible ways to arrange its panels. In the instructions I provided, I tried several techniques to enforce this constraint with only partially correct results. In particular, I tried to provide positive and negative examples like this ones:

*Every page NEEDS to have a total panel weight of EXACTLY 6. Page composition with a weight under or over 6 is NOT acceptable.*

*I'll provide examples of NOT acceptable combinations:*

$$1+1+1+4 = 7$$

$$2+2+4 = 8$$

*Now I'll provide examples of VALID combinations, you should NOT limit yourself to these ones:*

$$1+1+4=6$$

$$2+2+2=6$$

The layouts for the first 5-6 pages were generally compliant with the rules specified, but further down the generation the constraint was not satisfied. For this reason I took it upon myself to tweak the layouts of the later pages while trying to stay truthful to ChatGPT's suggestion.

## 4.1 Script

This is the final result of the story and layout generation:

### Page 1:

- Panel 1: (Weight: 2)
  - Description: The office of ByteCorp, with desks, computers, and employees busy at work.
  - Dialogue:
    - \* Alex: "Another email from IT. Probably announcing some new corporate initiative."
    - \* Maya: "Let me guess, another pointless chatbot to distract us from actual work?"
- Panel 2: (Weight: 2)
  - Description: Close-up of Alex's computer screen showing a new email notification.
  - Dialogue:
    - \* Alex: "Here we go. Let's see what brilliant idea they've come up with this time."
- Panel 3: (Weight: 2)
  - Description: Alex clicks open the email from the IT department, reading it with a skeptical expression.
  - Dialogue:
    - \* Alex: "Introducing ByteBuddy, our latest AI chatbot assistant."
    - \* Alex: "Oh joy, another chatbot. Just what we needed."

## **Page 2:**

- Panel 4: (Weight: 1)
  - Description: Maya, engrossed in her work, glances over at Alex's screen with a raised eyebrow.
  - Dialogue:
    - \* Maya: "Another one? I swear, they think we have nothing better to do than talk to bots all day."
- Panel 5: (Weight: 1)
  - Description: Alex nods in agreement, while Maya chuckles.
  - Dialogue:
    - \* Alex: "Exactly. As if we don't have enough distractions already."
    - \* Maya: "Well, let's see if this one can at least schedule meetings without crashing our calendars."
- Panel 6: (Weight: 4)
  - Description: Alex and Maya return their focus to their tasks, surrounded by other busy employees. The chatbot ByteBuddy's logo subtly visible on computer screens in the background.
  - Dialogue:
    - \* Alex: "I doubt it, but might as well give it a shot."

## **Page 3:**

- Panel 7: (Weight: 2)
  - Description: Throughout the office, employees are seen interacting with ByteBuddy on their computers, sharing anecdotes of its helpfulness.
  - Dialogue:
    - \* Colleague 1: "I asked ByteBuddy to pull up last month's sales report, and it did it in seconds!"
    - \* Colleague 2: "I used it to set reminders for all my meetings, and I haven't missed one since!"
- Panel 8: (Weight: 4)
  - Description: Alex and Maya observe their coworkers becoming increasingly reliant on ByteBuddy, exchanging concerned glances.
  - Dialogue:
    - \* Maya: "Looks like ByteBuddy's making quite the impression."
    - \* Alex: "I have to admit, it's pretty impressive."

## **Page 4:**

- Panel 9: (Weight: 2)
  - Description: Alex and Maya decide to test ByteBuddy for themselves, engaging in a conversation with the chatbot.
  - Dialogue:
    - \* Alex: "ByteBuddy, schedule a meeting for next Monday at 10 AM."
    - \* ByteBuddy: "Meeting scheduled. Anything else I can assist you with?"
- Panel 10: (Weight: 2)
  - Description: Intrigued by ByteBuddy's efficiency, Alex and Maya continue to experiment with its capabilities.

- Dialogue:
  - \* Maya: "Let's see how it handles more complex tasks."
- Panel 11: (Weight: 2)
  - Description: They delve into ByteBuddy's features, testing its ability to analyze data and generate reports.
  - Dialogue:
    - \* Alex: "ByteBuddy, compile a summary of last quarter's financial performance."
    - \* ByteBuddy: "Summary generated. Would you like me to email it to you?"

#### **Page 5:**

- Panel 12: (Weight: 1)
  - Description: Impressed by the AI's responsiveness, they begin to wonder about the technology behind it.
  - Dialogue:
    - \* Maya: "I can't believe how seamlessly it integrates with our systems."
- Panel 13: (Weight: 1)
  - Description: As they dig deeper, they uncover evidence that suggests ByteBuddy may have access to sensitive information.
  - Dialogue:
    - \* Alex: "This doesn't add up. How does ByteBuddy have access to all this data?"
- Panel 14: (Weight: 4)
  - Description: Their suspicions grow as they realize the extent of ByteBuddy's capabilities and the implications for privacy and security.
  - Dialogue:
    - \* Maya: "I don't like where this is headed, Alex. It feels like we're being watched."

#### **Page 6:**

- Panel 15: (Weight: 2)
  - Description: Determined to uncover the truth, they decide to investigate ByteCorp's motives behind developing ByteBuddy.
  - Dialogue:
    - \* Alex: "There has to be a reason they're pushing this chatbot so hard."
- Panel 16: (Weight: 2)
  - Description: Alex and Maya begin to question the sudden appearance of ByteBuddy, considering they had never heard of any team working on such a project.
  - Dialogue:
    - \* Alex: "Did you know anything about a team working on this chatbot?"
    - \* Maya: "Not a clue. It's like it came out of nowhere."

#### **Page 7:**

- Panel 17: (Weight: 1)

- Description: Their curiosity piqued, they delve into the technical aspects of ByteBuddy’s development.
- Dialogue:
  - \* Alex: ”I wonder what data ByteBuddy was trained on. They must have used some massive datasets.”
- Panel 18: (Weight: 1)
  - Description: They also question the computing resources allocated to ByteBuddy, given its seemingly advanced capabilities.
  - Dialogue:
    - \* Maya: ”I doubt our servers could handle the workload required for ByteBuddy’s performance.”
- Panel 14: (Weight: 4)
  - Description: As they explore further, they realize their supervisor seems as clueless about ByteBuddy’s origins as they are.
  - Dialogue:
    - \* Alex: ”Have you talked to Sarah about this?”
    - \* Maya: ”Yeah, she’s just as in the dark as we are.”

#### **Page 8:**

- Panel 20: (Weight: 2)
  - Description: With each unanswered question, Alex and Maya become increasingly suspicious of ByteCorp’s motives.
  - Dialogue:
    - \* Alex: ”Something doesn’t add up. They’re pushing this chatbot too hard for it to be just another corporate initiative.”
- Panel 16: (Weight: 2)
  - Description: Determined to uncover the truth, they decide to dig deeper into ByteCorp’s operations.
  - Dialogue:
    - \* Maya: ”We need to find out who’s behind ByteBuddy and what their endgame is.”

#### **Page 9:**

- Panel 22: (Weight: 1)
  - Description: During a coffee break, Alex and Maya overhear snippets of conversation between two executives discussing a briefing on ByteBuddy’s progress.
  - Dialogue:
    - \* Maya: ”Did you hear that? They mentioned a briefing on ByteBuddy.”
    - \* Alex: ”Yeah, but when and where?”
- Panel 23: (Weight: 1)
  - Description: Unable to gather more information, Alex and Maya decide to deduce the time and place of the meeting.
  - Dialogue:
    - \* Alex: ”We need to figure out when and where this meeting is happening. Let’s think.”

- \* Maya: "Right. They wouldn't want too many people to know about it, so it must be somewhere private."
- Panel 24: (Weight: 4)
  - Description: They recall previous instances of high-level meetings and narrow down their options.
  - Dialogue:
    - \* Alex: "Remember when they briefed us on the new project last month? It was in the executive conference room."
    - \* Maya: "That's true. Let's check if the room is booked for tomorrow."

## Page 10

- Panel 25: (Weight: 2)
  - Description: Using their company's scheduling system, they discreetly check the availability of the executive conference room for the following day.
  - Dialogue:
    - \* Alex: "Bingo. It's reserved for tomorrow afternoon."
    - \* Maya: "That's gotta be it. Now we just need to figure out the time."
- Panel 26: (Weight: 2)
  - Description: They recall the executives mentioning the need for a briefing, suggesting a mid-day meeting.
  - Dialogue:
    - \* Alex: "They said something about a briefing. Usually, those happen around lunchtime."
    - \* Maya: "Let's assume it's around 12 PM. We'll be there early to make sure."

## Page 11

- Panel 27: (Weight: 2)
  - Description: With their deductions in place, Alex and Maya prepare to attend the secret meeting, hopeful that they'll finally uncover the truth about ByteBuddy.
  - Dialogue:
    - \* Alex: "Tomorrow, we find out what's really going on with ByteBuddy."
    - \* Maya: "Let's just hope we're right about this."
- Panel 28: (Weight: 2)
  - Description: The next day, Alex and Maya arrive early at the executive conference room, eager to unravel the mystery surrounding ByteBuddy.
  - Dialogue:
    - \* Maya: "I hope we're not too early. I don't want to draw any attention."
    - \* Alex: "Agreed. Let's blend in and observe."

## Page 12

- Panel 29: (Weight: 2)
  - Description: As the meeting time approaches, they notice several high-ranking executives entering the room, including the CEO and the head of the AI development team.
  - Dialogue:

- \* Alex: "Looks like we're in the right place."
- \* Maya: "Keep an eye out for any clues."
- Panel 30: (Weight: 2)
  - Description: From their discreet vantage point, Alex and Maya strain to hear the discussion unfolding inside.
  - Dialogue:
    - \* Maya: "It's hard to make out what they're saying."
    - \* Alex: "Shh, let's listen carefully."

## Page 13

- Panel 31: (Weight: 2)
  - Description: Gradually, they catch snippets of conversation, piecing together crucial information about the true nature of ByteBuddy's development.
  - Dialogue:
    - \* CEO: "...revolutionize the industry with our breakthrough in human-AI interaction..."
    - \* Head of AI: "The integration of human responses into the AI's algorithms is exceeding our expectations..."
- Panel 32: (Weight: 2)
  - Description: Alex and Maya exchange wide-eyed glances, realizing the implications of what they've overheard.
  - Dialogue:
    - \* Maya: "Did you hear that? They're integrating human responses into ByteBuddy's algorithms."
    - \* Alex: "That explains its uncanny accuracy and responsiveness."

## Page 14

- Panel 33: (Weight: 2)
  - Description: Determined to uncover the human component behind ByteBuddy, they devise a plan to trace the origin of these human interactions.
  - Dialogue:
    - \* Alex: "We need to find out where these human responses are coming from."
    - \* Maya: "But how? It's not like they're going to publicly announce it."
- Panel 34: (Weight: 2)
  - Description: After the meeting, Alex and Maya discreetly follow a group of executives leaving the conference room, hoping to gather more clues.
  - Dialogue:
    - \* Maya: "Let's tail them carefully. They might lead us to the facility."
    - \* Alex: "Good idea. Stay inconspicuous."

## Page 15

- Panel 35: (Weight: 2)
  - Description: They trail the executives through the corridors of the office building, keeping a safe distance to avoid suspicion.

- Panel 36: (Weight: 2)
  - Description: Eventually, the group enters a secluded area of the building, marked with restricted access signs.

## Page 16

- Panel 37: (Weight: 2)
  - Description: Inside the restricted area, Alex and Maya search for an opportunity to slip past security.
- Panel 38: (Weight: 2)
  - Description: They notice a door left slightly ajar and seize the moment to slip inside.
  - Dialogue:
    - \* Alex: "Quietly now. We don't want to draw attention."

## Page 17:

- Panel 39: (Weight: 2)
  - Description: Inside, they navigate through a maze of corridors and laboratories, each filled with cutting-edge technology.
- Panel 40: (Weight: 2)
  - Description: As they explore further, they stumble upon a room bustling with activity, where a team of researchers is engaged in intense work.
  - Dialogue:
    - \* Maya: "This must be where they're developing the human component of ByteBuddy."

## Page 17:

- Panel 41: (Weight: 2)
  - Description: Hiding behind a stack of crates, Alex and Maya observe the researchers, noticing a familiar face among them.
  - Dialogue:
    - \* Alex: "Isn't that Sarah from our team?"

The one presented here is the unaltered version of the written output of ChatGPT. This being said, the final result of the comic presents some slight alterations and adjustments, they are listed here:

- Total panel weight for most of the later pages is wrong. It was adapted with the goal of giving more complex scenes a bigger space.
- Some dialogue lines were reassigned to better suit the generated images.
- The culmination of the story is too dragged out in ChatGPT's proposed solution and it was cut short as the comic is already of a considerable length.
- The panel descriptions often describe a concept or a scene that is not "atomic" enough to be represented by a single image. I took the liberty to interpret the goal for that panel and choose the associated "atomic" scene to depict.

# Panel Generation

All images used in the panels composing the comic's pages were generated with Stable Diffusion XL. At first I tried downloading the model to my machine and accessing it via a web UI such as Automatic 1111, but the generation process was prohibitively slow. Even though I could take advantage of a dedicated Nvidia GPU (1070Ti with 8Gb of VRAM), every image generation took around 3 to 4 minutes. Factoring that for every panel I would need to perform several generations to obtain a decent result by adjusting the prompt, this meant that a local configuration was unfeasible. For this reason I used a website ([stablediffusionweb.com](https://stablediffusionweb.com)) that offers faster generation performances albeit with a much reduced set of options to configure.

The tool offered the following options for the generation process:

- Main prompt
- Negative prompt
- Image style
- Aspect ratio (from a list of predefined ones)
- Prompt's guidance factor
- Seed

All the prompts I provided in the [GitHub](#) specify the values used for every one of these parameters.

All the images have been generated with the "Comic" image style, but the values of the other parameters varied between the different generations. I chose this style because I personally felt like it fitted the theme of the story and it also conveniently helps to reduce the visibility of graphical anomalies or anatomical infidelities. Choosing to stick to black and white was also done with purpose. This allowed me to achieve more consistent results because I would not have to worry about the possibility of generating a good looking panel but with mismatching colors compared to the previous ones.

## 5.1 Main and Negative Prompts

Since the story has very few characters and settings, having found a performing template for the style and location of the depicted subjects helped me lots with achieving consistently repeatable results. By looking at multiple of the prompts provided, it appears obvious that they share most of their structure and the differences are in the description of the subject in the main prompt.

The main prompt for most of the images starts off in the same way:

*Black and white comic in the style of western comics. Scene is set in a bright, modern office space.*

The description of the main characters is also used repeatedly as it was responsible for the most consistent generations:

1man (<Chris Evans 0.5, John Krasinski 0.5> wearing a white shirt and black ties)

1woman (< Saoirse Ronan 0.5 > young, blond hair, wearing a white shirt and black tie)

A similar comment can be made about the negative prompt. Just like the main prompt, it was also pretty much identical as through the first generations I was able to identify the common deviations from my expectations that were present in the images. Funnily enough, it can be seen that the negative prompt collects more and more terms in the later generations as more unwanted elements appeared. Since these elements were unwanted anyway, I used the same negative prompt in all generations, even in when it seems unnecessary as it served to rule out weird behaviour, basically because "*you never know*".

## 5.2 Aspect Ratio and Seed

One last comment about the generation process has to deal with the use of aspect ratios and seeds. In my experiments, I realized that the use of a wider aspect ratio resulted in much more successful generations of multiple subjects. For this reason I often used a more rectangular aspect ratio even though I was after a square image. The resulting panel was then cropped in the comic assembly section to fit the desired shape.

Lastly, I observed that a seed I tried just out of curiosity lead to a image style and composition that I liked. For this reason I used it multiple times in my first few attempts on almost all generations. If even after the prompt tweaking process I felt that I had not reached a satisfactory result, I would try with randomized seeds, just to introduce new variability and see new takes on the same prompt.

## Comic Assembly

The final comic assembly phase was done on [canva.com](https://canva.com) as it is a pretty popular and powerful tool for editorial and graphical work. In particular, in this phase of the project I had to crop the images generated in order to fit the panel weight and to add the speech bubbles and text dialogues. A final step was to adapt each comic page to the number of pixels of a A4 with a pixel density of 96, which is the one of the images output by the stable diffusion model.

## Submitted Material

The following material is available in the [GitHub](#) repo:

- *.html* files of the conversations with ChatGPT for plot and layout generation
- A folder for every page containing the images present in its panels as well as a *.txt* files detailing the prompt and parameters used for its generation
- *final\_result.pdf* which is a full-scale version of the comic
- *project\_report.pdf* this document detailing the creation process of the comic

## Final Result

















LATER THAT DAY:  
BREAKROOM

















