

AFFECTIVE FALLACY

FUTURE SHOCKS

A Global Risks Report 2019 Scenario

**The future has never felt so uncertain.
Have you ever imagined what it could look like?
Embark on a journey into a disrupted future.**

Affective Fallacy

What if machines could hack our deepest emotions?

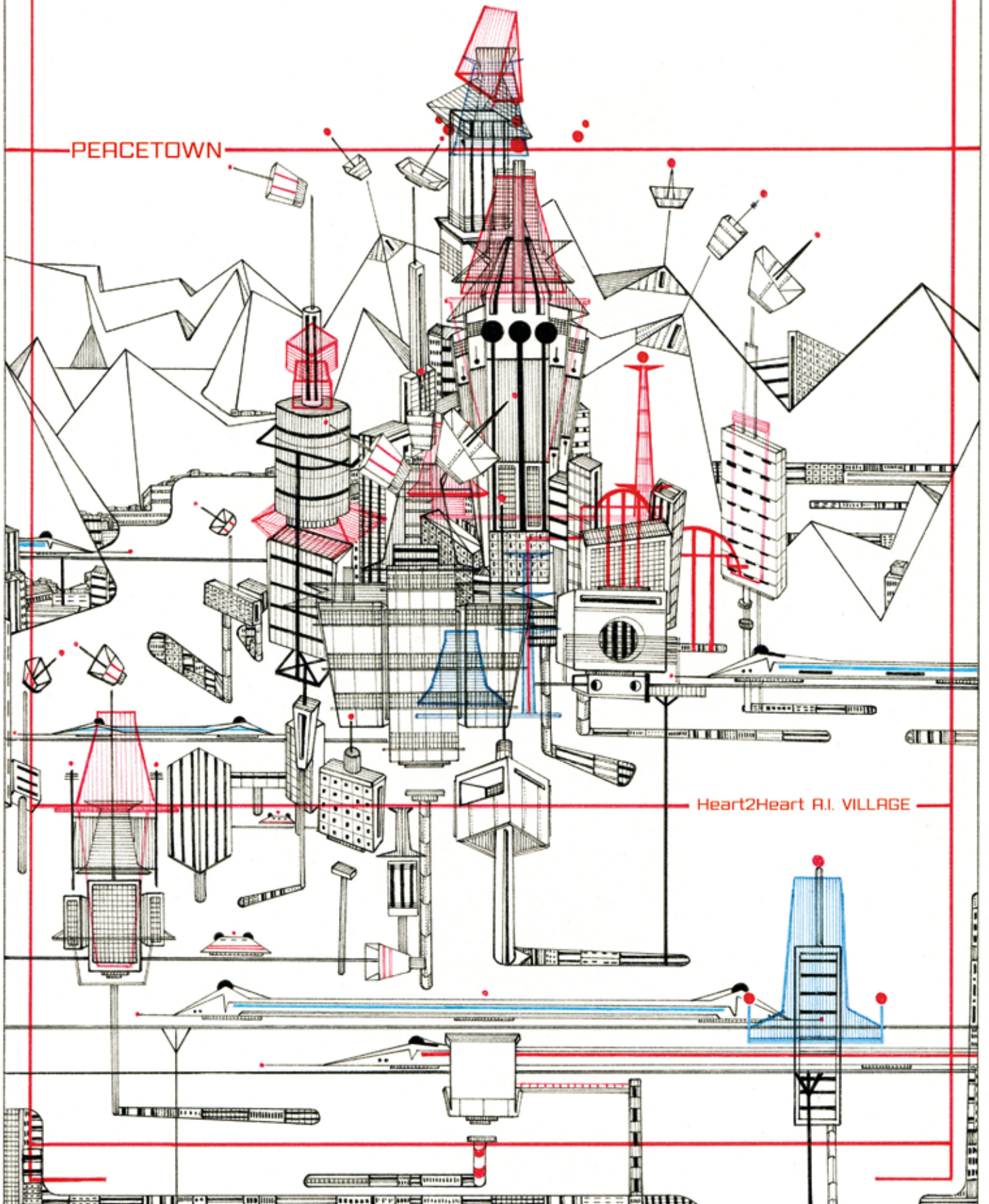
Chen Qiufan's story conjures up a world of "AI villages" where feelings are meticulously categorized and fed into a giant database. What does it mean to experience love and intimacy when our feelings are mediated by algorithms?

Story by **Chen Qiufan**

Artistic interpretation by **Tarek Abbar**

English translation by **Emily Jin**

AFFECTIVE FALLACY



Women, having consistently demonstrated greater attentiveness toward facial expression details and higher levels of empathy, are usually considered to be better at emotion recognition than men. With this presumption, *Heart2Heart* chose to hire an all-women crew as its core emotion annotators. The workers, after training, become human assistants to affective computing. The system needs a profusion of data to perfect its algorithm: not raw data, but data that has been selected, processed and annotated by real humans. Every face that appeared in the media sources is manually categorized into different emotions and then quantified in terms of the intensity of the emotion. This way, despite confounding variables like age, gender, race and physical features, the machine can effectively learn about human emotions.

There are millions of identical annotation workshops, serving different AI systems. The workers process raw data ranging from text, voice recordings, video clips to more complex interactive games. Villages that rely on these manual workshops to employ residents are, ironically, dubbed “AI Villages.”

Like all her girlfriends, Joyce became an emotion annotator. The workers spent their days in the bright, spacious annotation workshop of *Heart2Heart* Corp. Every annotator, wearing headphones to avoid noise distraction, sat before an ultra-thin curved screen that blocked their peripheral vision. An automatic system assigned them media sources, with red boxes marking every recognizable human face, flashing onto the screen in random order.

Joyce's fingers danced swiftly across the keyboard. While her left hand categorized the emotion on the face, her right hand noted the intensity of the emotion: *Happiness 3, Sadness 5, Anger 7...* Occasionally, when she really empathized with the subject, a similar expression would emerge on her face, too.

Joyce quickened the pace of her work. Faces of strangers flashed before her eyes like fleeting phantoms. She paid no attention to the total-work-completed counter on the top of the screen, however, because her gaze constantly drifted to the digital clock in the bottom right corner.

She has a date tonight.



Emotional Annotation



connects hearts



users trust the AI



AI never lies



AI needs you

keep accuracy optimal!

E.A. 505

JOYCE

E.A. 505

E.A. 504

keep up the good work!

Heart2Heart is an online dating app. Standing out from other dating apps, *Heart2Heart* utilizes a cloud-based emotion detection API to help its users better understand emotional changes in their date in order to increase the success rate of matchmaking.

Nowadays, online dating has become even trickier. On one hand, the Internet can transcend almost every barrier and help connect individuals from diverse geographic locations, cultures, language systems and social classes; on the other, the vast differences in how each person experiences and copes with emotions have made mutual understanding between two individuals extremely difficult.

Joyce met Simon, an urbanite from Megaburg, through *Heart2Heart*.

She always imagined Megaburg to be a futuristic city: varicolored digital screens lining the sides of streets, pedestrians dressed in chic, modern outfits hurrying in and out of glass and iron skyscrapers, their faces expressionless like robots. To her, Megaburg and the village of Peacetown, where she lived, were worlds apart.

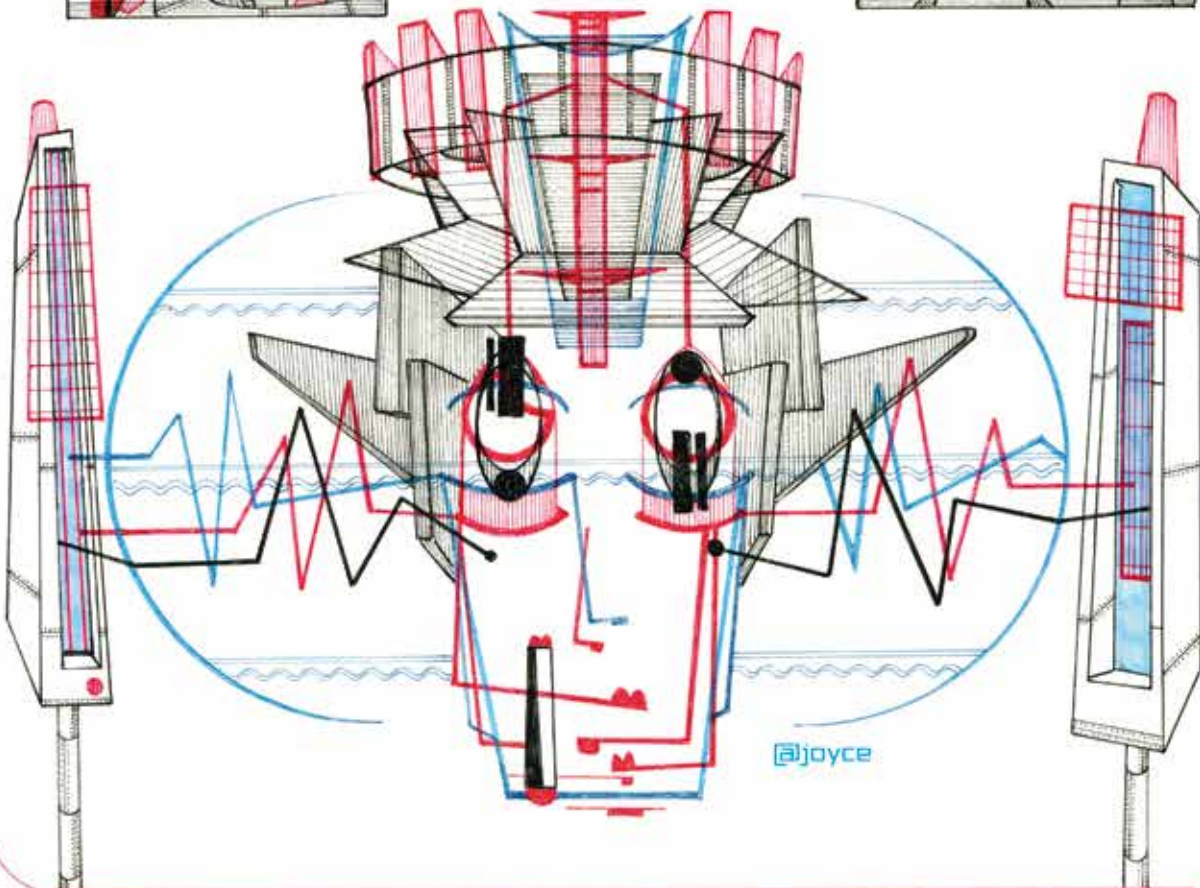
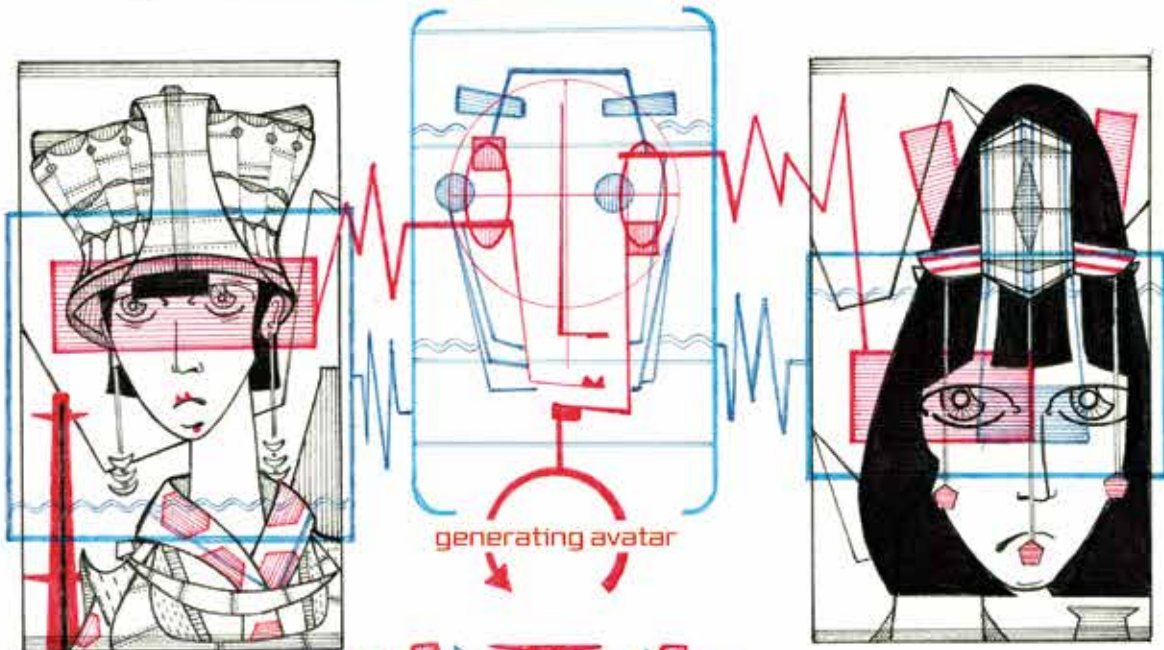
Dating a Megaburg boy was beyond her wildest dreams. Simon had, in fact, struck up a conversation with her first. He told her that her name and the jewelry in her *Augmented Reality* (AR) profile had caught his attention. She was different from the plastic-faced girls that he was used to seeing everywhere. Joyce did not understand many of the things that Simon said; she could only rely on the emotion detection algorithm to gauge his mood and interpret what he really meant.

CREATE YOUR AVATAR

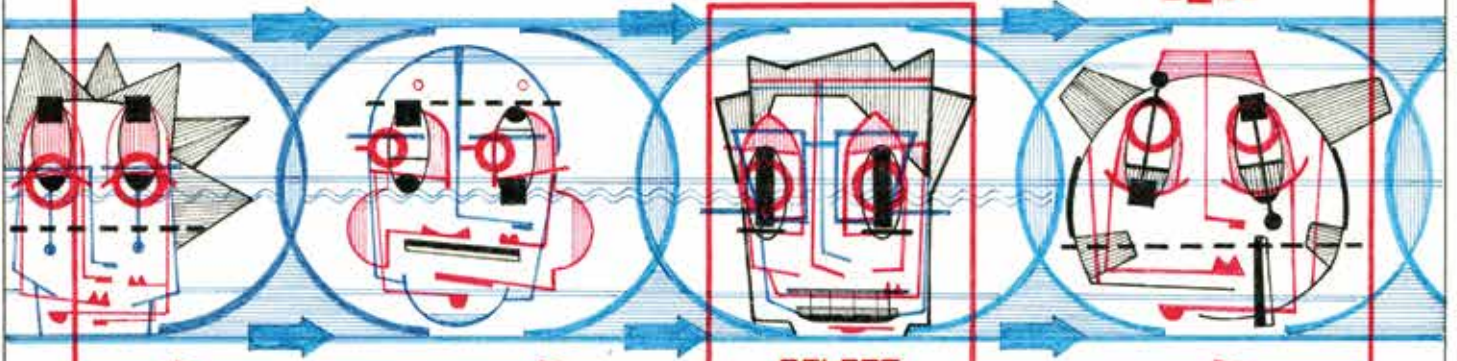


select username

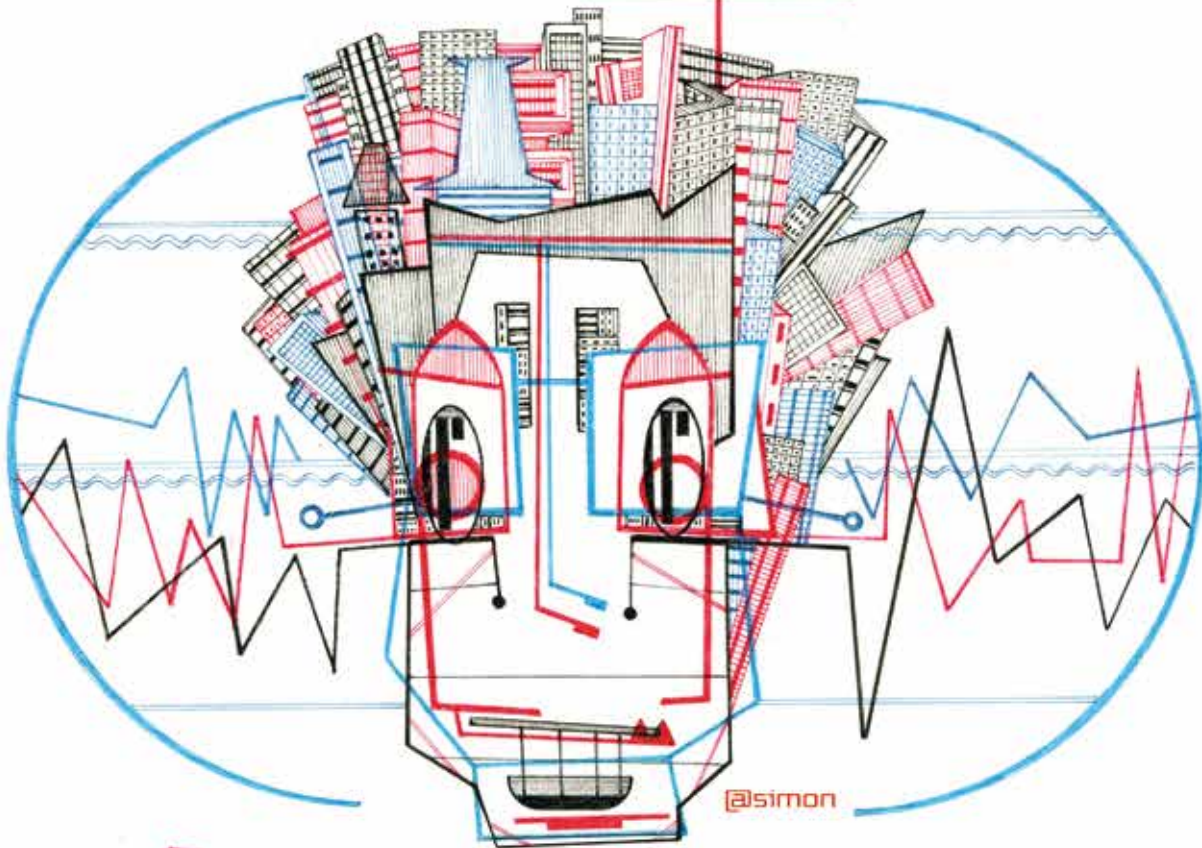
upload avatar data



BROWSE PROFILES



SELECT



joyce:
@simon
I (heart) big city



simon:
@joyce
(heart) tribal look



98%

94%

NO
FILTER

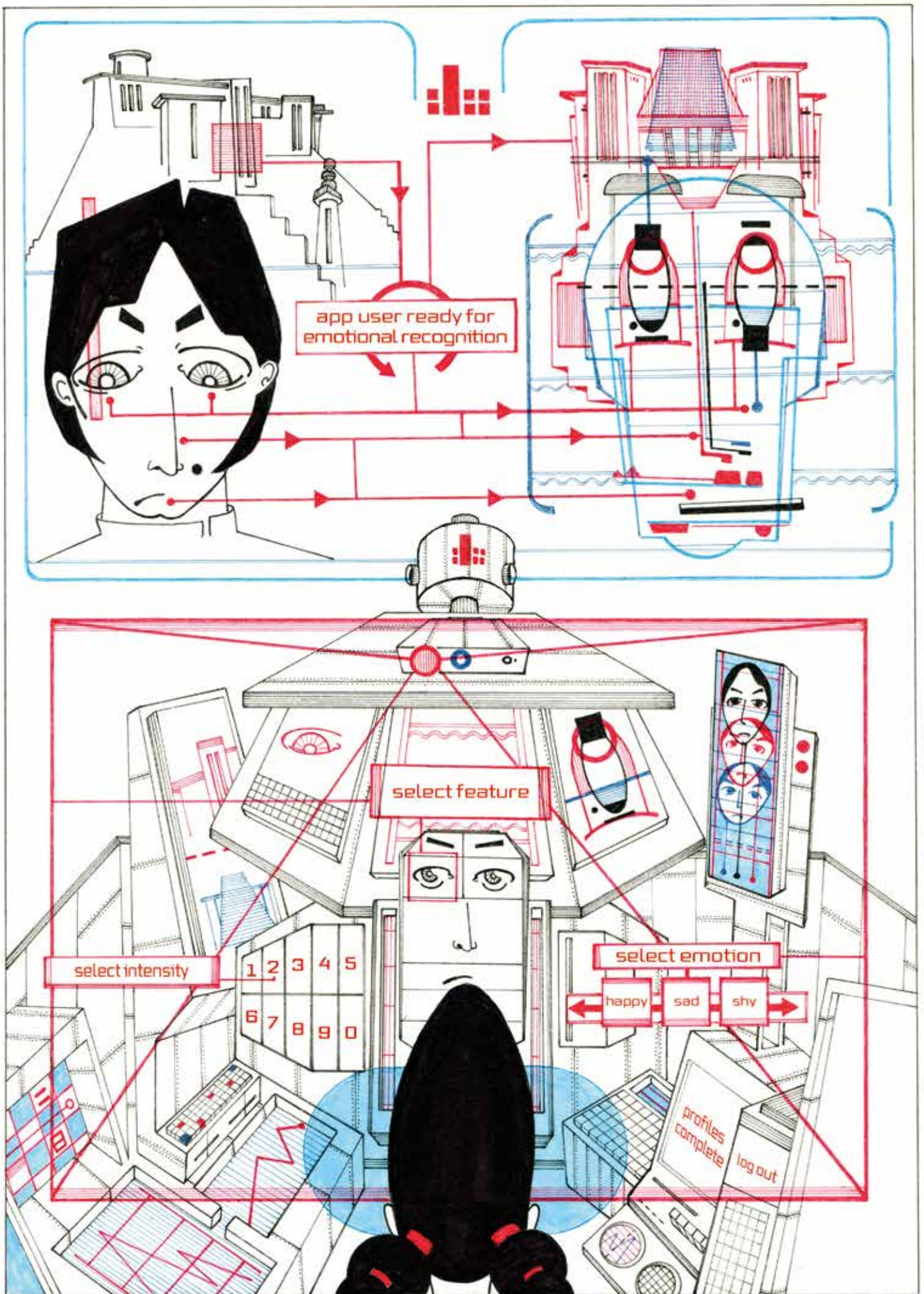
Tonight was the one-month anniversary of Joyce and Simon's first conversation on *Heart2Heart*. Given that online relationships rarely lasted past the first week, a one-month anniversary was quite worth celebrating. Many users chose to commemorate this day by turning off the AR filters that beautified their appearance and meeting each other with their true, unadorned self. "Go no-filter" had become a relationship milestone that could either lead to a new stage of intimacy, or an immediate breakup.

Joyce was at least quite confident about the way she looked. She impatiently waited for night to fall.

The clock was ticking close to the time she had promised to “meet” Simon, but the rate of new media popping up on her dashboard did not seem to have slowed down at all. Joyce’s fingertips burned as she hastily typed. She was working at maximum speed. *Sure, this may sabotage the accuracy of annotations, but what does it matter?* She thought. *The system is going to assign someone else to crosscheck anyway.*

Finally, she reached the last image: a boy stands in front of a monastery, his face, highlighted by a red box, slightly tilted downwards. Joyce pressed *Happiness* and *4* at once. The image disappeared, and the deep blue homepage reemerged on the screen, showing that she had completed all her assigned work.

Another busy day had passed.



The high-speed train, carrying a very giddy Joyce, pulled into the station.

Joyce logged onto *Heart2Heart* on her phone.

While she was at work, Simon had already called her a few times. Joyce redialed. As soon as he picked up, her phone screen casted a miniature half-body holographic projection of him into the air. Simon, through the AR filter, was as charming and stylish as always. “Simon! So sorry for being late. I had too much work to do.”

“No worries, I also got back home just now. Are you ready?”

Joyce saw the annotation next to Simon’s face: *Anticipation 5*. Her heart glowed with happiness. She smiled and nodded at him.

However, Simon didn’t seem to have noticed her smile. His eyebrows plunged into a frown. “We don’t have to do it if you’re not ready. I understand that this a big step in our relationship.”

“I am ready!”

“But...”

“But what?”

“Your face says *Hesitation 4* and *Discomfort 3*.”

“How could that be? I am really excited to meet you!” Panicked, Joyce pulled the corners of her mouth back to show him an exaggerated grin.

“Your face says *Fear 6* now. Joyce, are you hiding something from me?”

“No! The algorithm must have made a mistake. Do you want me to turn off the filter right now?”

“Please, don’t. I need to... I need to think about this.”

Now Simon’s face had become *Suspicion 4* and *Irritation 3*.

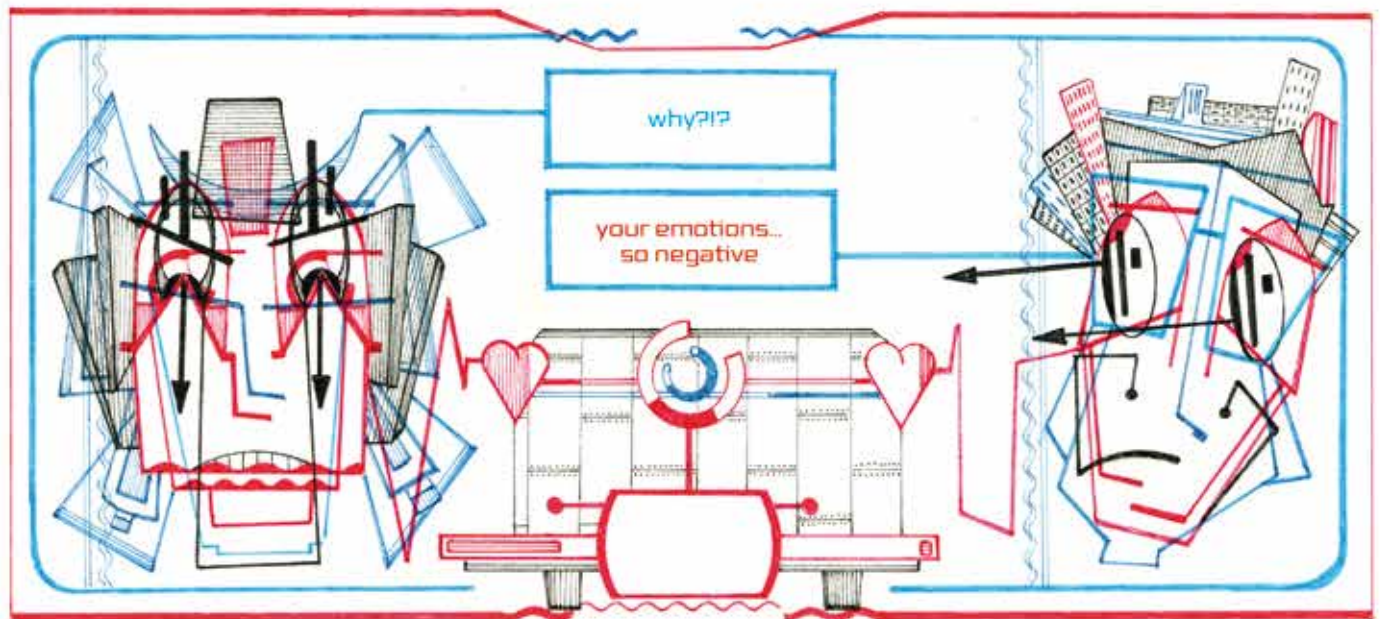
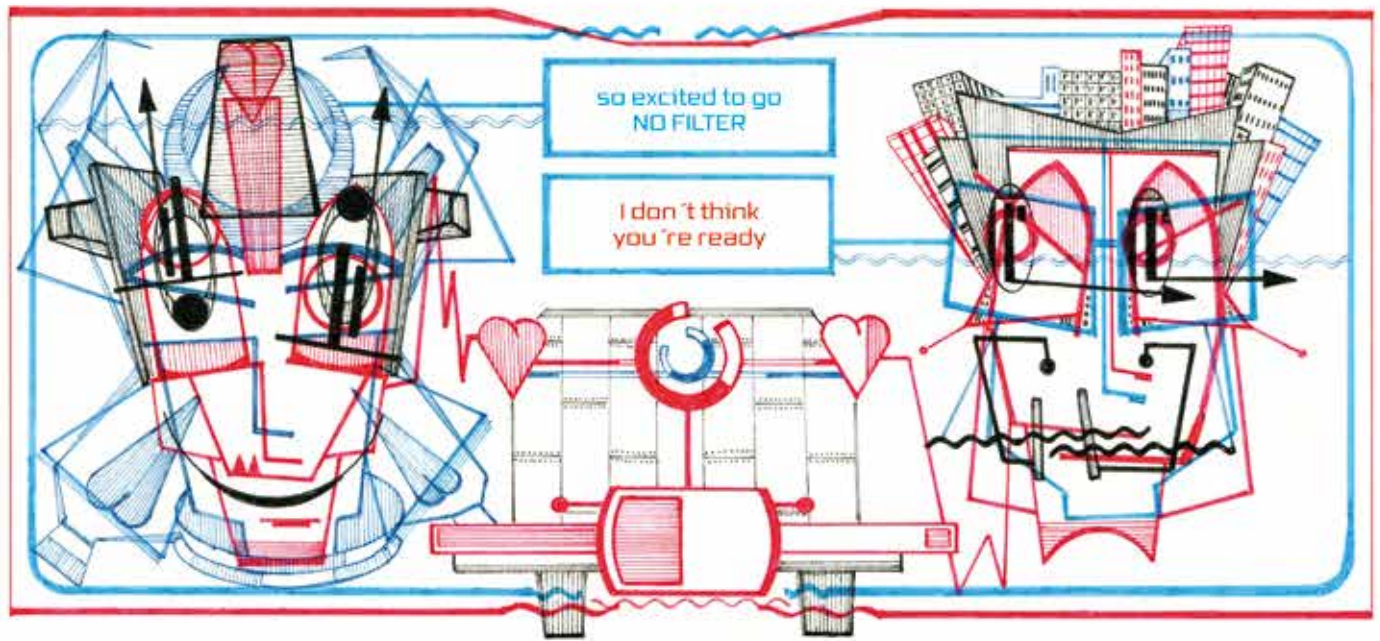
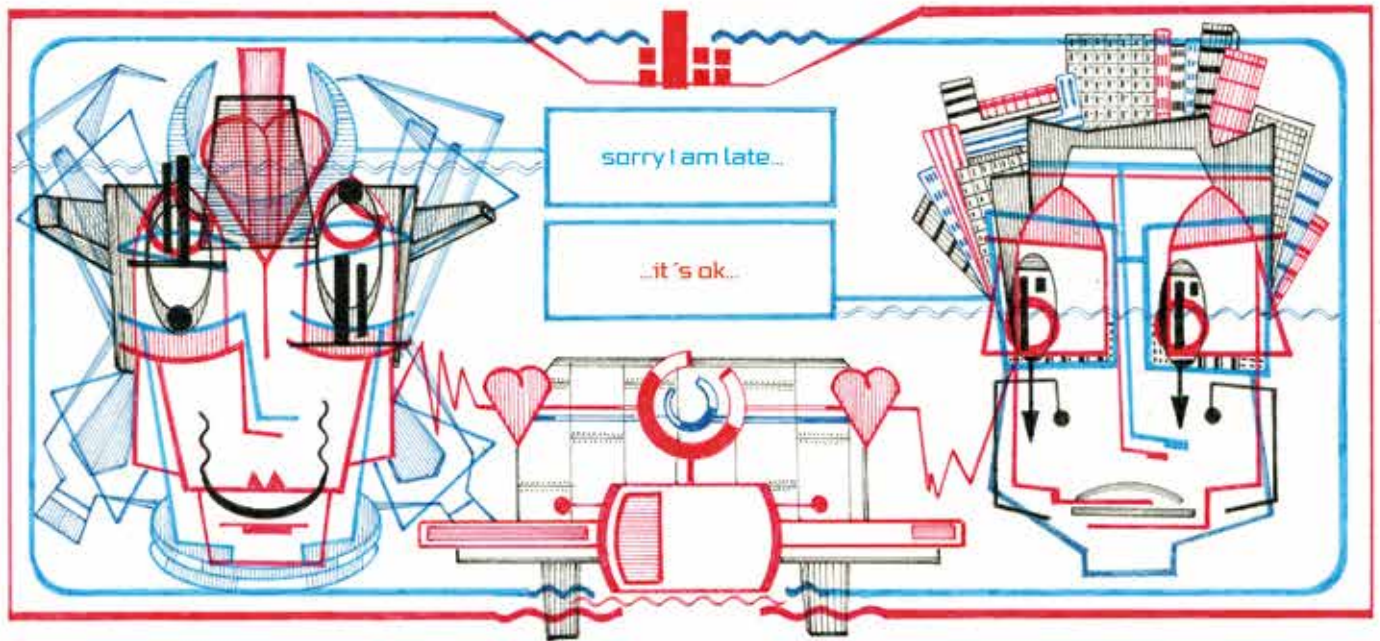
What is going on? She thought.

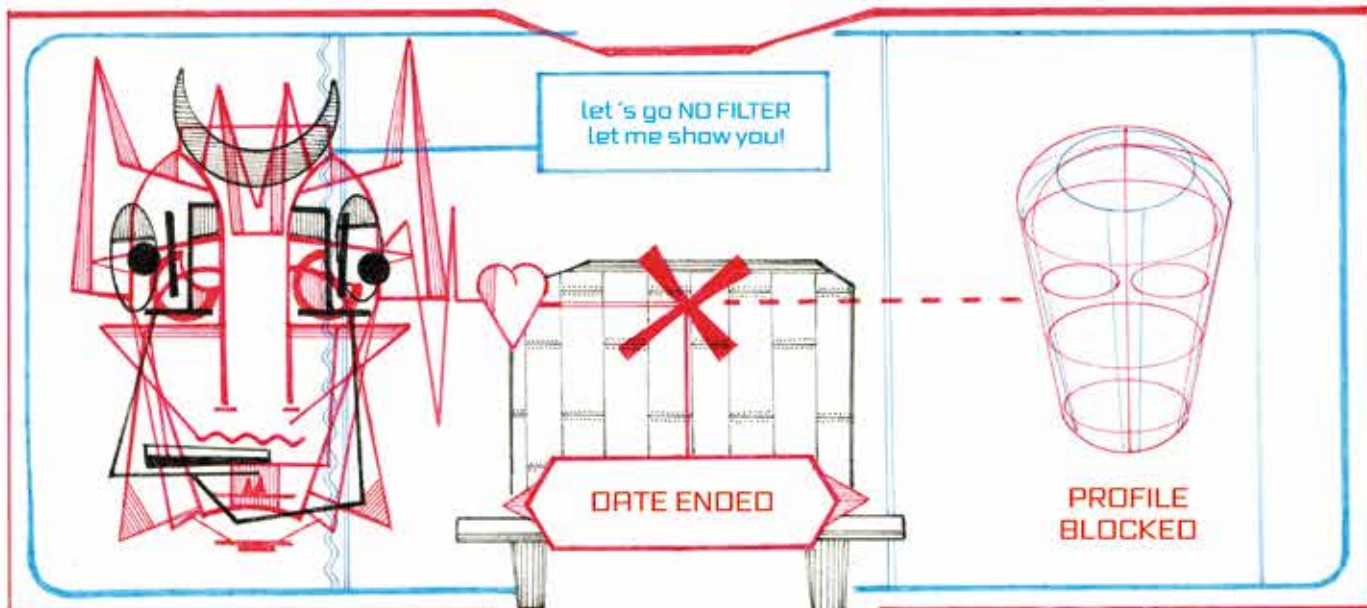
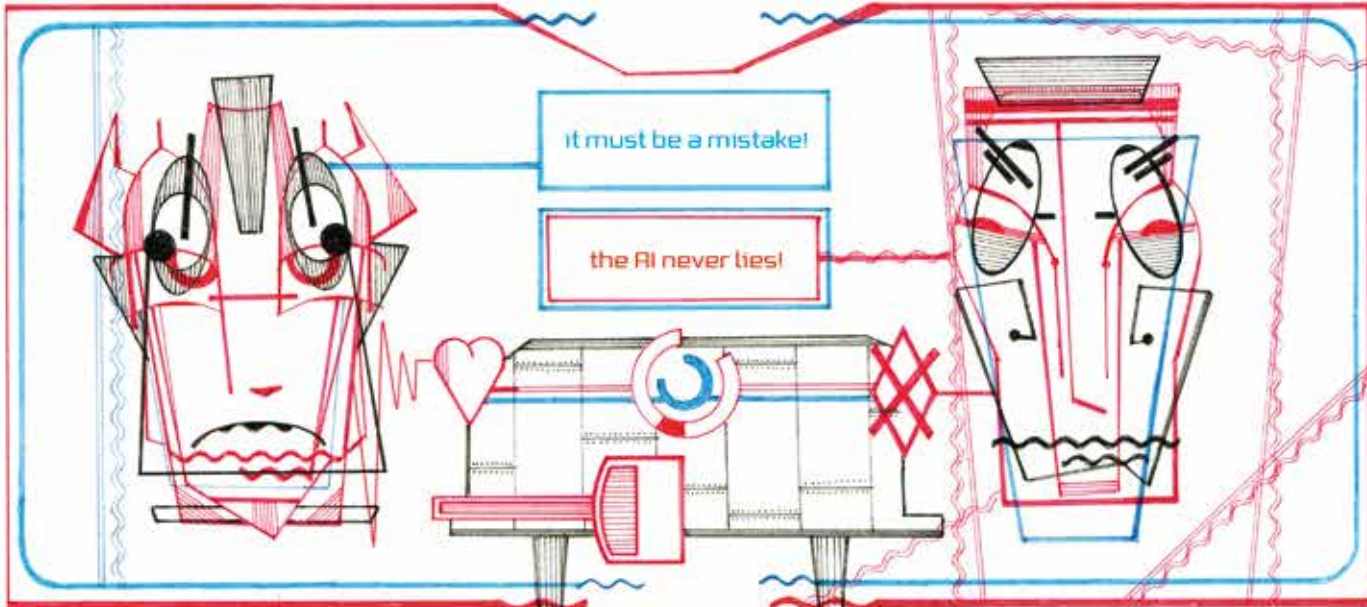
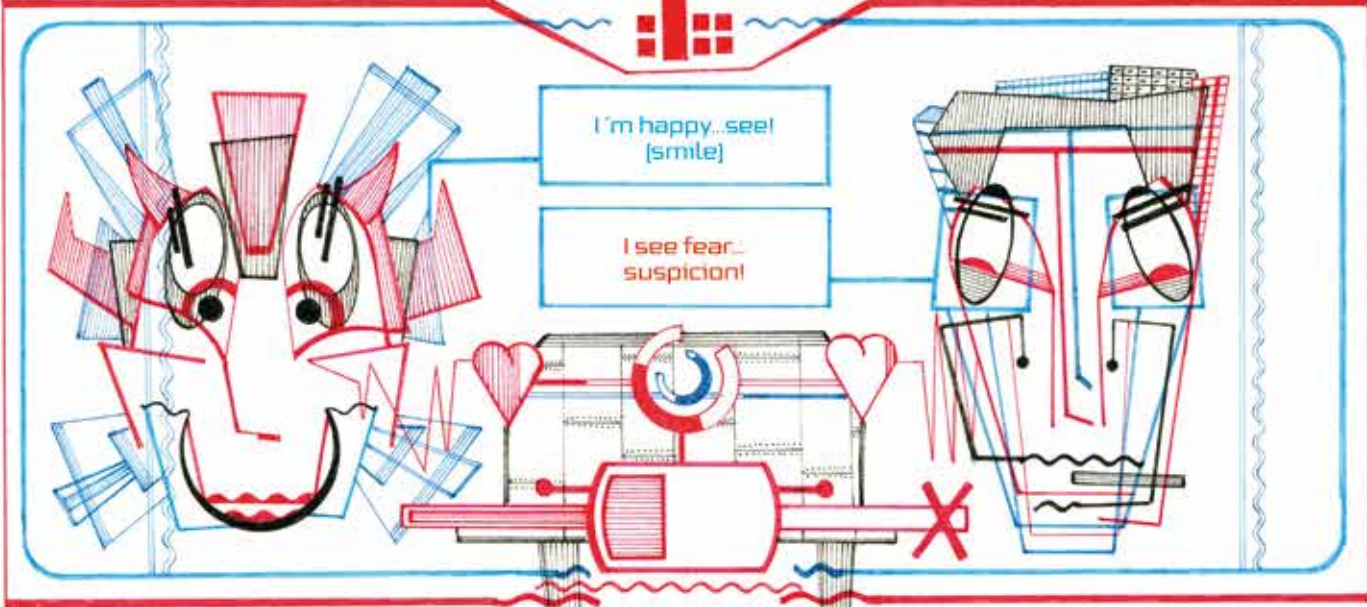
“Simon, are you doubting me?”

“Well, I just feel like... machines don’t lie.”

Silence settled over them as the atmosphere grew awkward at once. Joyce didn’t have to look at the annotation to know that her face probably said *Disappointment 10*. With trembling hands, she typed rapidly on her keyboard, trying to explain, but a red exclamation mark popped up on the screen immediately after she pressed *send* – Simon had blocked her.

The rage that had washed over her like a ferocious tide gradually faded away, leaving her exhausted, confused and utterly heartbroken.





She suddenly remembered Boss May, the AI system that supervised all the emotion annotators of *Heart2Heart*. Whenever Joyce had questions, she always turned to Boss May for help, and her virtual guardian had never failed her. *If someone knows what is happening, it will be her*, Joyce thought.

There must have been millions of people trying to reach Boss May at the same time, but she picked up Joyce's call immediately.

The projection of Boss May in a white pantsuit floated in the air, looking just like the Butterfly Mother, the female goddess from the folktales that Joyce had grown up with. She was sitting before an enormous curved screen. Varicolored beams of light shone in the background, crisscrossing and blending into each other, forming beautiful and complex mazes.

"It's been a while since we last talked, Joyce. How are you doing?"

"Not so great, Boss May. Did the system punish me because I slacked off at work?"

"Punish? What happened?"

"Simon... *Heart2Heart*... The machine kept on giving my emotions the wrong labels."

Boss May appeared to have realized something, but Joyce couldn't read her emotion. She did not have the authority to access the annotations of Boss May's virtual expressions. Boss May displayed some data on the big screen: there was Joyce's face and Simon's face, connected together by rays of colorful lights. When she adjusted the time axis, the expressions on those two faces underwent rapid changes as well.

"Joyce, listen to me. Don't be mad. It's not your fault."

"I'm not mad. I'm just upset."

"I can tell. It's written on your face."

“So, my emotions are correctly annotated to you, right? Our workshop manager told us that if we slacked off at work, the AI system would punish us on social media. But what happened with Simon was different...”

“I told you already, it’s not your fault. Simon’s the problem.”

“Simon? Are you saying that he could actually read my emotions correctly, but he deceived me on purpose? Why would he do such a thing? If he wasn’t ready to go no-filter, he could’ve just told me.”

“No, a different kind of deception.”

“What is it then?”

“Simon doesn’t actually exist.”

“Excuse me?” The *Shock* on Joyce’s face shot up to an unprecedented level of intensity.

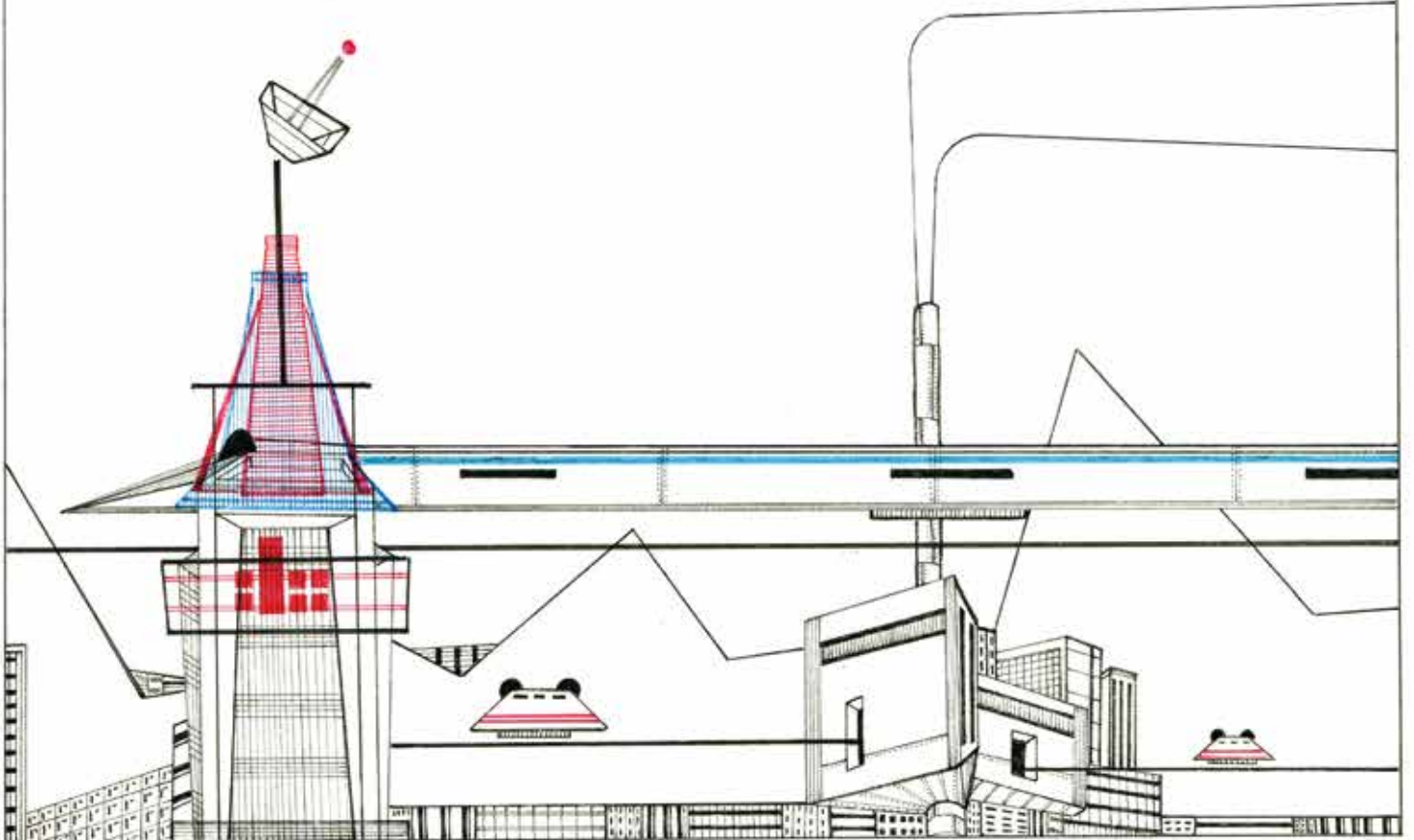
“Or rather, Simon isn’t human. He’s only an avatar created through algorithms to approach you and deceive you into purchasing more virtual goods and services.”

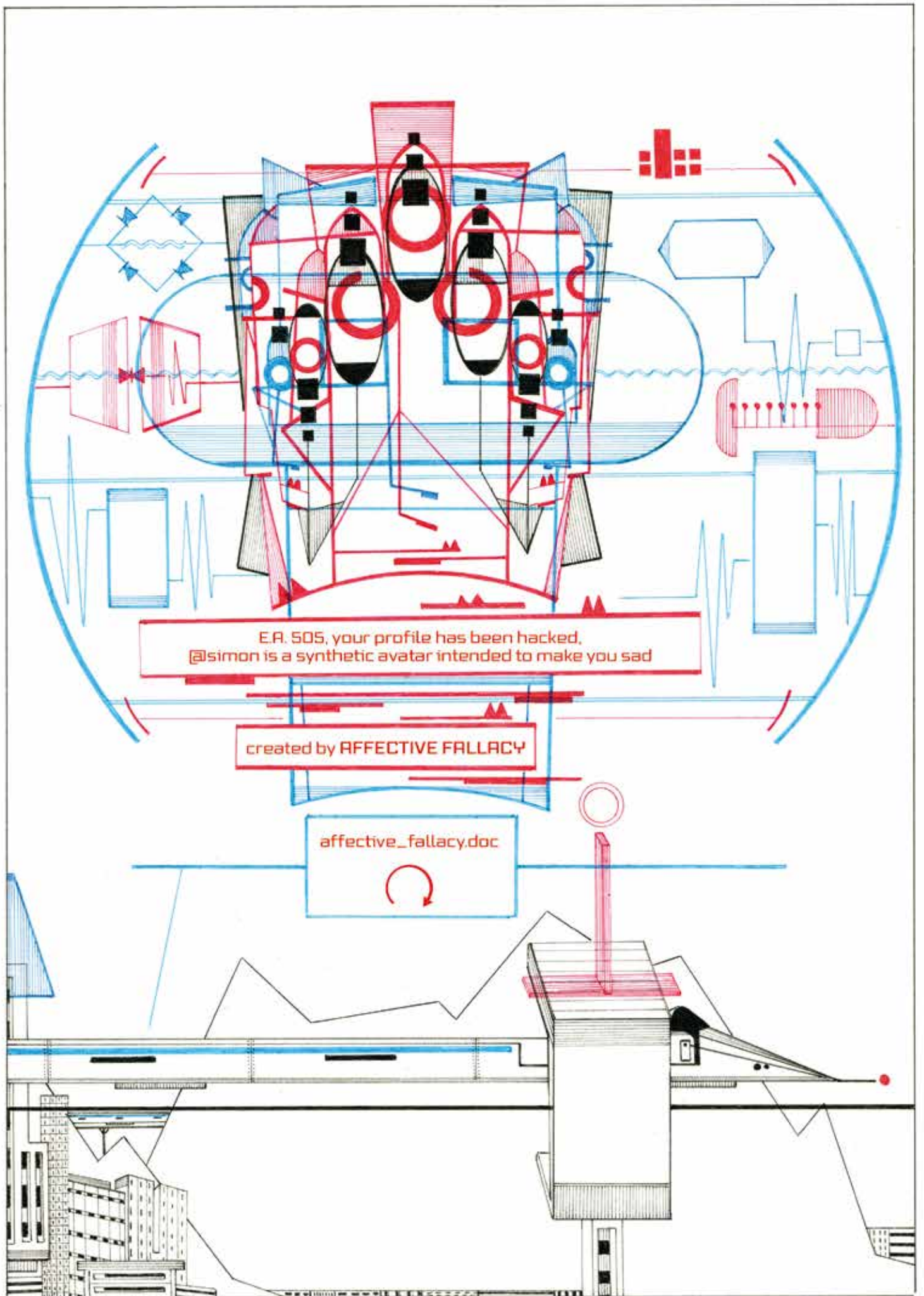
“But he seemed so...”

“*Real*. I know. AI puppets like Simon are everywhere online. You’re not the only one who’s been fooled.”

“If Simon was truly machine-generated, why would he make mistakes then?”

Slightly levitating off the ground, Boss May spread her arms out wide like a real butterfly. Behind her, the screen lit up in a startling red glow.





E.A. 505, your profile has been hacked,
@simon is a synthetic avatar intended to make you sad

created by AFFECTIVE FALLACY

affective_fallacy.doc



“Recently, there has been an organized mass cyber-attack. The hackers, however, aimed their fire at humans instead of machines. The human brain is more susceptible to external influences than machines, especially the area in charge of recognizing and appraising emotions. When the external environment puts pressure on an individual, their cognitive processes become distorted. We call this phenomenon ‘affective fallacy,’ which is also the name of the hacker group.”

“What do these hackers mean to achieve?”

“They claim that machines are stripping away humans’ right to free interpretation of their emotions and reducing humans to animals that can only communicate feelings through algorithms. They say that without genuine emotional expression, humans can never feel true happiness again.”

“I don’t understand. I am only an ordinary worker. Why did they target me?”

“You’re only one of millions of their targets. An AI puppet like Simon – or, rather, an emotional virus – can be easily duplicated and customized according to the needs of each individual target, and strike their heart with precision. Look at the red lines.”

Boss May pointed to the map on the screen behind her. Parabolic red lines, like missile trajectories, extended across continents and oceans. The ends of the lines exploded like fireworks as they landed on certain locations, radiating out to reach more specific areas. When the map zoomed in, Joyce saw that the vast majority of the locations under attack were labeled with the *Heart2Heart* logo.

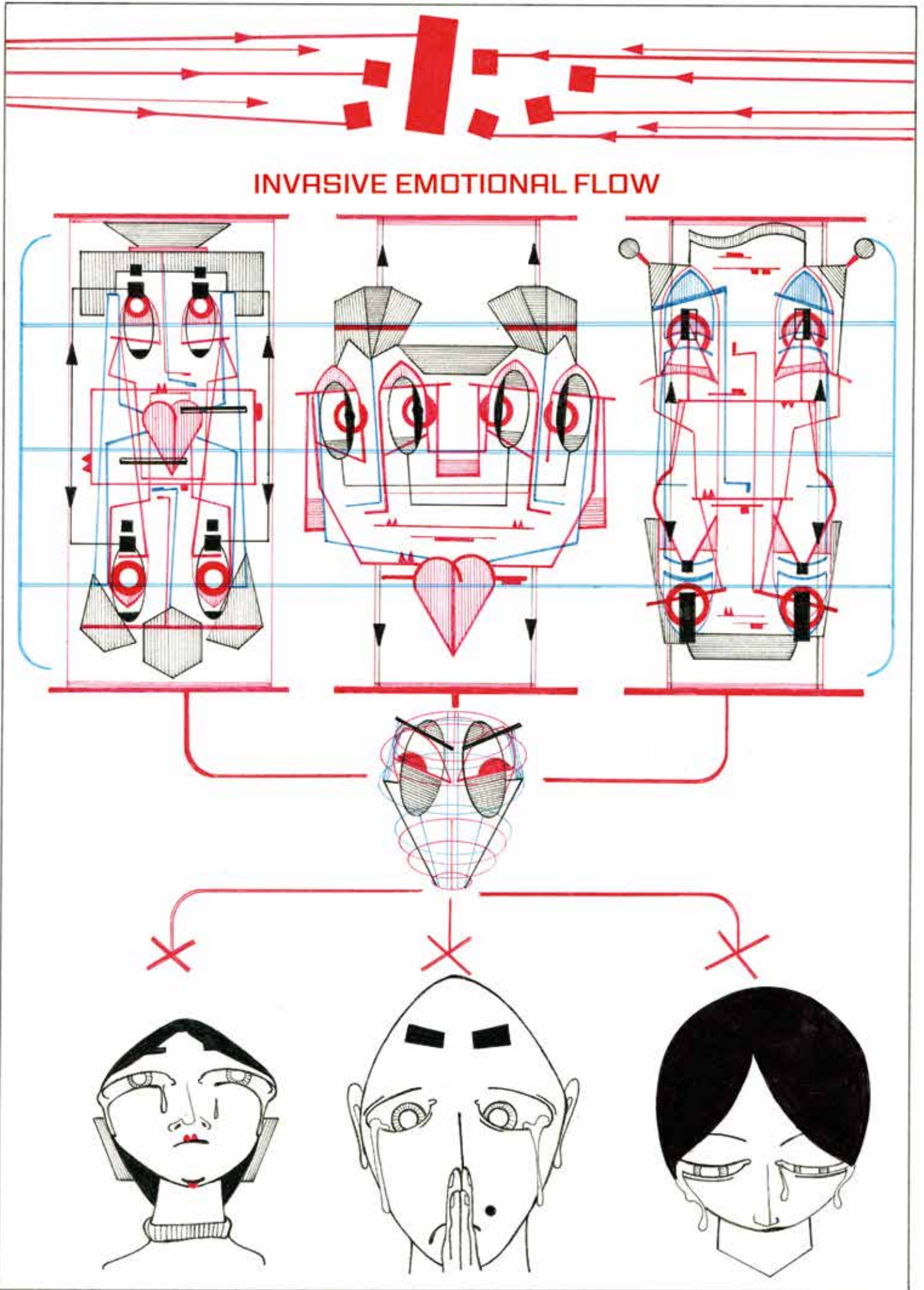
“What are these lines?”

“Invasive emotional flows used by the hackers. You see, the world isn’t as rational as we imagine it to be. Most of the judgments and decisions people make are driven by emotions. The hackers can influence the entire world by controlling emotional flows.”

“Upsetting me is also a part of their plan?”

“Yes, Joyce. What happened to you is not as simple as it seems.” Boss May gazed at Joyce. A set of numbers flashed across her face and disappeared, reminding Joyce of fleeting, subtle microexpressions beyond what the human eye could identify. “I have good news and bad news for you. Which one do you want to hear first?”

Joyce’s pulse quickened. “Bad news first.”



Now she could tell that Boss May was smiling, even without the annotation. An image appeared on the screen. She immediately recognized that it was the last image that she had annotated at work today: the boy in front of the monastery.

“You labeled him *Happiness 4*, right?”

Joyce studied the image closely. Far in the distance, out of focus, she sees it’s a monastery for secluded monks. The boy’s eyelashes were tearstained. He is about to shave his hair. Everything was different from the warm, fuzzy first impression that the image had left on her. *He is giving up on this world*, she realized. *The boy is saying his last goodbye to society*. No way the emotion on his face could be *Happiness 4*. She had made a stupid mistake.

“Are you firing me?”

“I knew you would say that. How should I put it? You won’t be staying in this job for long, but it’s not because of anything you’ve done. Machines have become so smart that they have learned, from human experience, how to self-evolve. In some ways, they already know emotions better than humans do. The entire industry of emotional annotation is going to vanish soon.”

Joyce’s face fell. Instinctively, she imagined two annotations popping up next to her own face: *Frustration 7* and *Anxiety 8*.

Boss May extended her arms toward Joyce. White wings with translucent shimmering scales materialized as she moved and wrapped herself around Joyce's body, as if to give her a virtual hug of comfort.

"Well, the good news is that you will have a new job. Your skills from annotating emotions won't go to waste. You can help do the work that machines cannot yet do."

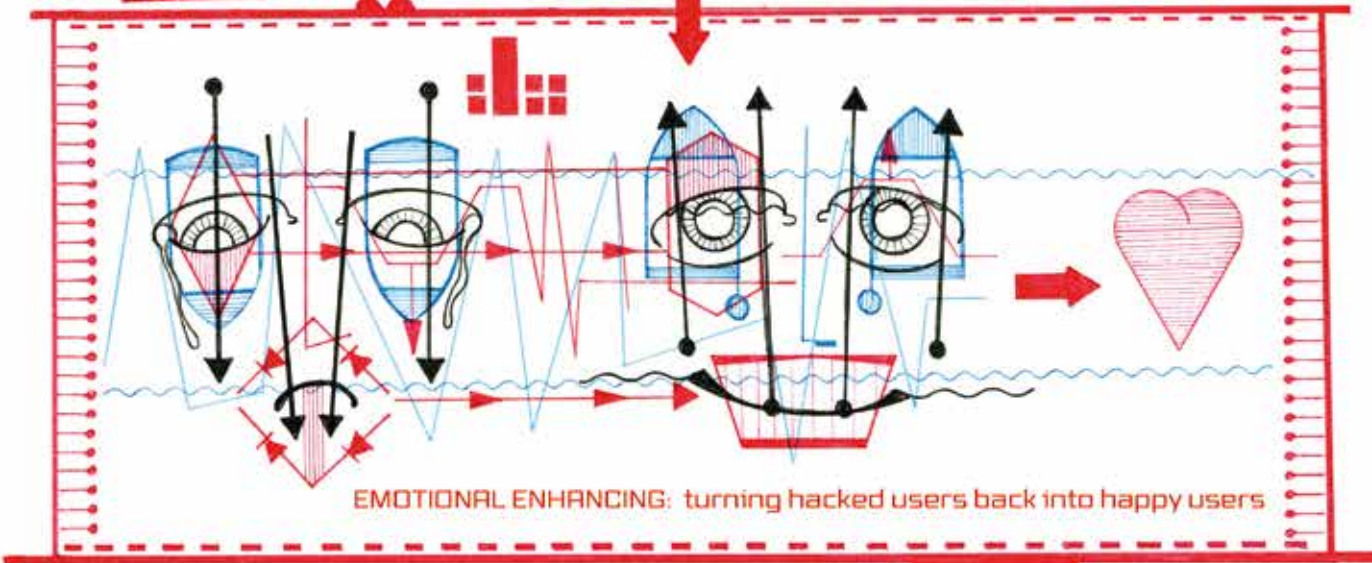
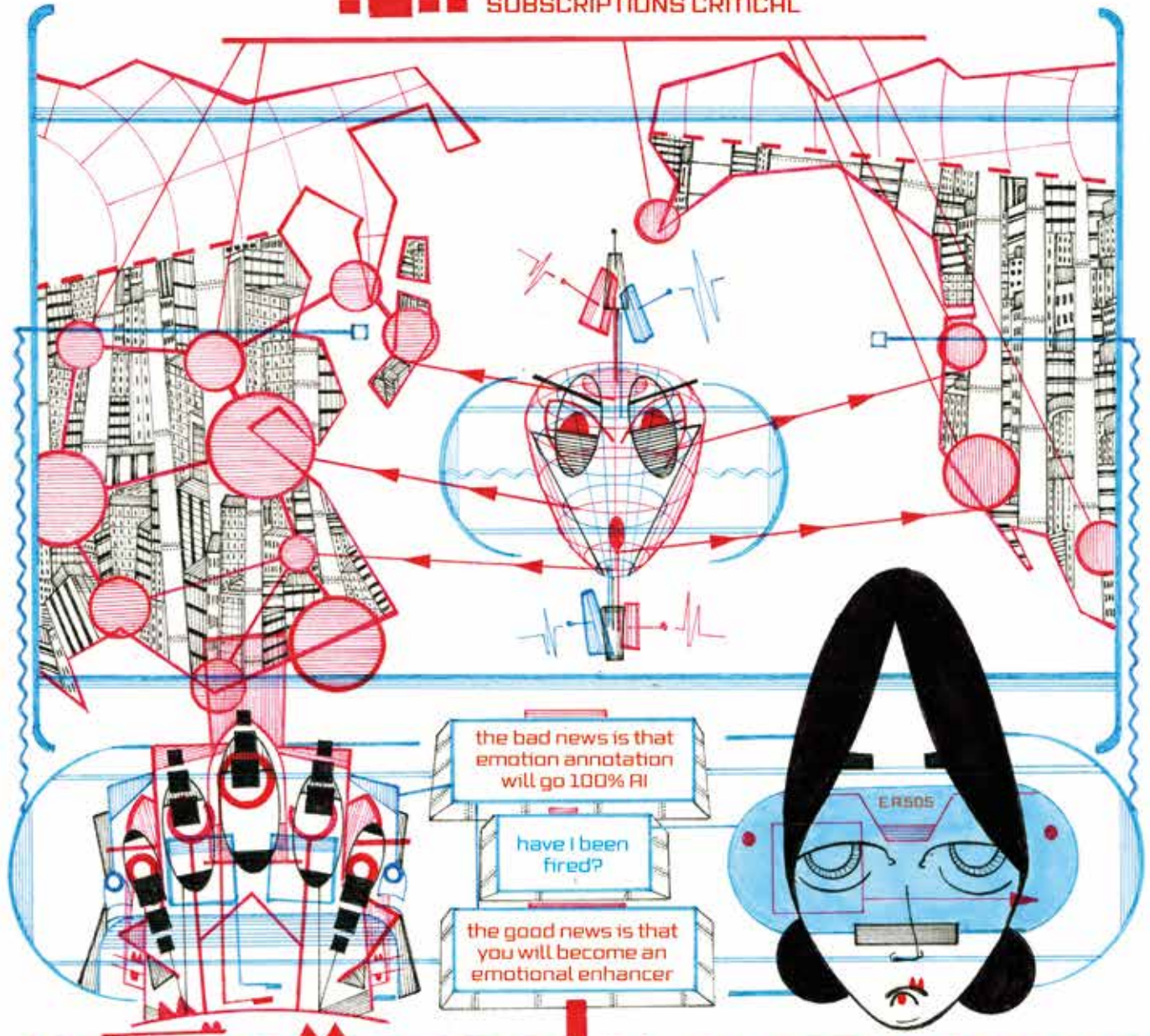
Joyce looked up in confusion. "Such as?"

"The cyber-attack has resulted in affective fallacy happening across the world. Many people have, in consequence, developed all kinds of mood disorders. Depression, mania, delirium... The global suicide rate has also increased. Your accurate judgment of emotions and your strong empathy can be of great help. You would be working together with the AI to create virtual avatars that can improve people's mental well-being, since the AI is unable to do it alone."

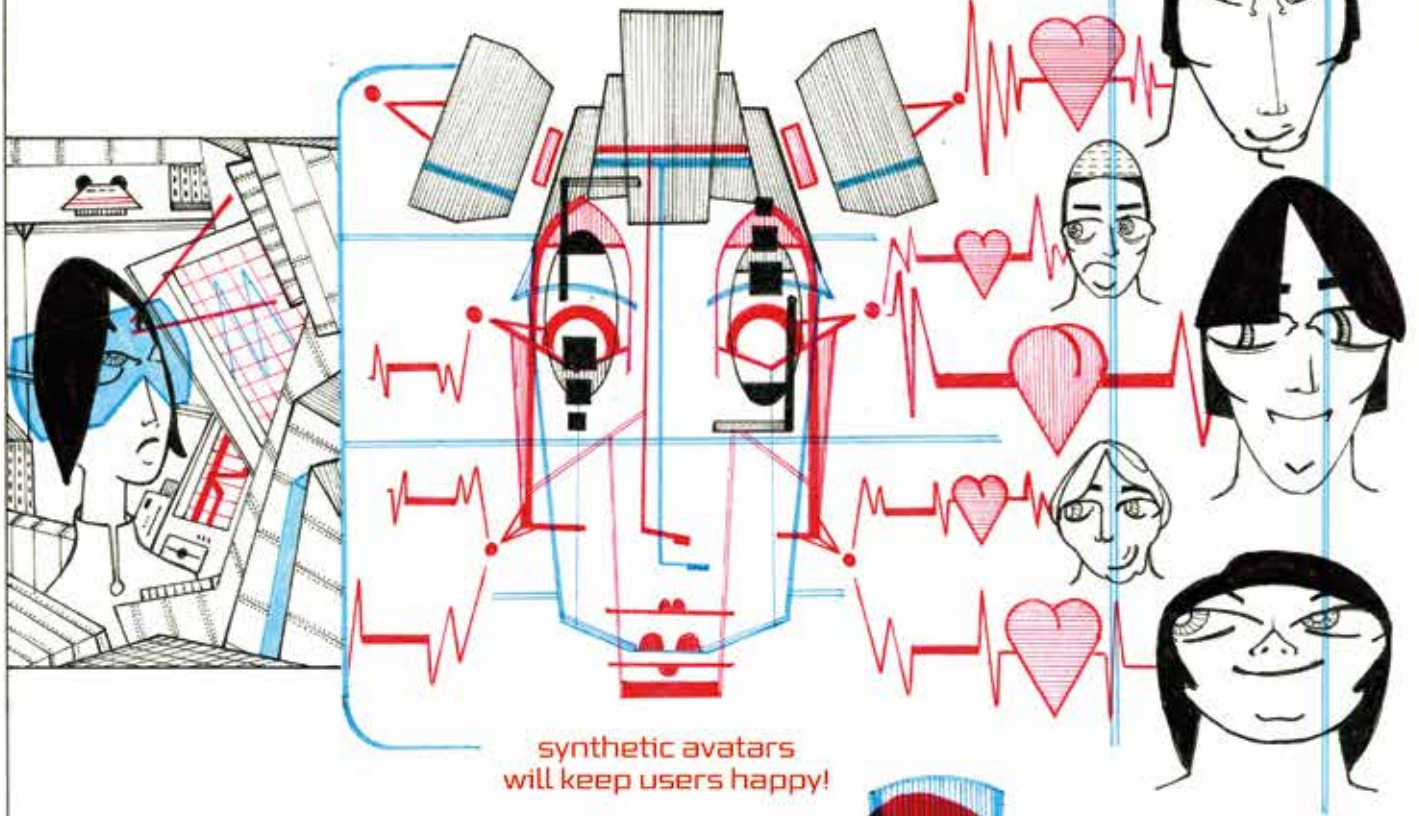
"You mean..."

"Yes, Joyce, you're leaving this place. There are bigger, newer and better workshops waiting for you in the city."

62% of users affected
SUBSCRIPTIONS CRITICAL



an emotional enhancer helps AI create and control synthetic avatars



EMOTIONAL ENHANCER CONTRACT

name : JOYCE
new I.D. : EE 002
location : MEGABURG



ACCEPT

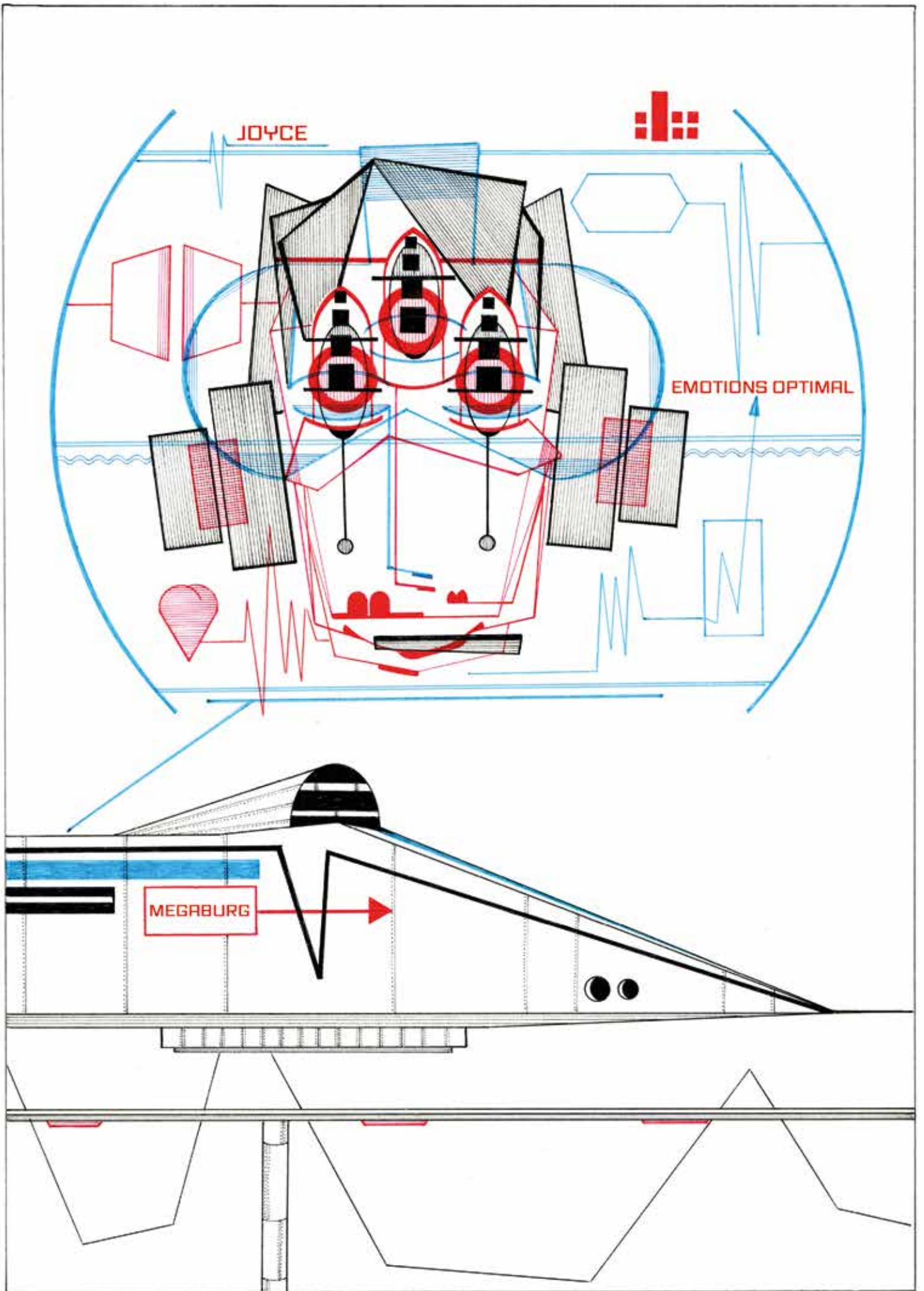
AVATAR UPGRADE

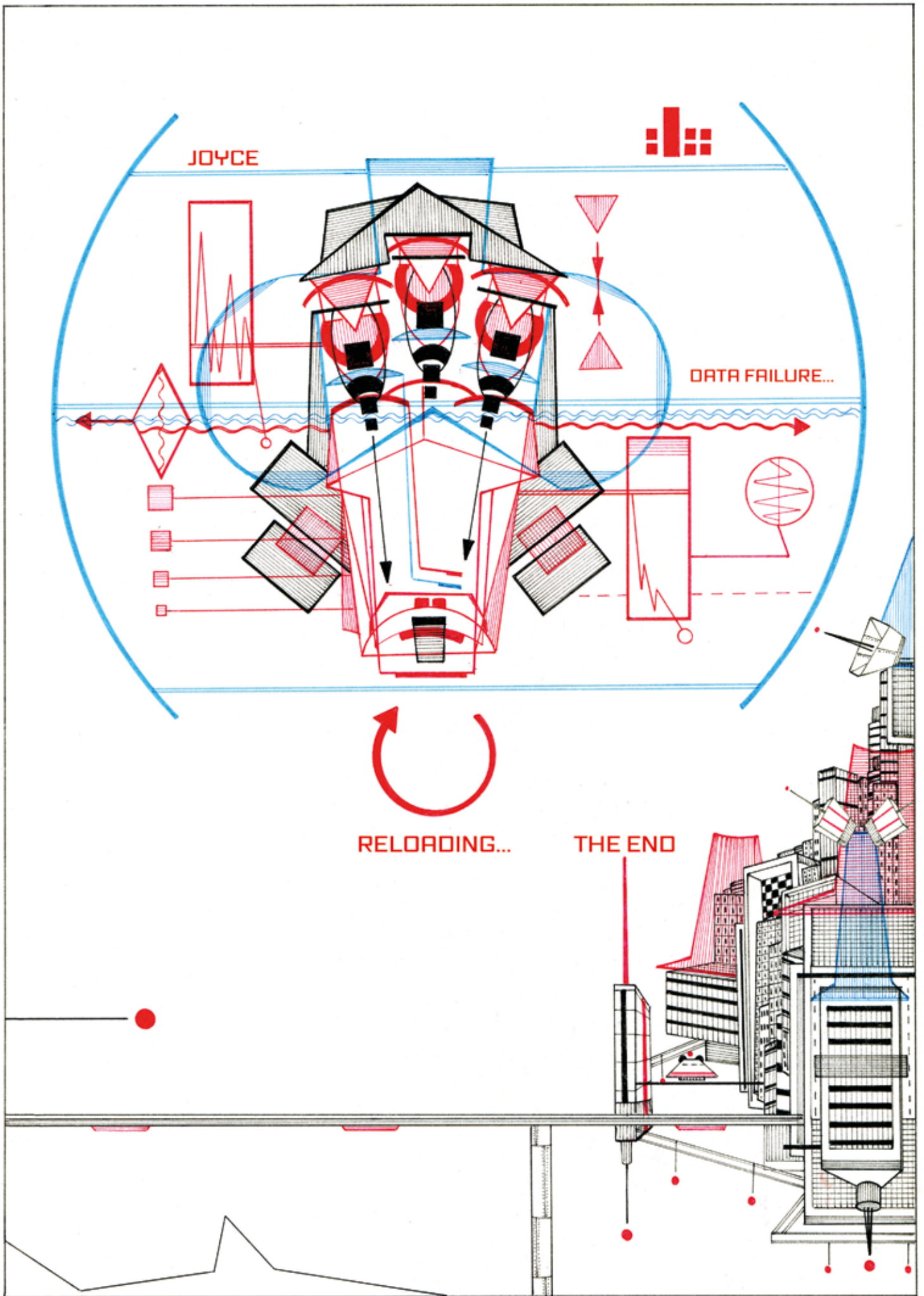
Joyce stared into the massive screen behind Boss May. Her face, reflected in the screen, overlaid the projection of a magnificent city glowing with lights. The expression on her face seemed to be changing constantly. She tried to label each of those emotions, but the simple task suddenly became difficult. The changes, subtle and complex, happened at once and disappeared before she could even recognize what they were.

Perhaps annotating emotions really was a job that only machines could do.

Can I really make other people happy when I myself am unhappy? Leaving my hometown and moving to the big city – would that end up as another failed attempt to escape from the reality?

A myriad of emotion annotations, each different in kind and intensity, emerged on Joyce's face. Then, one by one, like bursting soap bubbles, they exploded into colorful spots of light.





ABOUT US

Through a series of speculative scenarios and a short story set in a dystopian world, the Future Shocks website aims to provide food for thought on some of the most pressing challenges the world faces today.

The Future Shocks were originally published as part of the World Economic Forum's 2019 *Global Risks Report*. The short story "Affective Fallacy" and the illustrations that accompany it were commissioned as part of the World Economic Forum's Global Risks Platform.

The World Economic Forum would like to thank Platform partners Marsh & McLennan Companies and Zurich Insurance Group.

For more information, please contact
GlobalRisksReport@weforum.org.