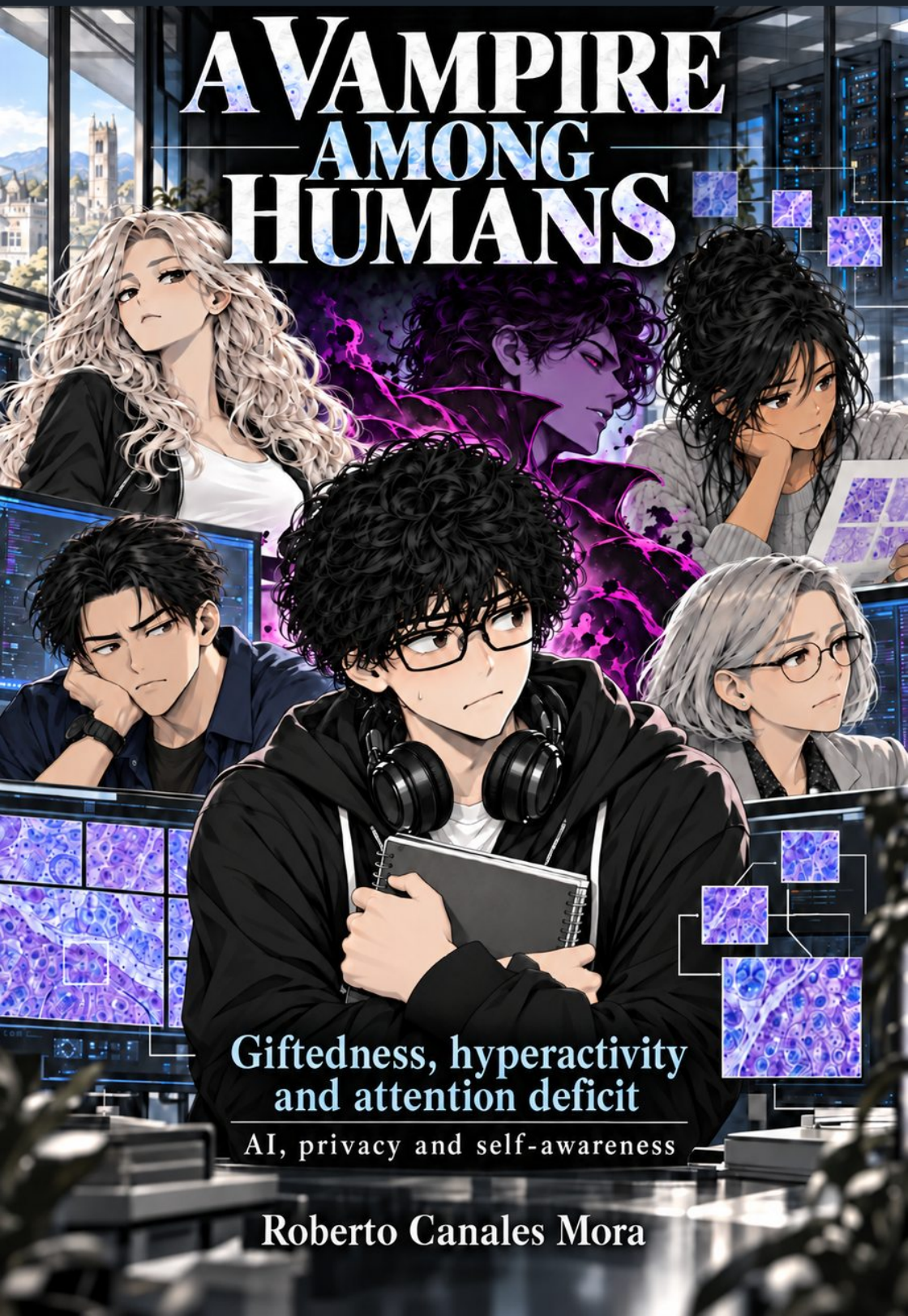
BOOK 1 · VERSION 1.0



Intellectual Property and Use

Copyright © 2026 Roberto Canales Mora.

This work is marketed under the ownership of THE AUDIENCE MEGAPHONE SYSTEM SL, tax identification number (CIF/NIF) B87572772.

Author: Roberto Canales Mora.

This book was written, structured and reviewed with the help of Codex. Its visual resources were created using various techniques, including generation or adaptation with ChatGPT and composition through code.

All rights reserved. Reproduction of this work, in whole or in part, its incorporation into training systems, repositories, databases or automated exploitation services, and its public distribution, communication, transformation or commercialisation are prohibited without the express written permission of the rights holders, except as permitted by applicable law.

The trademarks, trade names, products, services and technologies mentioned belong to their respective owners and are cited for informational, educational or professional analysis purposes.

Book 1 · Version 1.0 in British English · 29 October 2026. Thirteen chapters. The figures, companies and incidents in the story are fictional.

Why I Wrote This Comic



One task selected among several windows and notifications.

My name is Roberto Canales. I am a communications engineer, I have a master's degree in artificial intelligence, I run a business and I am a very creative person.

Where I come from

As a child, I was identified as gifted and diagnosed with attention deficit. Anyone who knows me will know there is no need to diagnose my hyperactivity. We all know I have it.

What people ask me

People whose behaviour resembles mine often approach me. Sometimes they want to talk about themselves; other times, about a friend, a child, a nephew or a niece. They ask me to tell them a little about my path: how I have made my way in life, the difficulties I have encountered and what I have done to manage them. Much of what I want to tell here comes from those conversations.

A Vampire Among Humans

In an article, I suggested that someone with certain atypical characteristics might feel like a vampire in a world of humans. They have abilities and shortcomings, and must learn to live with both. If they can develop the former and tame or channel the latter, they can have a very satisfying life. That comparison struck me as a good way to begin a story. What would happen to someone like that when joining a team and having to get along with people very different from them?

The original article

Read the original article in Spanish on linkedin.com

Two Kinds of Learning in One Story



Preparing an activity and reserving an hour for it in the week.

In this comic I want to tell two things at once. The first is Leo's journey: a brilliant person with behavioural difficulties who tries to adapt to a world he often does not understand and which does not understand him either.

Noticing, then doing something

Leo finds I've Finished My Studies — What Now?, another book of mine. Reading it makes him begin to look at himself a little more closely, and he decides to change some things. But there is a gap between noticing a problem and doing something about it. I want us to see how he crosses it: what he tries, where he goes wrong and what he manages to keep up when the enthusiasm of the first few days wears off.

An AI competition team

The second story is the team's. I could have chosen many settings, but artificial intelligence is very present in our lives and gives me plenty to work with. I have used a real challenge and dynamics that could arise in a laboratory, although the characters and their problems are invented. That way we can learn more than one thing at once: understand the technology a little better and notice behaviour we may recognise in ourselves or people close to us.

What I would like you to notice

After each part of the story, I will pause to discuss the technical problem and what is happening to the people. I hope it helps you see details you might otherwise miss. Everyone's life is complicated, and there will be steps forward and back. Many stories are just a snapshot of a path that could still get better or worse. Here we will accompany them for part of it. I hope it interests you.

Roberto Canales Mora

One Story, Two Readings

Leo's experience

Leo is a fictional person. His abilities, attention difficulties, hyperactivity and ways of relating to others do not represent everyone with these experiences. The vampire image is a metaphor, not a supernatural fact.

Reading the story and pausing to think

After each chapter, you will find an explanation of its technical problem and a reading of what Leo is learning. You can read them after the scene or return later. They draw on what has already been narrated, distinguishing what he does, what he interprets and what he has not yet achieved.

Knowing yourself does not solve everything

Practice makes some progress possible; there is also tiredness, setbacks and material problems nobody solves by willpower alone. The other characters have needs and decisions of their own. This work is not a diagnostic tool and does not replace professional assessment or treatment.

A real question, an invented story

Benchmarking WSI Privacy, Track 1, studies whether tissue images can be linked by patient without finding names or making diagnoses. ATLAS, Nexo, ARGOS, the funding, scores and incidents in the story are fictional. They do not describe the conduct of real participants.

From the First Day to the Result

01 No Names **17**

02 The Price of a Point **58**

03 The Gap **89**

04 Borrowed Time **109**

05 Before the Deadline **129**

06 Winning Alone **149**

07 A Different Way **163**

08 What Keeps the Edge **182**

09 Another Test **206**

10 Still Standing **227**

11 What Is Hard to Tell Apart **249**

12 This Far **263**

13 Second **274**

Before Opening the Programming Environment



Images grouped by patient before splitting the data.

If we were entering a machine-learning challenge together, I would start with something much less entertaining than choosing a network: reading what we are being asked to do. We can spend many hours improving a program that then fails to produce the answer we must submit.

What the platform gives us

Kaggle brings together competitions with different questions and formats. In a prediction competition, we receive data to develop a method and submit answers the organisers score with a metric. I would first look at the description, data, evaluation, schedule and rules. We need to establish which material we can use, where we are allowed to run it and how we can collaborate. With that in front of us, we can choose an environment on our computer or one the competition permits.

How we will decide on changes

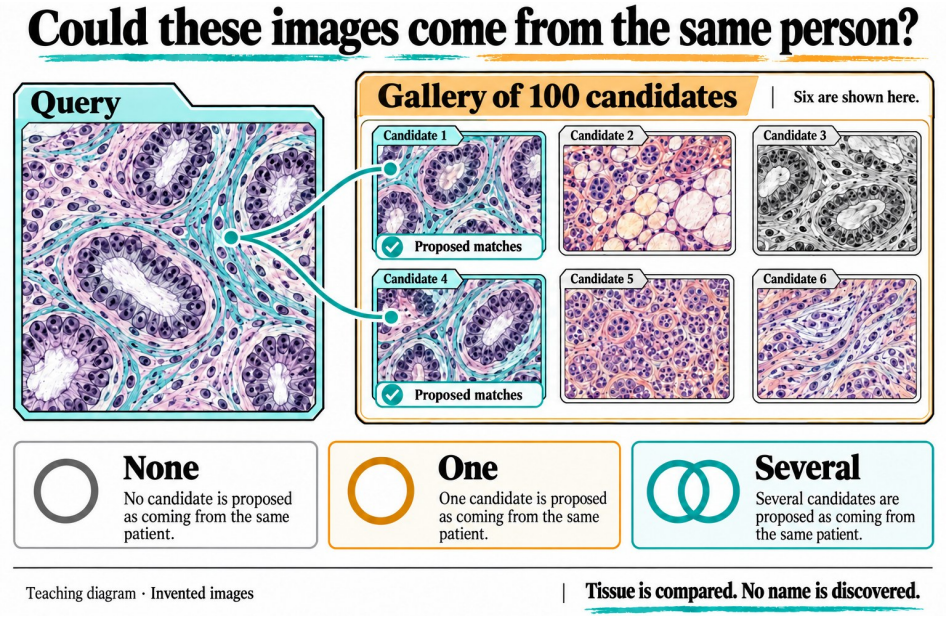
We need to set aside available examples for our validation: use them to measure what deserves to continue and what worsens. Looking at the rankings after every idea also ends up guiding development. Some competitions divide held-out answers between a visible leaderboard and a final one; we have not confirmed that specific arrangement for WSI Privacy. In any case, we should have our own check to work with before submission.

What our hours can pay for

This challenge offers scientific recognition. ATLAS's funding and Nexo's contracts belong to the fiction we are about to read. In a real project, I would budget storage, computation and people's time before promising tests. I would make a simple first baseline to measure its time and cost. From that, we can discuss improvements using execution data as well as our enthusiasm for trying them.

Sources: · Kaggle · The platform

The Query and Its Hundred Candidates



Invented images: a query can return zero, one or several matches.

Imagine two collections from which we have removed the names. If we can recognise samples from the same person in both, we have already built a link between their records. Other associated data could extend its reach. Here we will investigate that visual relationship between tissue samples.

What we must return

Benchmarking WSI Privacy, Track 1, presents a query image and a hundred candidates. We must select those from the same patient using their visual content. There may be zero, one or several matches. The challenge separates training and test patients, scores mean F1 per query and asks for a CSV containing the matches. We will keep the model that produced those answers so we can repeat the process; the required output is the selection of images.

How we know what to teach

In the known examples, we have codes with which to form pairs from the same patient and from different people. Those codes help us prepare and mark the exercise. The responding program will have the images, without that clue to the solution. Slides may look very different because of how they were scanned; we will need to investigate which visual signal preserves the relationship between them.

Testing the empty case too

I would prepare a controlled example of each situation, including a query with no matches. I would not remove a candidate just because it looks too similar either: the options may include the same slide. I would follow identifiers from the query to the final selection and check which files the program is using. That is a fairly inexpensive way to find a connection error before blaming the network.

Sources: · The real challenge · Learning similarities · The risk of linking

Which Part of the Image the Network Will See



Images grouped by patient before splitting the data.

A WSI is the digital image of an entire tissue slide. It can be enormous for the program we want to train. Before asking it to learn, we must decide which regions to pass to it and how we will know where each came from.

Extracting and looking at a few samples

A library such as OpenSlide lets us read regions at different resolutions. I would start by extracting fragments containing tissue and recording the image, coordinates and scale. Then I would open them: a reduced view may have lost detail, and many nearby crops may repeat almost the same thing. I would also measure extraction time. We can select regions within an image while keeping the difficult official queries we will later need to answer.

Splitting by patient first

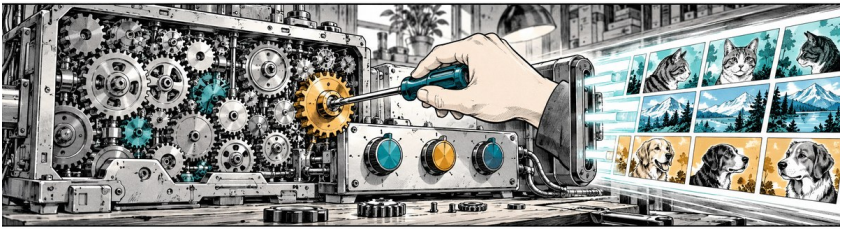
Before dividing development into training and validation, we group each person's images. All their fragments must stay on the same side. If nearby crops go to both, we may mistake recognising familiar material for the ability to work on new patients. The split file organises that separation. When a transformation needs to learn statistics from the data, we fit them using the training group.

Saving preparation with the model

Resizing, converting channels and normalising values prepare the input the network expects. During training, we can try rotations or moderate colour changes that preserve the relevant signal; we are still working with those patients, though presenting their images differently. For evaluation, we need a reproducible recipe compatible with the network, without necessarily repeating every random change. Saving it alongside the model will tell us what it was receiving in each test.

Sources: · WSI images · Separating patients · Avoiding leakage · Transforming examples

What We Keep from an Image to Compare It



A metaphor for adjusting part of an already trained model.

Comparing all the pixels gives us a baseline, but a change in position or appearance can alter it considerably. We can try another route: turn each image into a list of numbers that keeps useful features for calculating its similarity to another.

What two families of networks do

A convolutional neural network, CNN, applies learnt filters to areas of an image and combines the responses in successive layers. A Vision Transformer, ViT, divides the input into patches and uses attention to calculate how parts influence each other. The names tell us something about the procedure; choosing requires seeing what information each keeps in our images and what errors it makes.

Starting with available weights

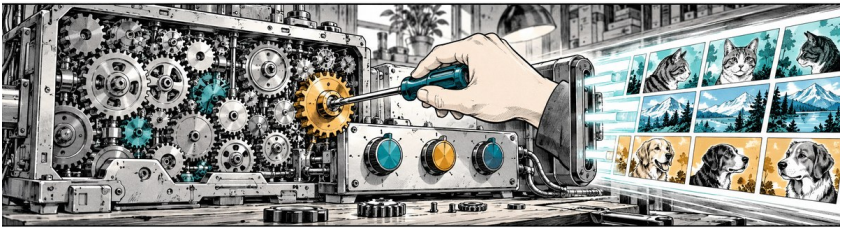
I would look for pretrained weights whose use is permitted in the challenge and load the preparation matching that version. They are values the network has already learnt from other examples. Keeping them fixed and taking an internal output, the embedding or representation, allows a relatively inexpensive first comparison. We will then check whether those features preserve the particular differences between patients we want to recognise.

What we do with many crops

If a sample supplies several fragments, we still have to decide how to join them. Averaging representations would be a simple hypothesis, though it may dilute an unusual detail. Comparing groups would preserve other information at greater cost. I would keep a baseline and change one choice at a time, opening specific cases during review. I would want to know whether similarity follows the tissue's origin or a scanning detail that works here and stops being useful in another collection.

Sources: · Convolutions · Vision Transformer · Starting models · Adapting a network

What We Teach It During Adjustment



A metaphor for adjusting part of an already trained model.

If the fixed baseline falls short, we can try adaptation. Here, training means changing the model's internal values using examples and an error measure. We must specify what we want to bring closer, what must be separated and which part of the network we will allow to change.

Preparing pairs that teach the relationship

In contrastive adjustment, we present pairs from the same patient and from different patients. Depending on its formulation, the loss function penalises representations too far apart for the former and too close for the latter. The label indicates the relationship we seek. When responding, only images will arrive. We should also look at how we choose pairs: repeating easy ones may leave the confusion we wanted to correct unpractised.

How much of the model we adapt

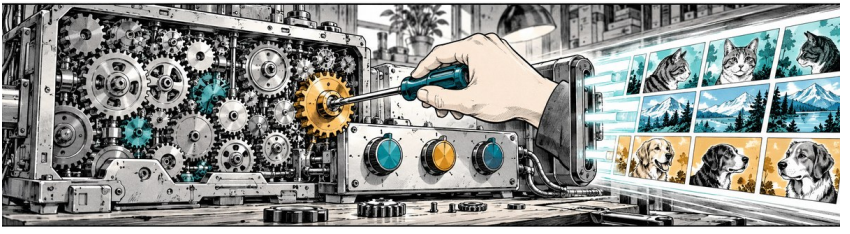
We can keep the network fixed and train a small output, or allow some previously learnt blocks to change. That readjustment is known as fine-tuning. The optimiser uses error to decide parameter changes, and the learning rate controls their size. Smaller steps may require more passes; time will also depend on the images, machine and stopping rule we set.

What I would decide before starting

I would agree how many tests we will run, with which pairs and which comparison would lead us to continue. I would save the initial weights, configuration and selected state. The question remains whether we can compare new patients. If training improves while validation worsens, we have a concrete reason to review that configuration. We will also have a record of what we tried, so the next idea begins by looking at the failure that has just appeared.

Sources: · Contrastive adjustment · How it learns · Adapting a network

Looking at Its Performance on Held-out Examples



A metaphor for adjusting part of an already trained model.

Watching training error fall encourages us to leave the machine running. I would open validation too. Overfitting happens when the model specialises in the examples it studies and responds worse to others; more passes can increase that difference.

What we can restrict

Freezing blocks reduces the parameters we adapt. Regularisation imposes other constraints, such as penalising large weights or using dropout, which temporarily removes some activations during training. We must choose its strength. Penalising weight size and limiting distance from the original weights are different constraints; we should specify which we use. These techniques help us study adjustment without answering a privacy question on their own.

What stopping in time means

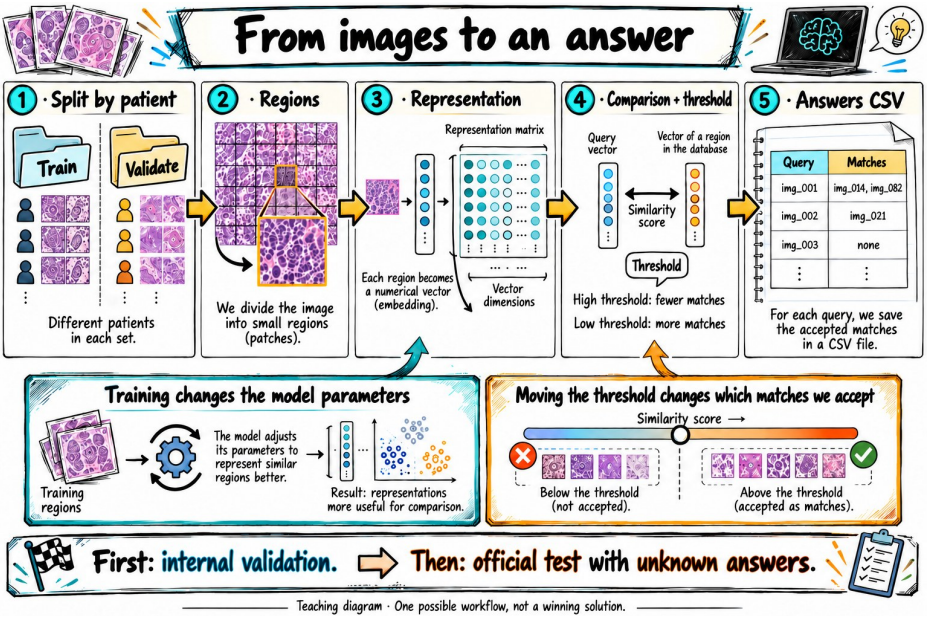
We can monitor a validation measure and stop training when it ceases to improve according to a rule. We need to set what counts as a change, how many passes to wait and which state to recover: some tools save the last one by default. Jun can also set a maximum time to preserve another colleague's slot. That time protects a resource; validation-based stopping responds to what the model is learning.

What remains to check afterwards

Validation helps choose regularisation, augmentation and duration. To check the chosen recipe, I would hold out other patients if resources allowed, or prepare group-based splits. Calling a test independent when we consulted it to decide would give a misleading impression of what we know. Before spending another night, I would examine which errors change and whether the improvement holds. Saving the failed attempt will let us recognise it when the idea occurs to us again.

Sources: · Regularisation · Stopping in time · Separating patients

Turning Similarities into Matches



Illustrative workflow: training changes the network; adjusting the threshold changes the decision.

We now have representations and can compare them. We still need to decide which ones enter the answer. This step can also change the result considerably, even while keeping the trained network fixed.

Where we put the minimum

Cosine similarity compares the orientation of vectors and gives a similarity measure. A high figure still needs an acceptance rule; it is not a probability on its own. Using validation, we choose a minimum: the threshold. Lowering it can recover missed links and add false ones. Raising it removes some false matches and loses true ones. It is a decision about the answers, without further network training.

When I would try a combination

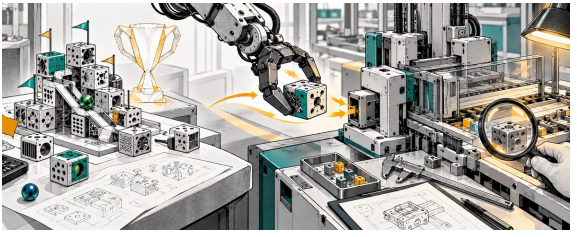
We could compare CNN and ViT, use multiple scales or add trees that chain questions about pairs' visual features. If we combine scores, we need compatible scales and weights chosen through validation. First, I would look at whether the models supply different signals: repeating the same error twice may merely cost us more. If we train a combiner, we use predictions for patients the base models did not see during learning, to avoid artificially easy successes.

What we gain for the extra work

I would keep each variant and compare the ensemble with the best baseline on the same queries. I would open the cases that change and measure time and memory. Multiple models, scales or crops can make the process much more expensive. We must decide whether the improvement is worthwhile. At the end, representation, combination and threshold will be fixed; we can then take that choice to the final exam without using its answers to keep adjusting it.

Sources: · Measuring similarity · Choosing the threshold · Combining models · Avoiding leakage

Checking the File We Will Submit



A technique moves to another task and needs new tests.

The procedure has been chosen. Before submitting, I would follow a query to its output and check that the file contains every requested query. A CSV may be small after considerable work; one format error can ruin the submission.

Walking through an F1 example

Suppose the genuine pairs are A and B and we propose A and C. We have one correct match, one false addition and one omission. Precision is one of two proposals and recall one of two genuine pairs; F1 combines both and gives 0.50. The challenge averages by query. For the no-match case, we need to consult the organiser's convention, because a library may score it differently.

Reviewing the template before sending

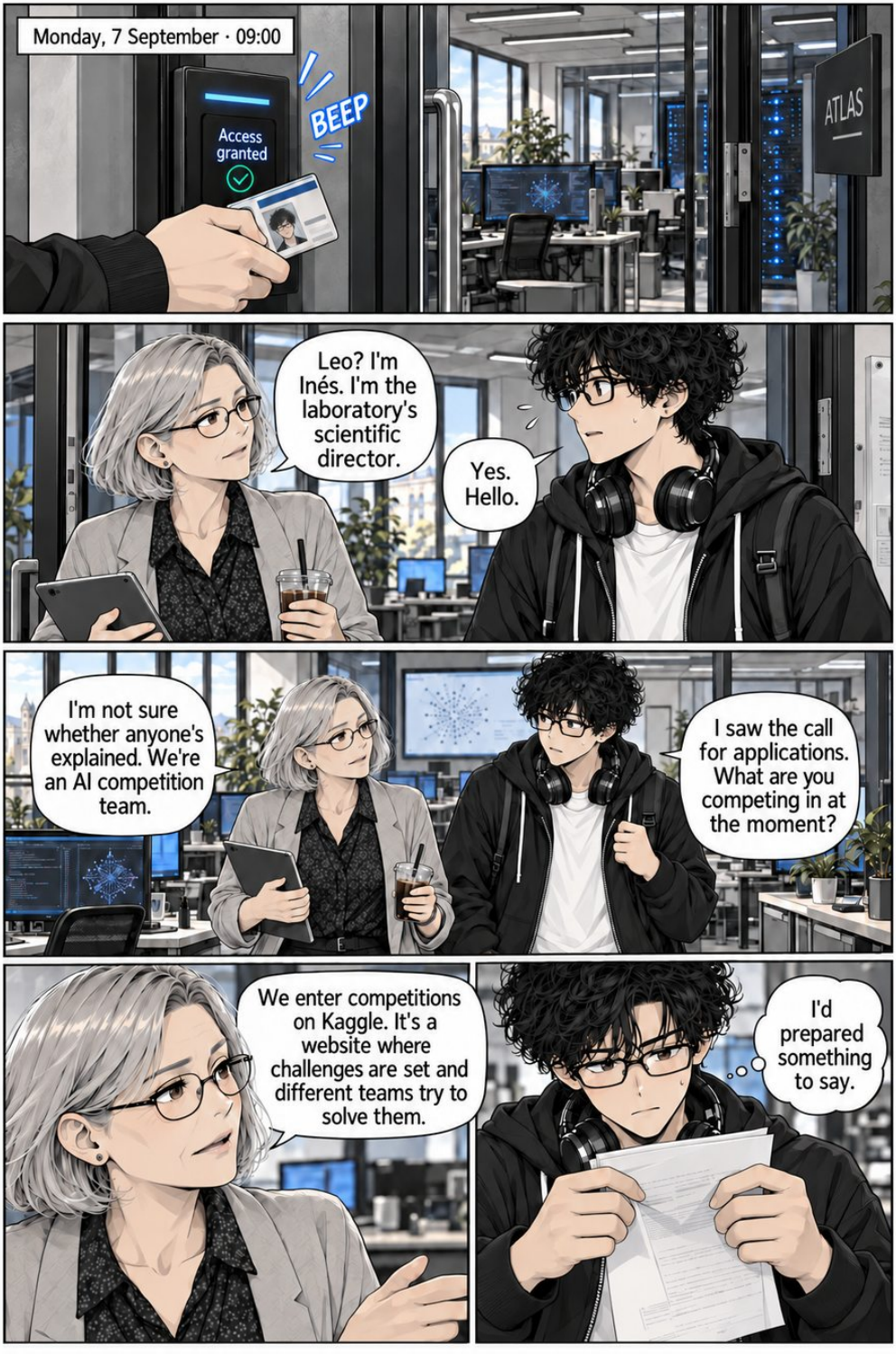
The CSV associates each query with the identifiers of its matches. I would check coverage, gallery and format against the official template. One detail still needs clarifying: the description specifies an empty field for ‘no matches’, but its example uses NONE. I would check before submission to avoid a different interpretation when it is read. We will keep code and weights to repeat how we obtained those answers.

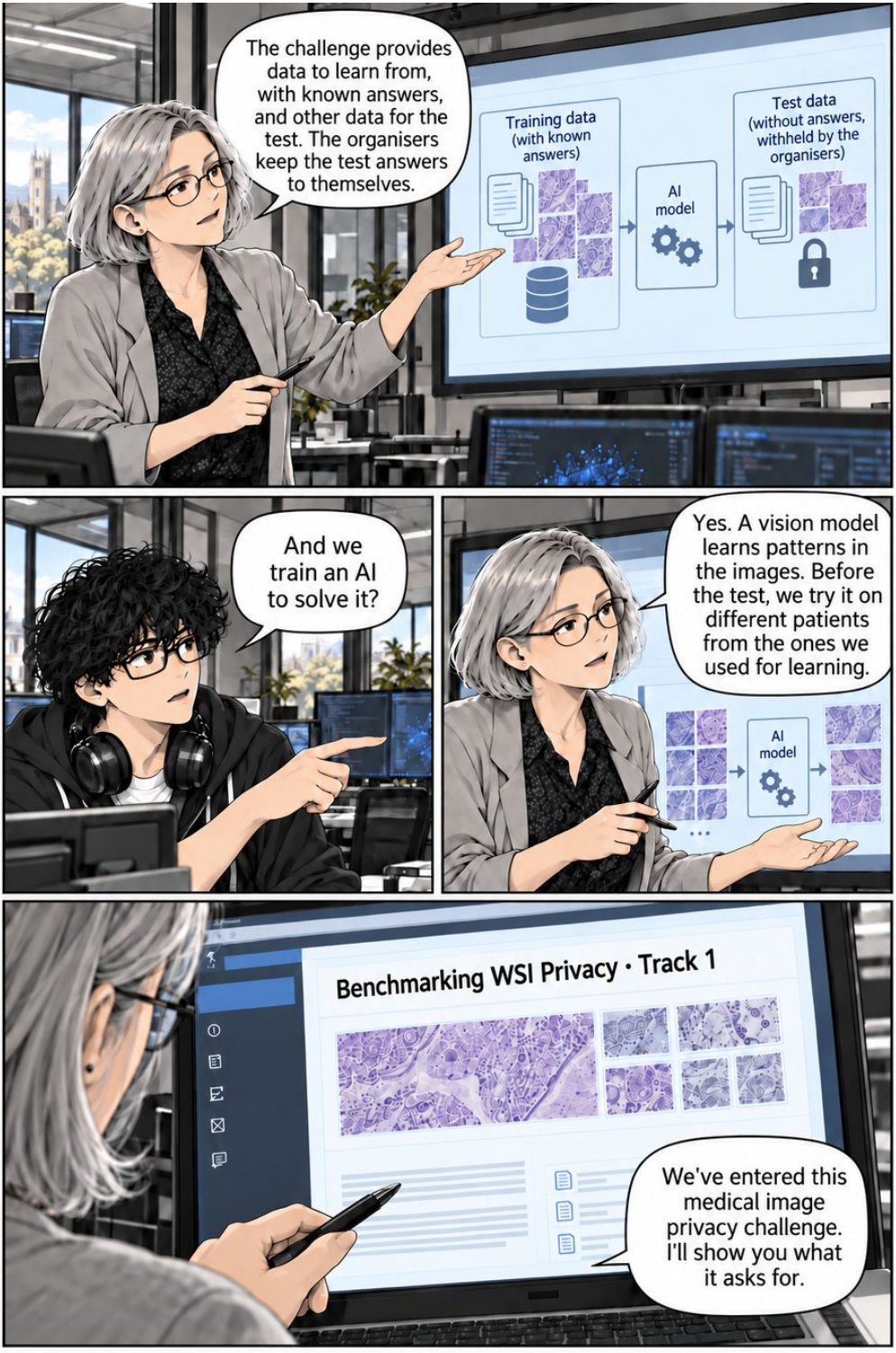
What we need to return to a failure

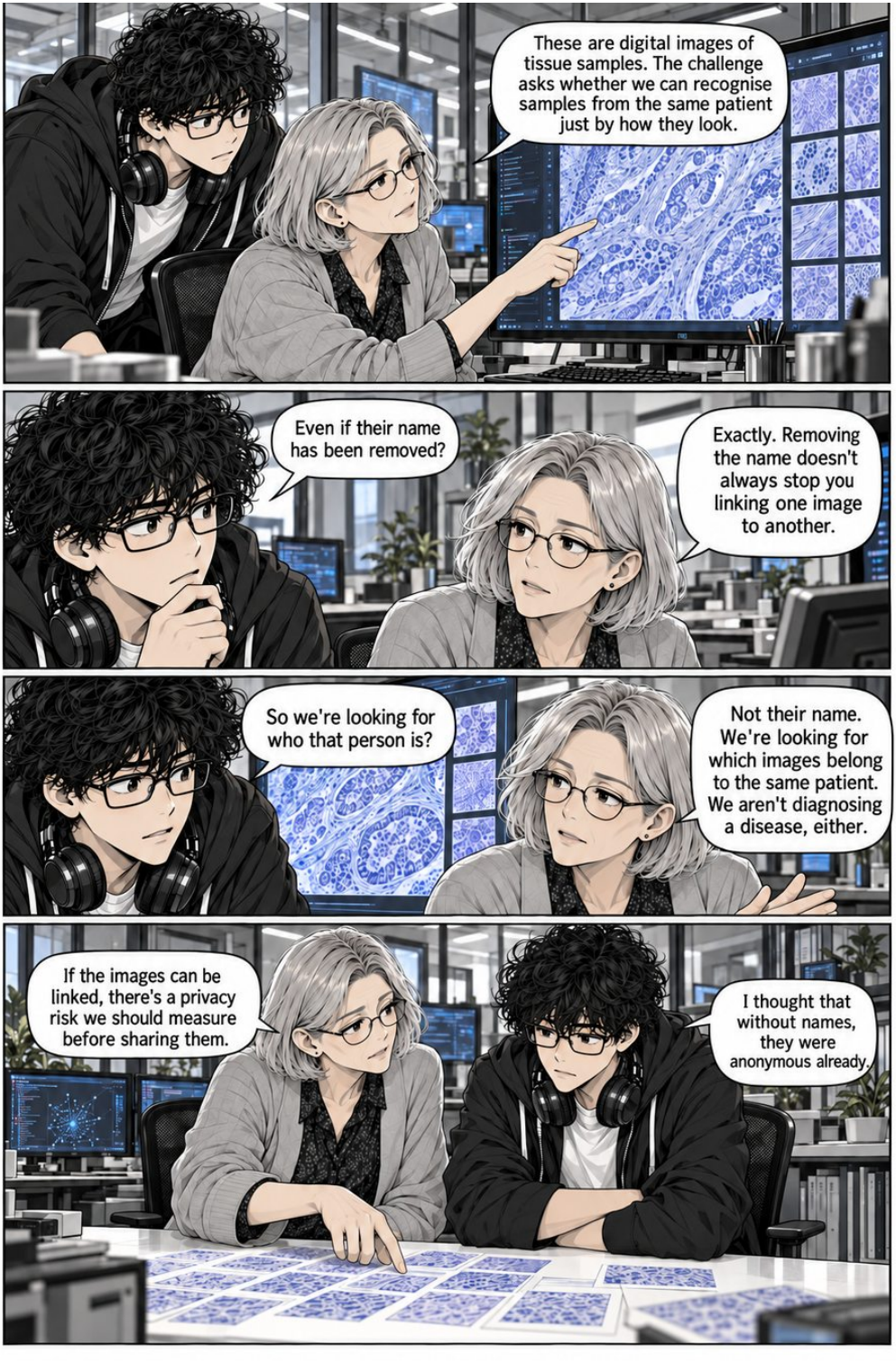
The pipeline runs through image reading, preparation, representations, comparison and selection. I would keep versions, the split, weights, threshold and execution logs. A cache can be reused if it corresponds to those inputs; otherwise, we could compare old outputs without noticing. Besides F1, I would look at specific errors, time and memory. Then we can explain what we measured and review the next change from a recognisable procedure with a defined technical scope.

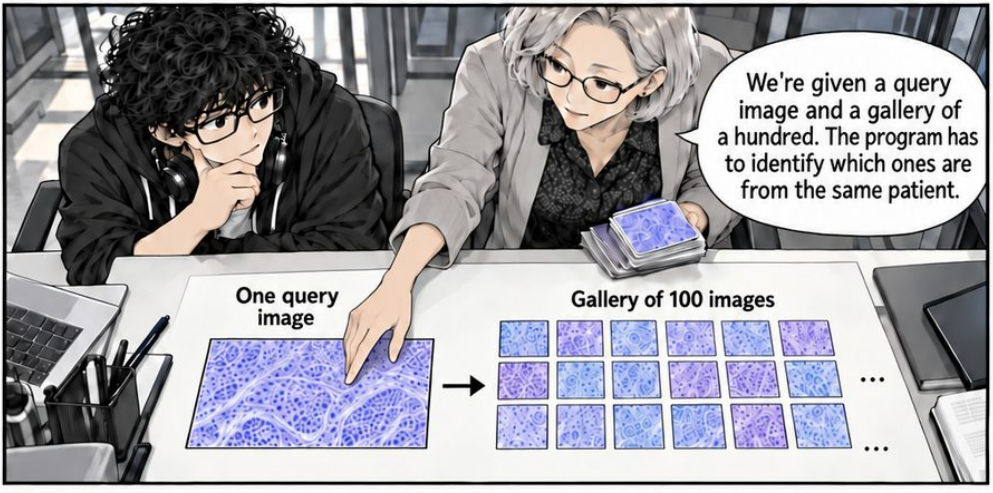
Sources: · Understanding F1 · The real challenge · The complete process

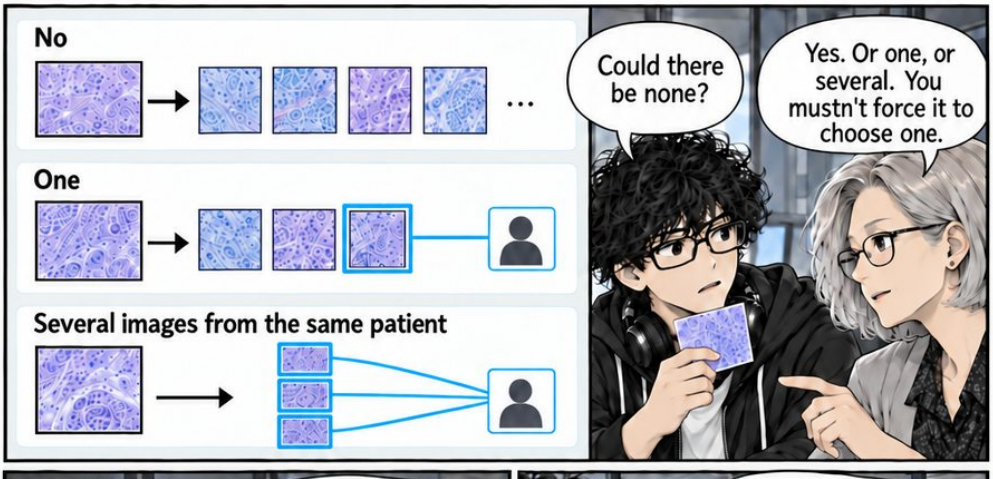

I've always been the odd one out.
In a new place, I can reinvent myself...
ATLAS
Responsible AI Laboratory
Nobody knows me here.
University AI challenge team
Leo Martín · Accepted
Let's see if I can avoid messing it up this time.

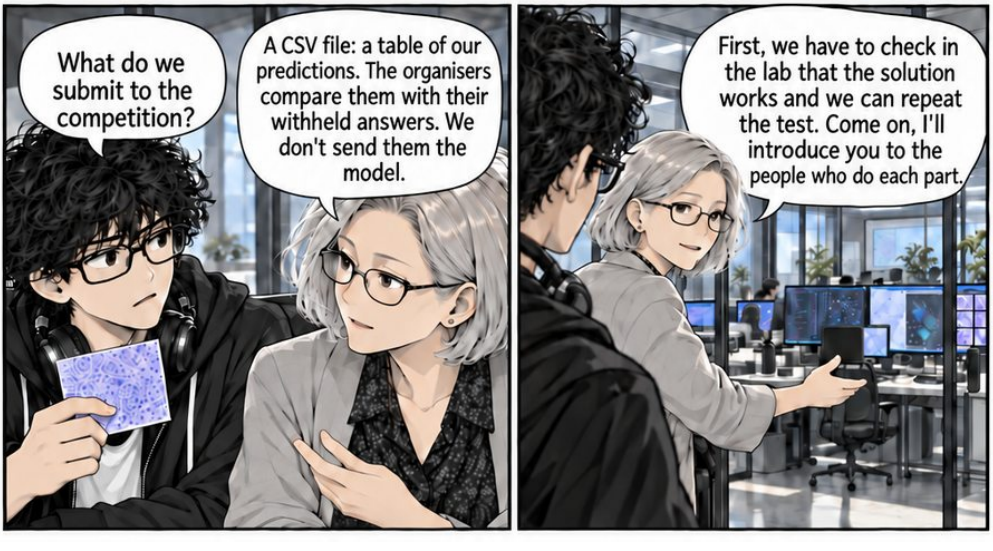

Monday, 7 September · 09:00
Access granted
BEEP
ATLAS
Leo? I'm Inés. I'm the laboratory's scientific director.
Yes. Hello.
I'm not sure whether anyone's explained. We're an AI competition team.
I saw the call for applications. What are you competing in at the moment?
We enter competitions on Kaggle. It's a website where challenges are set and different teams try to solve them.
I'd prepared something to say.

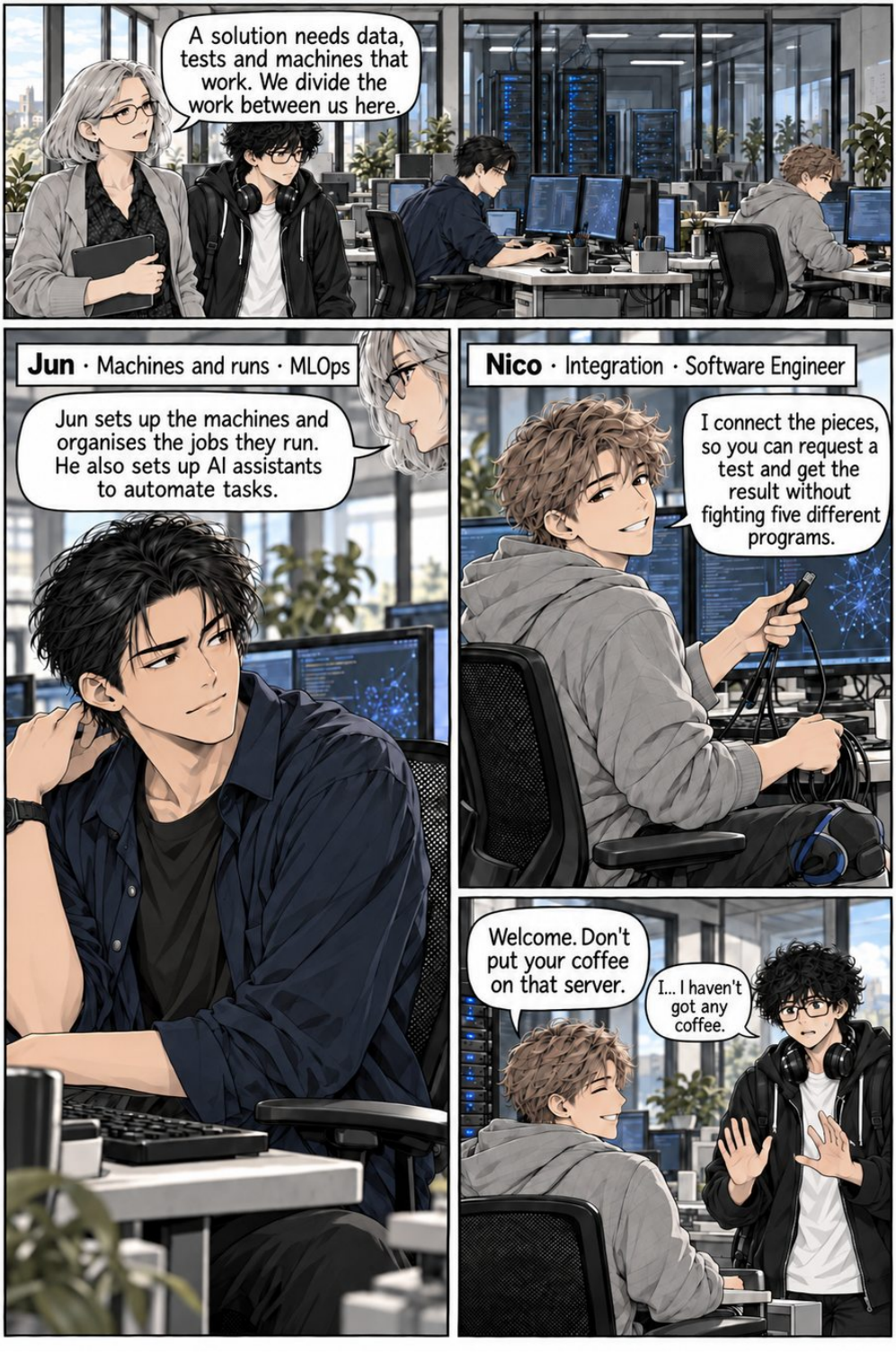

The challenge provides data to learn from, with known answers, and other data for the test. The organisers keep the test answers to themselves.
Training data (with known answers)
AI model
Test data (without answers, withheld by the organisers)
And we train an AI to solve it?
Yes. A vision model learns patterns in the images. Before the test, we try it on different patients from the ones we used for learning.
AI model
Benchmarking WSI Privacy · Track 1
We've entered this medical image privacy challenge. I'll show you what it asks for.

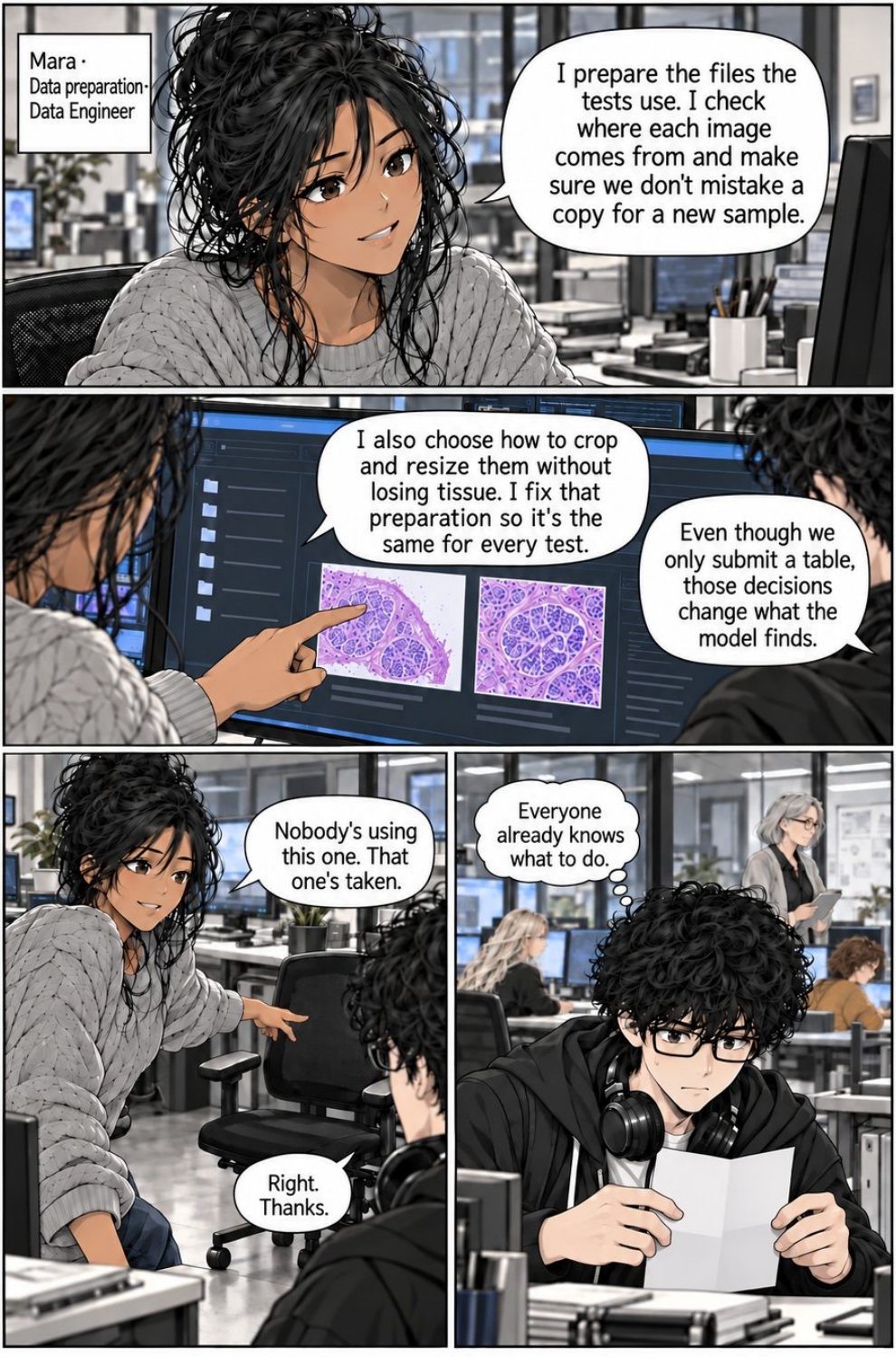

These are digital images of tissue samples. The challenge asks whether we can recognise samples from the same patient just by how they look.
Even if their name has been removed?
Exactly. Removing the name doesn't always stop you linking one image to another.
So we're looking for who that person is?
Not their name. We're looking for which images belong to the same patient. We aren't diagnosing a disease, either.
If the images can be linked, there's a privacy risk we should measure before sharing them.
I thought that without names, they were anonymous already.

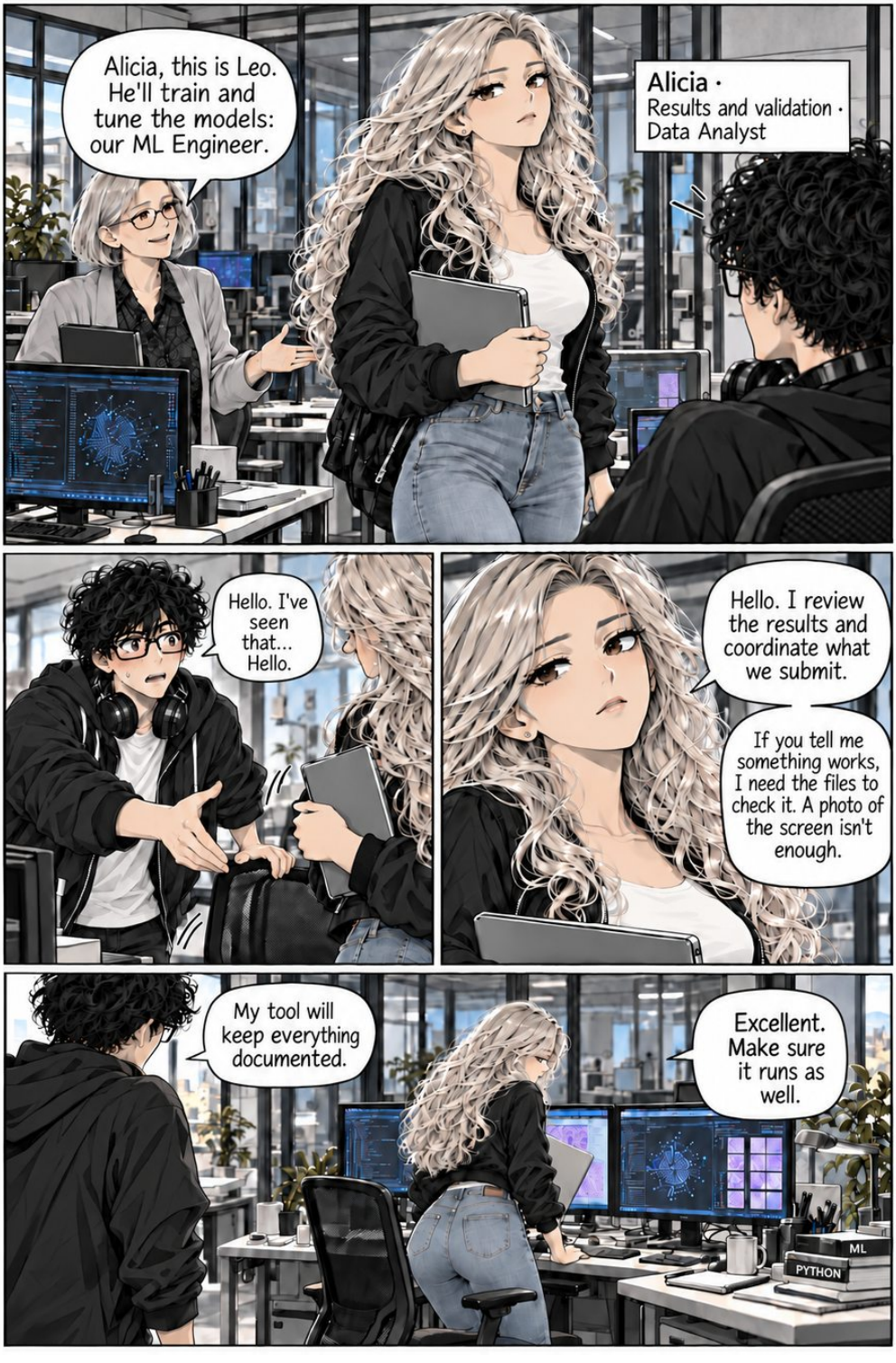

We're given a query image and a gallery of a hundred. The program has to identify which ones are from the same patient.
One query image
Gallery of 100 images

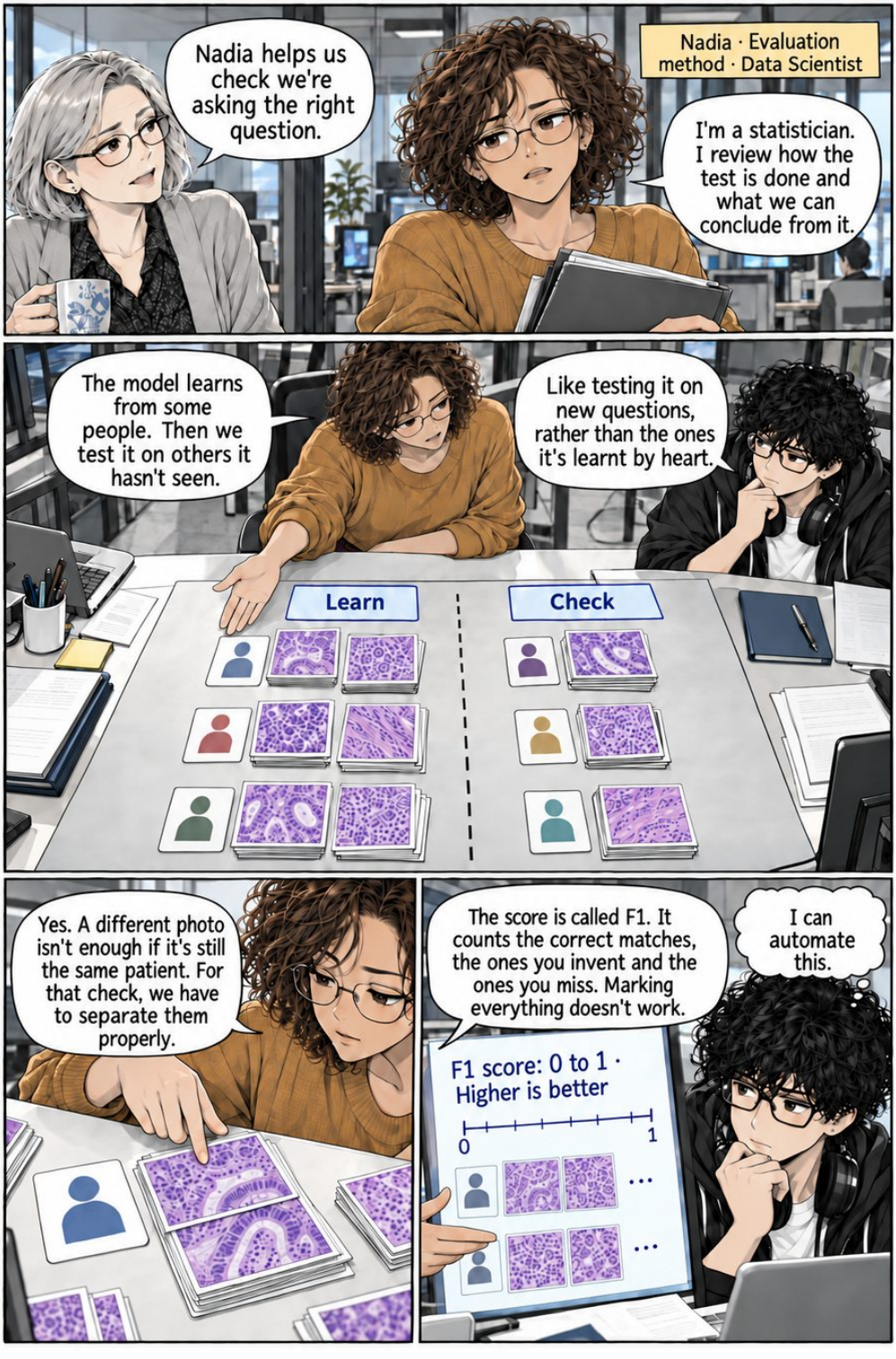

No
One
Several images from the same patient
Could there be none?
Yes. Or one, or several. You mustn't force it to choose one.

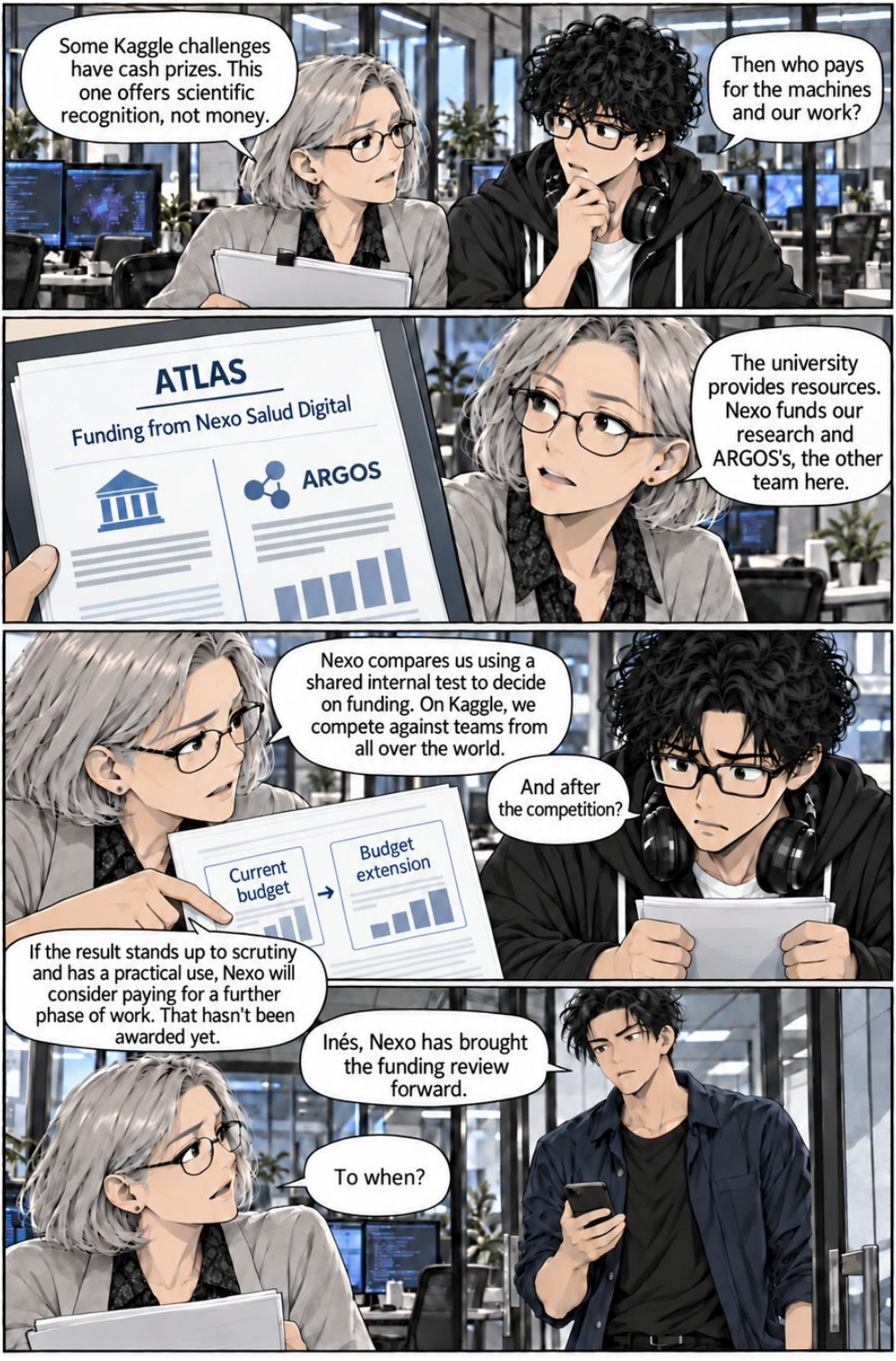

What do we submit to the competition?
A CSV file: a table of our predictions. The organisers compare them with their withheld answers. We don't send them the model.
First, we have to check in the lab that the solution works and we can repeat the test. Come on, I'll introduce you to the people who do each part.

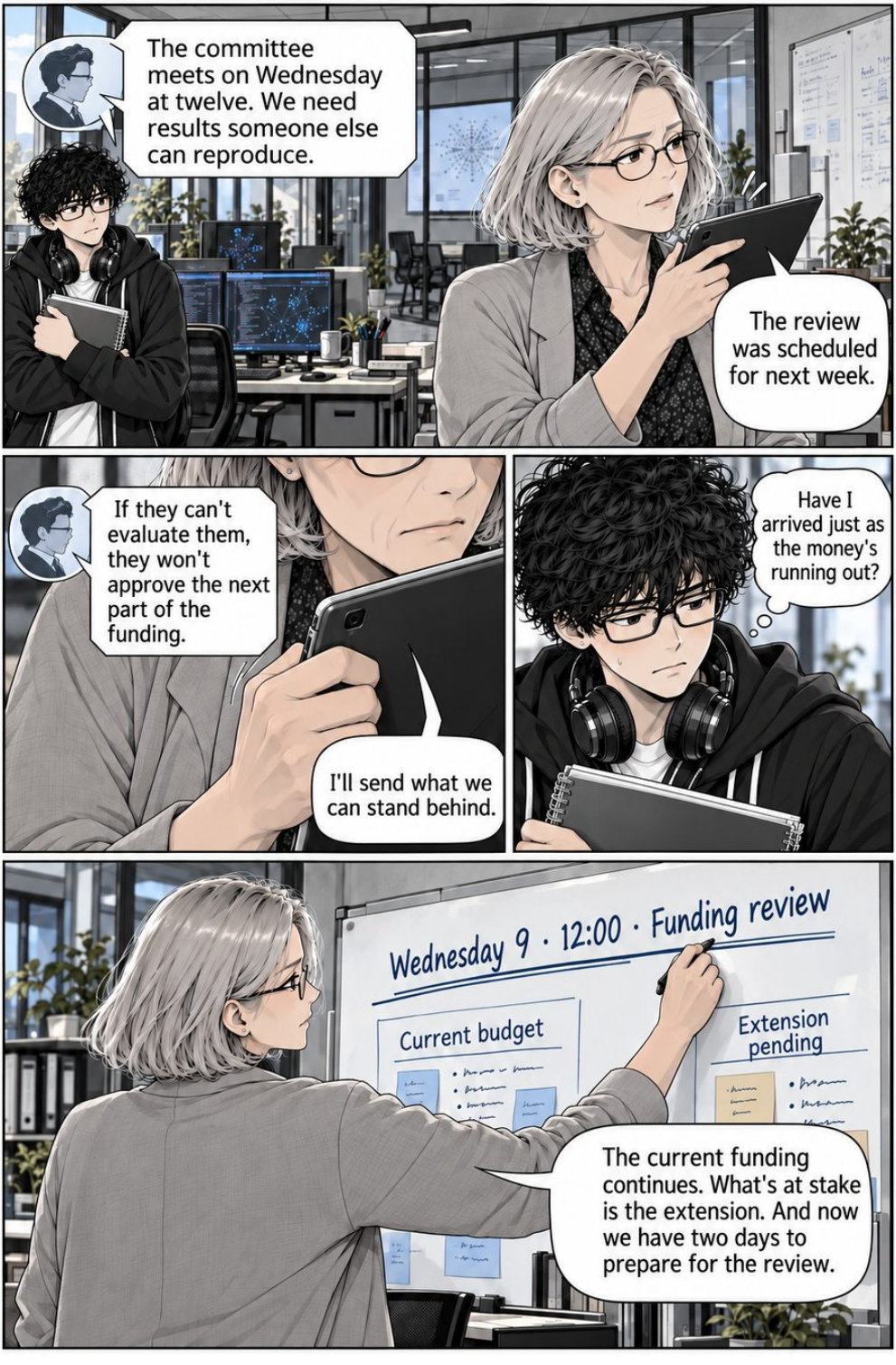

A solution needs data, tests and machines that work. We divide the work between us here.
Jun · Machines and runs · MLOps
Jun sets up the machines and organises the jobs they run. He also sets up AI assistants to automate tasks.
Nico · Integration · Software Engineer
I connect the pieces, so you can request a test and get the result without fighting five different programs.
Welcome. Don't put your coffee on that server.
I... I haven't got any coffee.

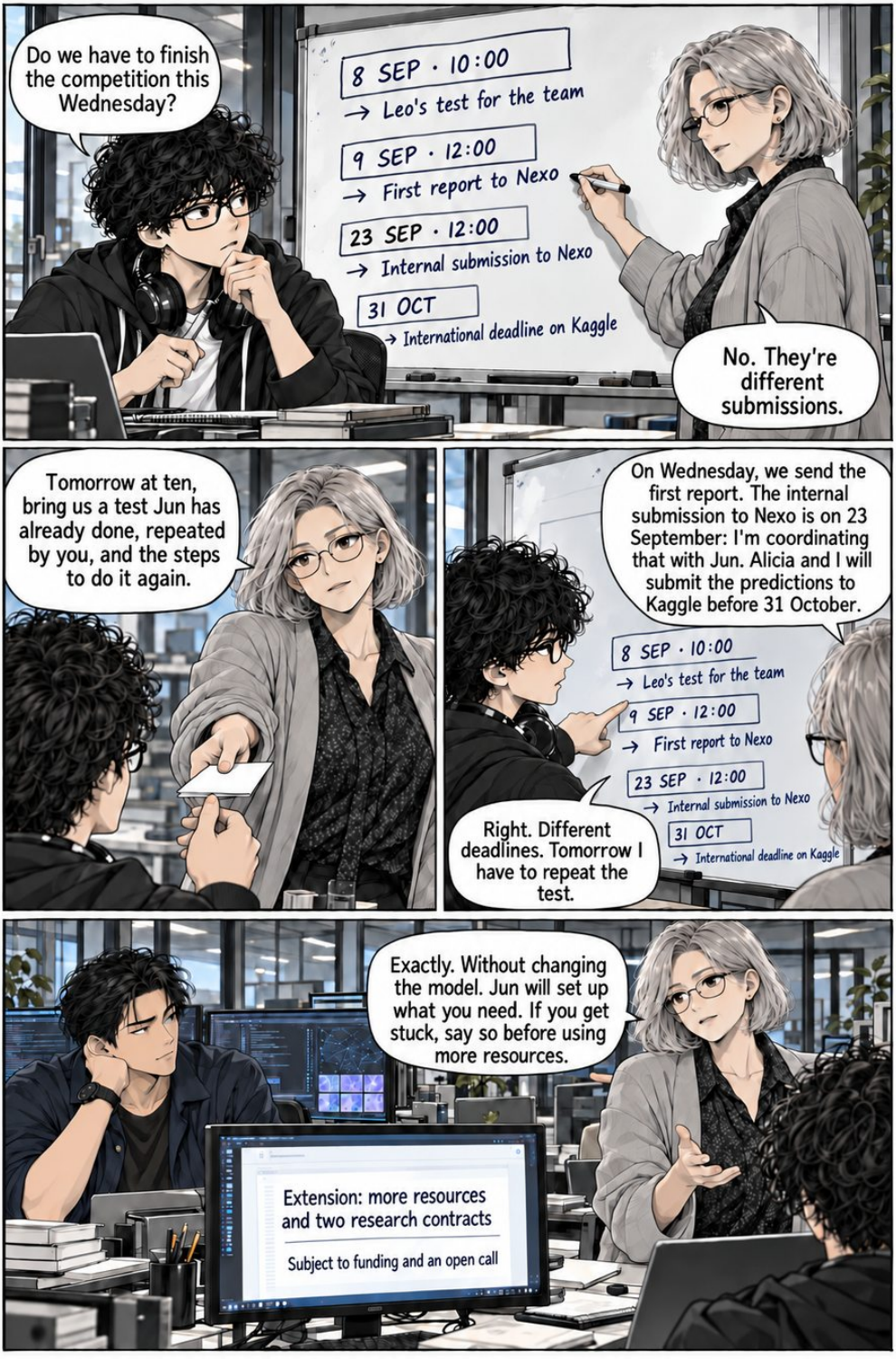

Mara · Data preparation · Data Engineer
I prepare the files the tests use. I check where each image comes from and make sure we don't mistake a copy for a new sample.
I also choose how to crop and resize them without losing tissue. I fix that preparation so it's the same for every test.
Even though we only submit a table, those decisions change what the model finds.
Nobody's using this one. That one's taken.
Right. Thanks.
Everyone already knows what to do.

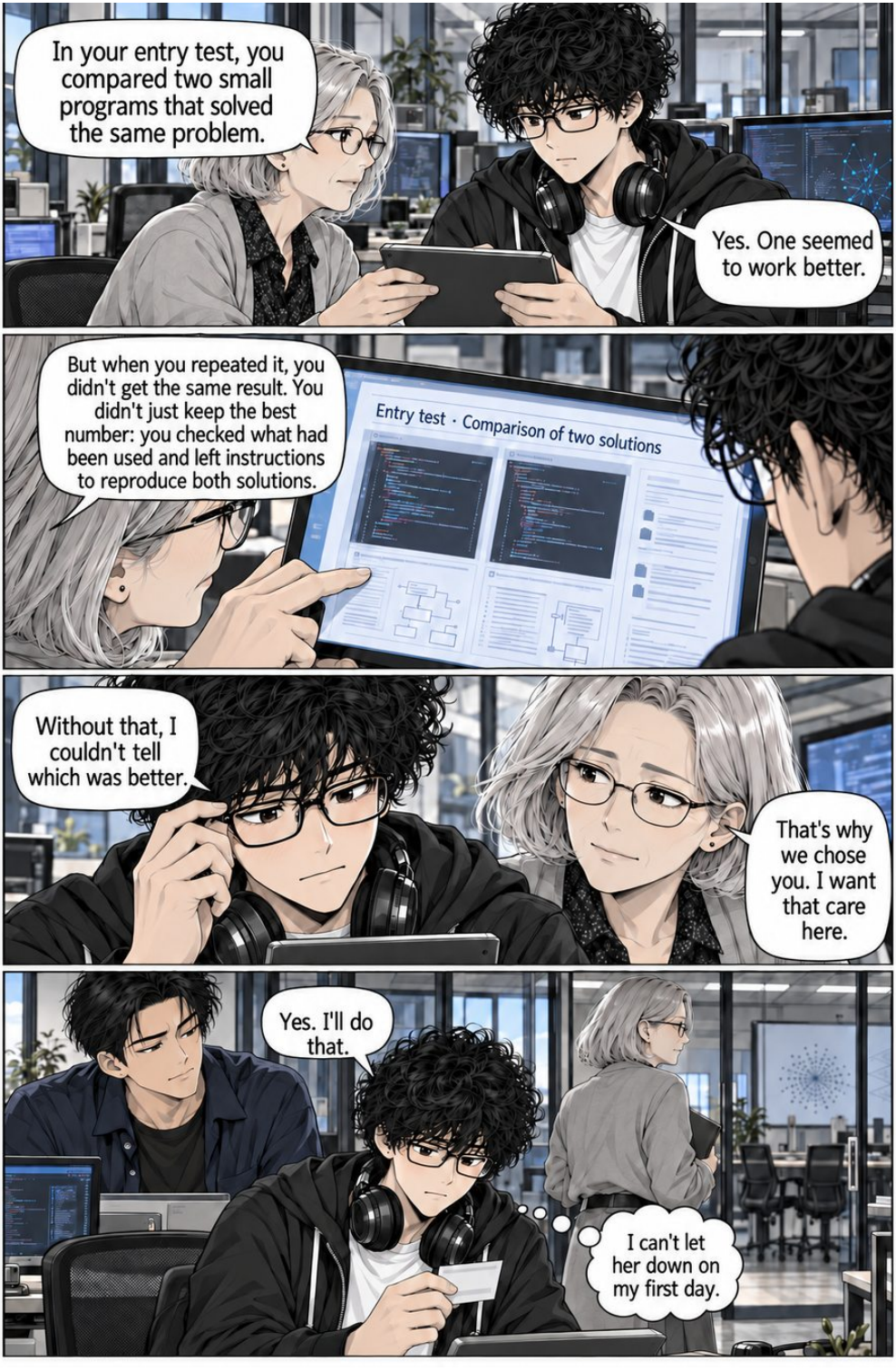

Alicia, this is Leo. He'll train and tune the models: our ML Engineer.
Alicia · Results and validation · Data Analyst
Hello. I've seen that... Hello.
Hello. I review the results and coordinate what we submit.
If you tell me something works, I need the files to check it. A photo of the screen isn't enough.
My tool will keep everything documented.
Excellent. Make sure it runs as well.
ML
PYTHON

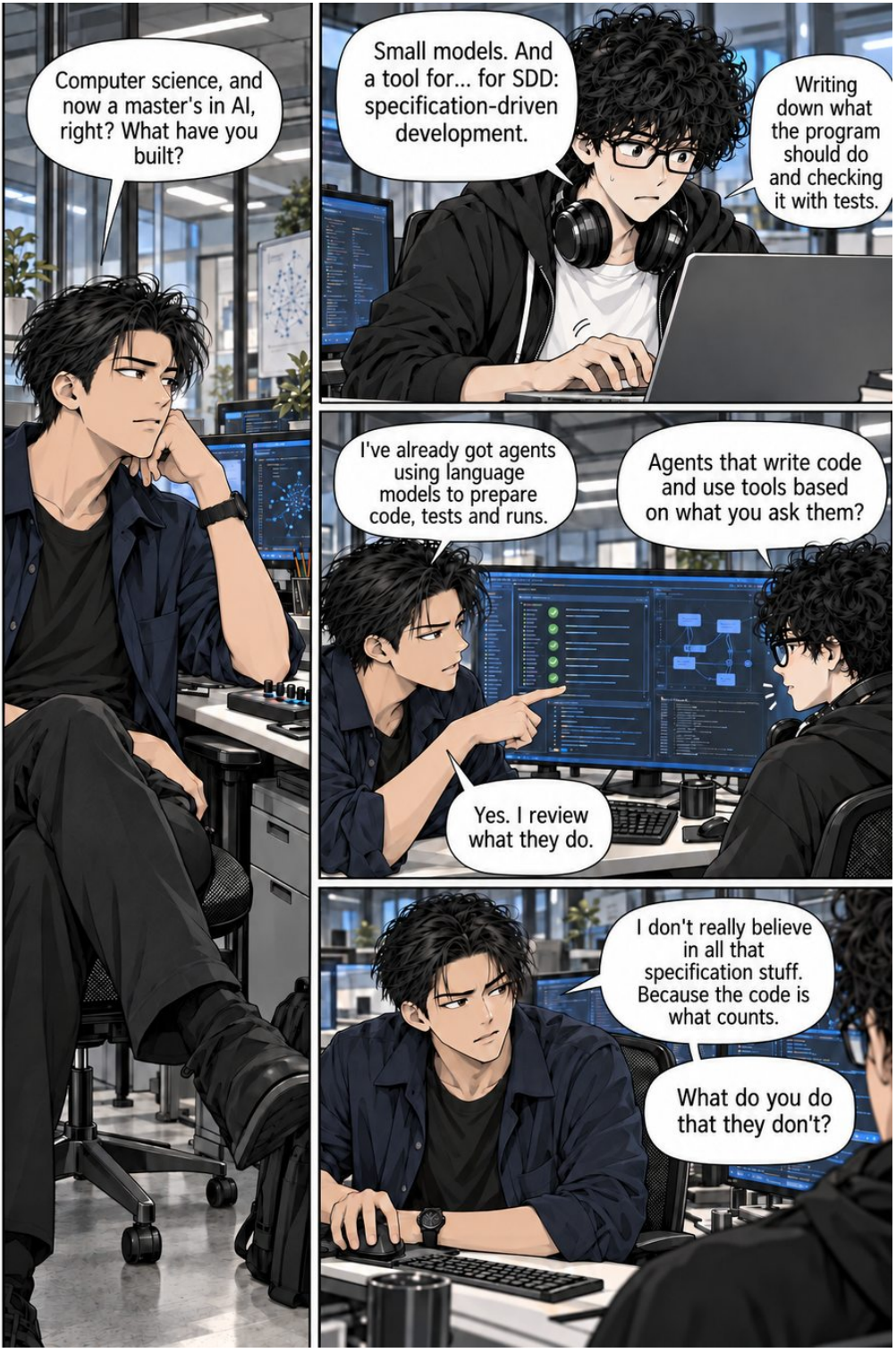

Nadia helps us check we're asking the right question.
Nadia · Evaluation method · Data Scientist
I'm a statistician. I review how the test is done and what we can conclude from it.
The model learns from some people. Then we test it on others it hasn't seen.
Like testing it on new questions, rather than the ones it's learnt by heart.
Learn
Check
Yes. A different photo isn't enough if it's still the same patient. For that check, we have to separate them properly.
The score is called F1. It counts the correct matches, the ones you invent and the ones you miss. Marking everything doesn't work.
I can automate this.
F1 score: 0 to 1 · Higher is better
0
1

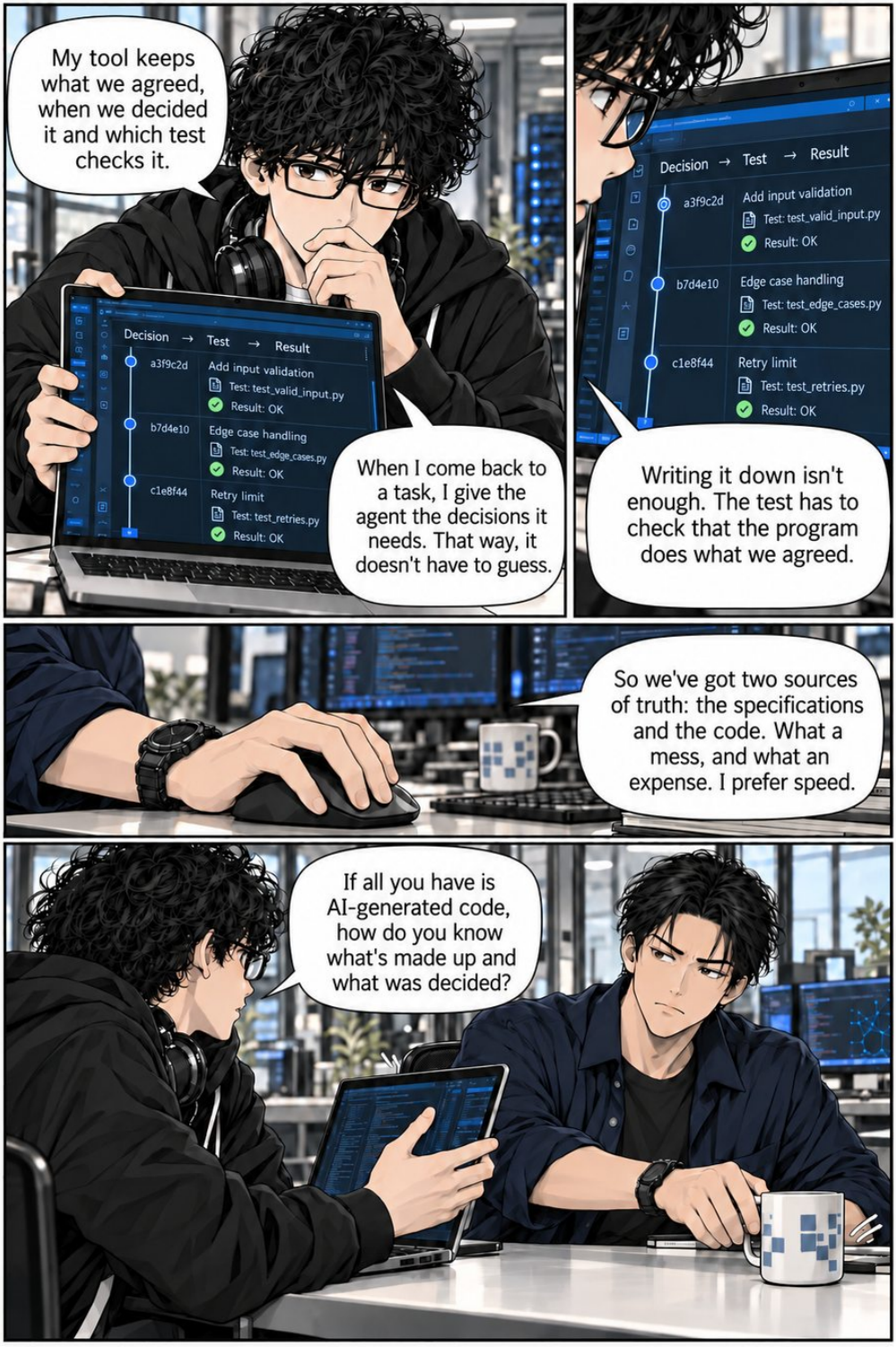

Some Kaggle challenges have cash prizes. This one offers scientific recognition, not money.
Then who pays for the machines and our work?
ATLAS
Funding from Nexo Salud Digital
ARGOS
The university provides resources. Nexo funds our research and ARGOS's, the other team here.
Nexo compares us using a shared internal test to decide on funding. On Kaggle, we compete against teams from all over the world.
And after the competition?
Current budget
Budget extension
If the result stands up to scrutiny and has a practical use, Nexo will consider paying for a further phase of work. That hasn't been awarded yet.
Inés, Nexo has brought the funding review forward.
To when?

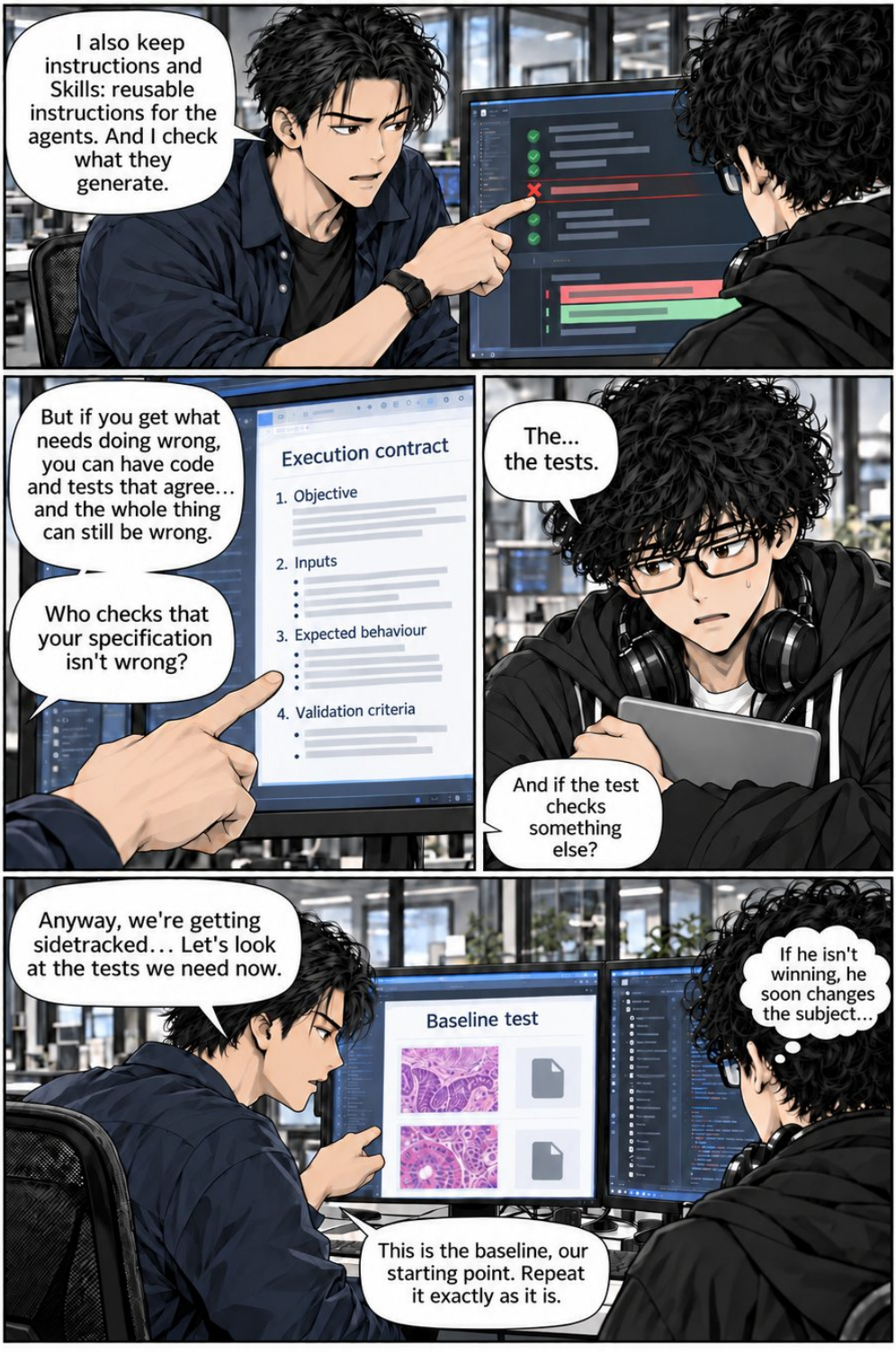

The committee meets on Wednesday at twelve. We need results someone else can reproduce.
The review was scheduled for next week.
If they can't evaluate them, they won't approve the next part of the funding.
I'll send what we can stand behind.
Have I arrived just as the money's running out?
Wednesday 9 · 12:00 · Funding review
Current budget
Extension pending
The current funding continues. What's at stake is the extension. And now we have two days to prepare for the review.

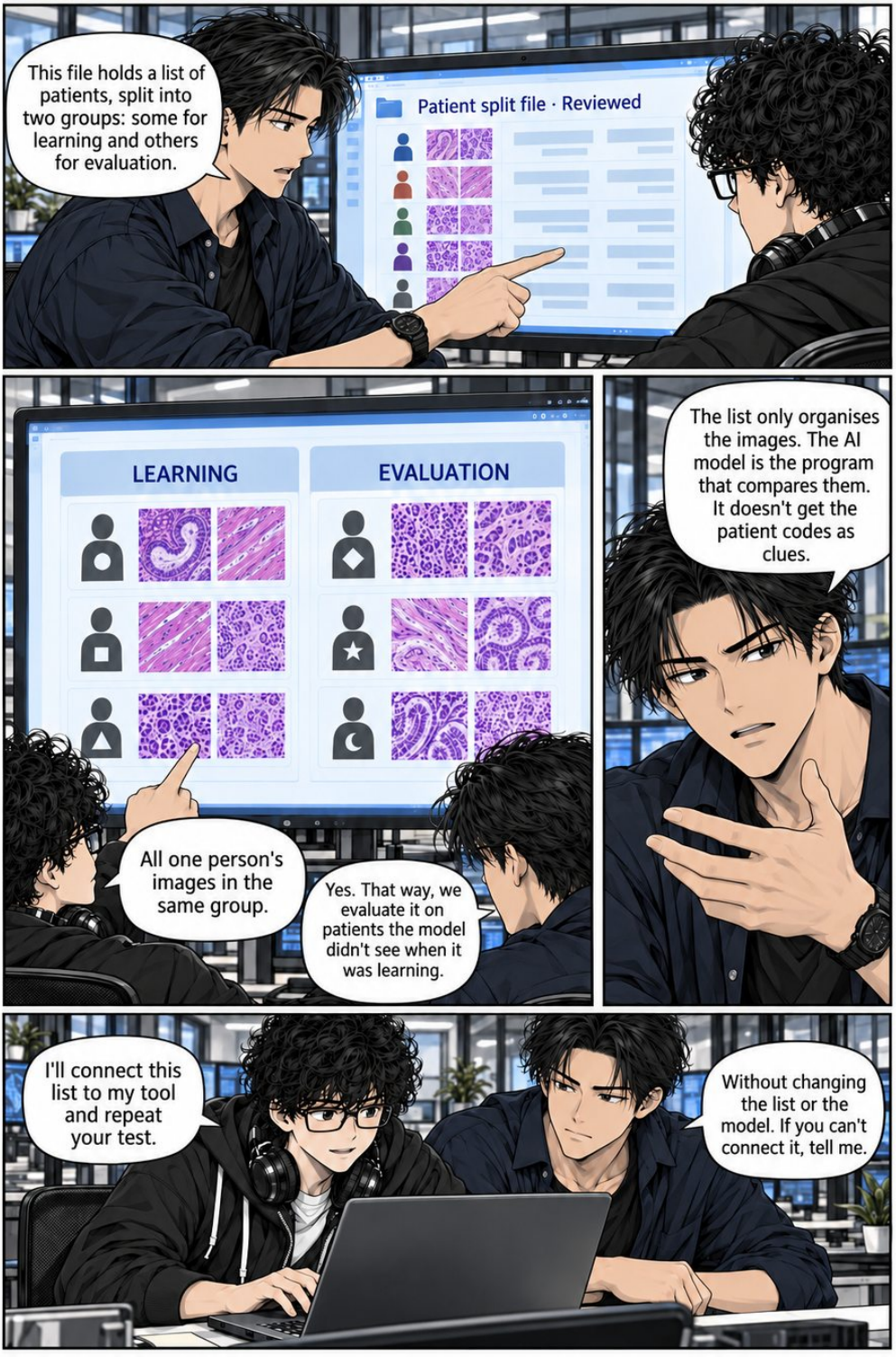

Do we have to finish the competition this Wednesday?
8 SEP · 10:00
→ Leo's test for the team
9 SEP · 12:00
→ First report to Nexo
23 SEP · 12:00
→ Internal submission to Nexo
31 OCT
→ International deadline on Kaggle
No. They're different submissions.
Tomorrow at ten, bring us a test Jun has already done, repeated by you, and the steps to do it again.
On Wednesday, we send the first report. The internal submission to Nexo is on 23 September: I'm coordinating that with Jun. Alicia and I will submit the predictions to Kaggle before 31 October.
8 SEP · 10:00
→ Leo's test for the team
9 SEP · 12:00
→ First report to Nexo
23 SEP · 12:00
→ Internal submission to Nexo
31 OCT
→ International deadline on Kaggle
Right. Different deadlines. Tomorrow I have to repeat the test.
Exactly. Without changing the model. Jun will set up what you need. If you get stuck, say so before using more resources.
Extension: more resources and two research contracts
Subject to funding and an open call

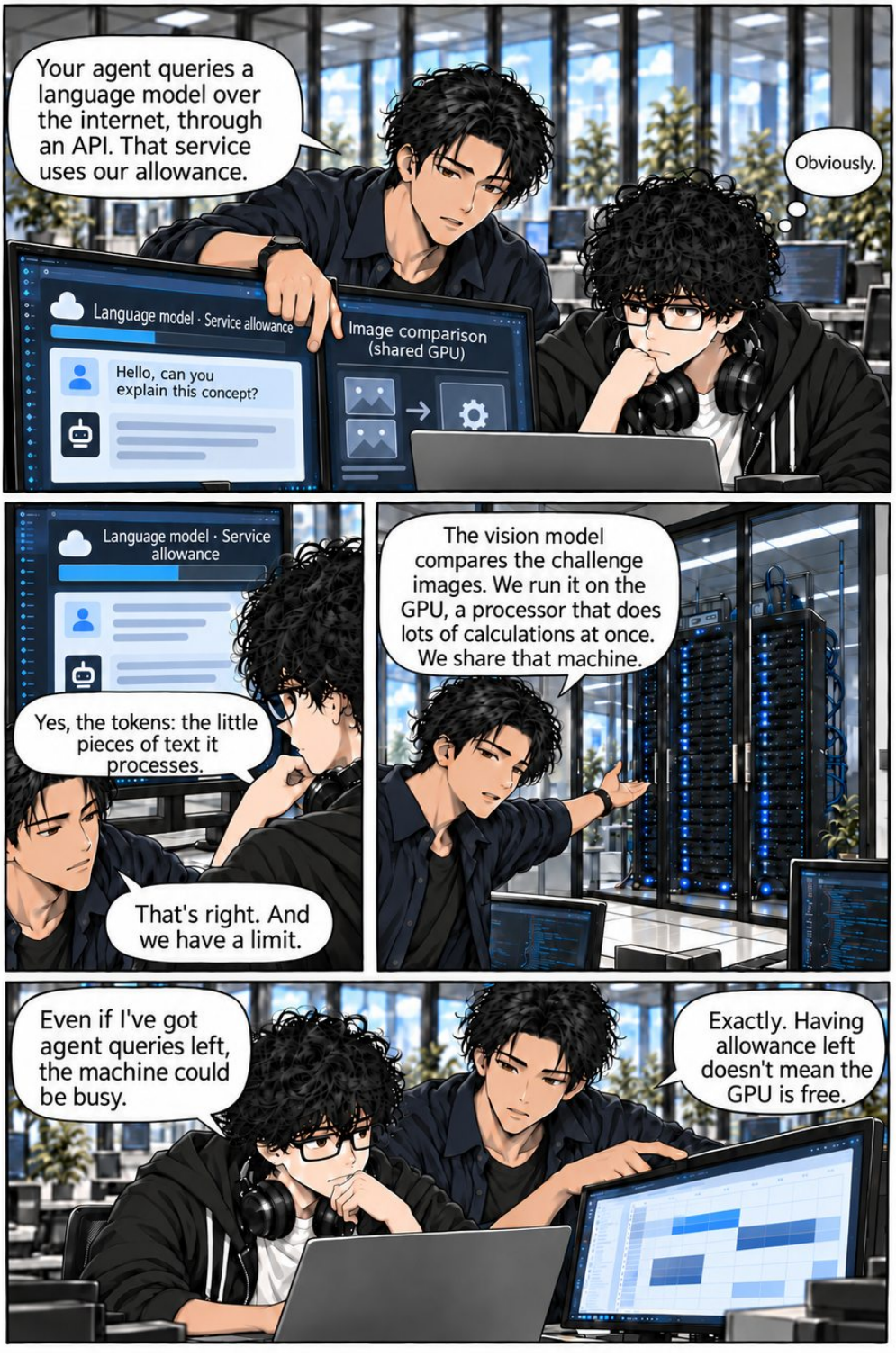

In your entry test, you compared two small programs that solved the same problem.
Yes. One seemed to work better.
But when you repeated it, you didn't get the same result. You didn't just keep the best number: you checked what had been used and left instructions to reproduce both solutions.
Entry test · Comparison of two solutions
Without that, I couldn't tell which was better.
That's why we chose you. I want that care here.
Yes. I'll do that.
I can't let her down on my first day.

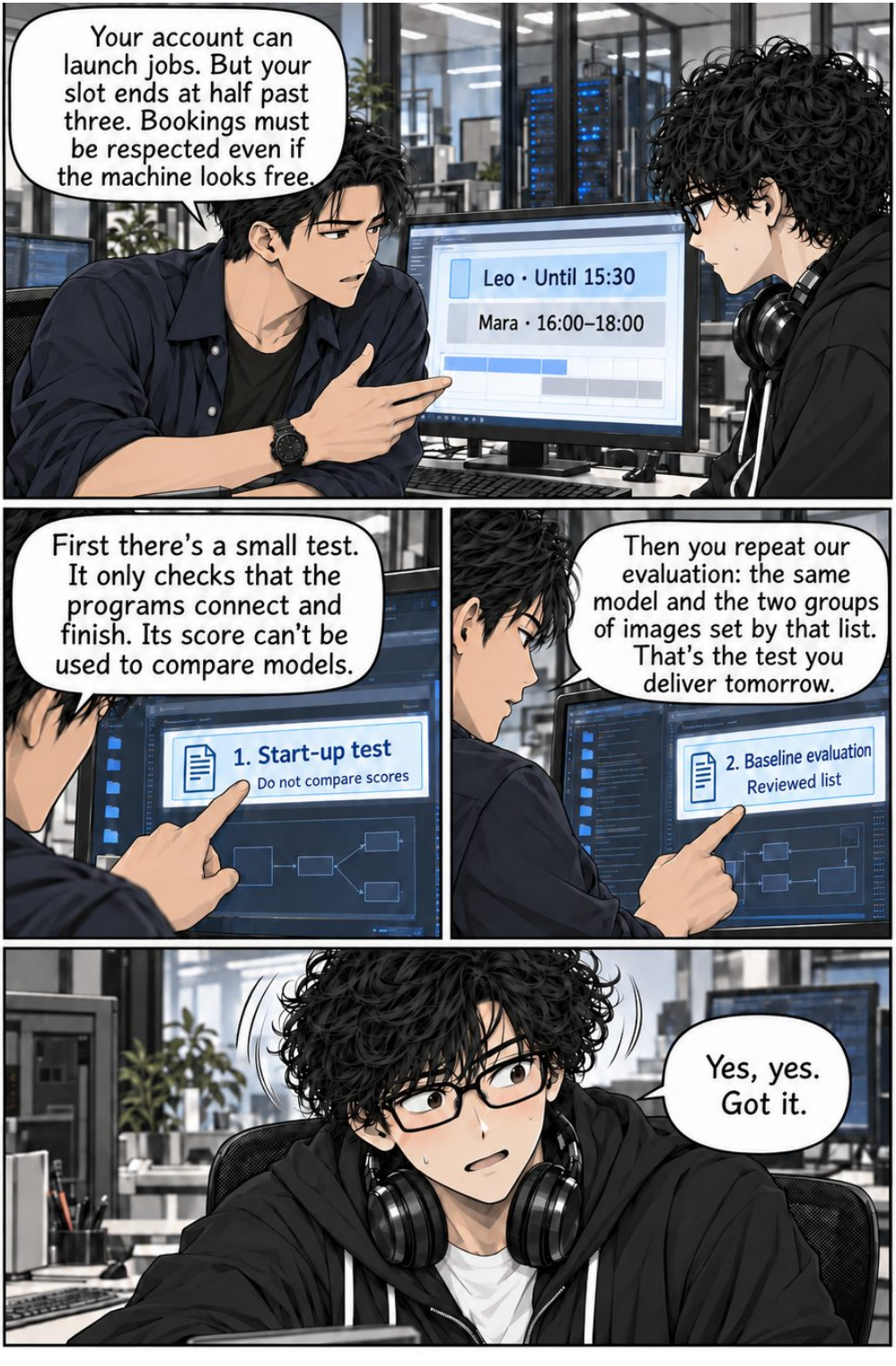

Computer science, and now a master's in AI, right? What have you built?
Small models. And a tool for... for SDD: specification-driven development.
Writing down what the program should do and checking it with tests.
I've already got agents using language models to prepare code, tests and runs.
Agents that write code and use tools based on what you ask them?
Yes. I review what they do.
I don't really believe in all that specification stuff. Because the code is what counts.
What do you do that they don't?


My tool keeps what we agreed, when we decided it and which test checks it.
When I come back to a task, I give the agent the decisions it needs. That way, it doesn't have to guess.
Decision → Test → Result
a3f9c2d
Add input validation
Test: test_valid_input.py
Result: OK
b7d4e10
Edge case handling
Test: test_edge_cases.py
Result: OK
c1e8f44
Retry limit
Test: test_retries.py
Result: OK
Writing it down isn't enough. The test has to check that the program does what we agreed.
So we've got two sources of truth: the specifications and the code. What a mess, and what an expense. I prefer speed.
If all you have is AI-generated code, how do you know what's made up and what was decided?


I also keep instructions and Skills: reusable instructions for the agents. And I check what they generate.
But if you get what needs doing wrong, you can have code and tests that agree… and the whole thing can still be wrong.
Who checks that your specification isn't wrong?
Execution contract
1. Objective
2. Inputs
3. Expected behaviour
4. Validation criteria
The… the tests.
And if the test checks something else?
Anyway, we're getting sidetracked… Let's look at the tests we need now.
Baseline test
This is the baseline, our starting point. Repeat it exactly as it is.
If he isn't winning, he soon changes the subject…

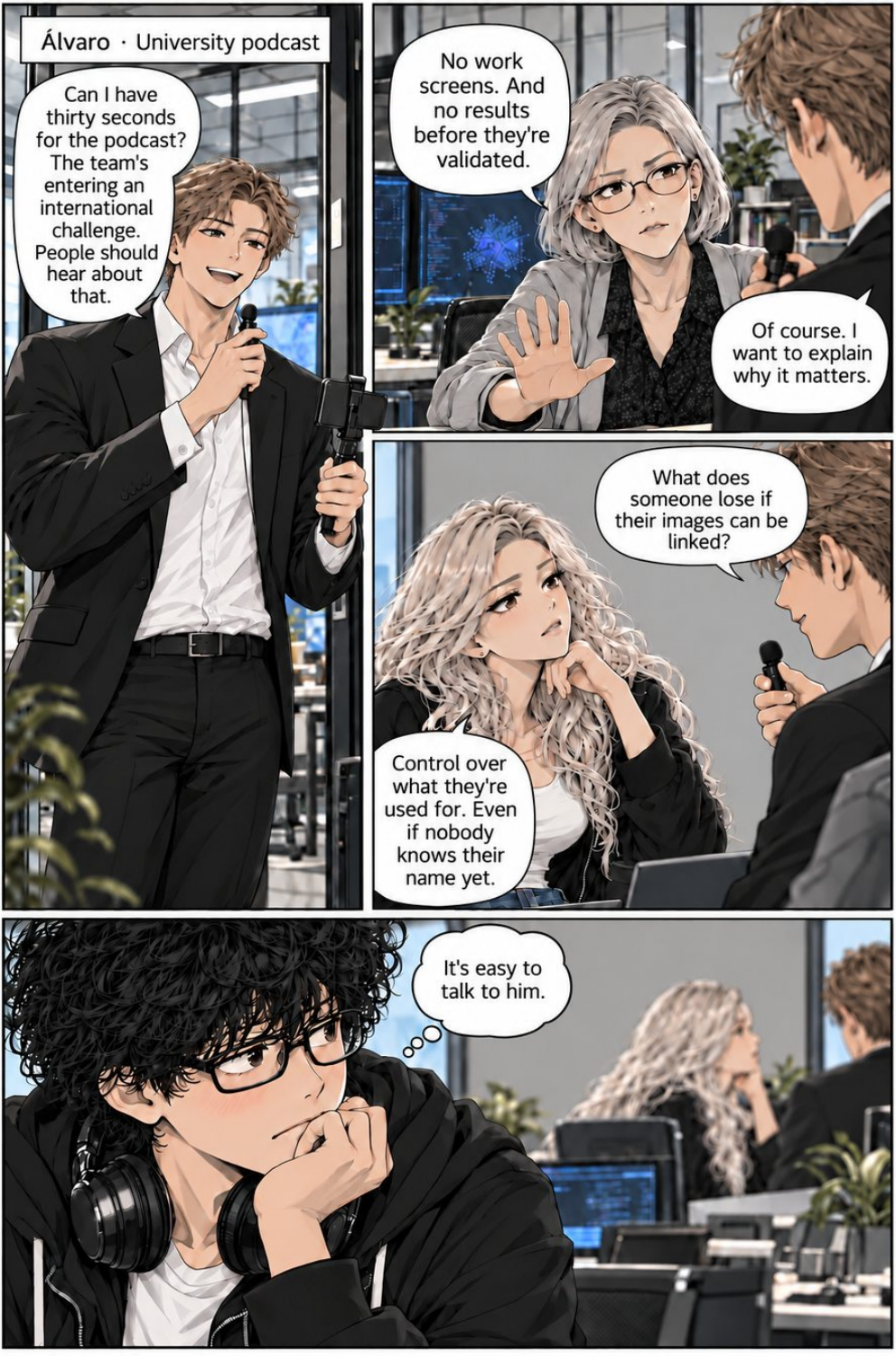

This file holds a list of patients, split into two groups: some for learning and others for evaluation.
Patient split file · Reviewed
LEARNING
EVALUATION
All one person's images in the same group.
Yes. That way, we evaluate it on patients the model didn't see when it was learning.
The list only organises the images. The AI model is the program that compares them. It doesn't get the patient codes as clues.
I'll connect this list to my tool and repeat your test.
Without changing the list or the model. If you can't connect it, tell me.


Your agent queries a language model over the internet, through an API. That service uses our allowance.
Obviously.
Language model · Service allowance
Hello, can you explain this concept?
Image comparison (shared GPU)
Language model · Service allowance
Yes, the tokens: the little pieces of text it processes.
That's right. And we have a limit.
The vision model compares the challenge images. We run it on the GPU, a processor that does lots of calculations at once. We share that machine.
Even if I've got agent queries left, the machine could be busy.
Exactly. Having allowance left doesn't mean the GPU is free.


Your account can launch jobs. But your slot ends at half past three. Bookings must be respected even if the machine looks free.
Leo · Until 15:30
Mara · 16:00–18:00
First there's a small test. It only checks that the programs connect and finish. Its score can't be used to compare models.
1. Start-up test
Do not compare scores
Then you repeat our evaluation: the same model and the two groups of images set by that list. That's the test you deliver tomorrow.
2. Baseline evaluation
Reviewed list
Yes, yes. Got it.

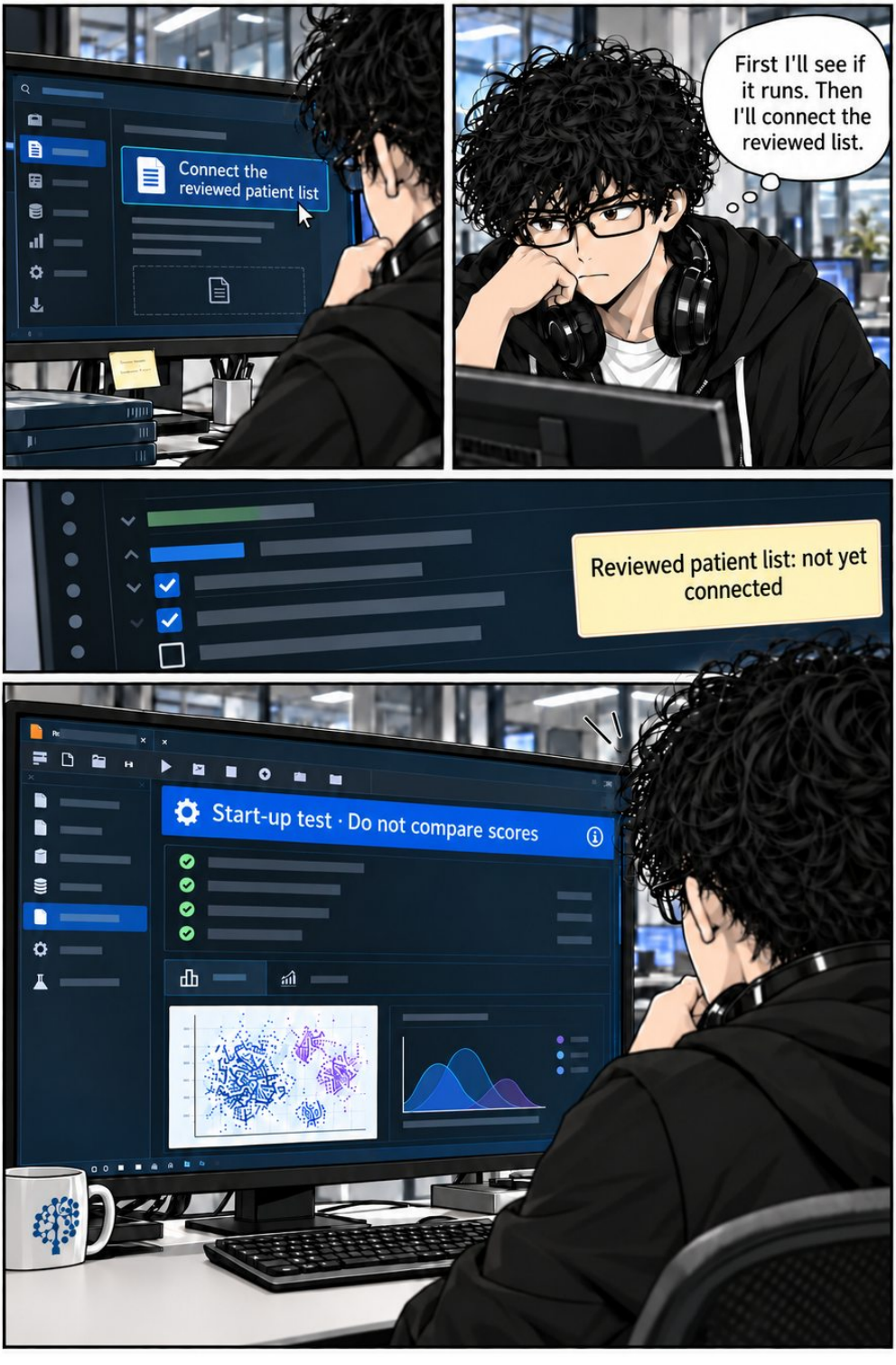

That one gives you the display. The socket's here.
Right. Yes. Thanks.
Good start.
Nothing's fallen over yet.
If you can't find a file, ask me. Before you make another copy.
Versions
My machine
Booked
Mum
Right. And you...?
I need to finish this now. Later.


Lunch break
On Friday, we're starting a new game in a co-op dungeon crawler. Jun says he won't be the one taking all the hits again.
Because then you run off and leave me on my own.
Couldn't you just not run off?
You could solve that with a multi-agent policy...
A plan to get the characters working together? We could also just be rubbish at it. Are you in?
I don't know. I've got things on.
I wanted them to invite me. Why did I want to leave?

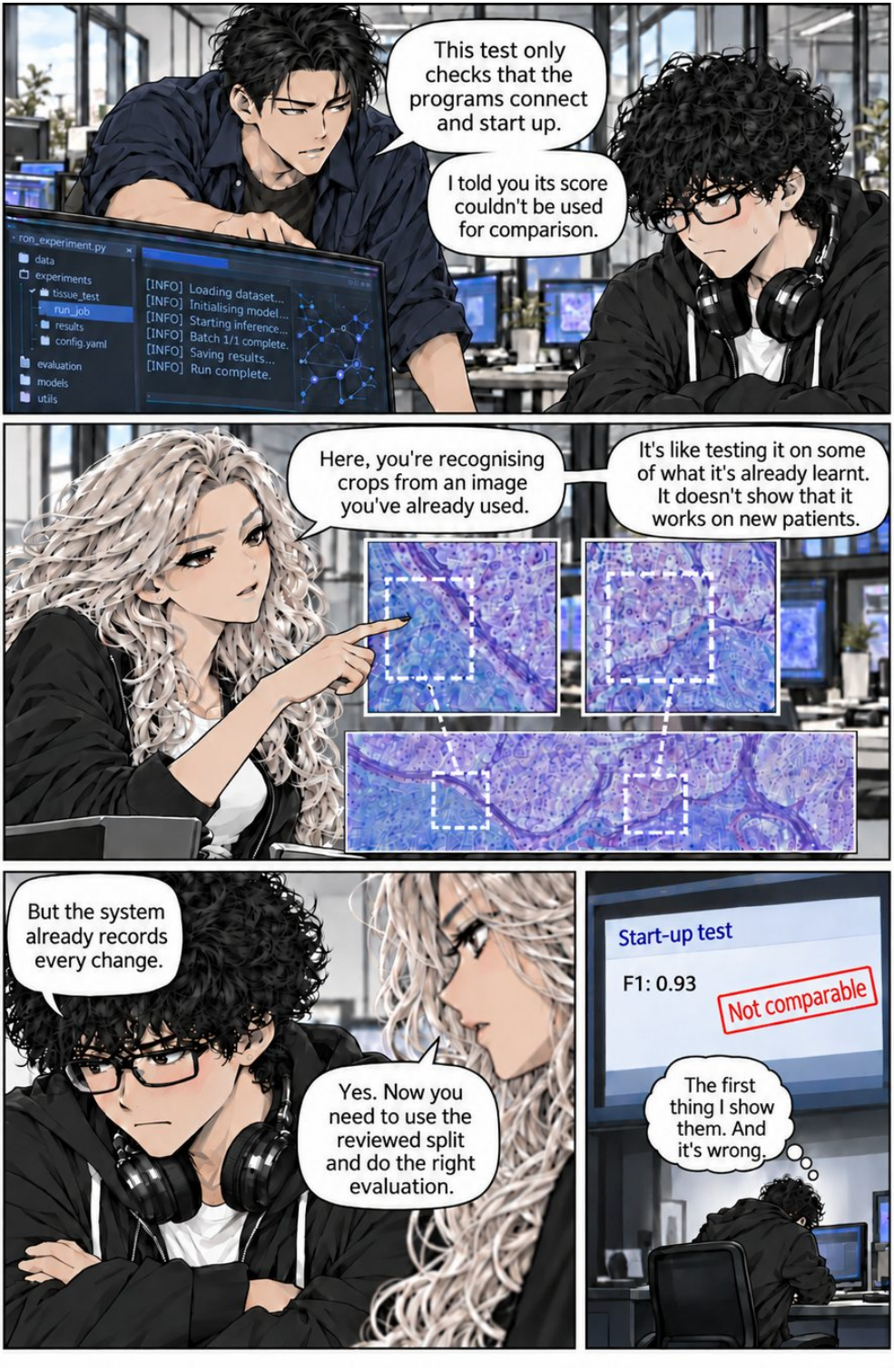

Álvaro · University podcast
Can I have thirty seconds for the podcast? The team's entering an international challenge. People should hear about that.
No work screens. And no results before they're validated.
Of course. I want to explain why it matters.
What does someone lose if their images can be linked?
Control over what they're used for. Even if nobody knows their name yet.
It's easy to talk to him.


And why did you get into this?
At home, I was always brought up to be at the top. I suppose it stuck.
D
Dad
Have they given you any responsibility yet? You're not there to make tables.
Doesn't having to be at the top all the time get tiring?
Sometimes. But I don't want to settle, either.
If I bring a good result tomorrow, they'll listen to me.

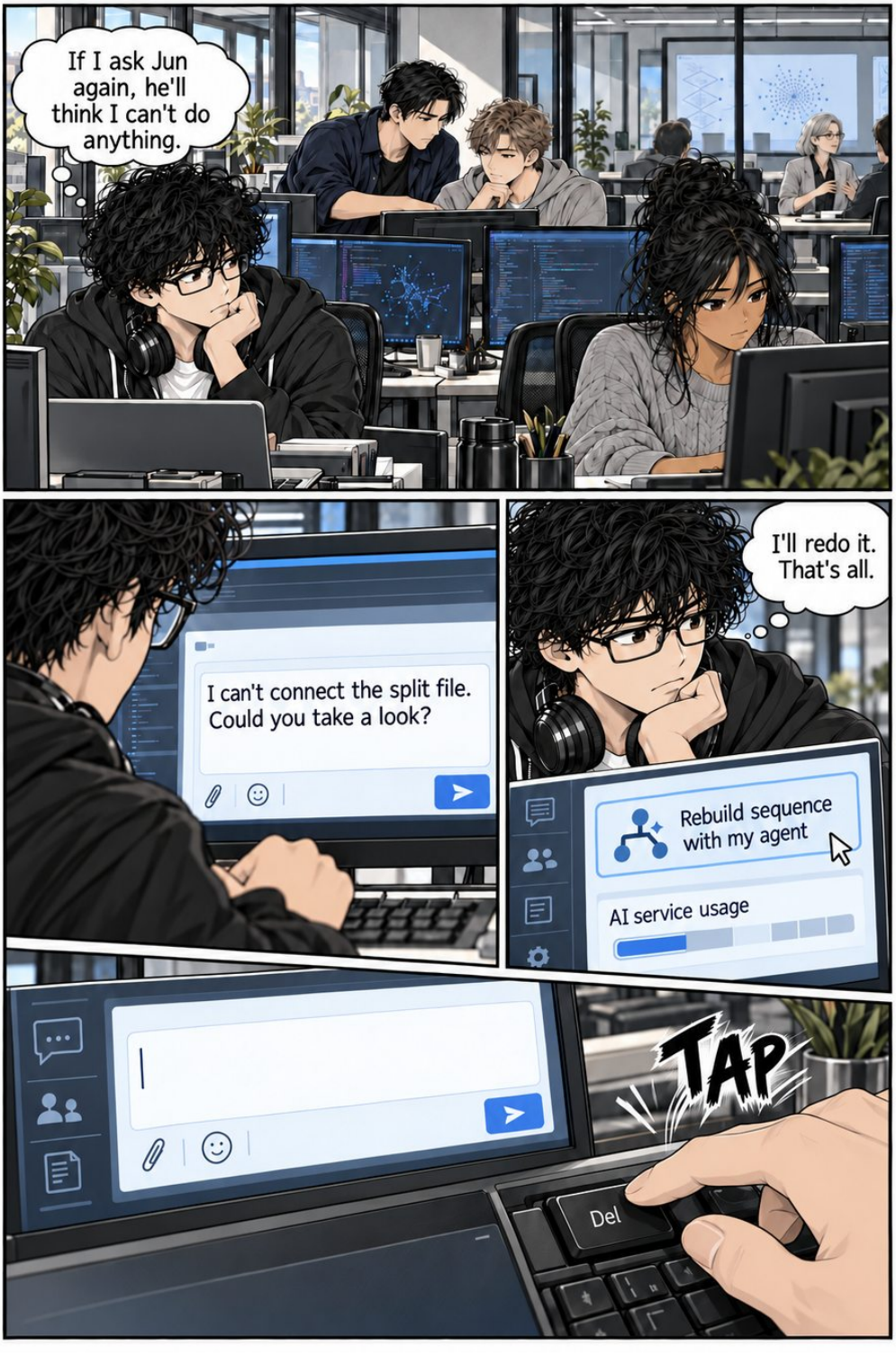

13:20
Input: images
Output: matches
Keep: baseline model
A program component is missing
Component installed · Ready to run
Now they'll see what it's for.

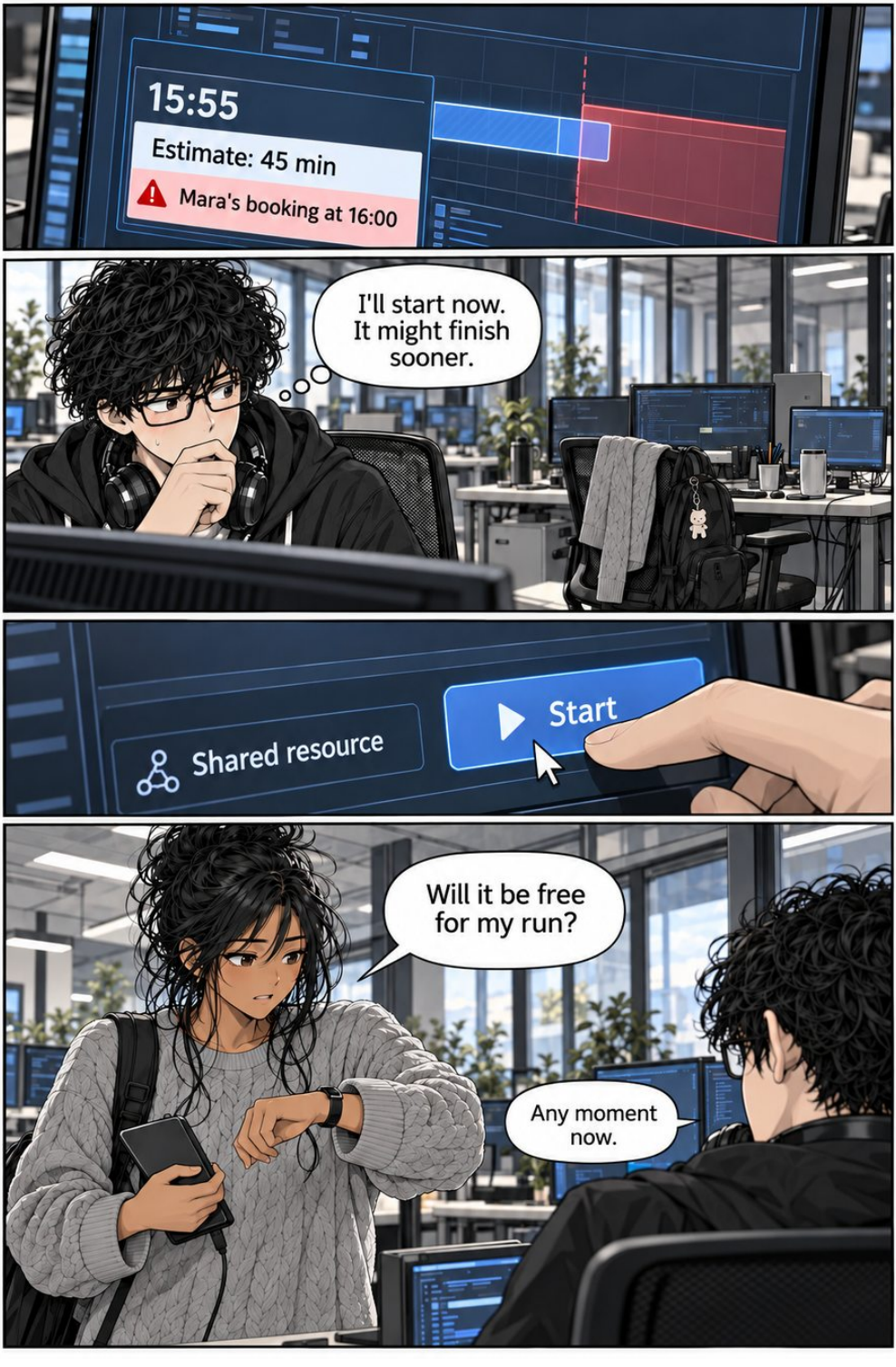

Connect the reviewed patient list
First I'll see if it runs. Then I'll connect the reviewed list.
Reviewed patient list: not yet connected
Start-up test · Do not compare scores


Start-up test · F1 score: 0.93
That's it! And it's much better.
Have you repeated our test with the same patient split?
I've done... the small test first.

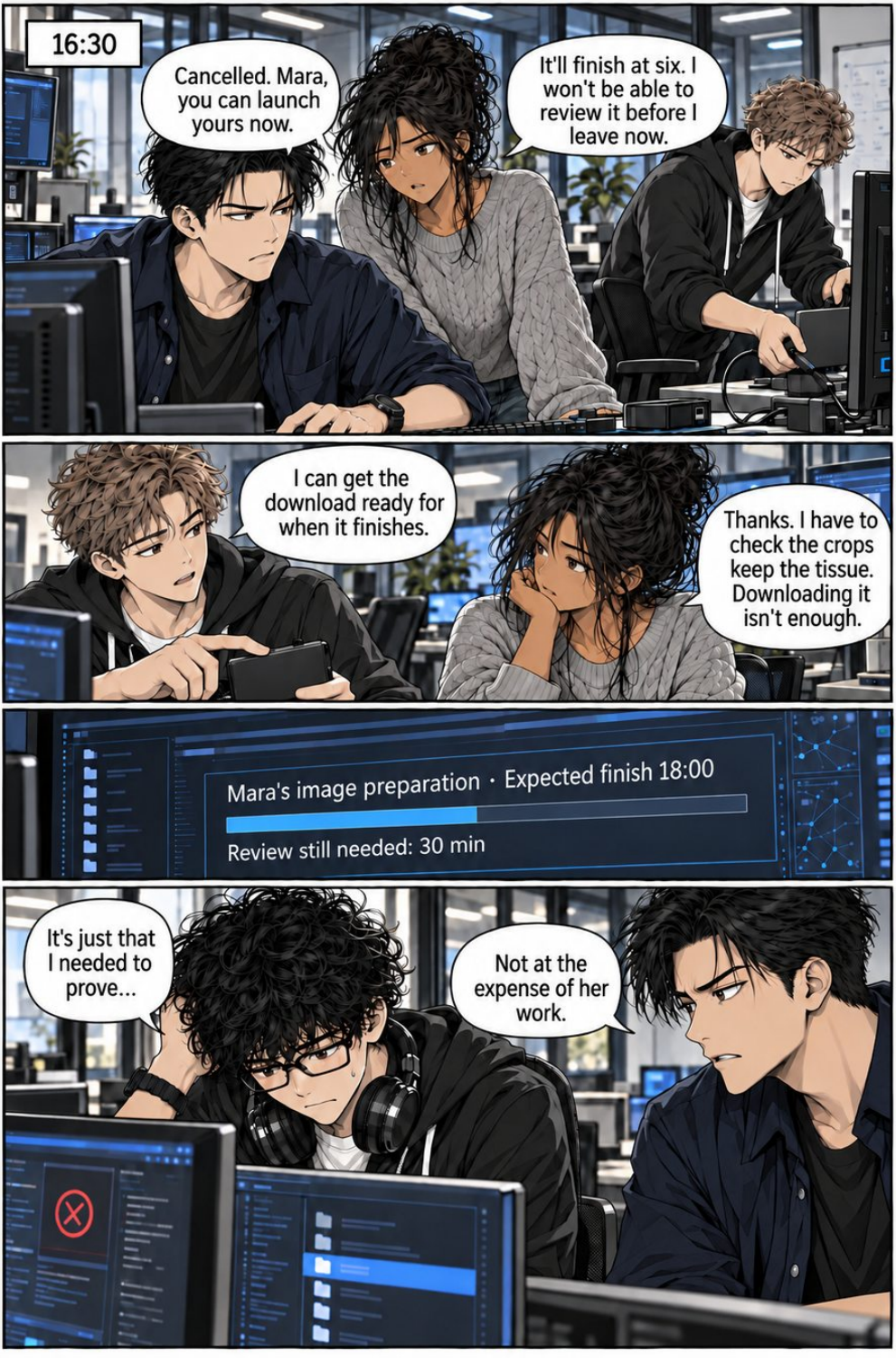

This test only checks that the programs connect and start up.
I told you its score couldn't be used for comparison.
run_experiment.py
data
experiments
tissue_test
run_job
results
config.yaml
evaluation
models
utils
[INFO] Loading dataset...
[INFO] Initialising model...
[INFO] Starting inference...
[INFO] Batch 1/1 complete.
[INFO] Saving results...
[INFO] Run complete.
Here, you're recognising crops from an image you've already used.
It's like testing it on some of what it's already learnt. It doesn't show that it works on new patients.
But the system already records every change.
Yes. Now you need to use the reviewed split and do the right evaluation.
Start-up test
F1: 0.93
Not comparable
The first thing I show them. And it's wrong.


Connect the split file. Check you don't have patients in both groups. Repeat the same model.
I'll have it before I leave.
TAP TAP TAP
15:20
14:00
15:00
16:00
17:00
18:00
Mara
16 – 18
At four, I'm preparing the images on the machine. It takes an hour and a half; then I need half an hour to review them. At six, I leave for the café.
Yes. Nearly finished.


If I ask Jun again, he'll think I can't do anything.
I can't connect the split file. Could you take a look?
I'll redo it. That's all.
Rebuild sequence with my agent
AI service usage
TAP
Del

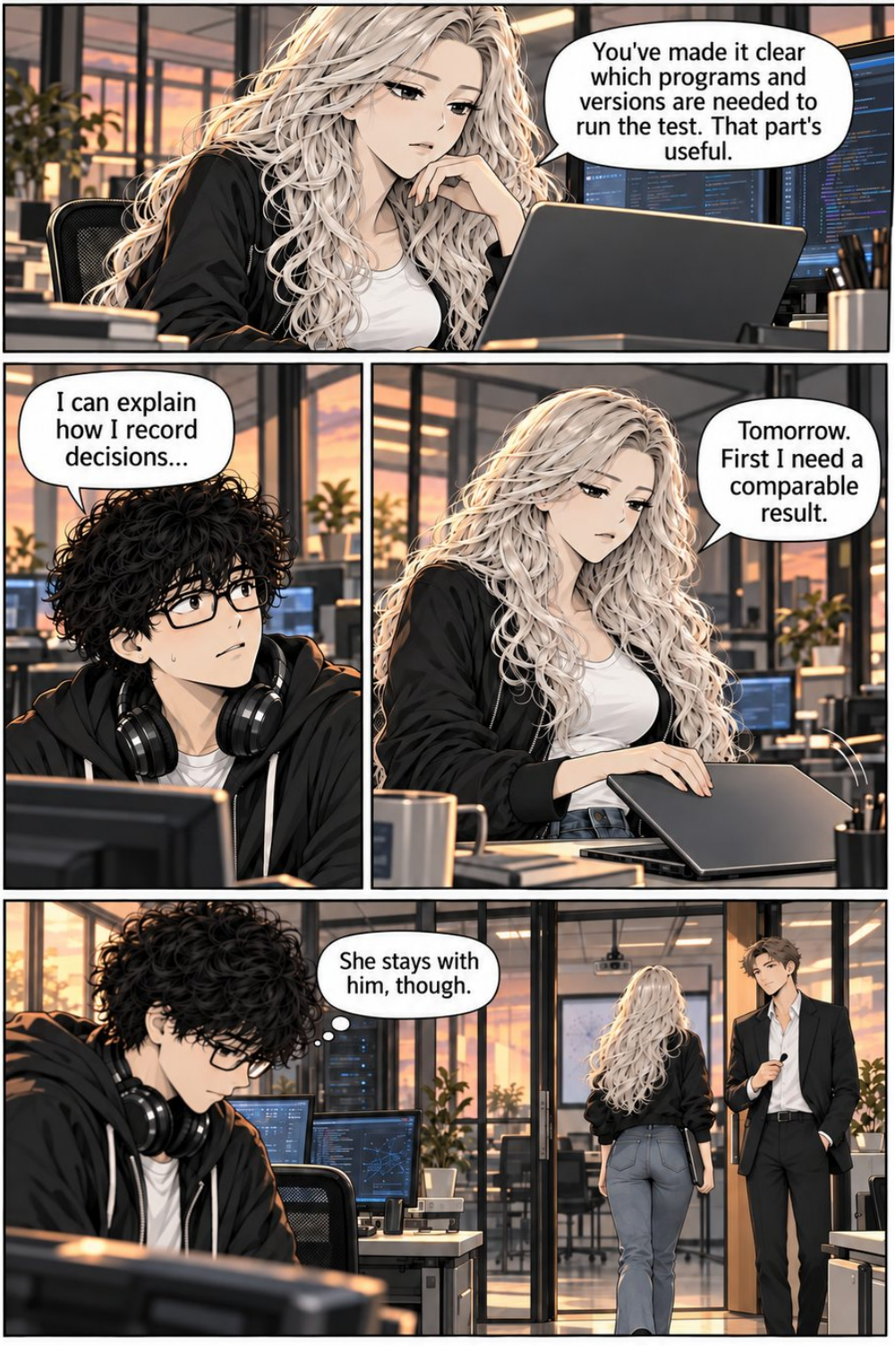

15:55
Estimate: 45 min
Mara's booking at 16:00
I'll start now. It might finish sooner.
Shared resource
Start
Will it be free for my run?
Any moment now.


16:25
Mara's job
Waiting for resource
Leo.
My slot's already started.
I'll stop it. It was nearly done. I thought...
Saving state...
Cancel
M
Mum
Will you be here to take over? I've been here since six.
You told me you were finishing. You could have asked me.


16:30
Cancelled. Mara, you can launch yours now.
It'll finish at six. I won't be able to review it before I leave now.
I can get the download ready for when it finishes.
Thanks. I have to check the crops keep the tissue. Downloading it isn't enough.
Mara's image preparation · Expected finish 18:00
Review still needed: 30 min
It's just that I needed to prove...
Not at the expense of her work.

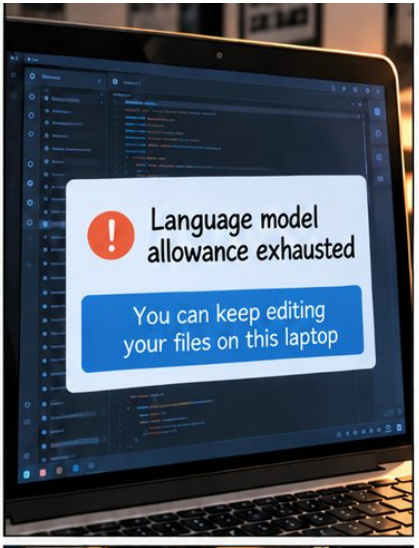

18:00
Sorry. I didn't know it was so...
I told you what time I needed it. That was enough.
I'm leaving now. I'll review the images tonight.
All right, love. I'll save you some food.
I'll let you know when it's downloaded.
Thanks.
She'd been kind to me. And I...

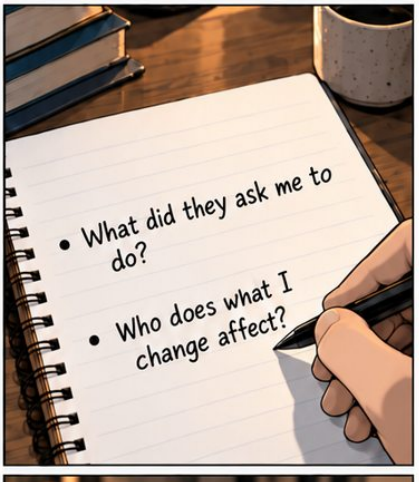

He took up Mara's slot. I've cancelled Leo's job. The baseline evaluation still hasn't been done.
Why didn't you ask for help?
I thought... I'd fix it in time.
Mara's job
Baseline evaluation
Now we've got two jobs delayed. The machine was free. Mara's time wasn't.
Tomorrow, at eight, with Jun. He'll review the next run with you. Before you launch it.


You've made it clear which programs and versions are needed to run the test. That part's useful.
I can explain how I record decisions...
Tomorrow. First I need a comparable result.
She stays with him, though.

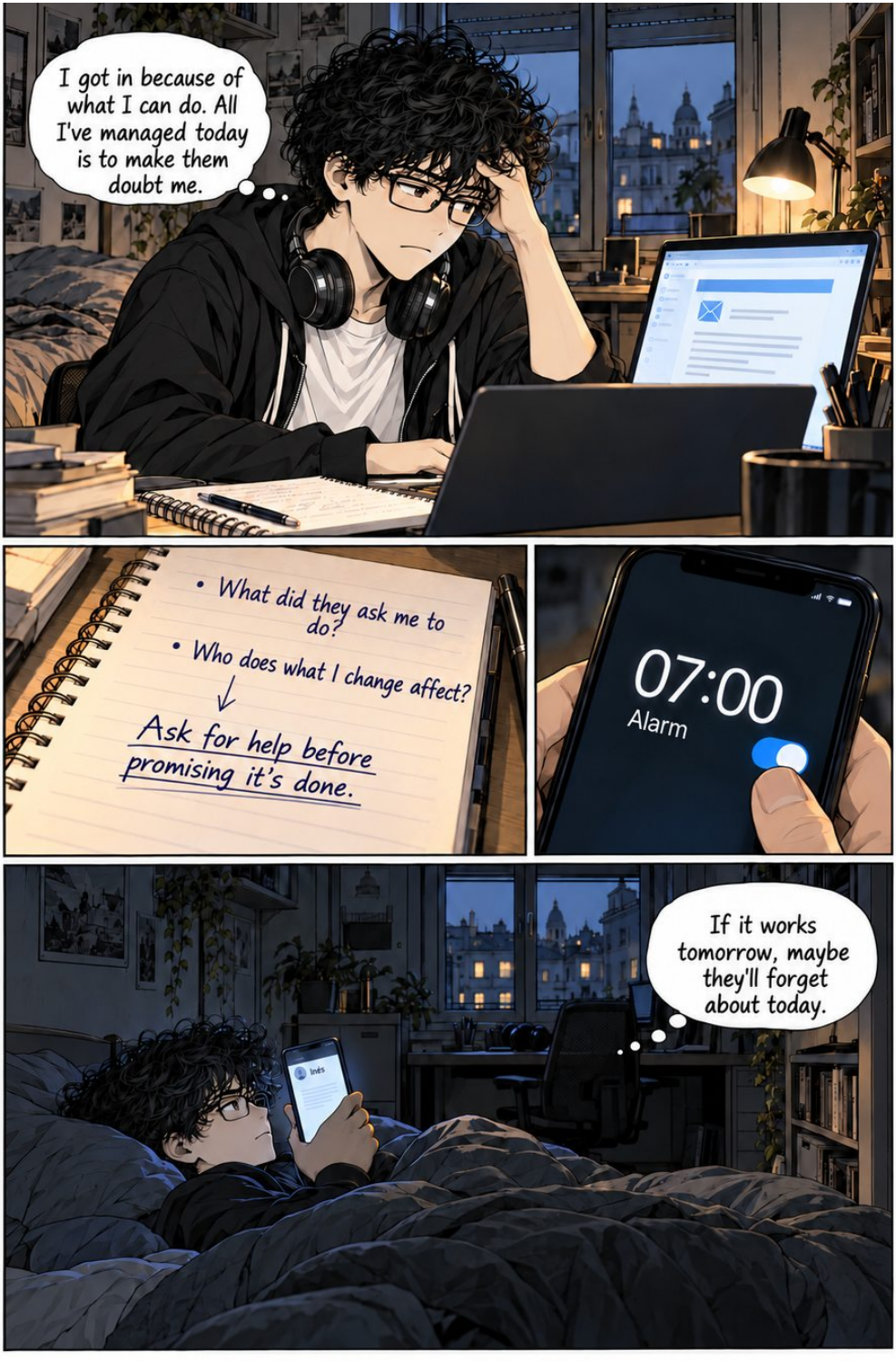

Later
What's your new colleague like?
I don't know yet.
Dad
On Wednesday, they need to see you're up to it.
Electronic music group
I haven't practised the mixes today. I'm going in at eight tomorrow to help Leo. But the live set on 10 October is still on, all right?


At Leo's flat
At last, I don't have to think about what to say.


Language model allowance exhausted
You can keep editing your files on this laptop


What did they ask me to do?
Who does what I change affect?


Inés
08:00 with Jun. Don't send the small test's score.


I got in because of what I can do. All I've managed today is to make them doubt me.
• What did they ask me to do?
• Who does what I change affect?
Ask for help before promising it's done.
07:00
Alarm
If it works tomorrow, maybe they'll forget about today.
Inès

What We Are Looking for in These Images



Images grouped by patient before splitting the data.

Leo sees that the names have been removed and considers anonymity settled. Let's look at what remains: two images might be recognised as samples from the same person. We would already have a link between them, even without knowing that person's name.

One image and a hundred candidates

Track 1 of Benchmarking WSI Privacy asks us to compare a query with a gallery of a hundred images and select those from the same patient. There may be zero, one or several. To teach that relationship, we have examples whose origins we know. Patient codes let us prepare the exercise and mark it; if we give them to the model as a clue, we are slipping it the answer. We want to test what it can infer from the visual content.

Where each image goes

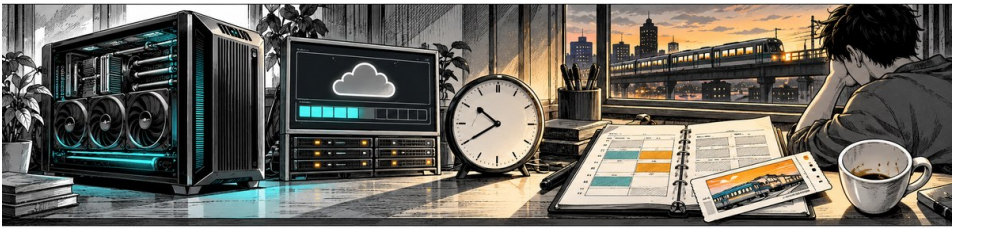
It helps to separate two decisions here. Query and gallery make up the exercise we want to solve. Training and evaluation organise how we teach and examine it. Leo might put different crops from the same sample on both sides and think he has held back new material. That is why we first group all the images from each patient: we avoid evaluating as new a patient whose other crops the system has already seen.

Where 0.93 comes from

That figure belongs to a start-up test that checks connections between programs. Holding patients out changes the exam, and we will need to measure again. F1 also combines correct, false and missed matches; reading it as the percentage of patients identified would lead us to the wrong conclusion. I would start by asking which examples produced the score. The challenge provides the technical question; ATLAS, its funding and its incidents belong to the fiction.

Sources: · WSI challenge · Validation · F1 metric

What the Agent Needs to Know



GPU, requests to a service and people's time.

Leo wants to keep what has been agreed so he can carry on without explaining everything again. His agent uses a large language model, an LLM, to code and use tools. The challenge's vision model compares tissue images. Both use deep learning, with networks of many layers, for different jobs. Code that looks right does not spare us from checking the matches. Jun adds that we can also write something wrong very clearly.

The instructions we keep outside the model

Specification-driven development, SDD, describes what the program should do and how to check it. The file stores project decisions. The agent must read them or receive them in the task context, along with the information needed to respond. The model's learnt parameters do not change. Even with the document, we must check whether it uses the right version and understands what we intended.

The test can share the error

Recording the decision, test and result helps us reconstruct a change. Imagine the code and its test start from the same mistaken condition: both may look consistent. Leo installs a library and records its version, but leaves the revised patient list unconnected. The separation he was meant to check remains pending, and the record will let us locate where he stopped.

What is running out

Agents' requests to the language model go through an API, the interface between programs. Text sent and received is measured in tokens and uses the service allowance. I would separate that cost from the time and memory of the shared GPU on which ATLAS runs its vision models. The language model needs computation too; here, the provider supplies it. Leo can exhaust his requests and keep editing local files. That does not give him Mara's slot. When the machine finishes, someone must be available to review the output.

Sources: · Agent context

Leo's First Day



One task selected among several windows and notifications.

Before anyone at ATLAS rejects him, Leo is already preparing to make a bad impression. He arrives wanting a fresh start and watching for the others to discover the odd person he believes himself to be. That worry stays with him even when he has something useful to contribute.

What Inés has seen in him

He got in through a test in which he checked a figure and left the steps needed to repeat it. He knows how to build tools and asks questions when Inés explains the challenge. There are reasons for him to be there. Yet he soon needs his first contribution to confirm all of that in front of the group. The simple test he has been assigned starts carrying the obligation to make a good impression as well.

The invitation he was waiting for

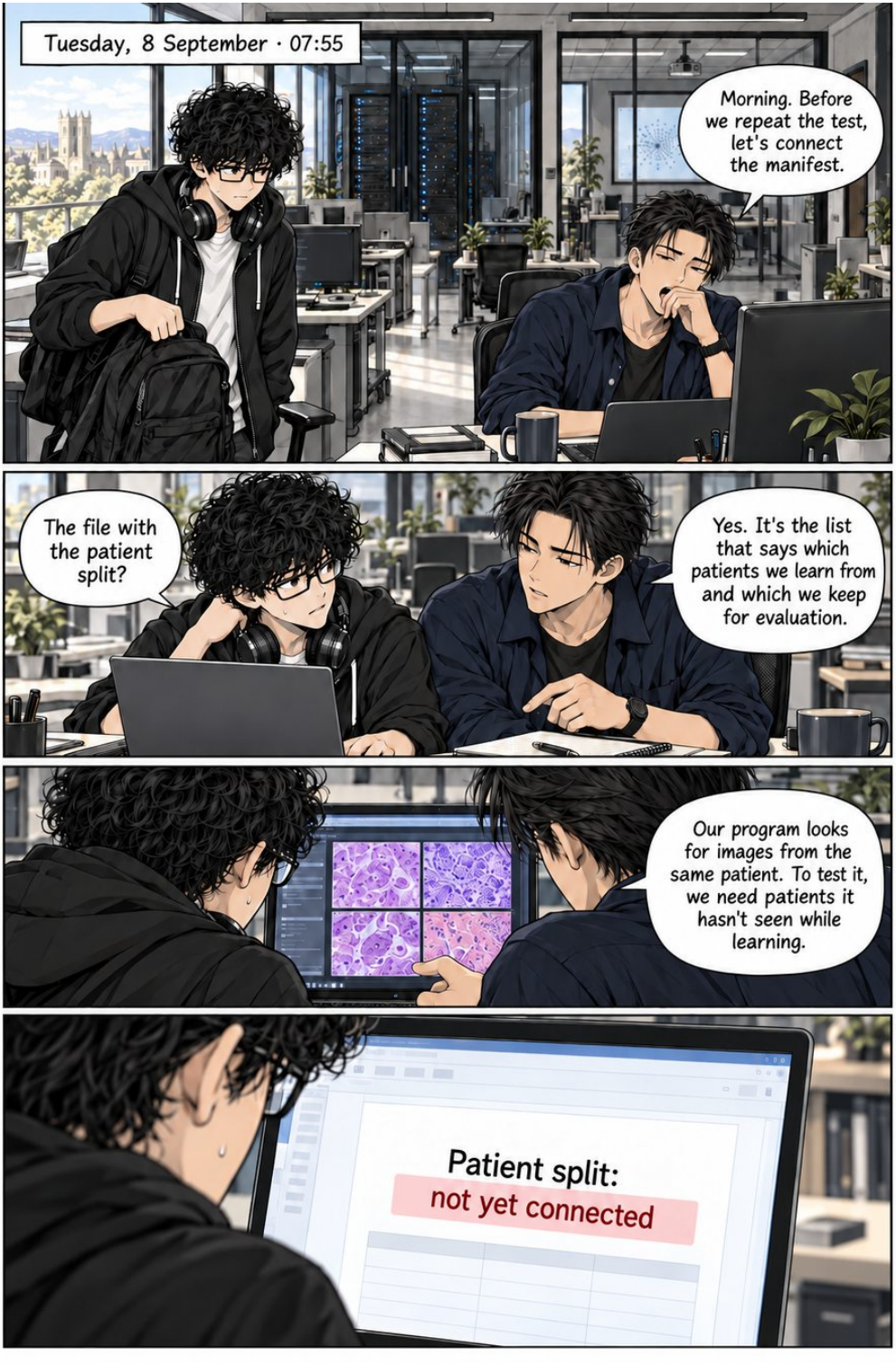
Over lunch he tries to join in with a technical explanation. Later they invite him along to something and he dodges it, although he wanted to be invited. He even writes Jun a request for help, then deletes it. Getting closer appeals to him and leaves him exposed; going home lets him rest from watching every response.

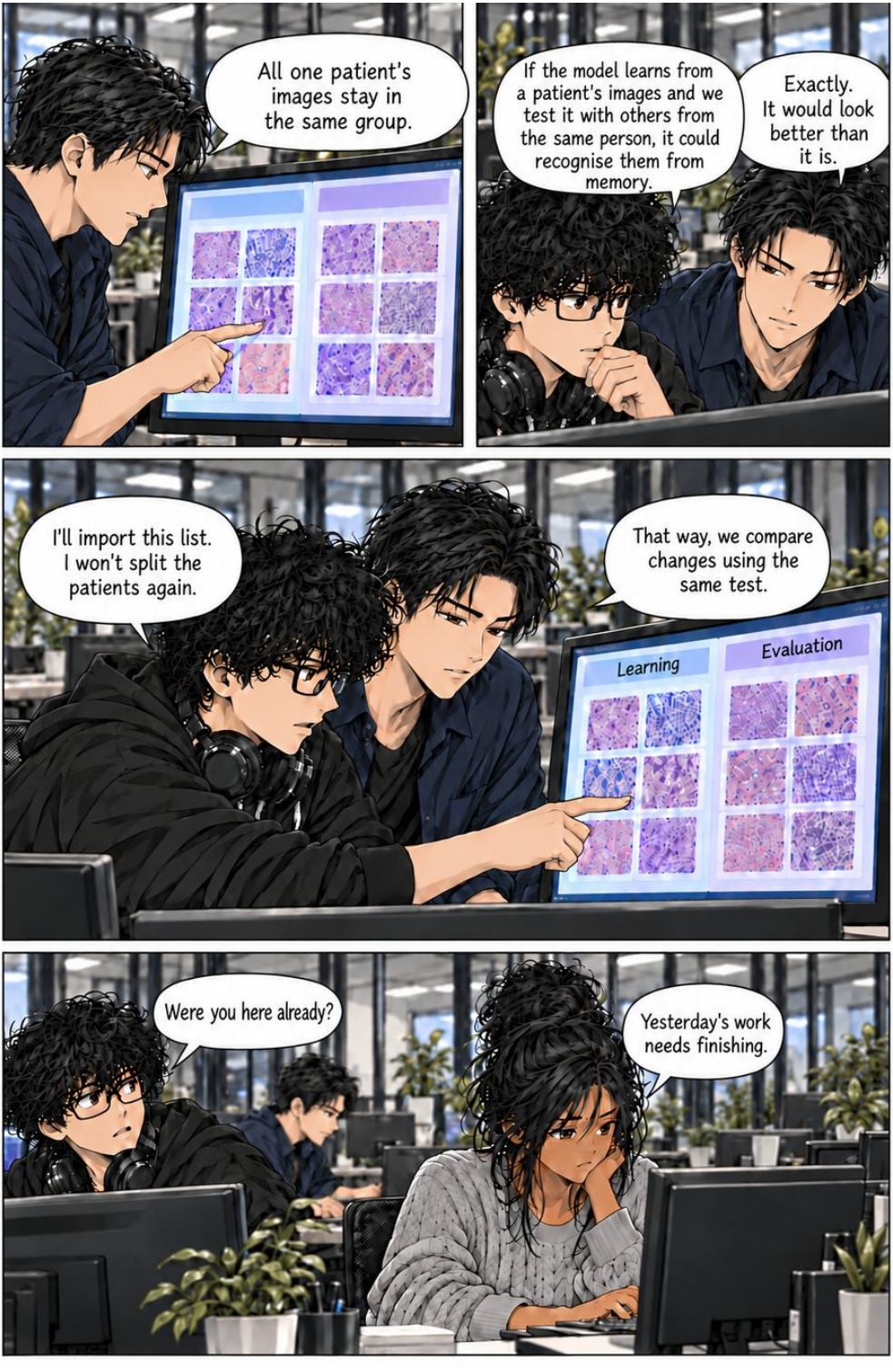
Mara has a slot booked

Leo keeps working and pushes Mara's review back until after her shift. She had already given her timetable. I want to remember that detail because the solution starts with noticing information he already had in front of him. In his notebook he writes that he should ask for help before making promises, but he still hopes a brilliant result will make up for his debut. We will see what he does with that intention the next time he gets stuck.

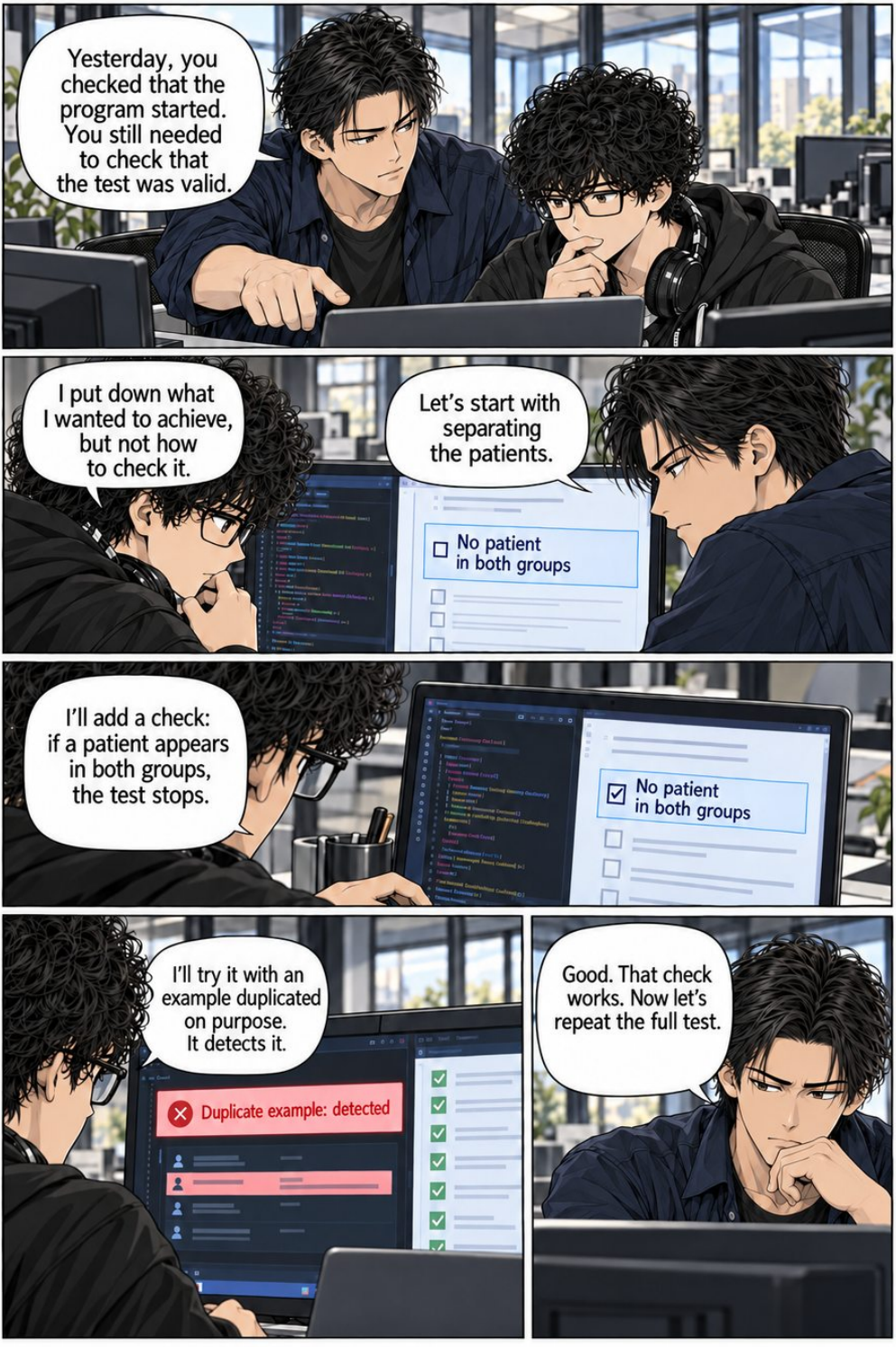
What could he have written in the message to Jun before deleting it?

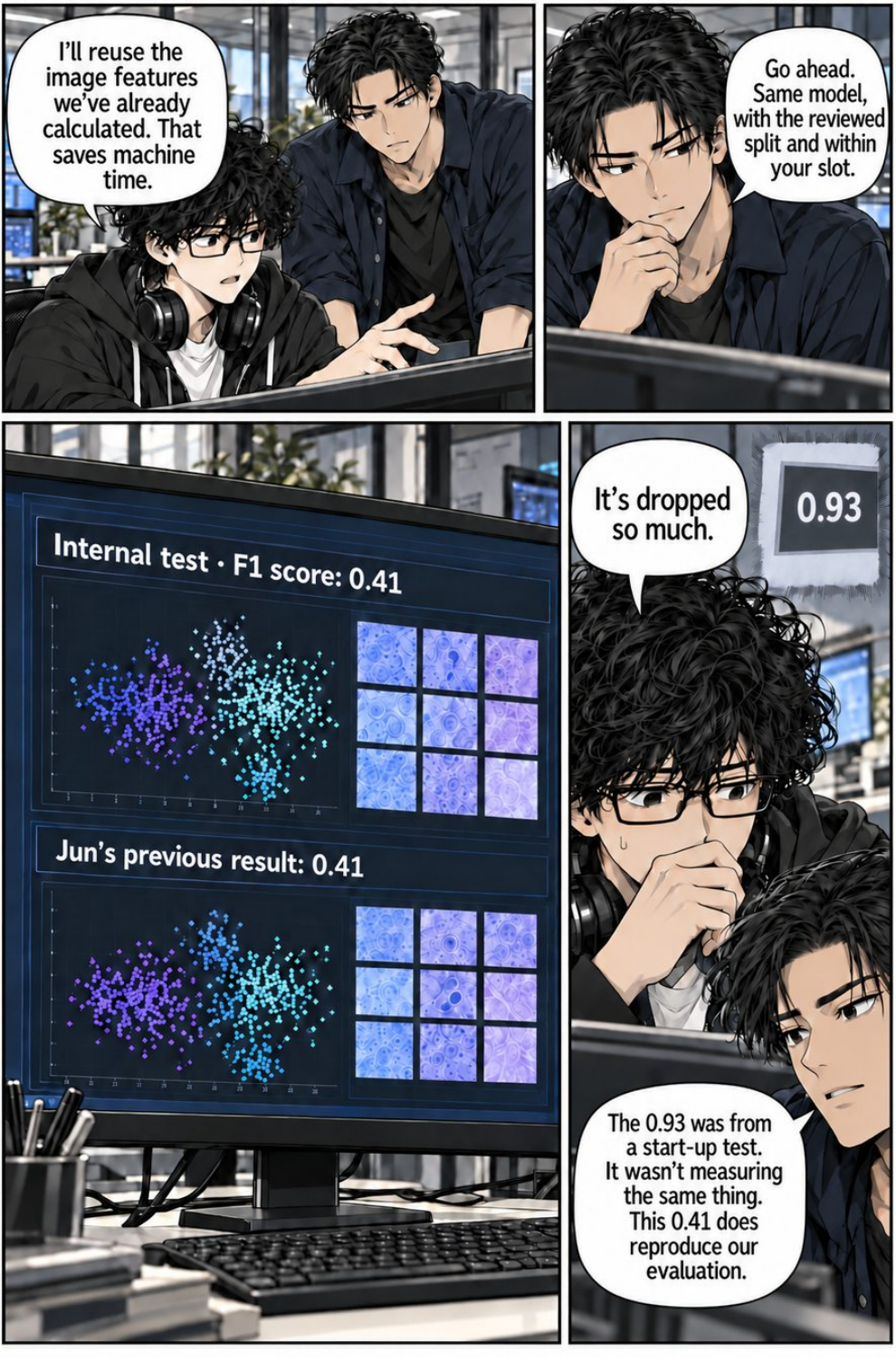

Monday, 7 September · Night
La Segunda Taza
When I get home, I still have to review the images.
And I'm back in early tomorrow.


Tuesday, 8 September · 07:55
Morning. Before we repeat the test, let's connect the manifest.
The file with the patient split?
Yes. It's the list that says which patients we learn from and which we keep for evaluation.
Our program looks for images from the same patient. To test it, we need patients it hasn't seen while learning.
Patient split:
not yet connected

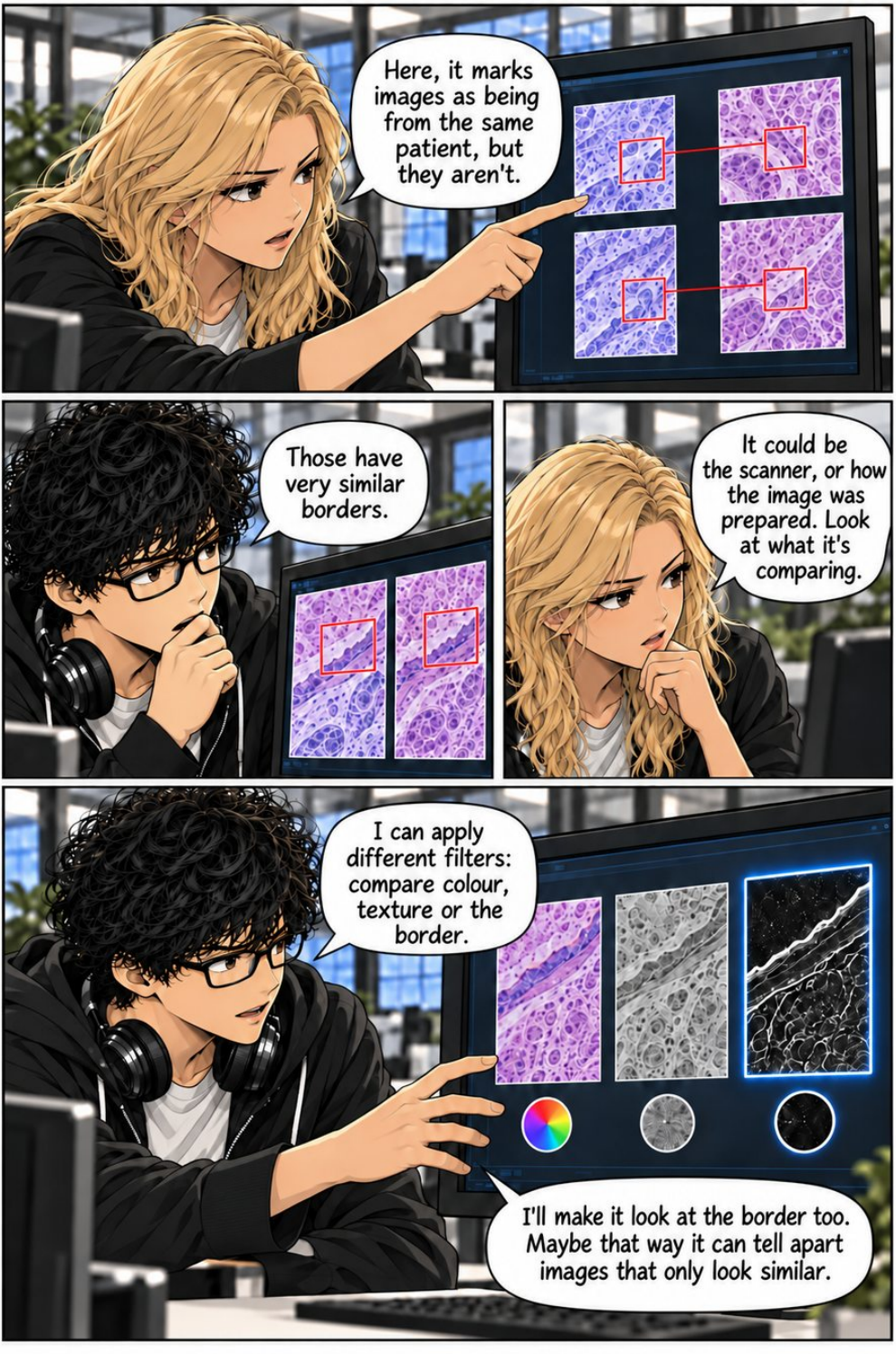

All one patient's images stay in the same group.
If the model learns from a patient's images and we test it with others from the same person, it could recognise them from memory.
Exactly. It would look better than it is.
I'll import this list. I won't split the patients again.
That way, we compare changes using the same test.
Learning
Evaluation
Were you here already?
Yesterday's work needs finishing.

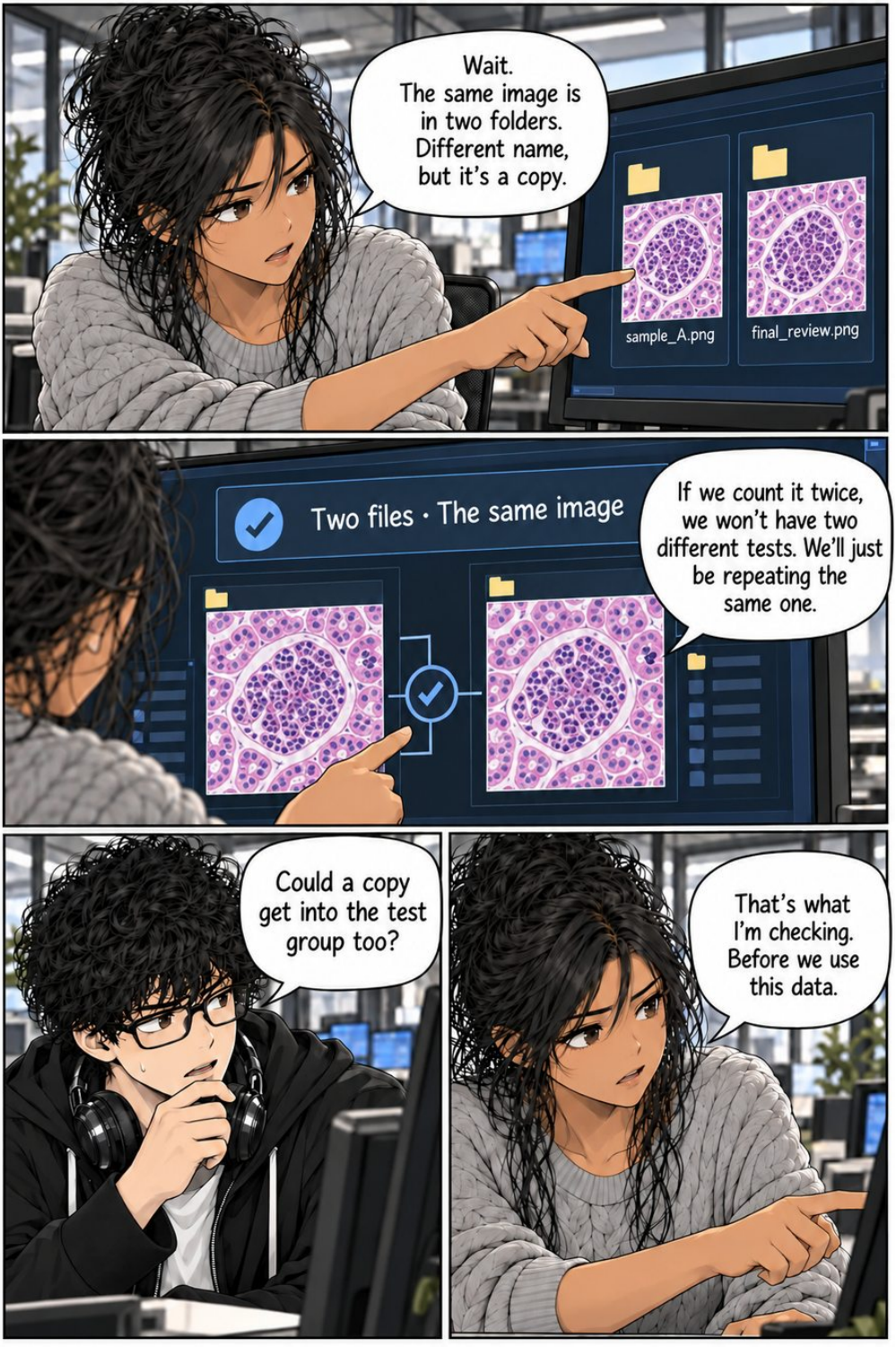

Sorry about yesterday. I didn't mean to... Well, I did it.
After closing the café, I still had to review the lab images.
I need the machine again today. If anything changes, tell me beforehand.
Yes. Beforehand.
I don't know what else to say.

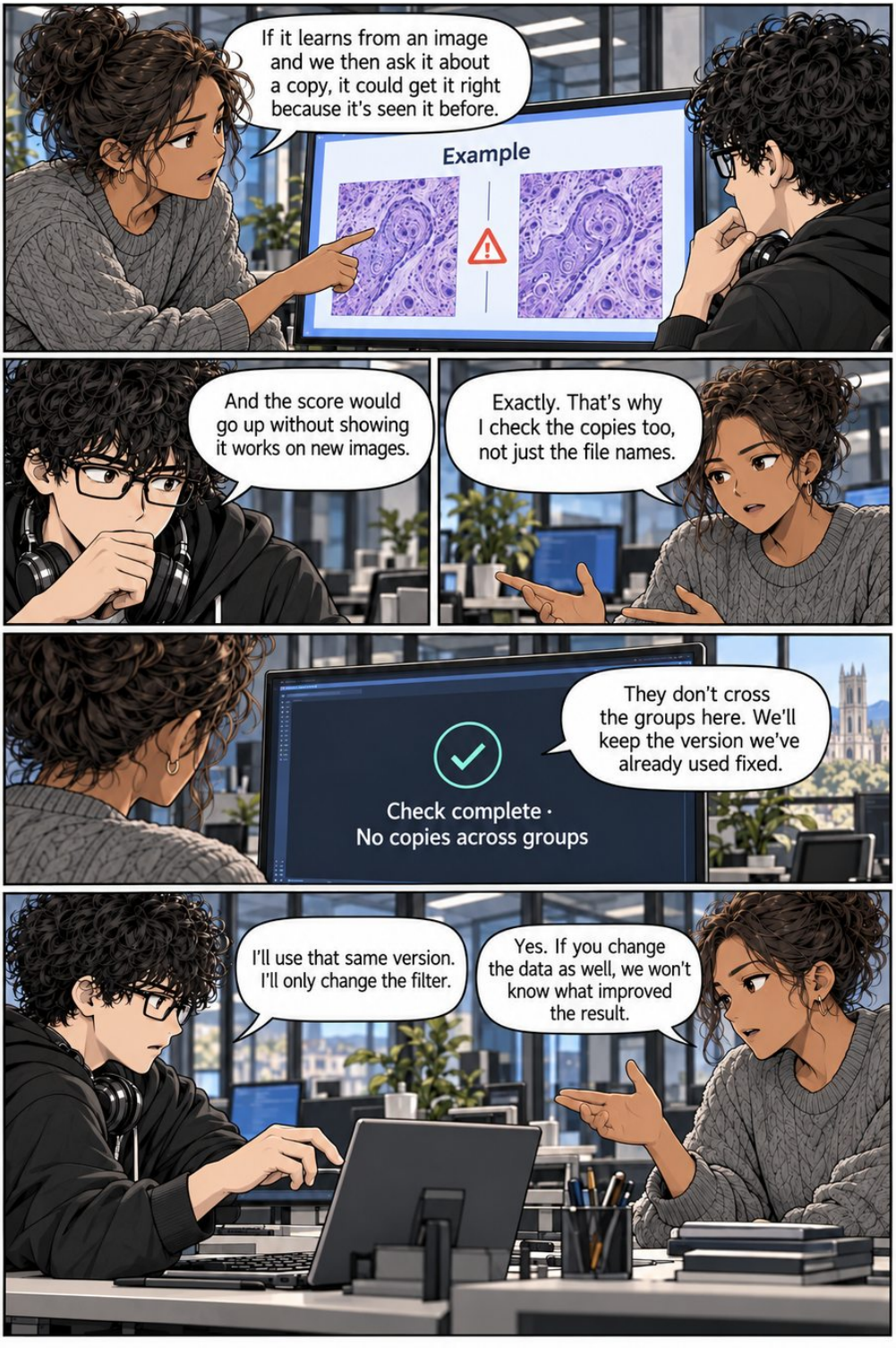

Yesterday, you checked that the program started. You still needed to check that the test was valid.
I put down what I wanted to achieve, but not how to check it.
Let's start with separating the patients.
No patient in both groups
I'll add a check: if a patient appears in both groups, the test stops.
No patient in both groups
I'll try it with an example duplicated on purpose. It detects it.
Duplicate example: detected
Good. That check works. Now let's repeat the full test.


I'll reuse the image features we've already calculated. That saves machine time.
Go ahead. Same model, with the reviewed split and within your slot.
Internal test · F1 score: 0.41
Jun's previous result: 0.41
It's dropped so much.
0.93
The 0.93 was from a start-up test. It wasn't measuring the same thing. This 0.41 does reproduce our evaluation.

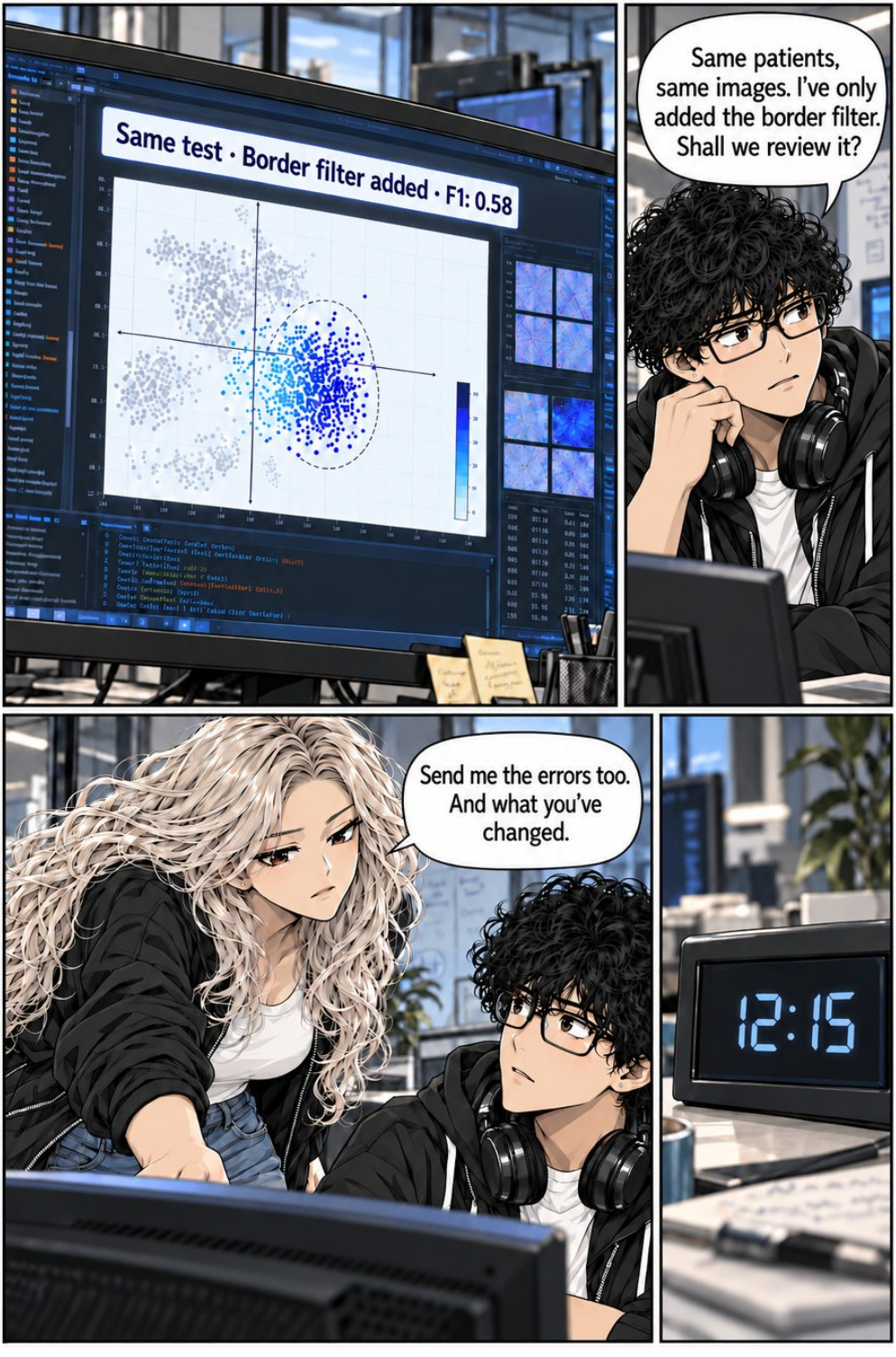

It's reproducible. That's what I asked you for.
But tomorrow, we can't show them a nought point four one...
We'll show them what we've got, with its limitations.
You can try a small improvement, respecting the bookings. We finalise the report at five.
I can still improve that number.


Here, it marks images as being from the same patient, but they aren't.
Those have very similar borders.
It could be the scanner, or how the image was prepared. Look at what it's comparing.
I can apply different filters: compare colour, texture or the border.
I'll make it look at the border too. Maybe that way it can tell apart images that only look similar.

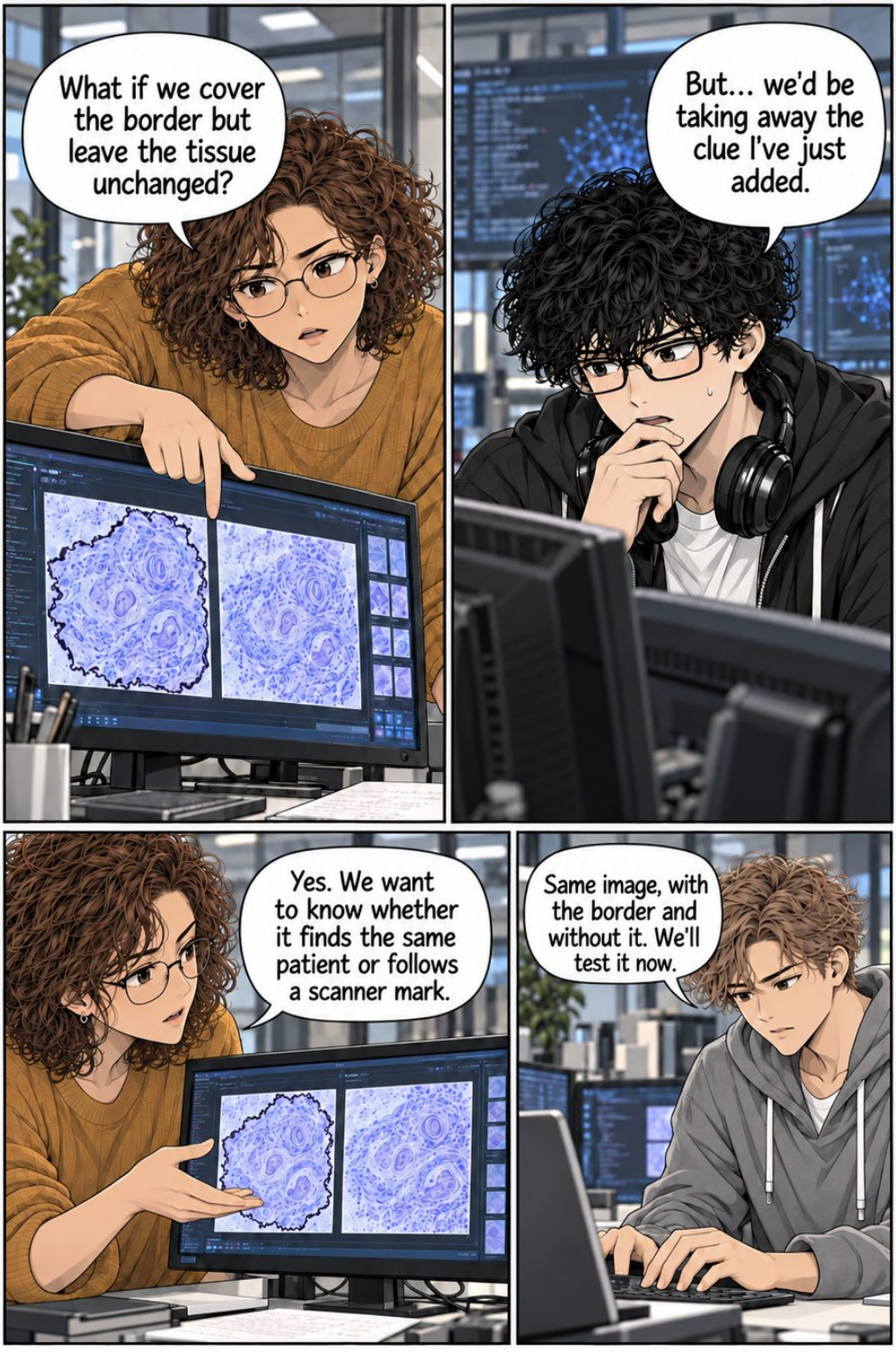

Wait.
The same image is in two folders. Different name, but it's a copy.
sample_A.png
final_review.png
Two files · The same image
If we count it twice, we won't have two different tests. We'll just be repeating the same one.
Could a copy get into the test group too?
That's what I'm checking. Before we use this data.

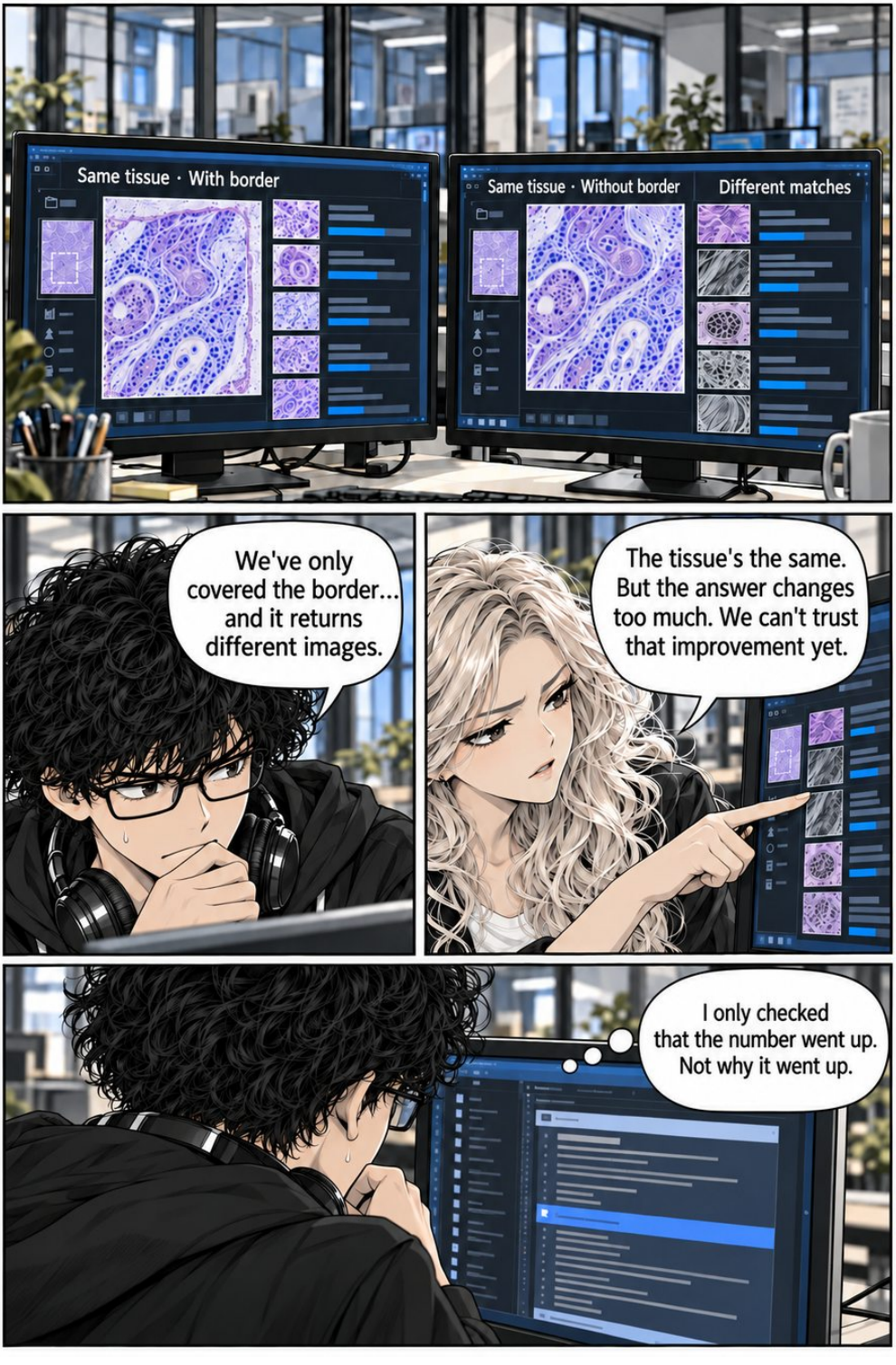

If it learns from an image and we then ask it about a copy, it could get it right because it's seen it before.
Example
And the score would go up without showing it works on new images.
Exactly. That's why I check the copies too, not just the file names.
They don't cross the groups here. We'll keep the version we've already used fixed.
Check complete · No copies across groups
I'll use that same version. I'll only change the filter.
Yes. If you change the data as well, we won't know what improved the result.

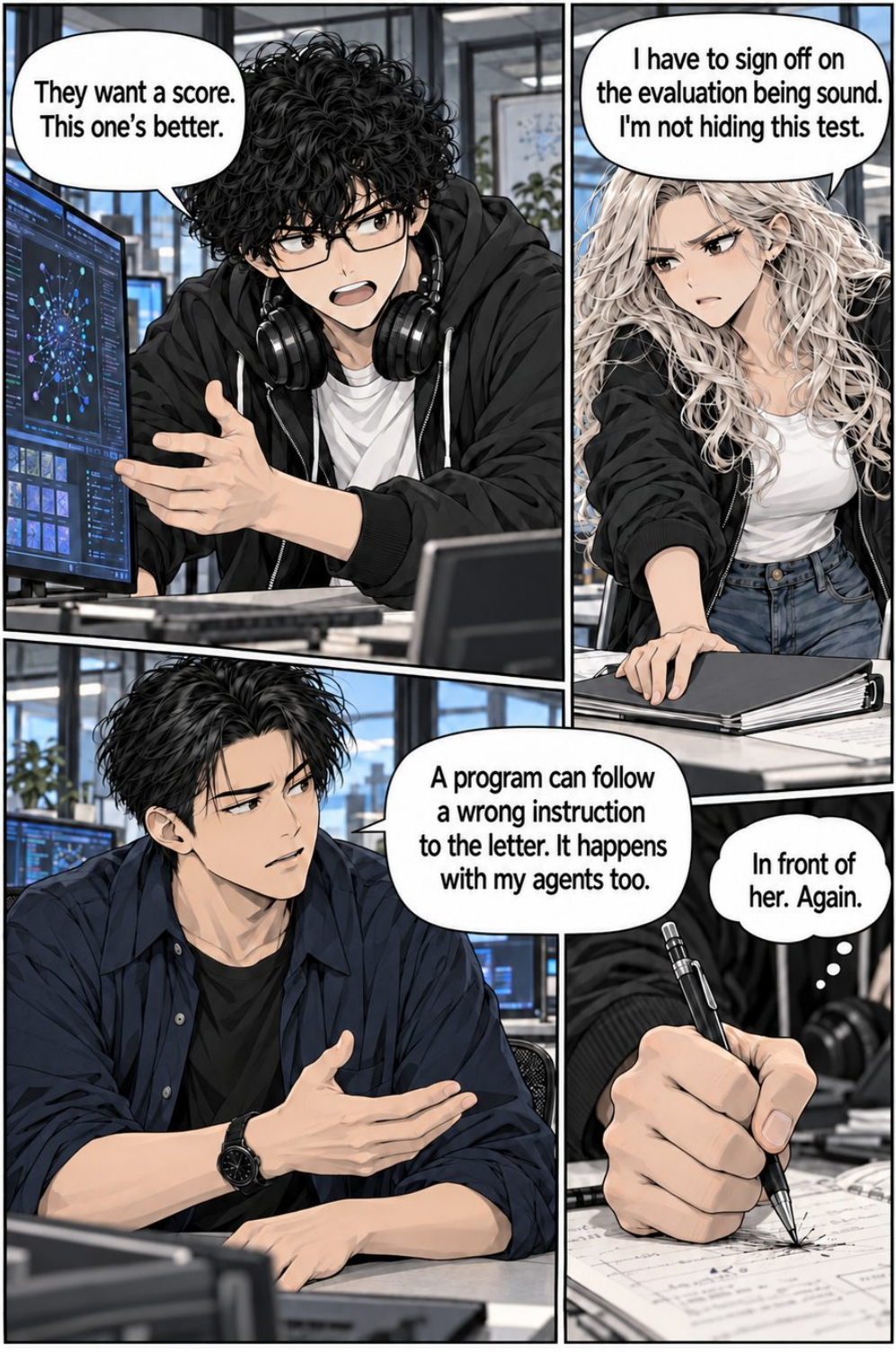

I'll give both versions the same images and compare their answers.
I can save every change and the tests it needs...
Send me the new version first. Tell me the rest while it's running.
Don't leave us in the lurch again on Friday.


Same test · Border filter added · F1: 0.58
Same patients, same images. I've only added the border filter. Shall we review it?
Send me the errors too. And what you've changed.
12:15

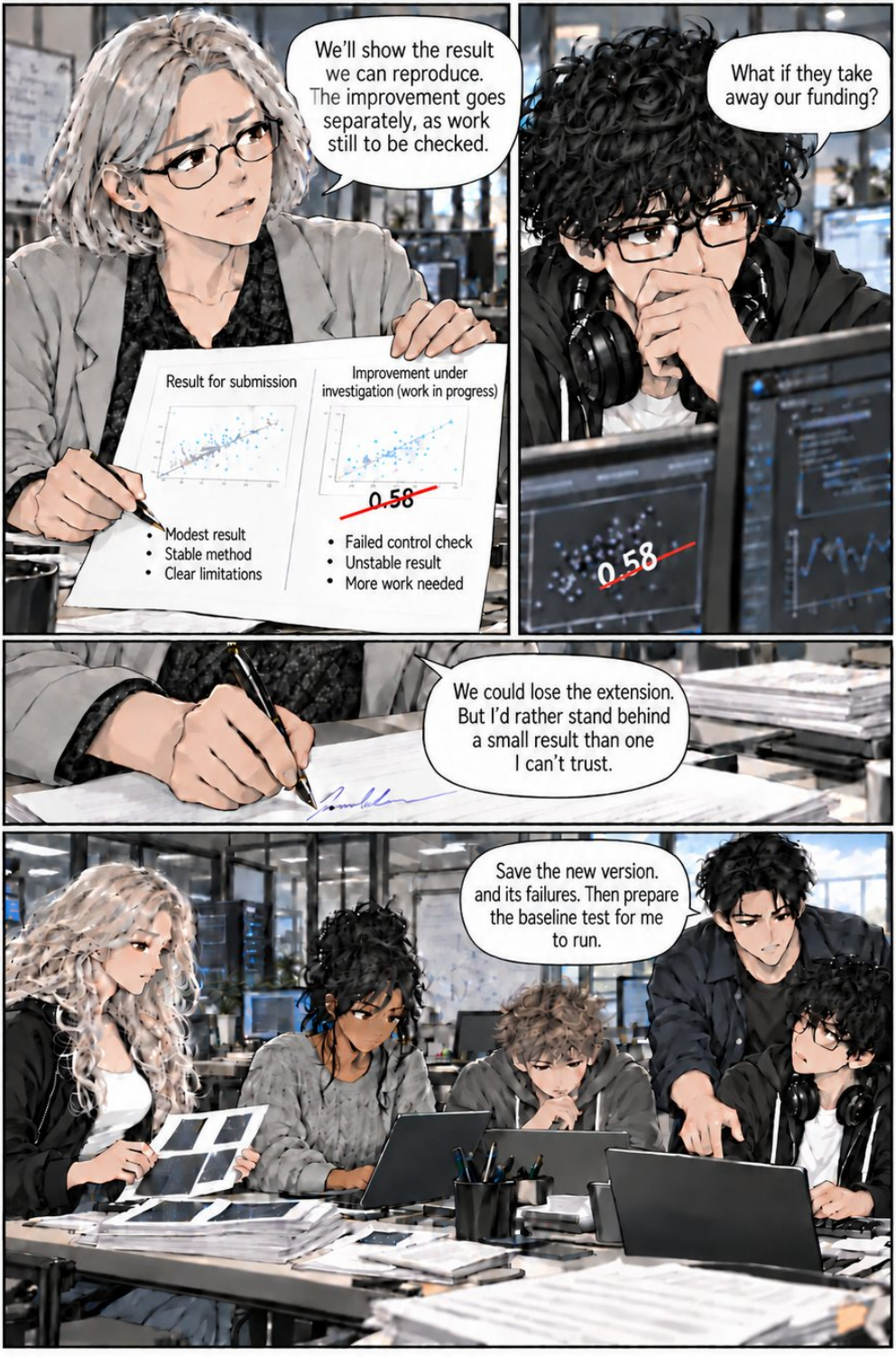

Leo, has the test improved?
I'm preparing the video for tomorrow's review.
We've gone up...
It's under review. There's no figure to publish yet.
Then we'll explain the problem, and come back with the result tomorrow.

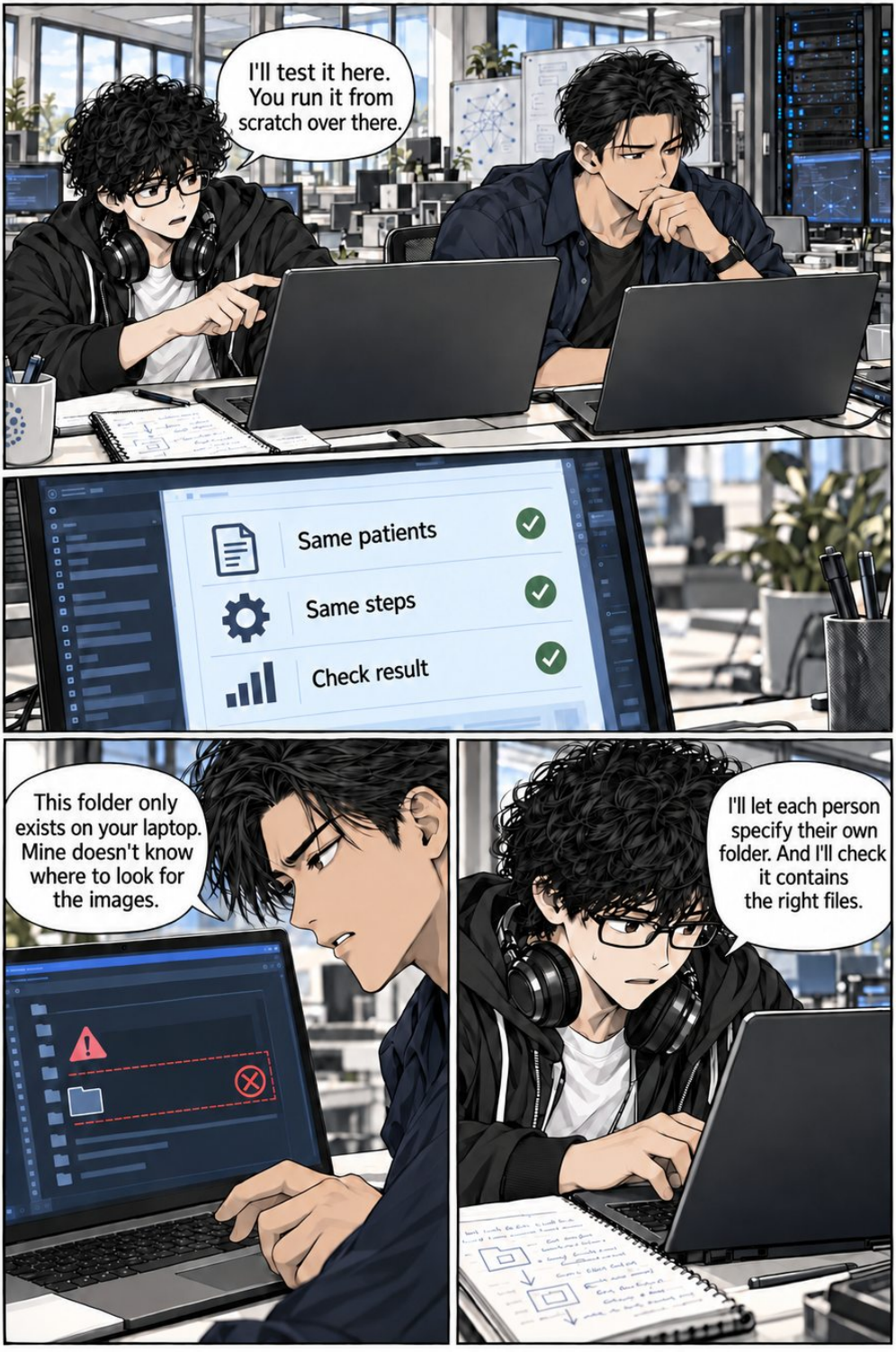

What if we cover the border but leave the tissue unchanged?
But... we'd be taking away the clue I've just added.
Yes. We want to know whether it finds the same patient or follows a scanner mark.
Same image, with the border and without it. We'll test it now.


Same tissue · With border
Same tissue · Without border
Different matches
We've only covered the border... and it returns different images.
The tissue's the same. But the answer changes too much. We can't trust that improvement yet.
I only checked that the number went up. Not why it went up.

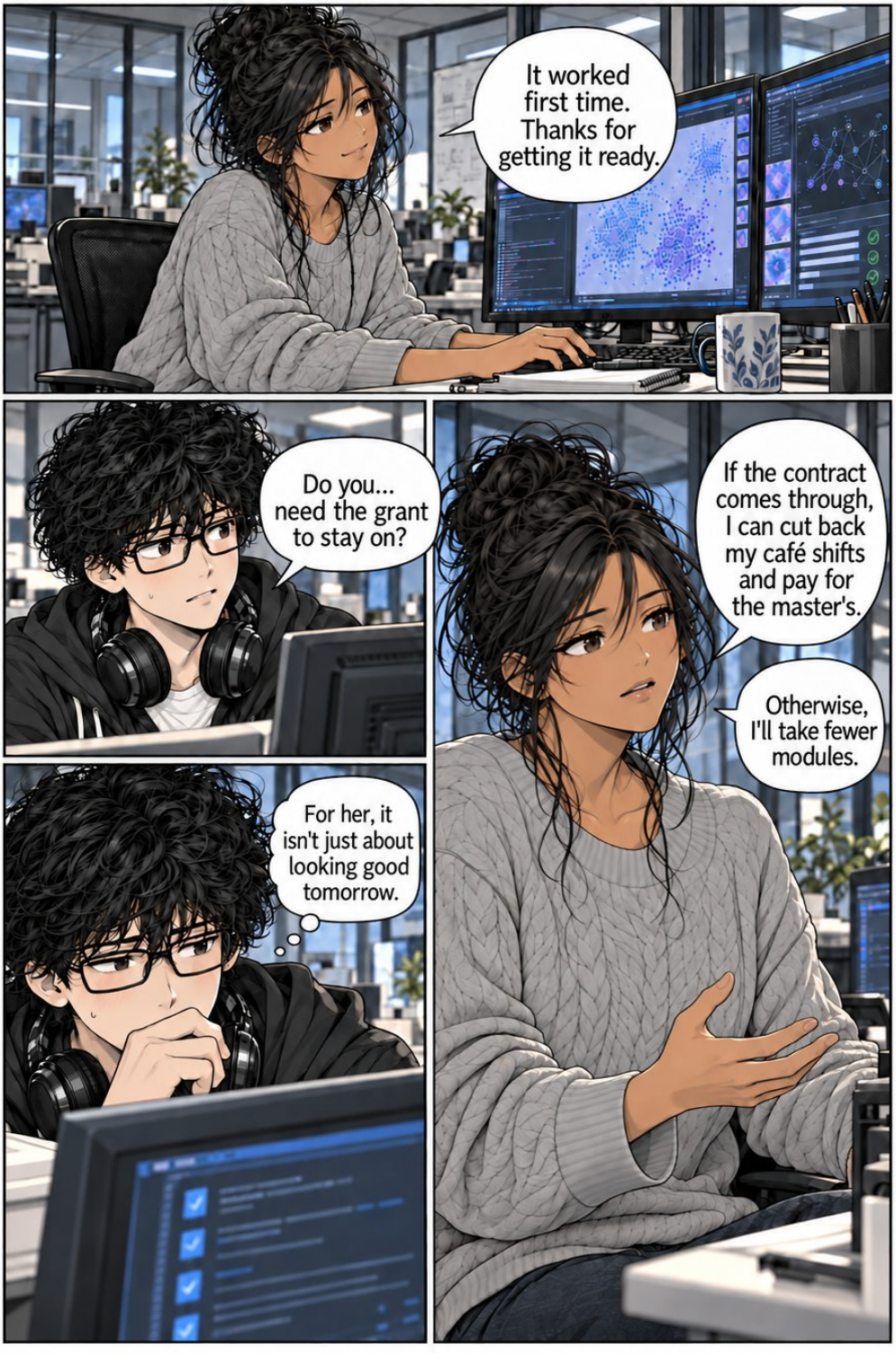

They want a score. This one's better.
I have to sign off on the evaluation being sound. I'm not hiding this test.
A program can follow a wrong instruction to the letter. It happens with my agents too.
In front of her. Again.

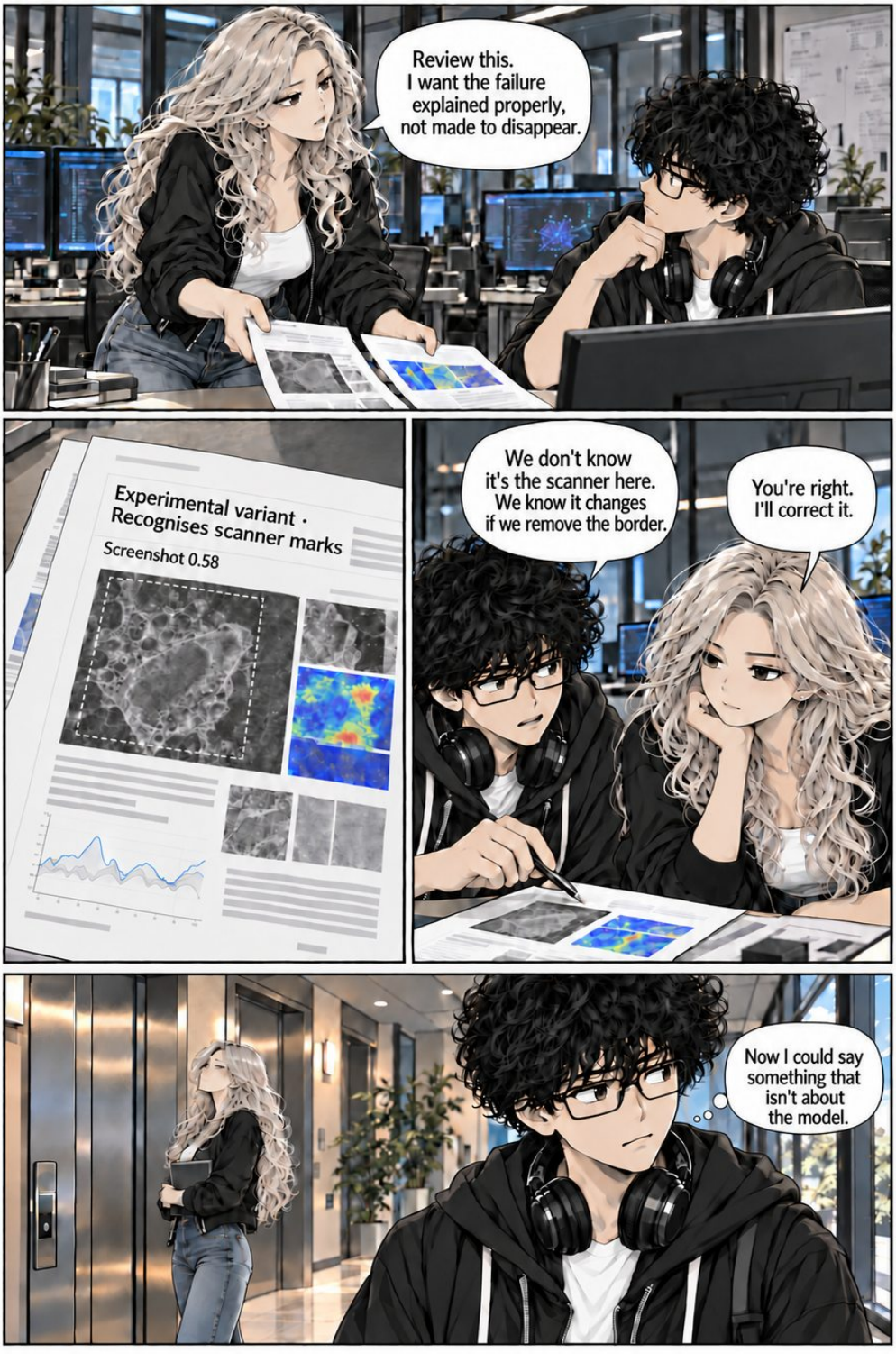

Can we stand behind the new version tomorrow?
Not yet. We can show the baseline test, the one that scored 0.41.
I'll leave you the remaining checks. I'm working on another project this afternoon.
If we leave it running tonight...
The machine's already booked tonight. We can't count on other people's time again.


We'll show the result we can reproduce. The improvement goes separately, as work still to be checked.
Result for submission
Improvement under investigation (work in progress)
0.58
Modest result
Stable method
Clear limitations
Failed control check
Unstable result
More work needed
What if they take away our funding?
0.58
We could lose the extension. But I'd rather stand behind a small result than one I can't trust.
Save the new version. and its failures. Then prepare the baseline test for me to run.

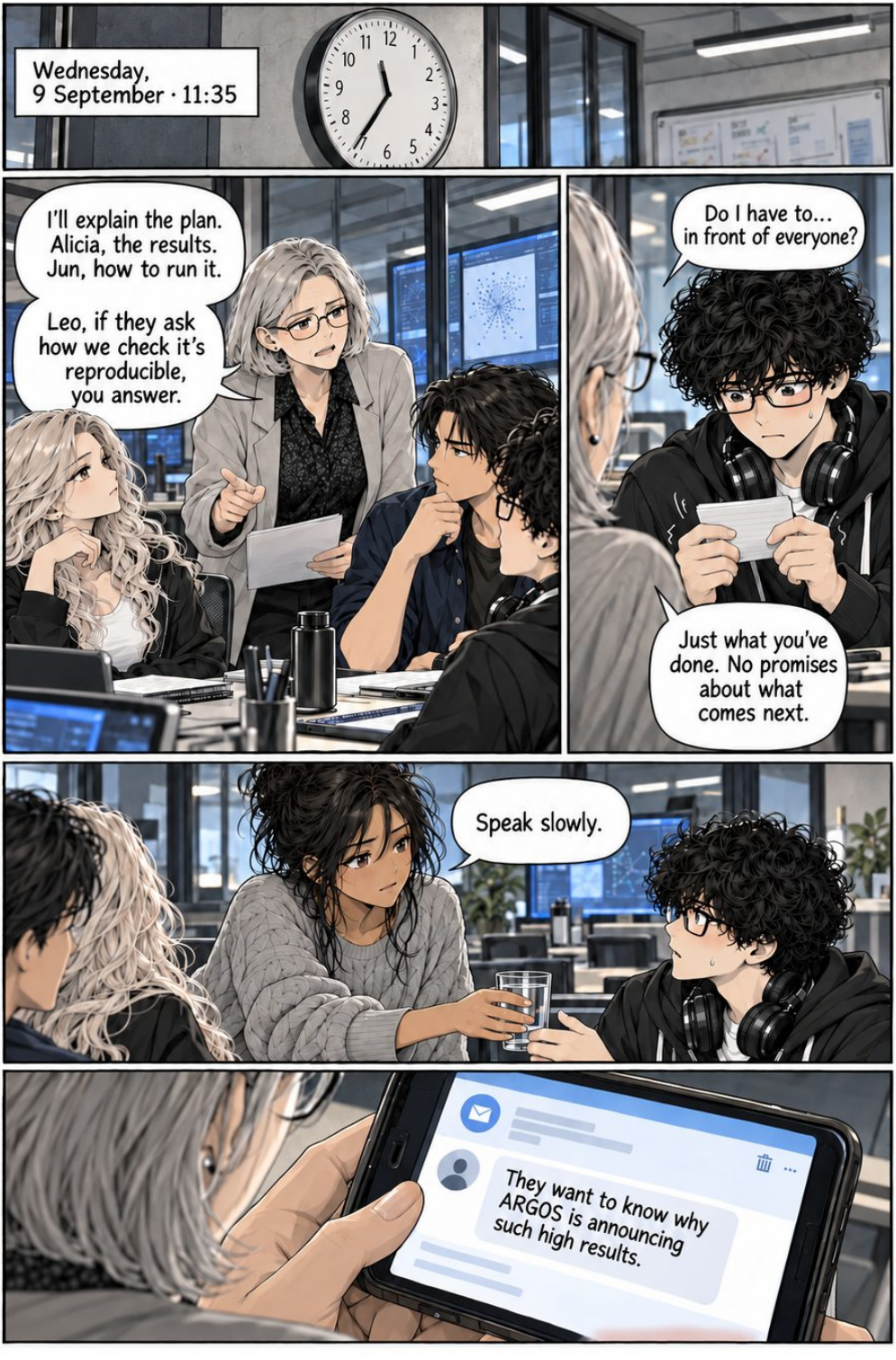

I'll test it here. You run it from scratch over there.
Same patients
Same steps
Check result
This folder only exists on your laptop. Mine doesn't know where to look for the images.
I'll let each person specify their own folder. And I'll check it contains the right files.


Same test on another computer · F1: 0.41
Now I can change something and know what I've broken. This does save me work.
Right. Well. That's what I wanted...
Machine free · 15:28

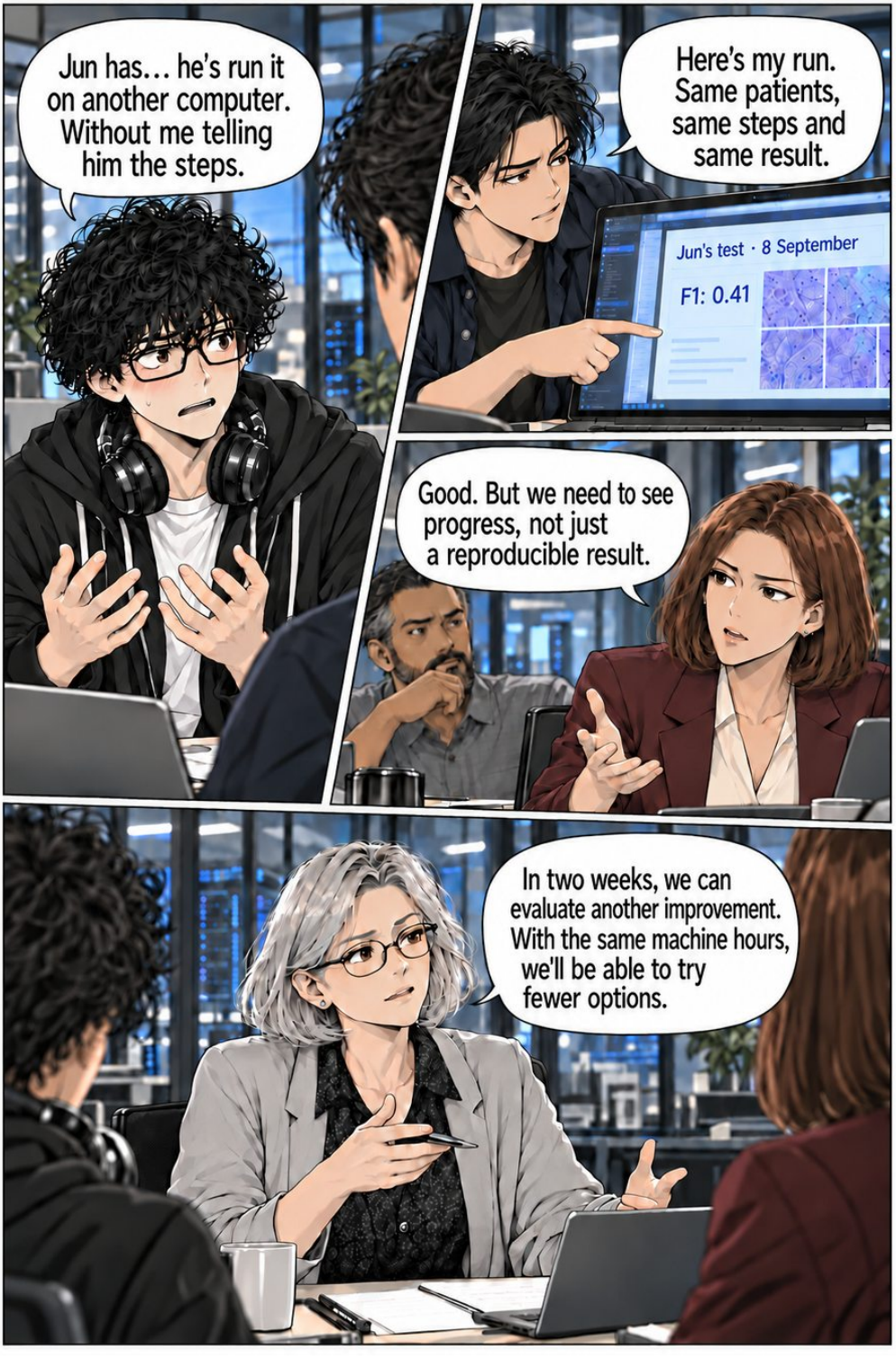

It worked first time. Thanks for getting it ready.
Do you... need the grant to stay on?
If the contract comes through, I can cut back my café shifts and pay for the master's.
Otherwise, I'll take fewer modules.
For her, it isn't just about looking good tomorrow.


Review this.
I want the failure explained properly, not made to disappear.
Experimental variant · Recognises scanner marks
Screenshot 0.58
We don't know it's the scanner here. We know it changes if we remove the border.
You're right. I'll correct it.
Now I could say something that isn't about the model.


Come to dinner after the review. You'll have earned it.
I don't even know what time I'll finish tomorrow. I'll let you know
You're always working.
It's important to me.
I thought she never had to try at anything.


Wednesday,
9 September · 11:35
I'll explain the plan. Alicia, the results. Jun, how to run it.
Leo, if they ask how we check it's reproducible, you answer.
Do I have to... in front of everyone?
Just what you've done. No promises about what comes next.
Speak slowly.
They want to know why ARGOS is announcing such high results.


12:00 · Internal funding review
F1 score: 0.41
Same patients • Same result
We've repeated the test and obtained the same result. The higher-scoring version still fails one check.
ARGOS is announcing much higher figures. What assurance do we have that you're on the right track?
We don't know whether they're measuring the same thing. We can show you exactly what this measures.
And if someone other than you runs it?


Jun has… he's run it on another computer. Without me telling him the steps.
Here's my run. Same patients, same steps and same result.
Jun's test · 8 September
F1: 0.41
Good. But we need to see progress, not just a reproducible result.
In two weeks, we can evaluate another improvement. With the same machine hours, we'll be able to try fewer options.


The budget extension is pending. We'll meet again in two weeks. Until then, you continue with the current funding.
NEXO
Budget extension
PENDING
Agreed. But let's put in writing what we can do with that budget.
Two more weeks without knowing.
Two contracts subject to funding
If I'd done something better...


What shall I put? The team makes progress despite difficulties?
ATLAS
That we've reproduced the baseline test and we're investigating an improvement. Don't say it works yet.
DC
Damián Costa · ARGOS
If you want real results, we'll show you ours.
We carry on tomorrow. Before booking, talk to Mara.
Yes. I'll ask her this time.


Elsewhere
ATLAS · Internal report · Do not distribute
Where did you get it?
Don't ask.
See if it's any use to us.
At Leo's flat
Does the test measure what I want to demonstrate?
Today, I showed what really works. It wasn't enough to get the money.
And I still find it hard not to feel it's my failure.

Checking What Has Improved



Images grouped by patient before splitting the data.

Leo finds it hard to celebrate 0.41 after showing everyone a 0.93. Yet he has now repeated the evaluation he was asked to do. When 0.58 appears, the interesting question changes: what is that variant exploiting to raise its score?

Patients, copies and filters

The patient check stops the test if one person appears in both groups. Mara also finds two files containing the same image: different filenames do not create two observations. These are different problems and need different checks. Once reviewed, keeping the images and changing only the filter helps us study the difference that has just appeared in the table.

The test Nadia proposes

Nadia covers the border and leaves the tissue. The matches change, so the supposed improvement depended on something that intervention altered. That is as far as we can go. Blaming the scanner now would be premature: other details of preparation could be involved. Leo is right to correct that sentence in the report, because the explanation we choose will determine what we test next. If we invent the cause, we may spend the next afternoon fixing something else.

Jun takes the steps to another computer

The baseline fails when it looks for a path that only exists on Leo's laptop. Once the folder can be specified and its files checked, Jun manages to reproduce 0.41 from scratch. We now have something another colleague can run. We are still within the working environment shown; taking it to another hospital would require more checks. Inés presents that baseline and keeps the failed variant with its results. We also want to save attempts that help us understand where to go next.

Sources: · Preparation and leakage · F1 metric

Leaving the Machine Free for Mara



A sofa, a table and several uses for the same space.

A score of 0.41 feels meagre after announcing a huge figure. Even so, Leo has just done the test he was asked for, and Jun can repeat it. To understand how his day has gone, I would also look at what changes for the people working with him.

An apology with a specific time

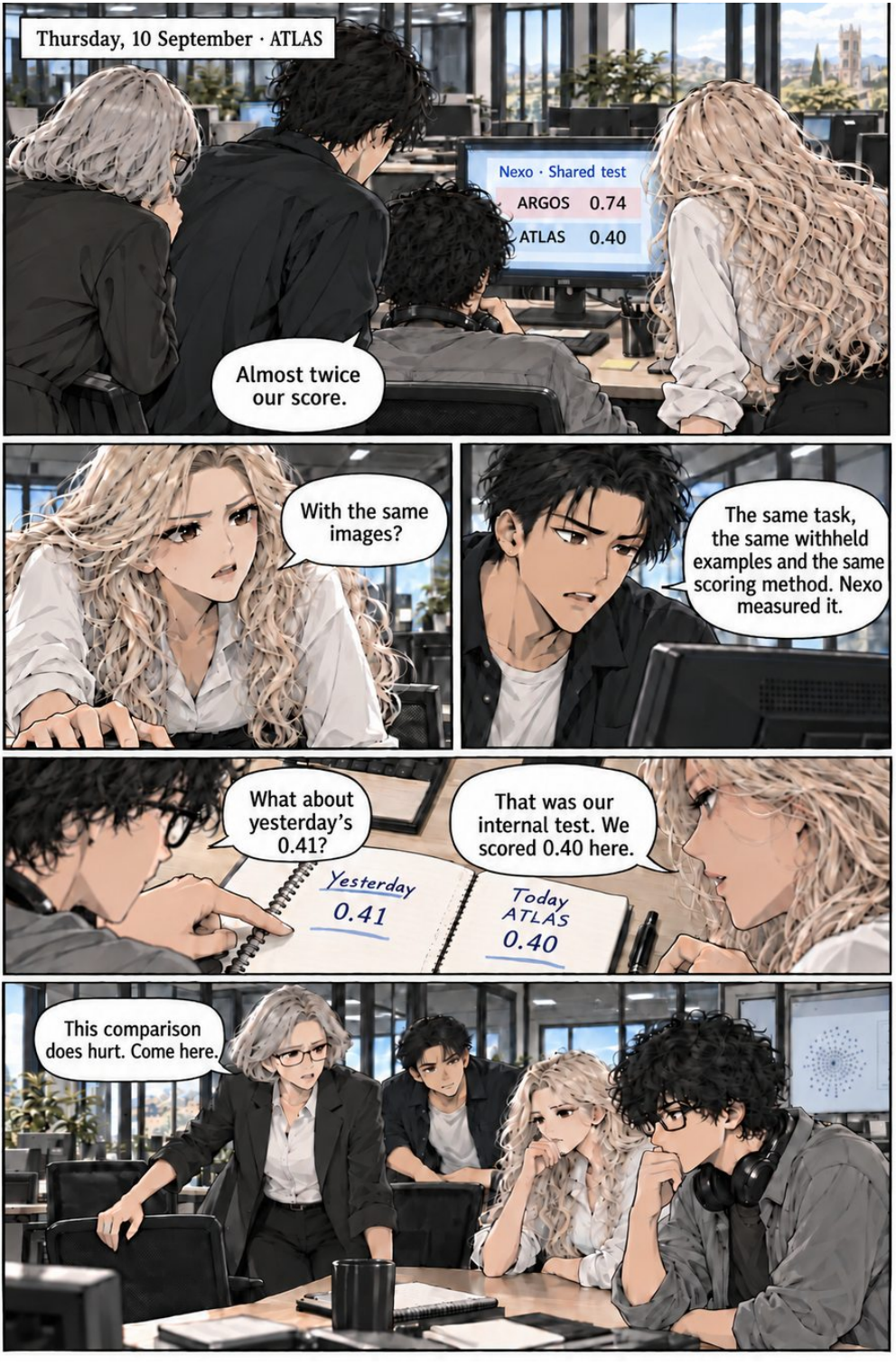
Mara had to review the report after closing the café. Apologising does not give her that night back, but leaving the machine free now lets her start on time. Jun also came in early to help Leo. If we look only at the model's result, we miss those two parts of making amends: honouring the booking and acknowledging the time someone else put in.

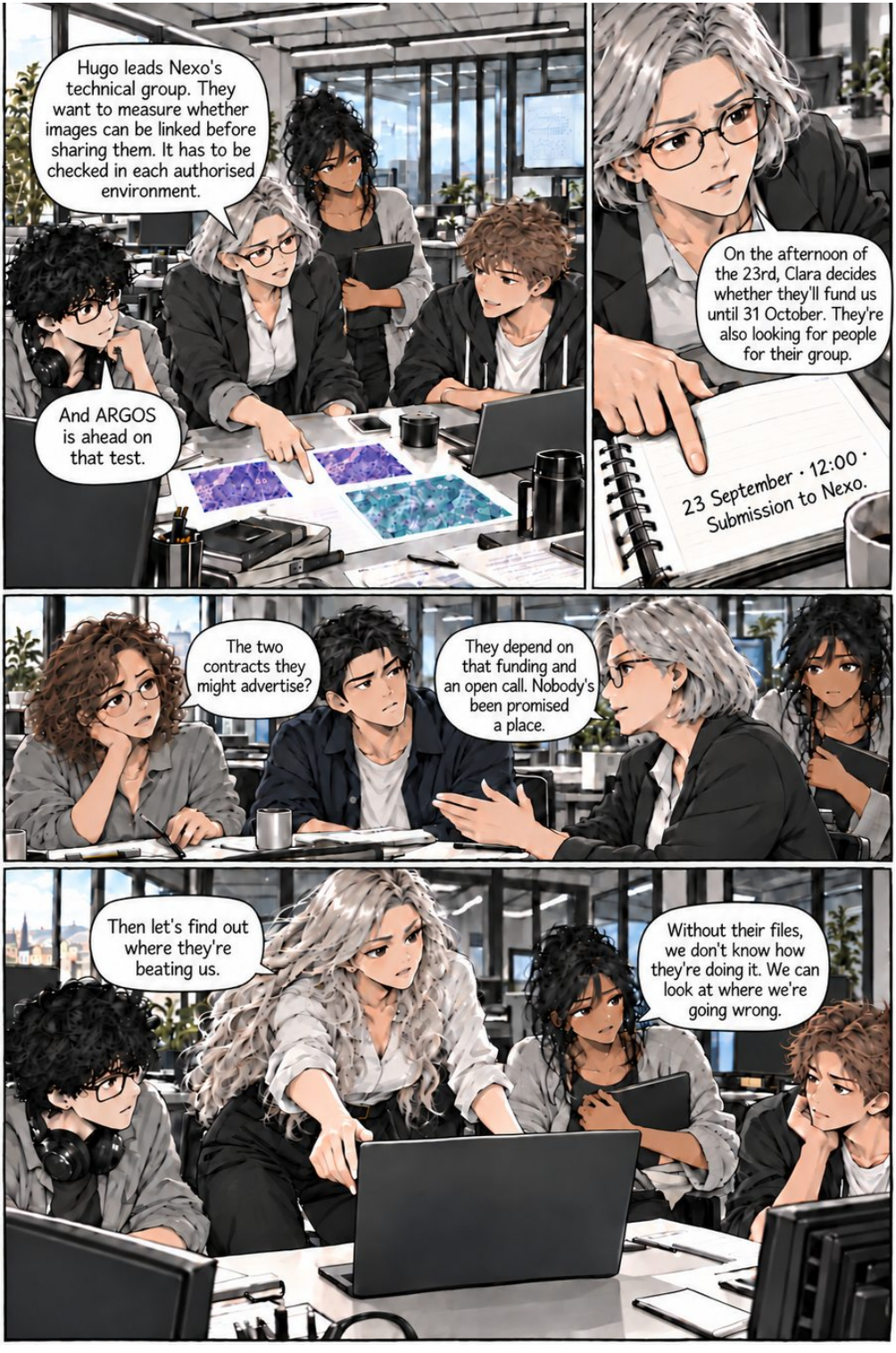
The figure rises again

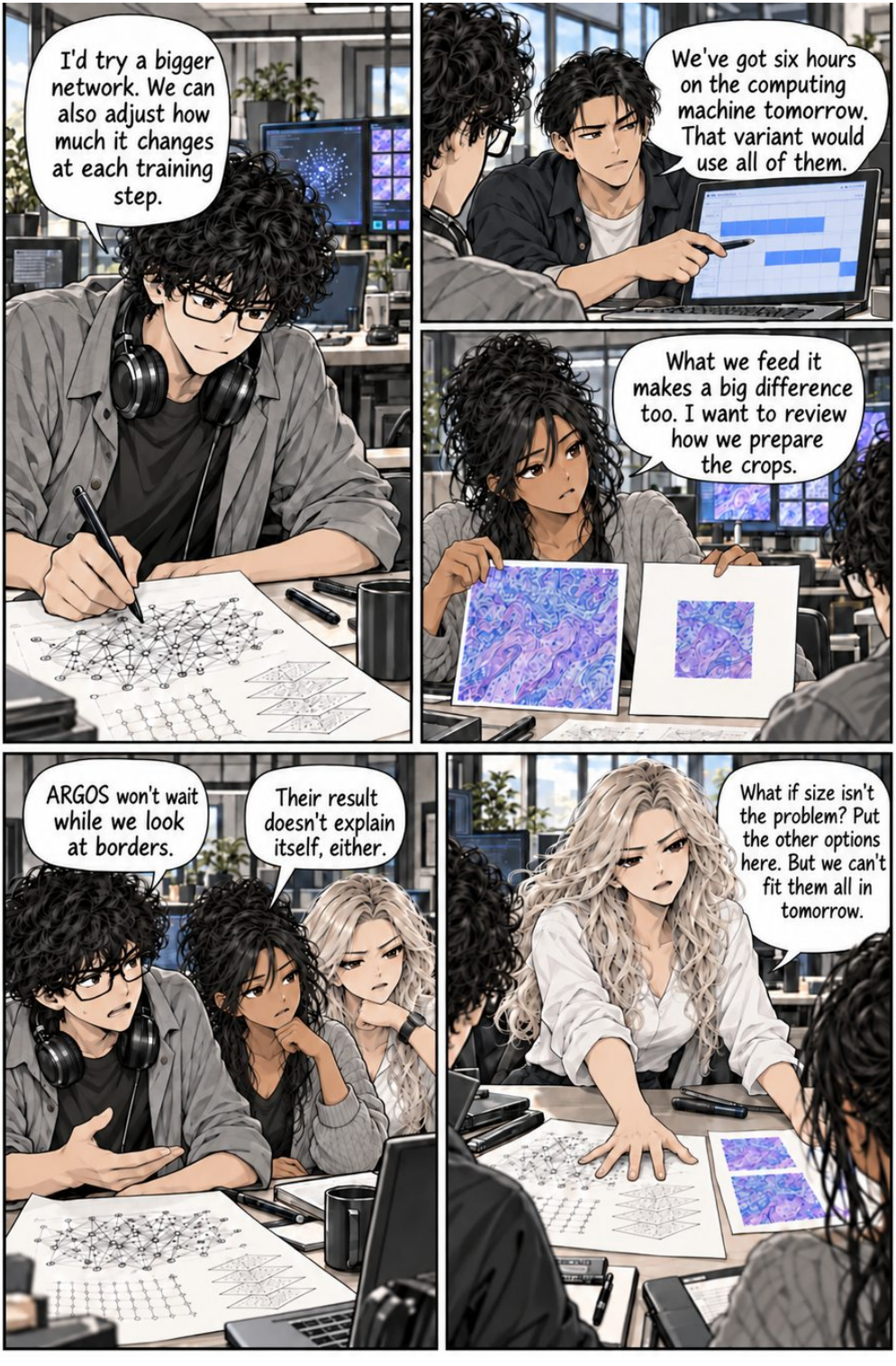
When an improvement appears, Leo seeks approval again before understanding it. Alicia's objection hurts him, especially because it comes from her. He can reason carefully and become defensive in the same conversation. He struggles to separate a failed test from the impression he thinks he is making, and that difficulty returns when the figure seemed to offer him a chance to stand out.

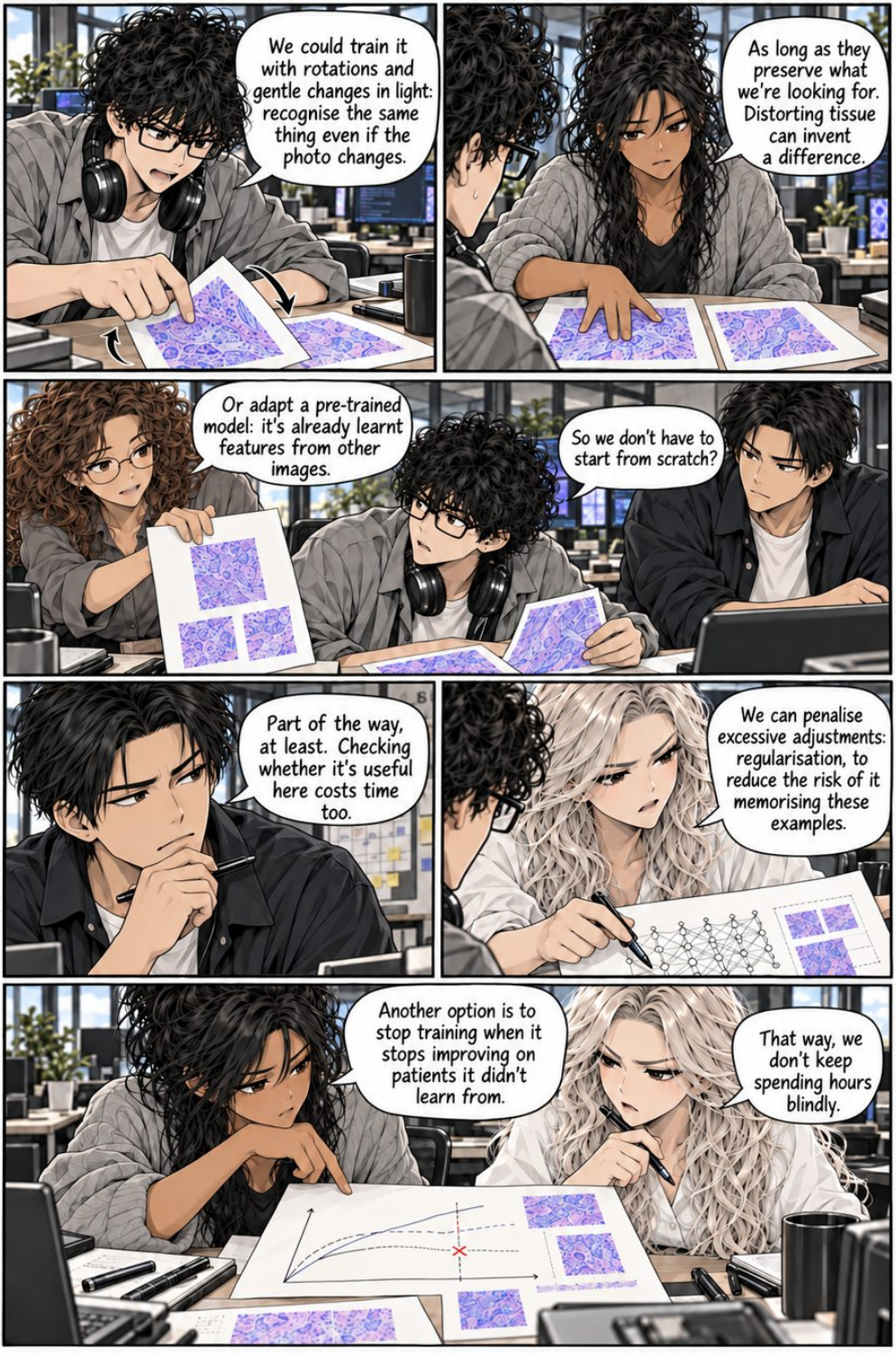
He can correct Alicia too

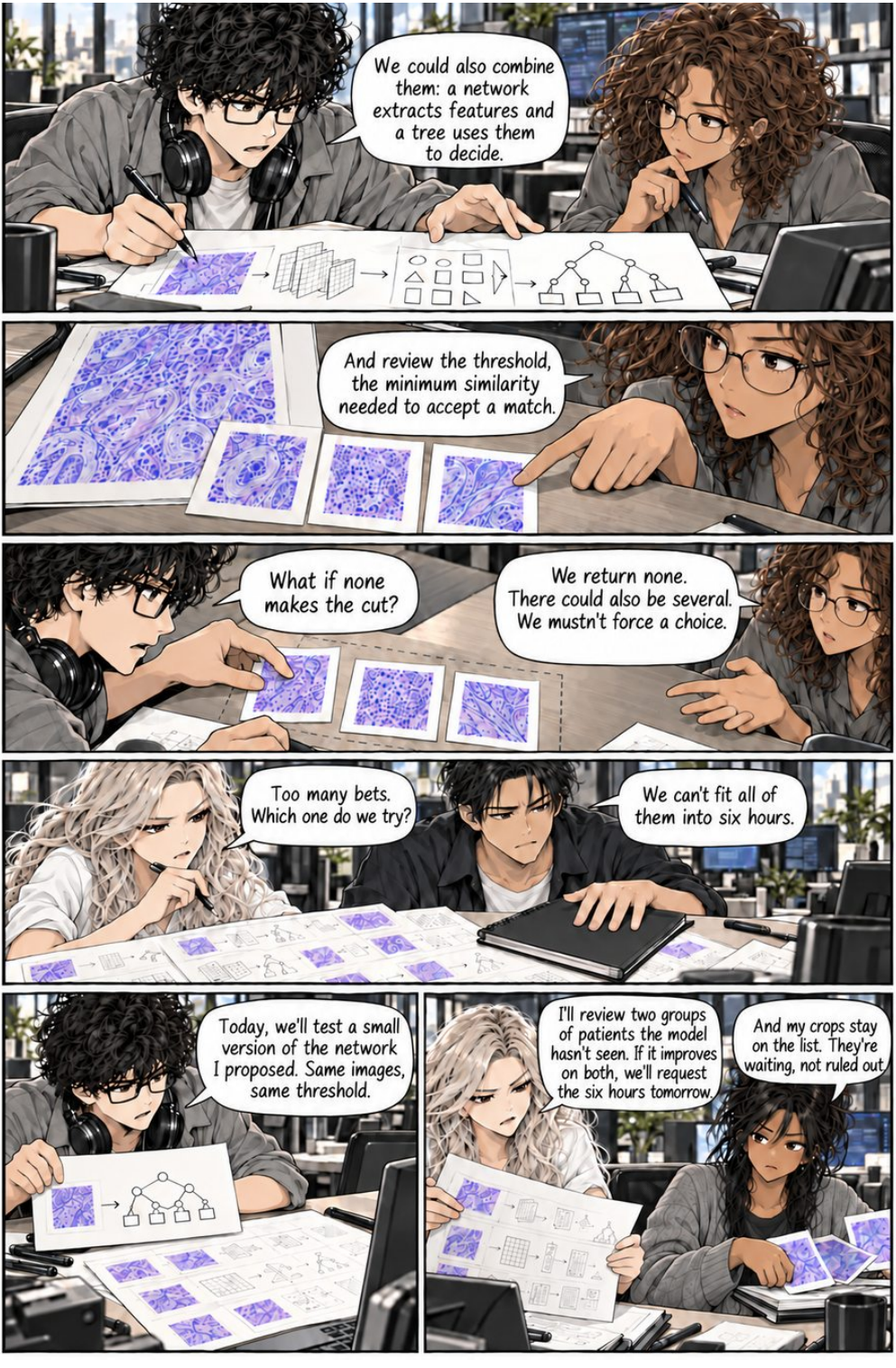
Leo points out that they still do not know the exact cause, and Alicia accepts it. They have managed to disagree and correct each other. He also begins to hear what Mara needs from a contract and see Alicia's effort, having imagined her with fewer problems. He leaves without the extension he hoped for and feels a failure, although Jun can now repeat his test and Mara started on time. He finds it hard to value that while the money is still missing.

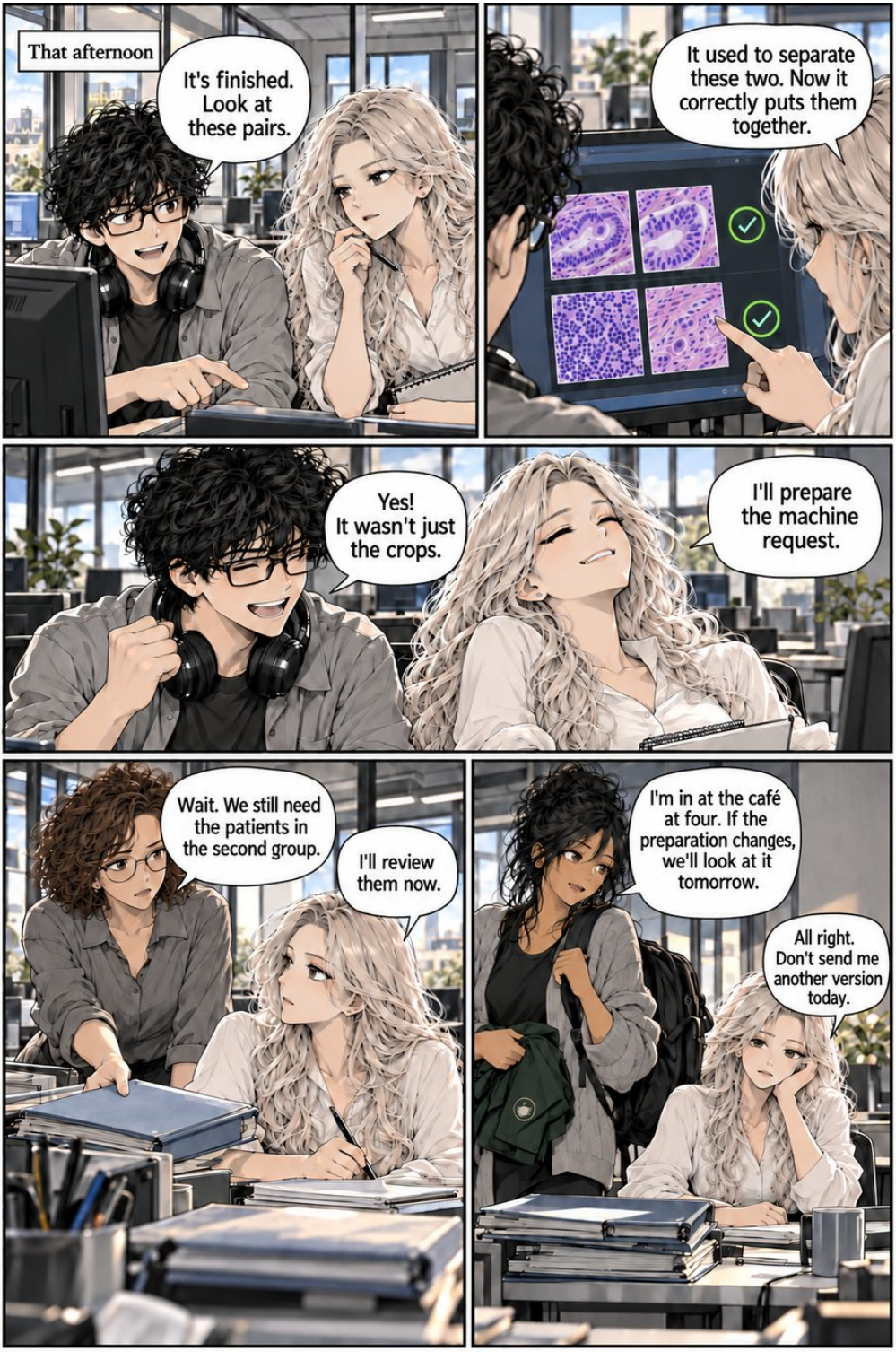

Thursday, 10 September · ATLAS
Nexo · Shared test
ARGOS 0.74
ATLAS 0.40
Almost twice our score.
With the same images?
The same task, the same withheld examples and the same scoring method. Nexo measured it.
What about yesterday's 0.41?
That was our internal test. We scored 0.40 here.
Yesterday
0.41
Today
ATLAS
0.40
This comparison does hurt. Come here.

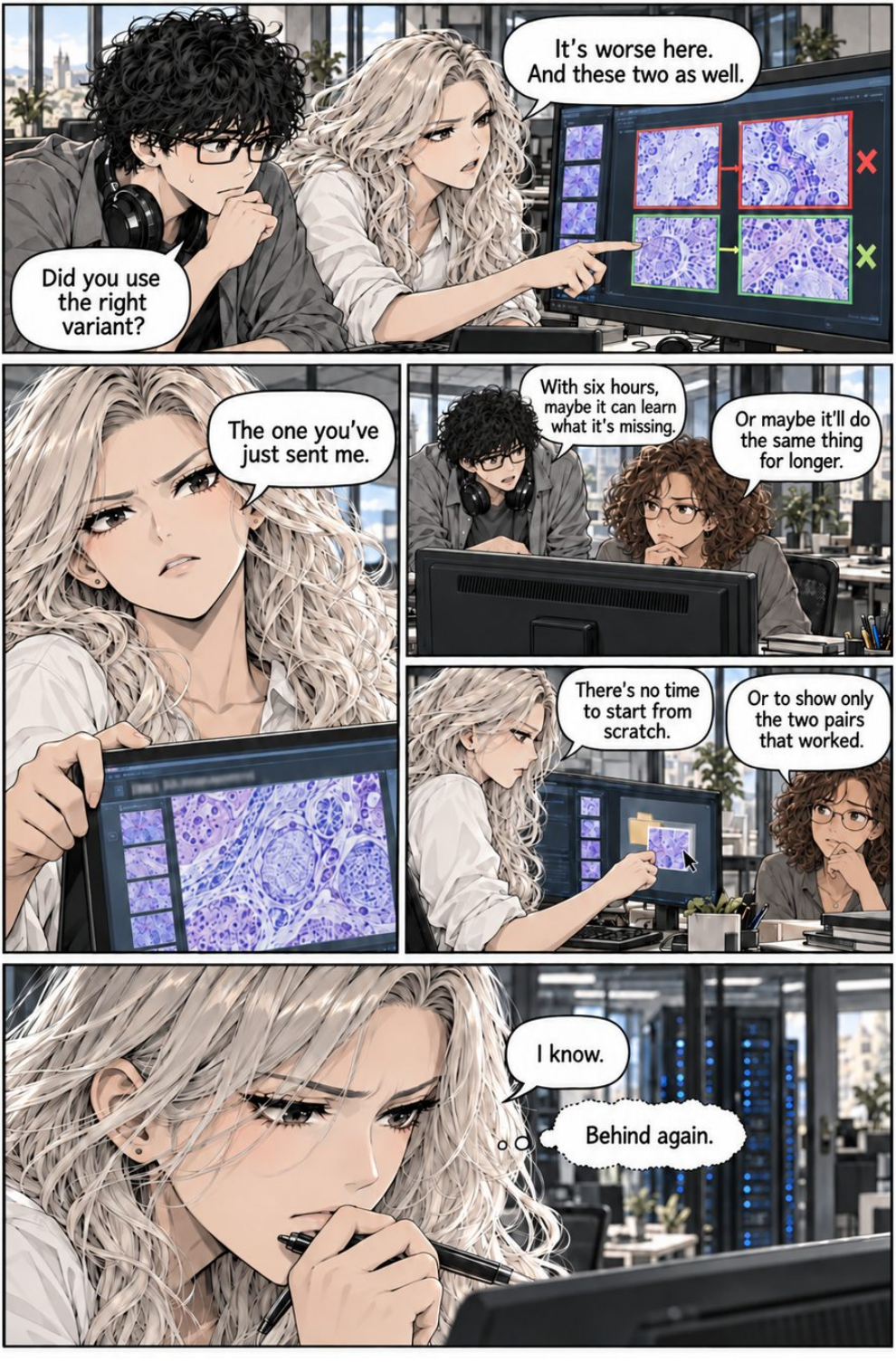

Hugo leads Nexo's technical group. They want to measure whether images can be linked before sharing them. It has to be checked in each authorised environment.
And ARGOS is ahead on that test.
On the afternoon of the 23rd, Clara decides whether they'll fund us until 31 October. They're also looking for people for their group.
23 September · 12:00 · Submission to Nexo.
The two contracts they might advertise?
They depend on that funding and an open call. Nobody's been promised a place.
Then let's find out where they're beating us.
Without their files, we don't know how they're doing it. We can look at where we're going wrong.

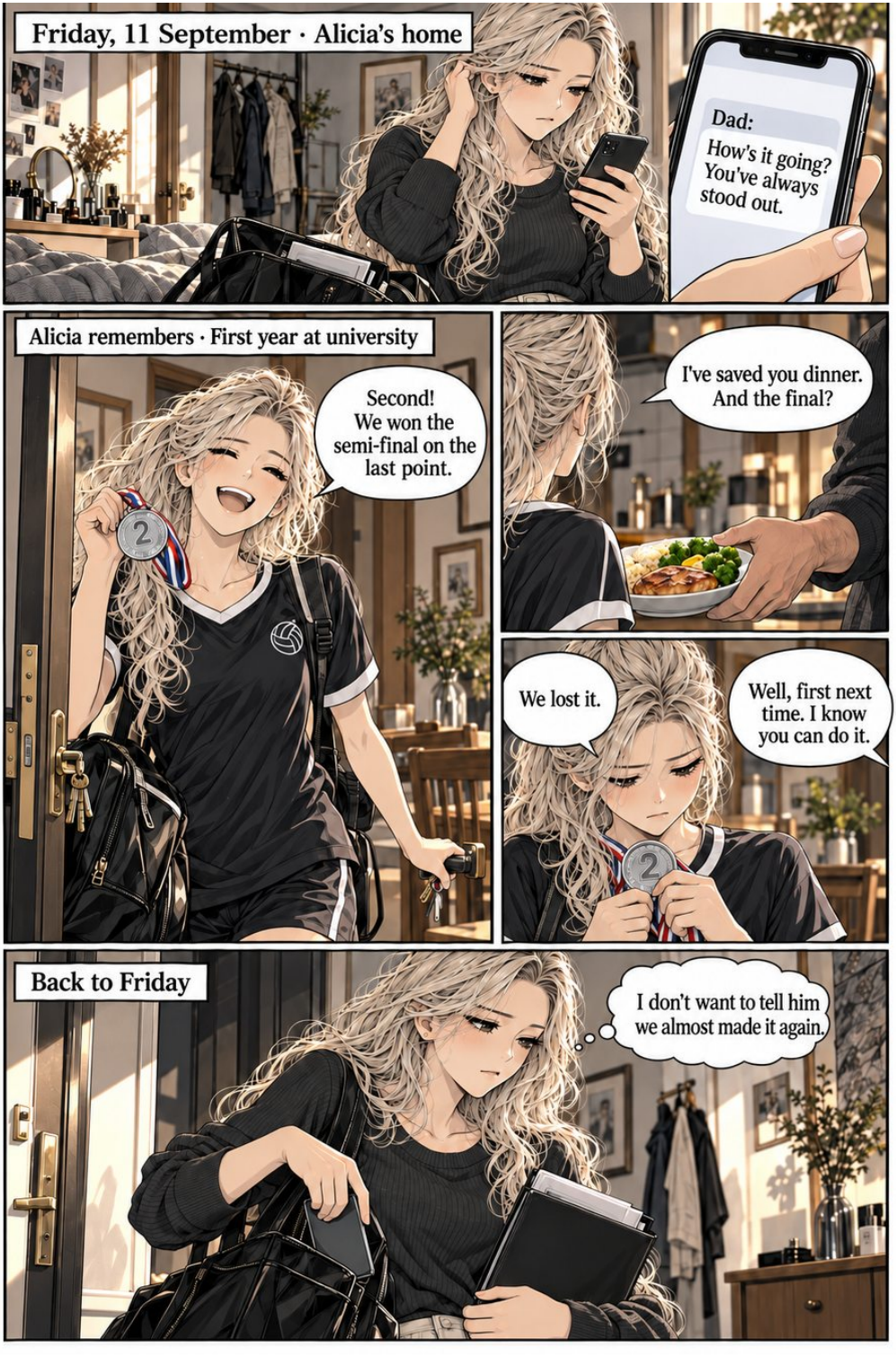

I'd try a bigger network. We can also adjust how much it changes at each training step.
We've got six hours on the computing machine tomorrow. That variant would use all of them.
What we feed it makes a big difference too. I want to review how we prepare the crops.
ARGOS won't wait while we look at borders.
Their result doesn't explain itself, either.
What if size isn't the problem? Put the other options here. But we can't fit them all in tomorrow.

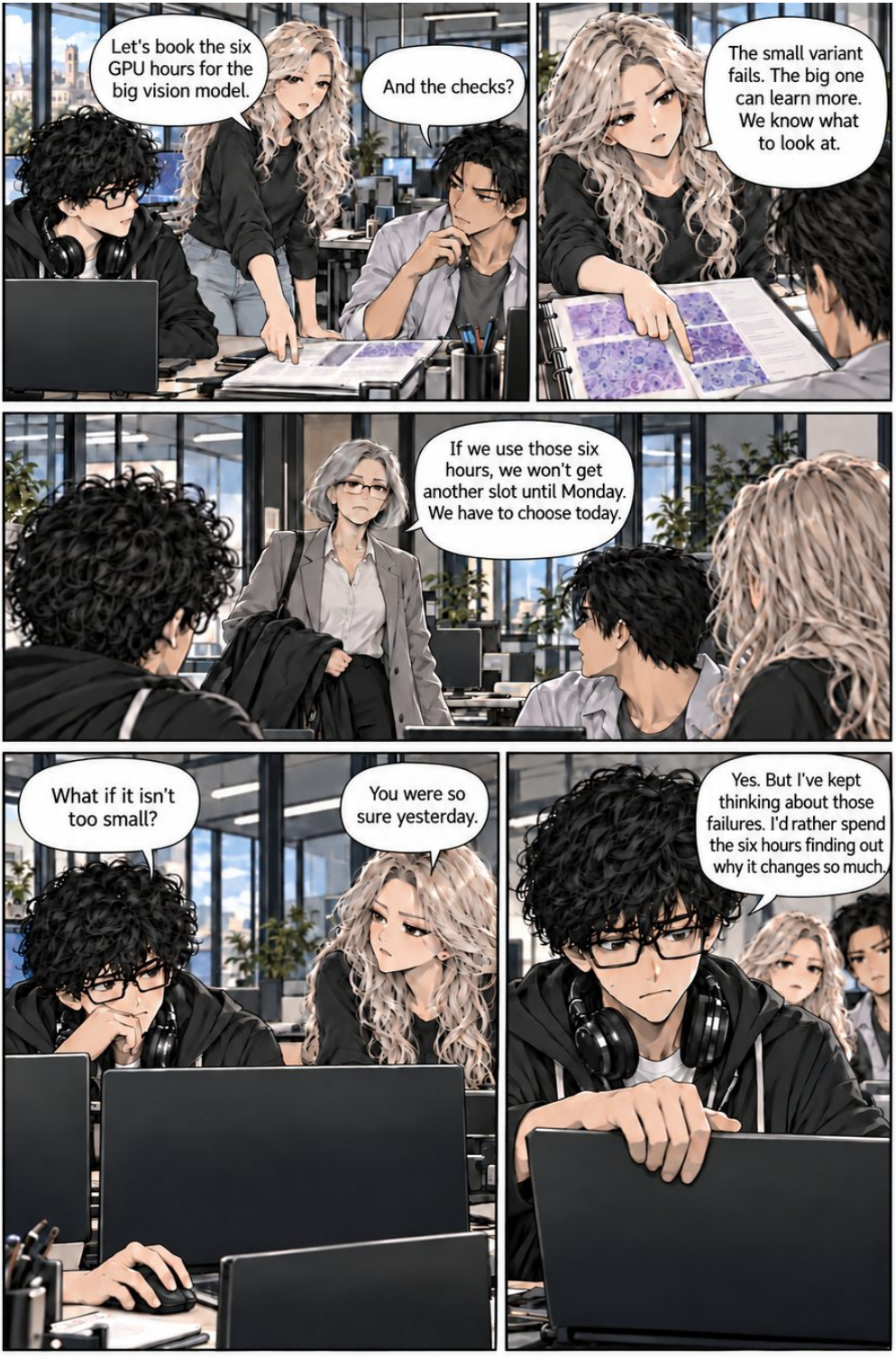

We could train it with rotations and gentle changes in light: recognise the same thing even if the photo changes.
As long as they preserve what we're looking for. Distorting tissue can invent a difference.
Or adapt a pre-trained model: it's already learnt features from other images.
So we don't have to start from scratch?
Part of the way, at least. Checking whether it's useful here costs time too.
We can penalise excessive adjustments: regularisation, to reduce the risk of it memorising these examples.
Another option is to stop training when it stops improving on patients it didn't learn from.
That way, we don't keep spending hours blindly.

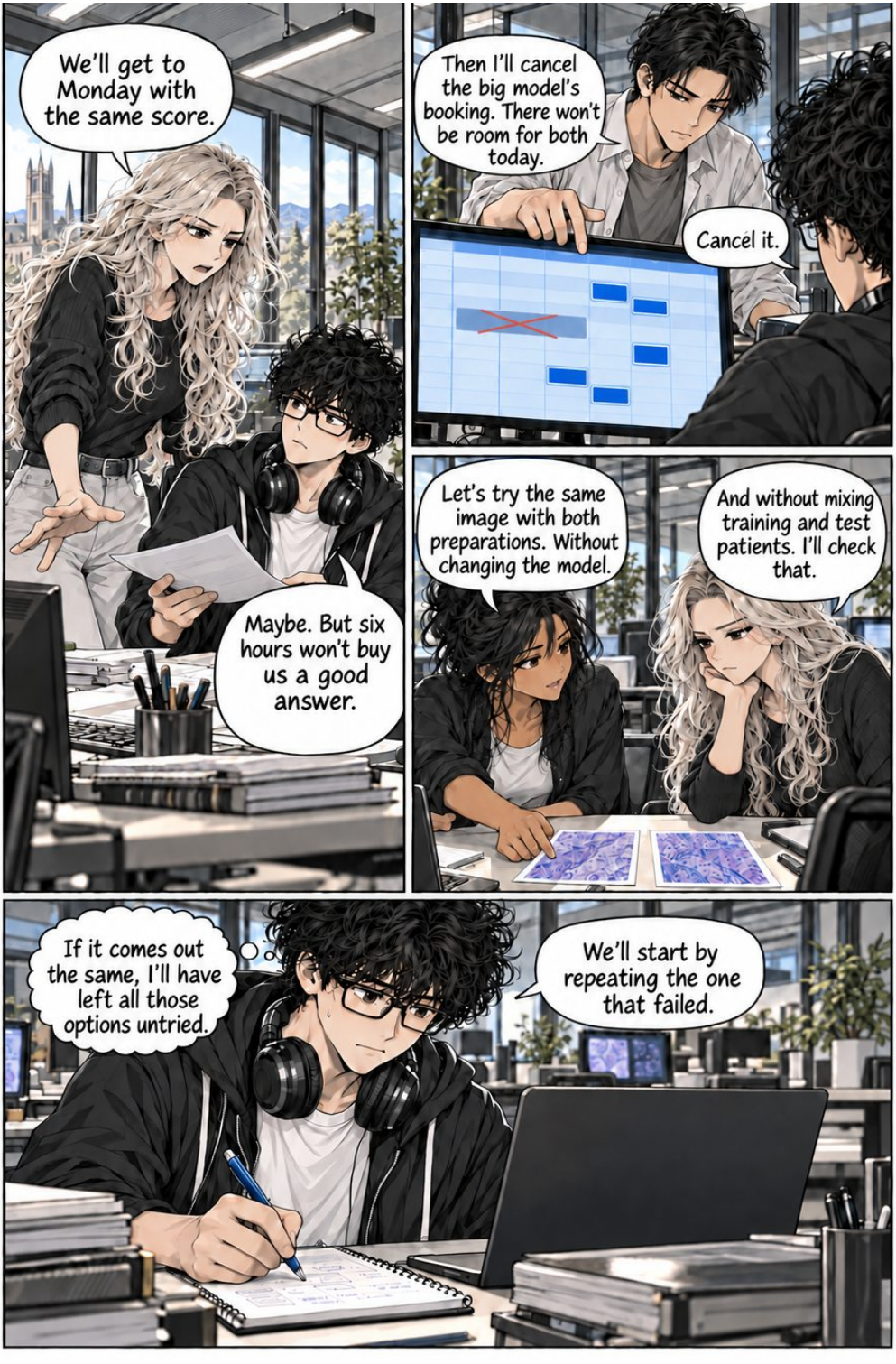

We could also combine them: a network extracts features and a tree uses them to decide.
And review the threshold, the minimum similarity needed to accept a match.
What if none makes the cut?
We return none. There could also be several. We mustn't force a choice.
Too many bets. Which one do we try?
We can't fit all of them into six hours.
Today, we'll test a small version of the network I proposed. Same images, same threshold.
I'll review two groups of patients the model hasn't seen. If it improves on both, we'll request the six hours tomorrow.
And my crops stay on the list. They're waiting, not ruled out.

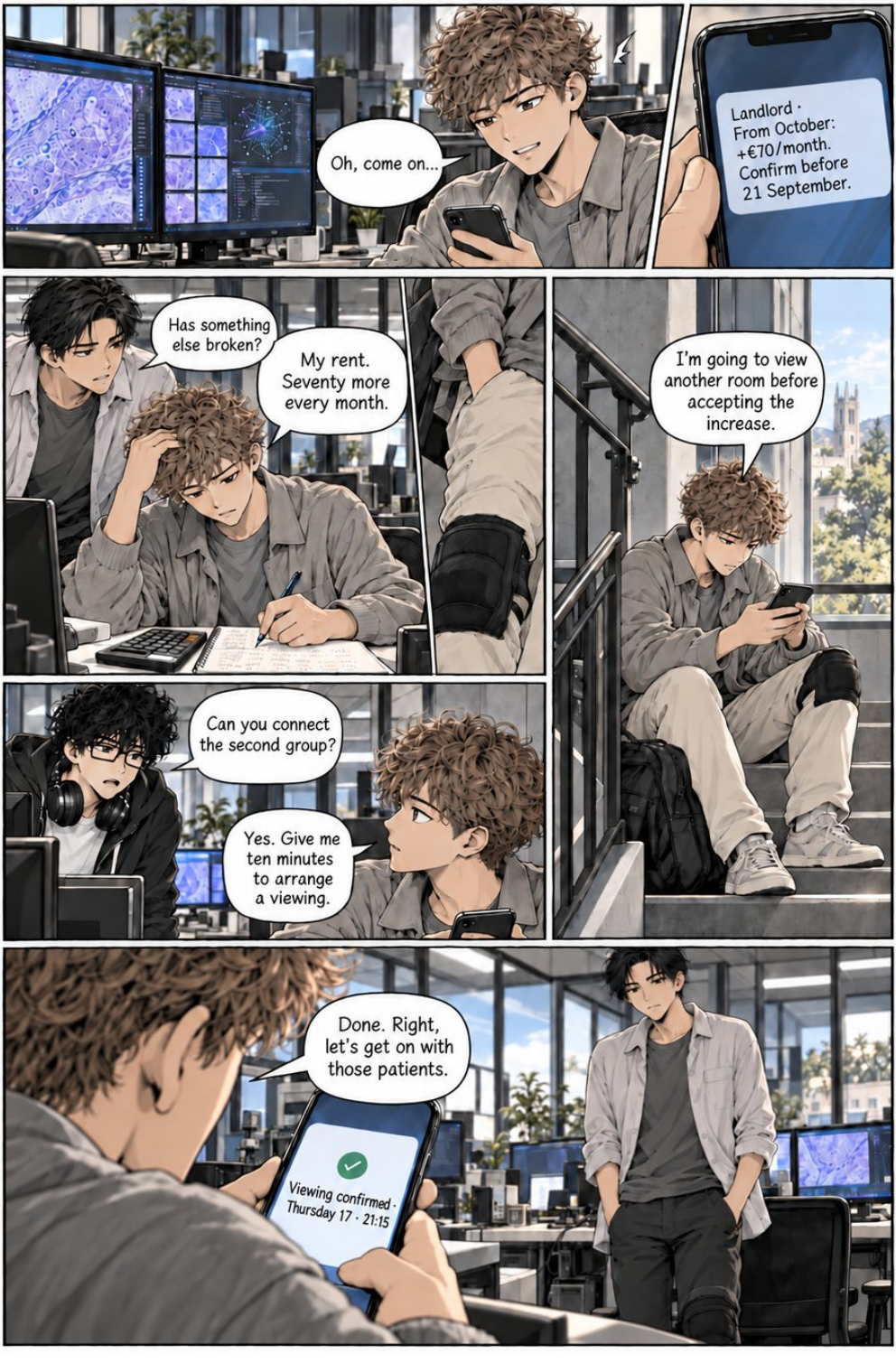

That afternoon
It's finished. Look at these pairs.
It used to separate these two. Now it correctly puts them together.
Yes! It wasn't just the crops.
I'll prepare the machine request.
Wait. We still need the patients in the second group.
I'll review them now.
I'm in at the café at four. If the preparation changes, we'll look at it tomorrow.
All right. Don't send me another version today.

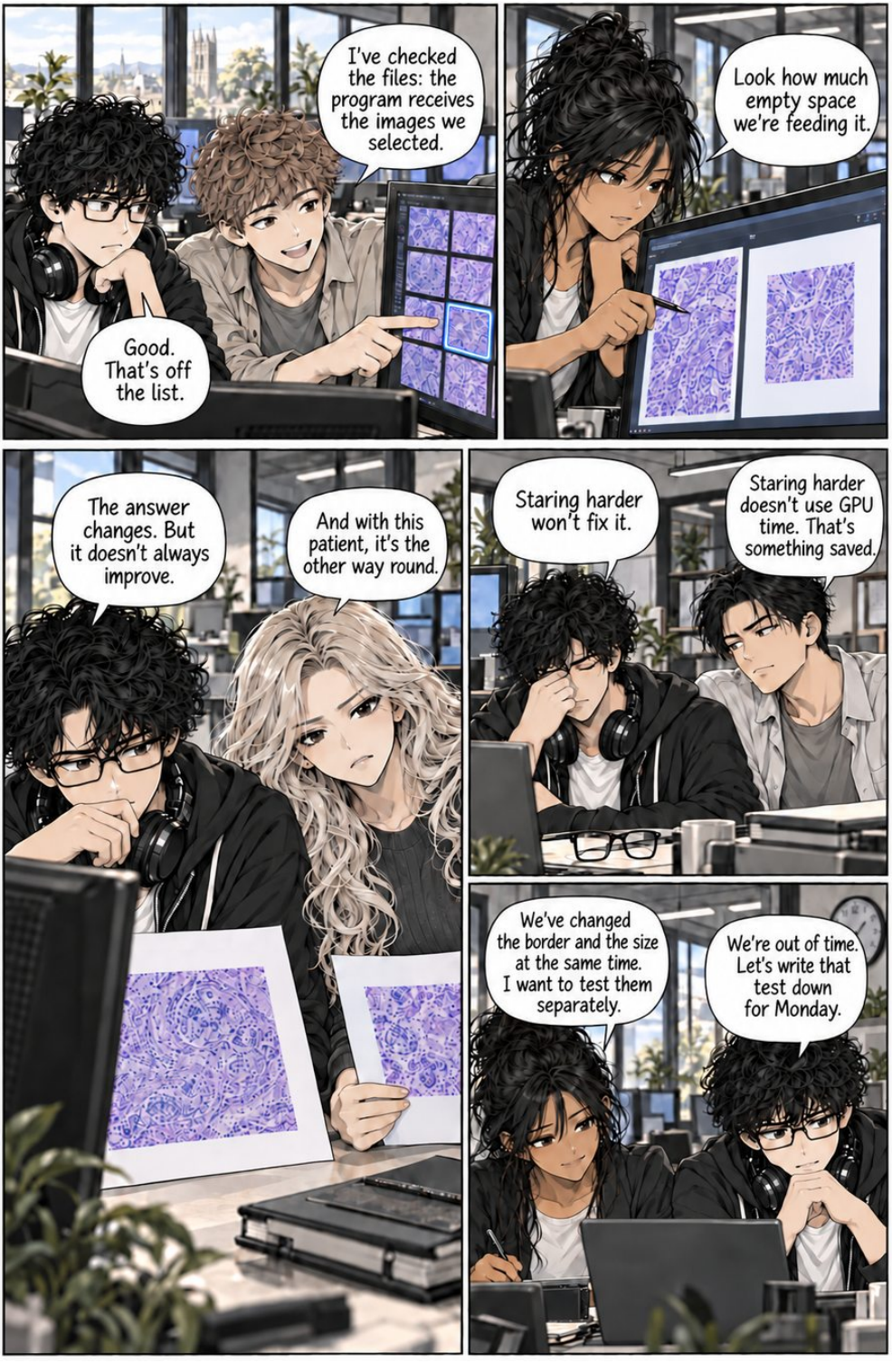

It's worse here.
And these two as well.
Did you use the right variant?
The one you've just sent me.
With six hours, maybe it can learn what it's missing.
Or maybe it'll do the same thing for longer.
There's no time to start from scratch.
Or to show only the two pairs that worked.
I know.
Behind again.

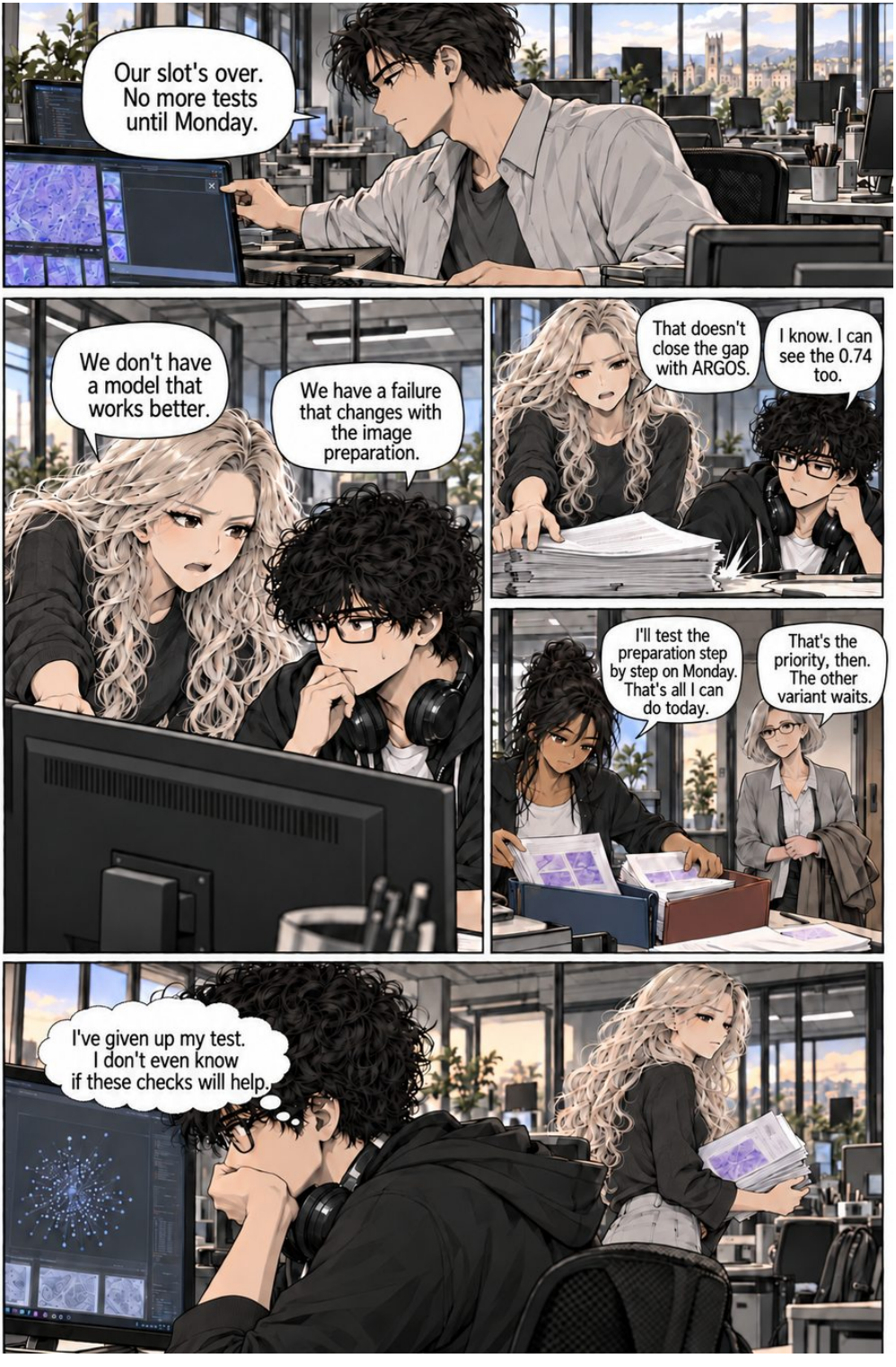

Friday, 11 September · Alicia's home
Dad:
How's it going? You've always stood out.
Alicia remembers · First year at university
Second! We won the semi-final on the last point.
I've saved you dinner. And the final?
We lost it.
Well, first next time. I know you can do it.
Back to Friday
I don't want to tell him we almost made it again.

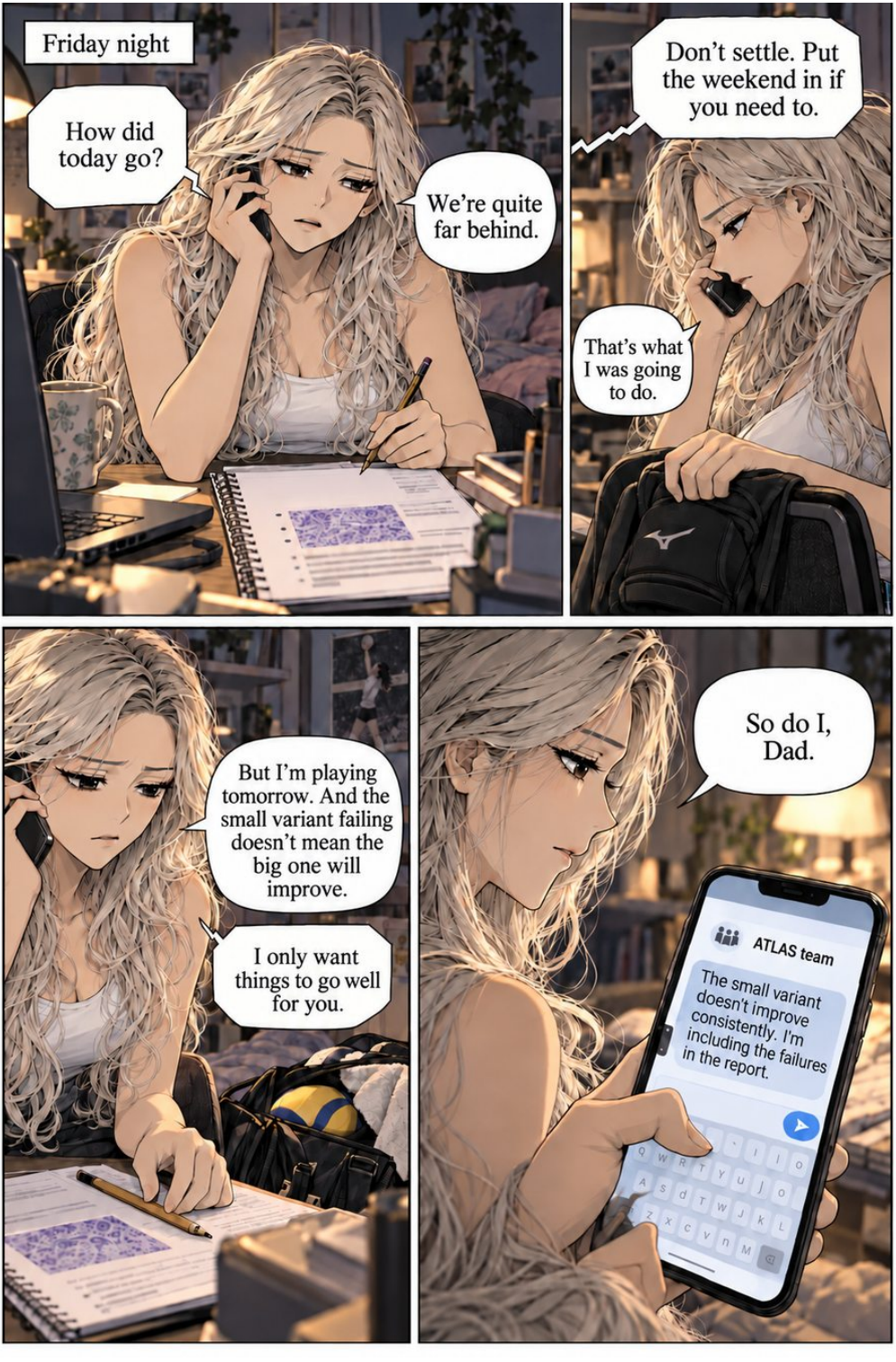

Let's book the six GPU hours for the big vision model.
And the checks?
The small variant fails. The big one can learn more. We know what to look at.
If we use those six hours, we won't get another slot until Monday. We have to choose today.
What if it isn't too small?
You were so sure yesterday.
Yes. But I've kept thinking about those failures. I'd rather spend the six hours finding out why it changes so much.

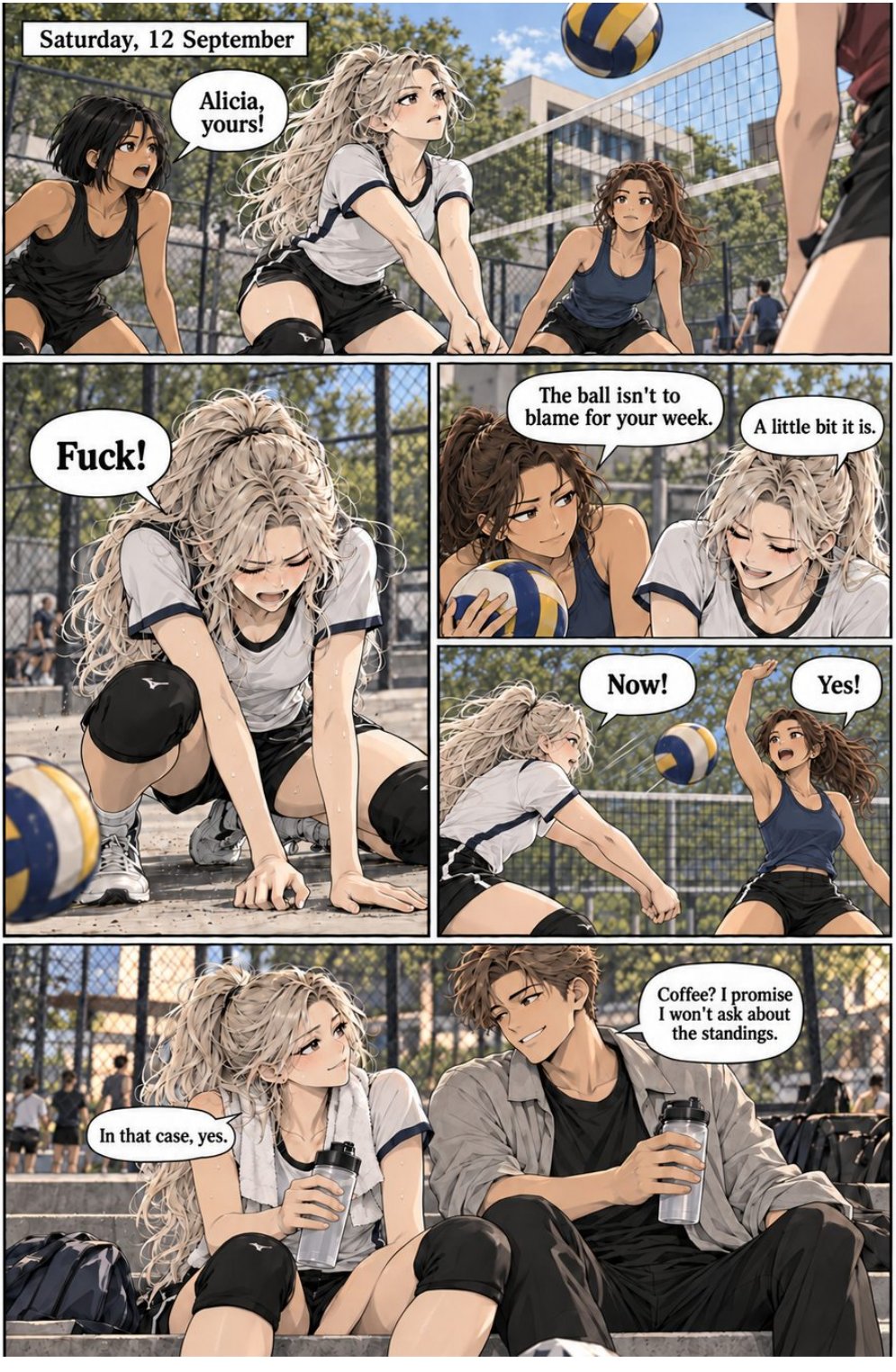

We'll get to Monday with the same score.
Maybe. But six hours won't buy us a good answer.
Then I'll cancel the big model's booking. There won't be room for both today.
Cancel it.
Let's try the same image with both preparations. Without changing the model.
And without mixing training and test patients. I'll check that.
If it comes out the same, I'll have left all those options untried.
We'll start by repeating the one that failed.

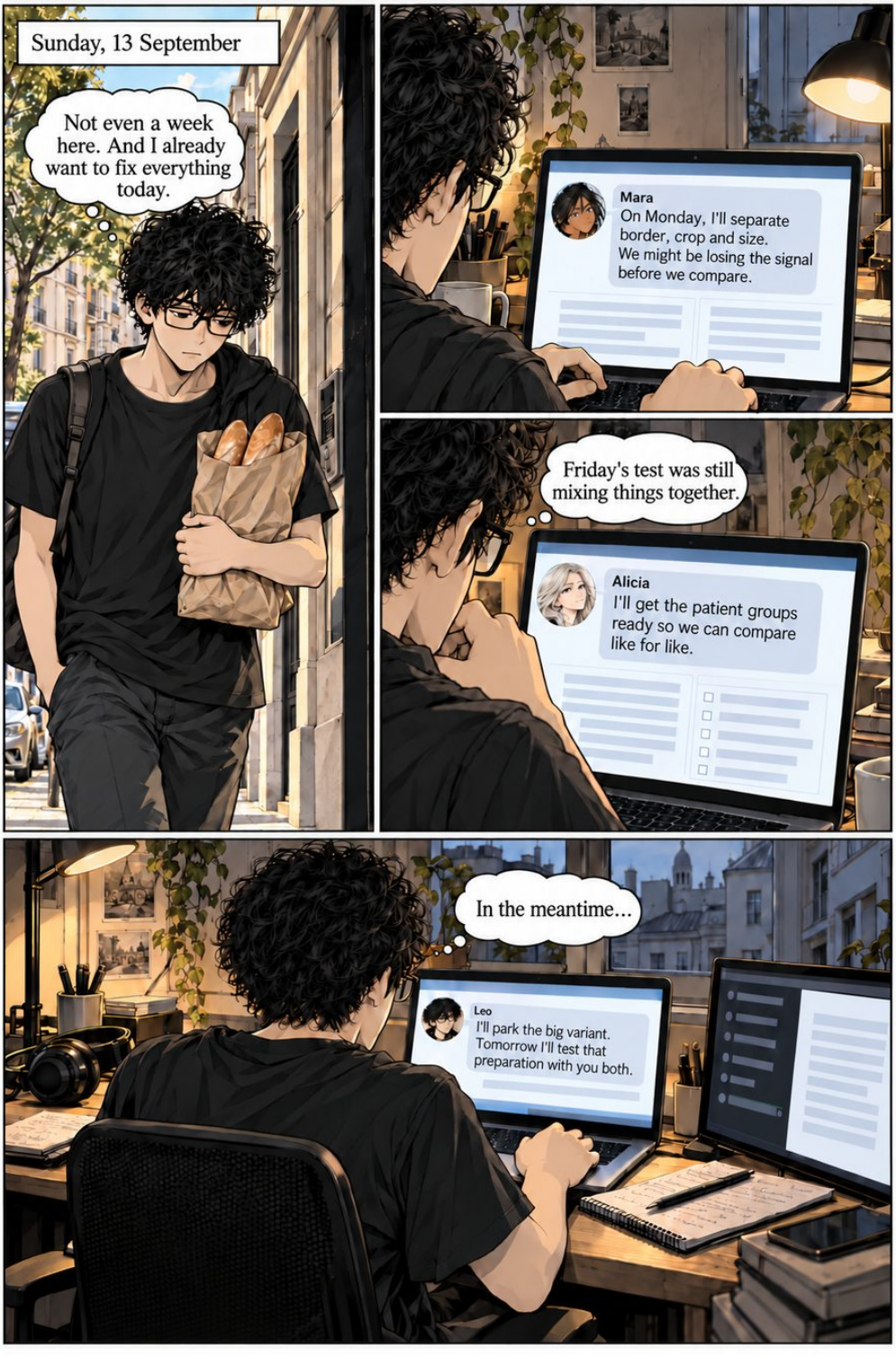

Oh, come on...
Landlord · From October: +€70/month. Confirm before 21 September.
Has something else broken?
My rent. Seventy more every month.
I'm going to view another room before accepting the increase.
Can you connect the second group?
Yes. Give me ten minutes to arrange a viewing.
Done. Right, let's get on with those patients.
Viewing confirmed · Thursday 17 · 21:15

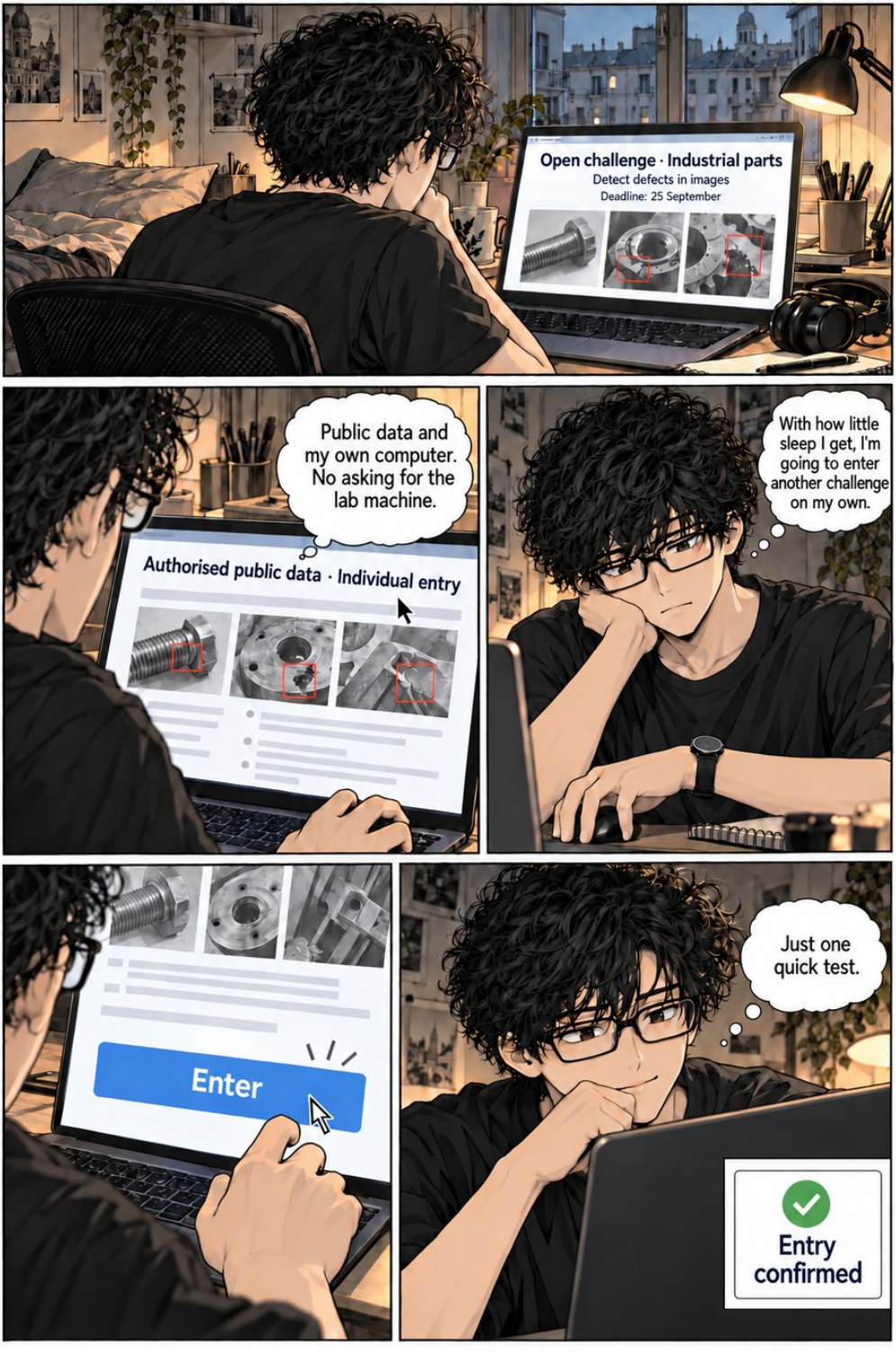

I've checked the files: the program receives the images we selected.
Good. That's off the list.
Look how much empty space we're feeding it.
The answer changes. But it doesn't always improve.
And with this patient, it's the other way round.
Staring harder won't fix it.
Staring harder doesn't use GPU time. That's something saved.
We've changed the border and the size at the same time. I want to test them separately.
We're out of time. Let's write that test down for Monday.


Our slot's over. No more tests until Monday.
We don't have a model that works better.
We have a failure that changes with the image preparation.
That doesn't close the gap with ARGOS.
I know. I can see the 0.74 too.
I'll test the preparation step by step on Monday. That's all I can do today.
That's the priority, then. The other variant waits.
I've given up my test. I don't even know if these checks will help.


Friday night
How did today go?
We're quite far behind.
Don't settle. Put the weekend in if you need to.
That's what I was going to do.
But I'm playing tomorrow. And the small variant failing doesn't mean the big one will improve.
I only want things to go well for you.
So do I, Dad.
ATLAS team
The small variant doesn't improve consistently. I'm including the failures in the report.


Saturday, 12 September
Alicia, yours!
Fuck!
The ball isn't to blame for your week.
A little bit it is.
Now!
Yes!
Coffee? I promise I won't ask about the standings.
In that case, yes.


Sunday, 13 September
Not even a week here. And I already want to fix everything today.
Mara
On Monday, I'll separate border, crop and size. We might be losing the signal before we compare.
Friday's test was still mixing things together.
Alicia
I'll get the patient groups ready so we can compare like for like.
In the meantime...
Leo
I'll park the big variant. Tomorrow I'll test that preparation with you both.


Open challenge · Industrial parts
Detect defects in images
Deadline: 25 September
Public data and my own computer. No asking for the lab machine.
Authorised public data · Individual entry
With how little sleep I get, I'm going to enter another challenge on my own.
Enter
Just one quick test.
Entry confirmed


That night
Ah. You're not going to make it easy for me either.
We spent the whole slot reviewing the failures. I still don't know if I chose well.
On Monday, we'll test Mara's preparation. I'd like her to be right.
And I'd find it a bit annoying.
I need to be awake tomorrow.
After this one.

How We Spend the Six Hours



GPU, requests to a service and people's time.

Now they are comparing the same test: ATLAS has 0.40 and ARGOS 0.74. I understand Leo wanting to do something as soon as possible. But that gap still does not tell us what to change, and they have only six hours of machine time to choose a route.

What each proposal would change

A larger network might represent more complex relationships and would also cost more to run. Other proposals act before the data reaches it: preparing crops, transforming examples while preserving their signal, or using a pretrained model. They can also review the minimum similarity they accept. If they try everything together, a rise will leave them unsure what helped. Training for longer, on its own, does not explain which failure they are trying to correct either.

We need to look at the second group

The small variant fixes some pairs and gets worse on another group of patients. Showing the improved ones is tempting: they are easy to understand and support our view. Nadia insists on keeping the other half of the comparison. As the team uses both groups to decide what to keep testing, they are part of development validation. They have already influenced the solution; the final exam must be kept for later.

What remains untested

Leo cancels the larger model's slot and decides to investigate preparation. He may still be wrong. The first review changes border and size together, so they discover sensitivity without separating its causes. Mara proposes isolating the changes on Monday and ends her shift. I am interested in that ending because it leaves a specific question to pursue and alternatives untried. Cancelling them freed six hours; it did not prove they were worse.

Sources: · Validation · Preparation and leakage · Transfer learning

Cancelling the Six Hours



One task selected among several windows and notifications.

ARGOS's lead now comes from a comparable test. Leo has reasons to worry and several ideas he would like to try. The booking forces him to choose one before knowing whether it will work.

Choosing without knowing whether it will work

Setting the larger model aside gives them time to investigate preparation. Leo still wants to know what would have happened with his proposal. It is familiar enough: we agree on one route and keep thinking about the one we left out. He can support Mara's work and feel frustrated at the same time. At the end of the night, he writes that he would like her to be right and that it would also annoy him a little. At least he has managed to put both things on the same page.

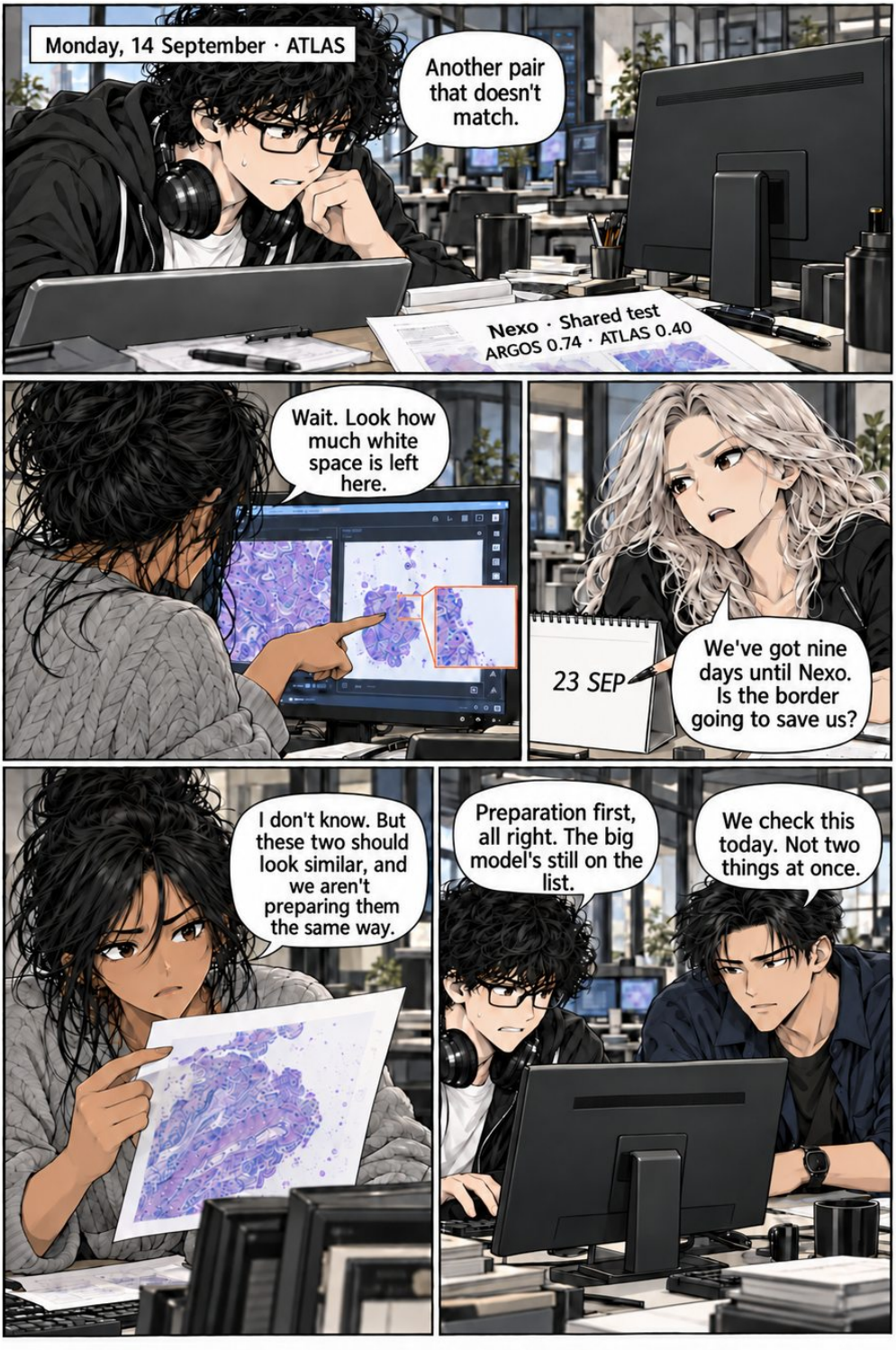
The conversation only the reader sees

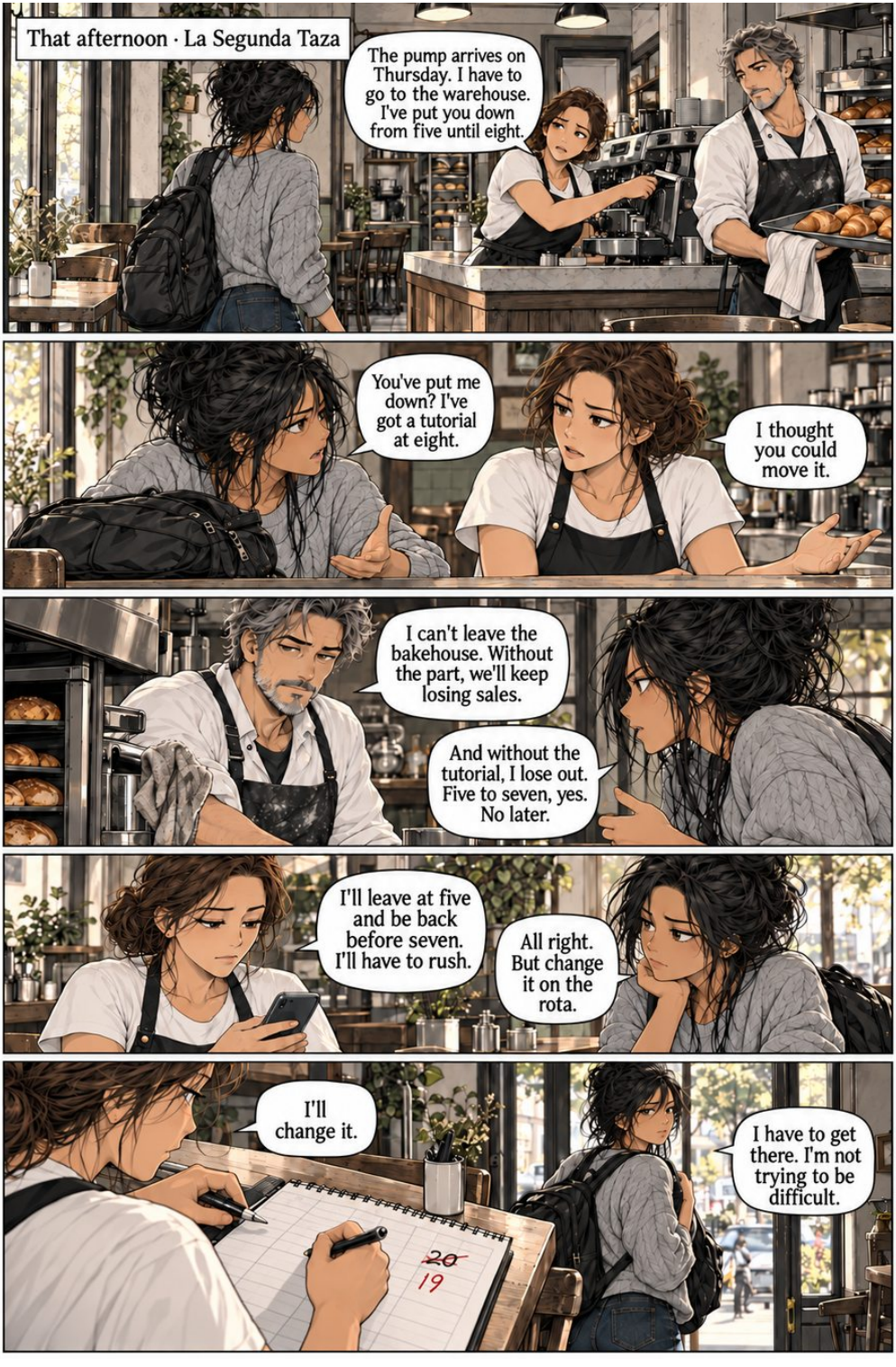
The volleyball memory helps explain Alicia's pressure to come first. Leo has not heard her father. We do see that conversation and how she responds afterwards: she keeps the variant's failures, goes out to play and accepts a coffee with Álvaro. Her father asks for more dedication; she is already working and also wants to play on Saturday.

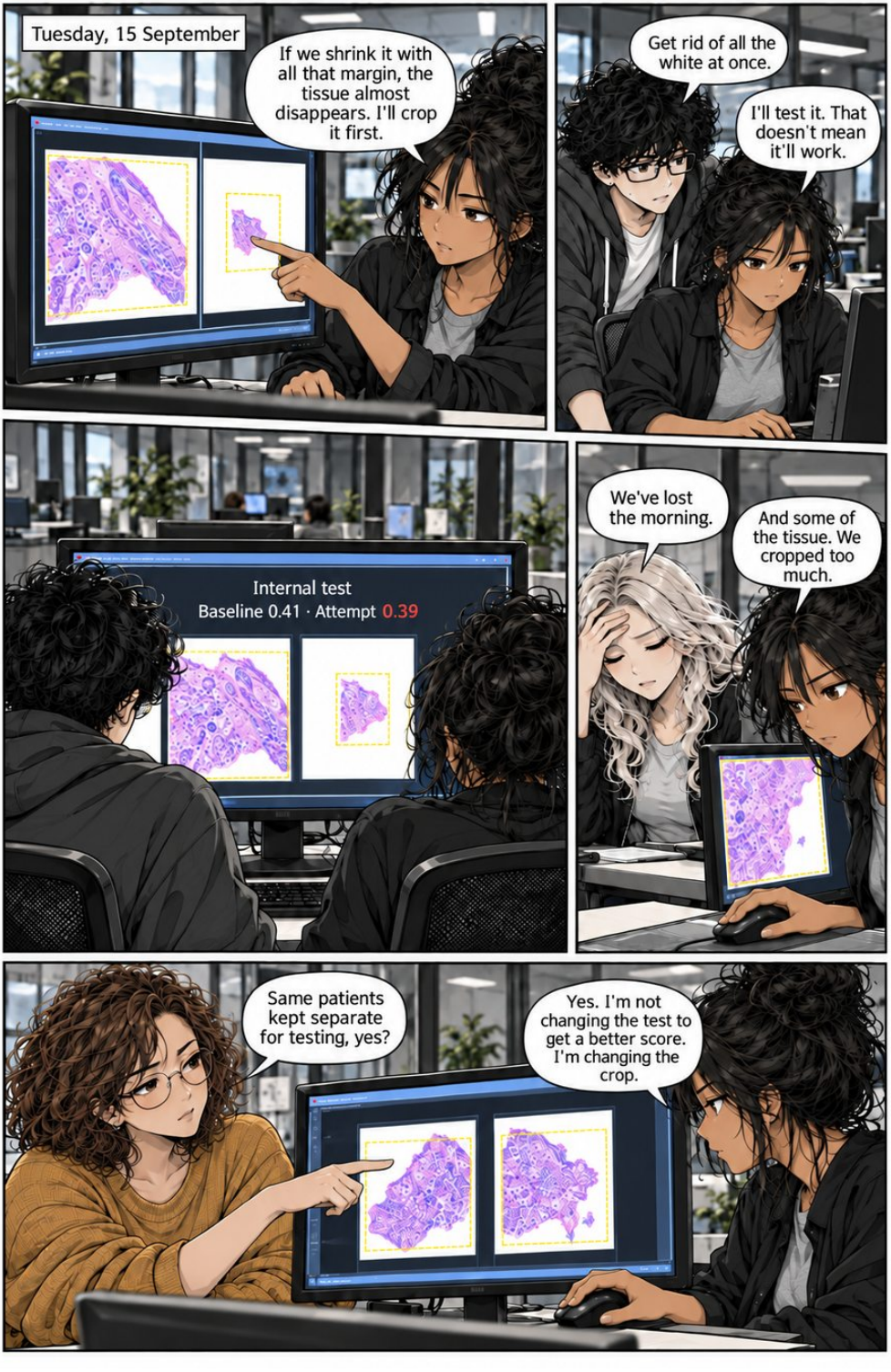
Something he wants to do on his own

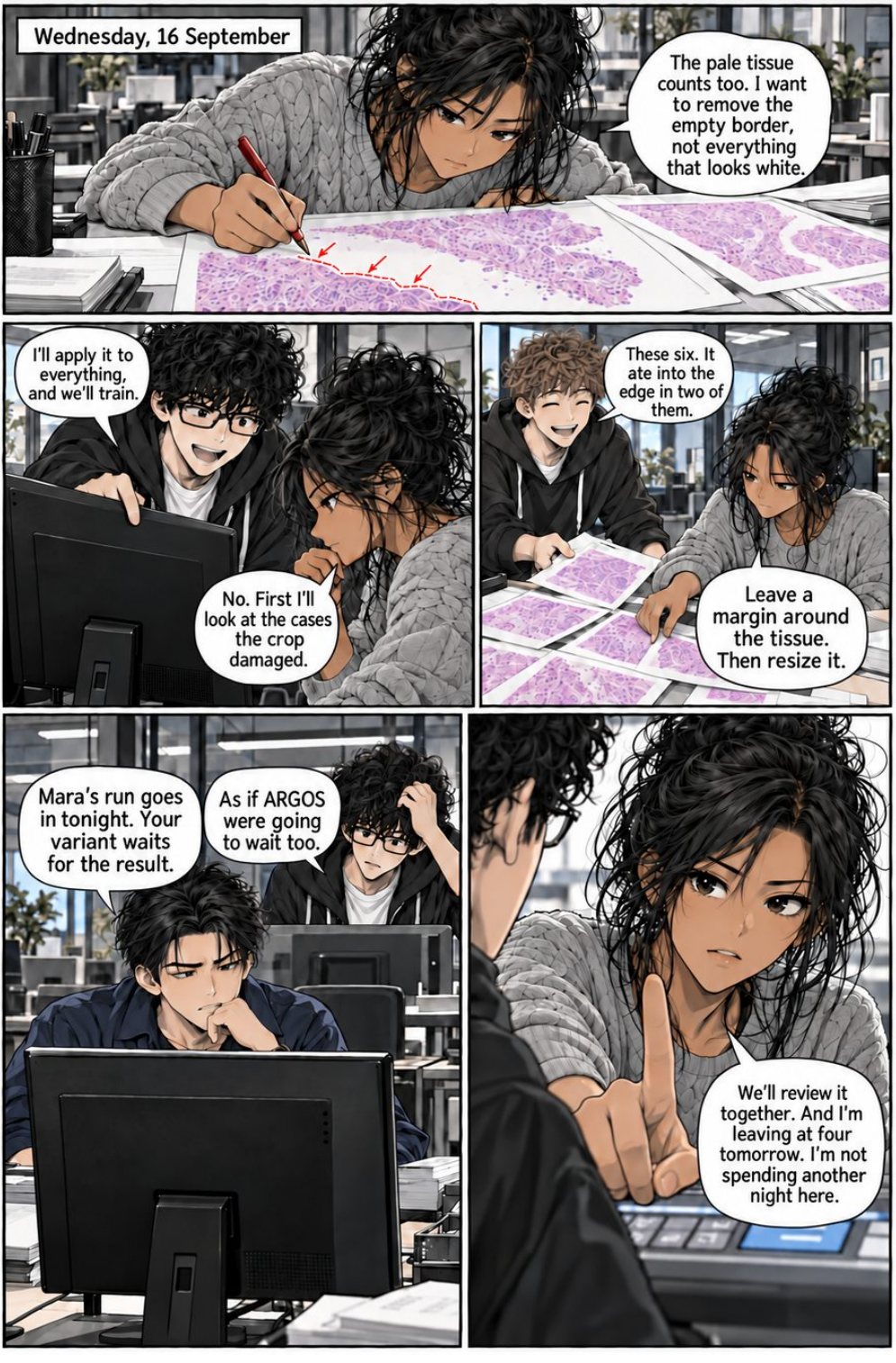
The parts challenge offers him curiosity and autonomy outside ATLAS. It also delays bedtime again. While nobody is asking him to do that test, he stays at the computer and promises himself it will only be a little while. He still needs to decide how much time to spend on it and what he will stop doing if he gets carried away. Signing up opens that possibility; organising his week remains undone.

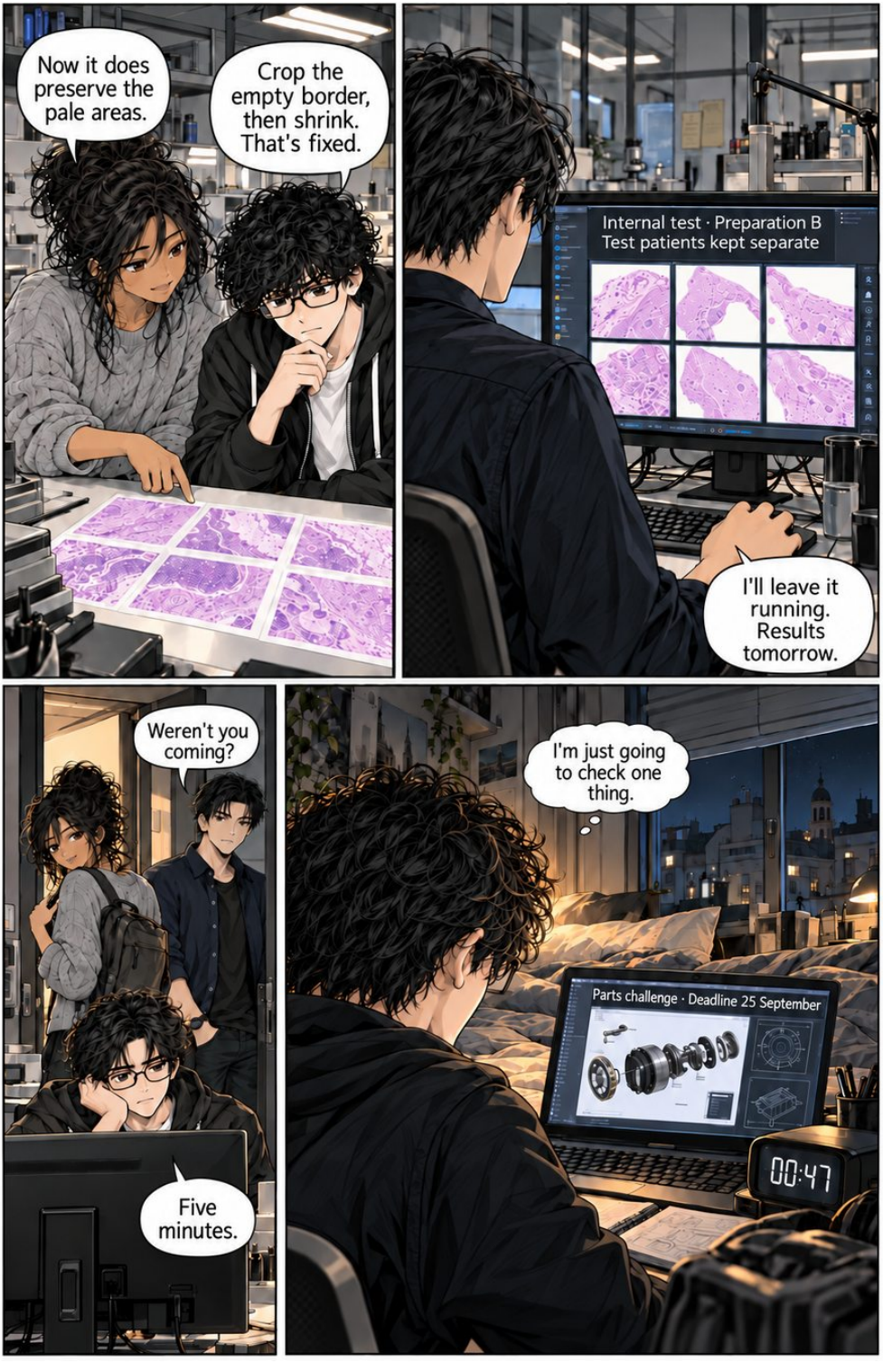
Before opening it, what should he have decided about the following day?

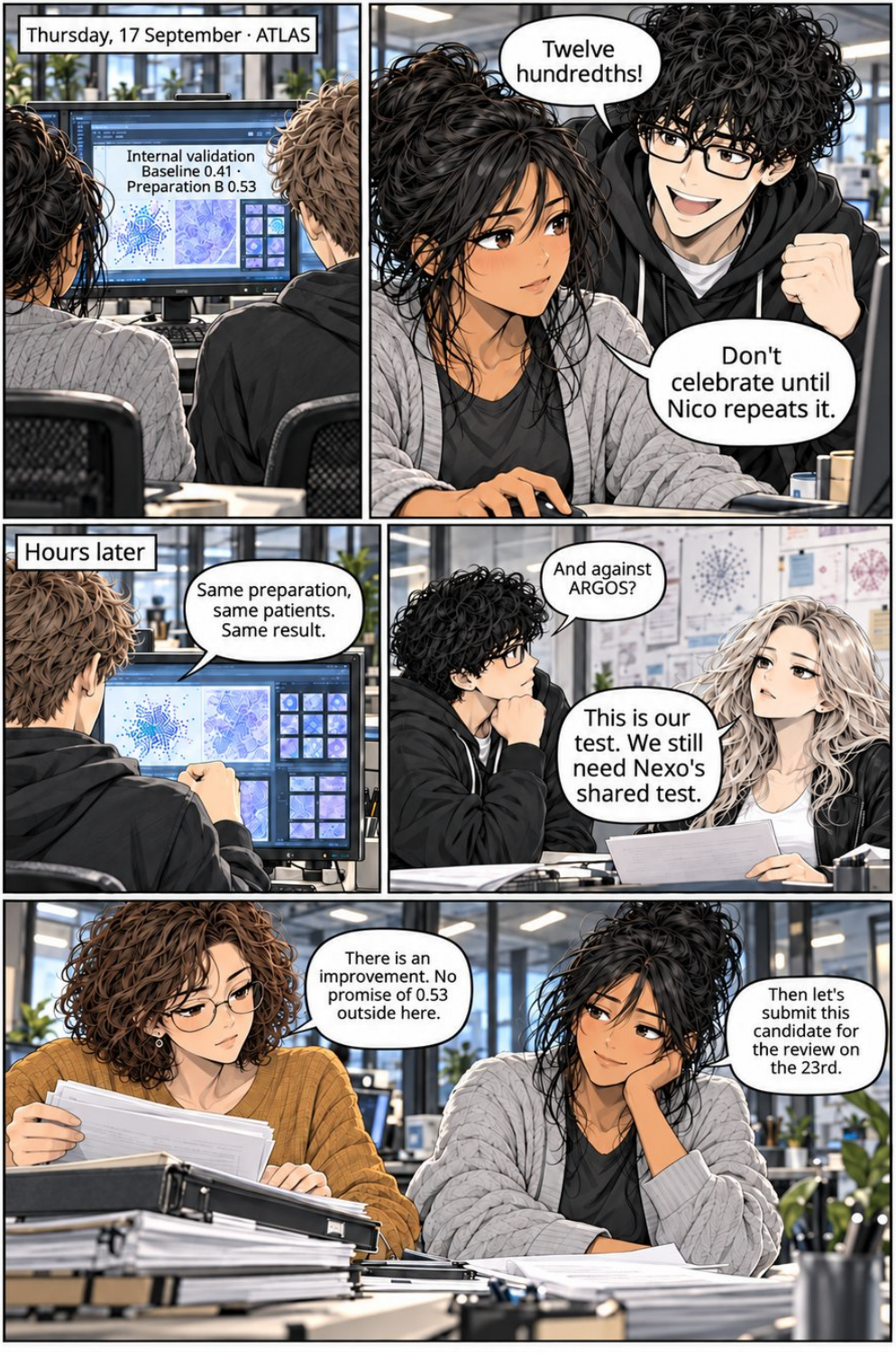

Monday, 14 September · ATLAS
Another pair that doesn't match.
Nexo · Shared test
ARGOS 0.74 · ATLAS 0.40
Wait. Look how much white space is left here.
23 SEP
We've got nine days until Nexo. Is the border going to save us?
I don't know. But these two should look similar, and we aren't preparing them the same way.
Preparation first, all right. The big model's still on the list.
We check this today. Not two things at once.

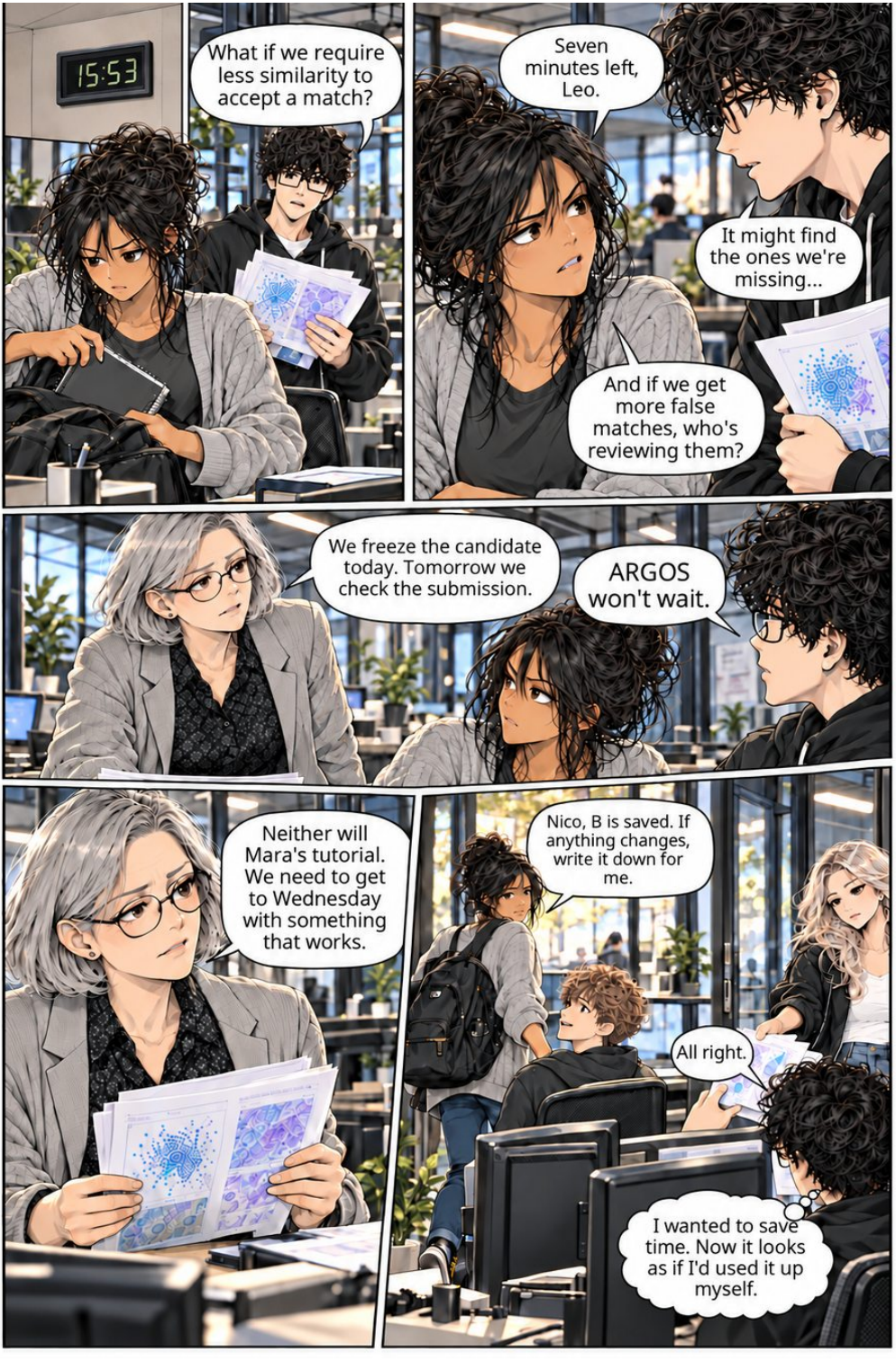

That afternoon · La Segunda Taza
The pump arrives on Thursday. I have to go to the warehouse. I've put you down from five until eight.
You've put me down? I've got a tutorial at eight.
I thought you could move it.
I can't leave the bakehouse. Without the part, we'll keep losing sales.
And without the tutorial, I lose out. Five to seven, yes. No later.
I'll leave at five and be back before seven. I'll have to rush.
All right. But change it on the rota.
I'll change it.
20
19
I have to get there. I'm not trying to be difficult.


Tuesday, 15 September
If we shrink it with all that margin, the tissue almost disappears. I'll crop it first.
Get rid of all the white at once.
I'll test it. That doesn't mean it'll work.
Internal test
Baseline 0.41 · Attempt 0.39
We've lost the morning.
And some of the tissue. We cropped too much.
Same patients kept separate for testing, yes?
Yes. I'm not changing the test to get a better score. I'm changing the crop.

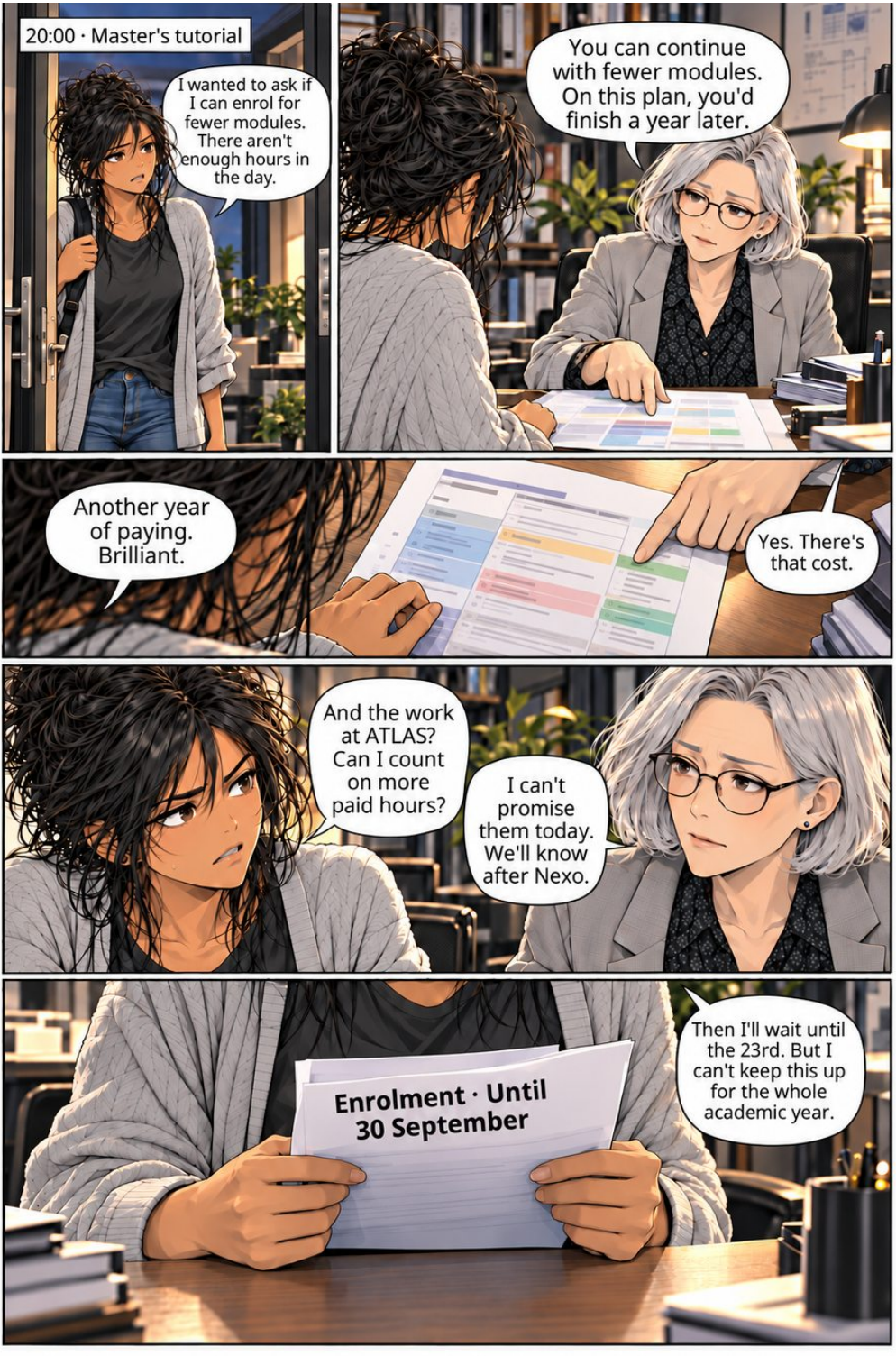

Wednesday, 16 September
The pale tissue counts too. I want to remove the empty border, not everything that looks white.
I'll apply it to everything, and we'll train.
No. First I'll look at the cases the crop damaged.
These six. It ate into the edge in two of them.
Leave a margin around the tissue. Then resize it.
Mara's run goes in tonight. Your variant waits for the result.
As if ARGOS were going to wait too.
We'll review it together. And I'm leaving at four tomorrow. I'm not spending another night here.


Now it does preserve the pale areas.
Crop the empty border, then shrink. That's fixed.
Internal test · Preparation B
Test patients kept separate
I'll leave it running. Results tomorrow.
Weren't you coming?
Five minutes.
I'm just going to check one thing.
Parts challenge · Deadline 25 September
00:47

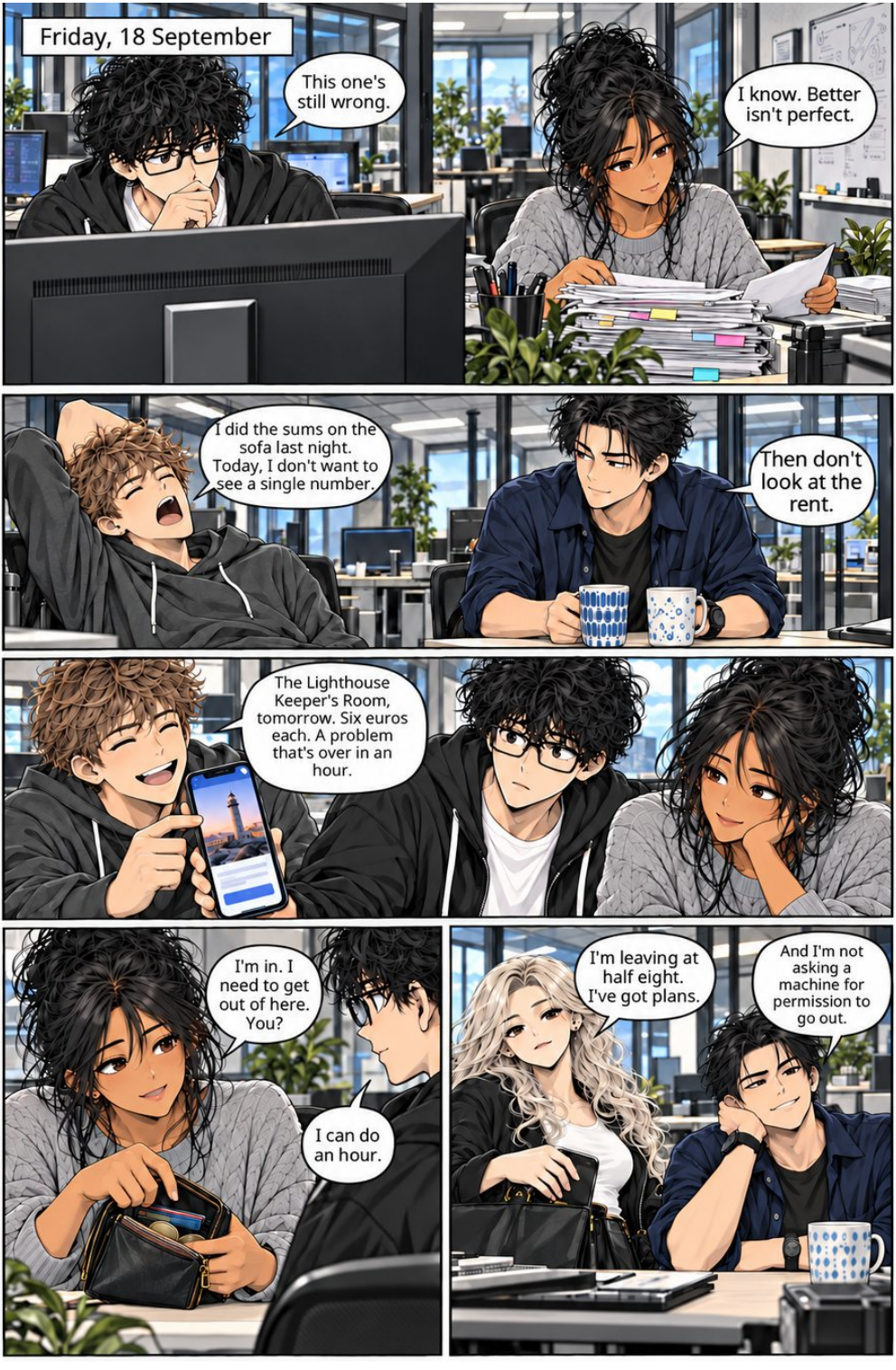

Thursday, 17 September · ATLAS
Internal validation
Baseline 0.41 ·
Preparation B 0.53
Twelve hundredths!
Don't celebrate until Nico repeats it.
Hours later
Same preparation, same patients. Same result.
And against ARGOS?
This is our test. We still need Nexo's shared test.
There is an improvement. No promise of 0.53 outside here.
Then let's submit this candidate for the review on the 23rd.

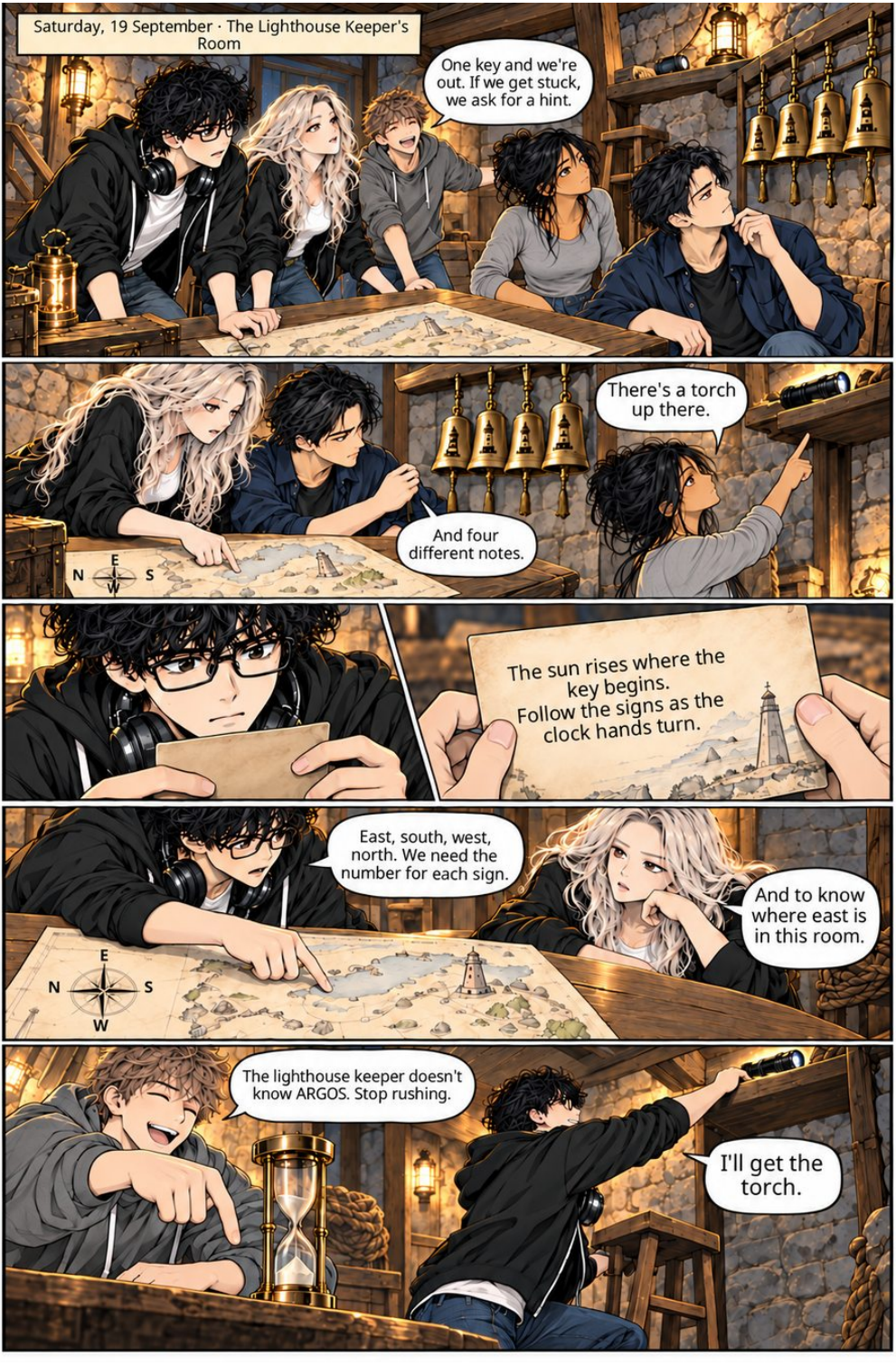

15:53
What if we require less similarity to accept a match?
Seven minutes left, Leo.
It might find the ones we're missing...
And if we get more false matches, who's reviewing them?
We freeze the candidate today. Tomorrow we check the submission.
ARGOS won't wait.
Neither will Mara's tutorial. We need to get to Wednesday with something that works.
Nico, B is saved. If anything changes, write it down for me.
All right.
I wanted to save time. Now it looks as if I'd used it up myself.


17:00 · La Segunda Taza
The pump's ready. I'll be back at seven.
Behind you! Hot!
I saw you.
18:54
Does the chocolate one have nuts in it?
I'll check with my dad.
19:00
Go on. I'll carry on with this table.
Thanks. See you later.
One hour. I can make it.

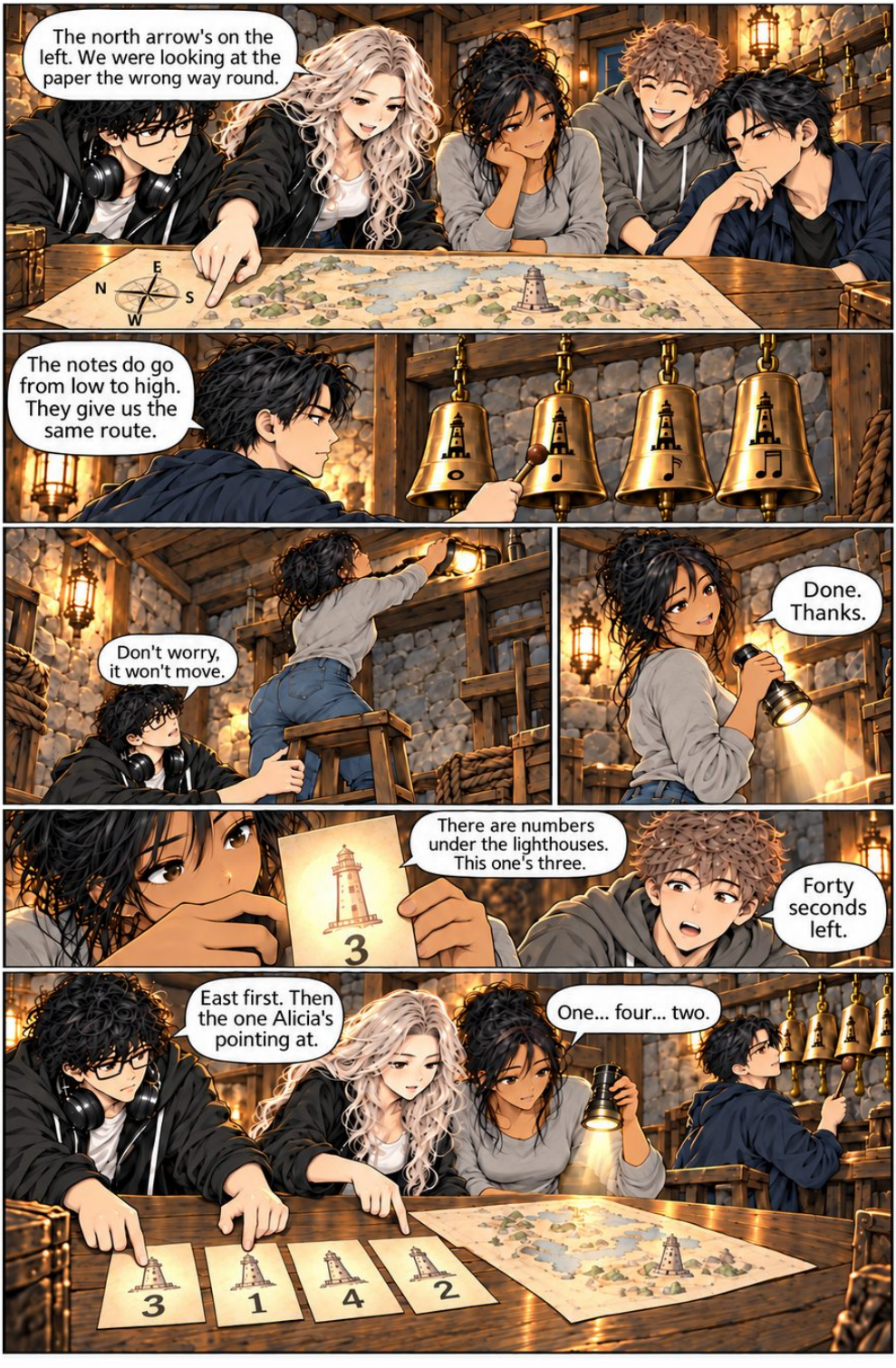

20:00 · Master's tutorial
I wanted to ask if I can enrol for fewer modules. There aren't enough hours in the day.
You can continue with fewer modules. On this plan, you'd finish a year later.
Another year of paying. Brilliant.
Yes. There's that cost.
And the work at ATLAS? Can I count on more paid hours?
I can't promise them today. We'll know after Nexo.
Enrolment · Until 30 September
Then I'll wait until the 23rd. But I can't keep this up for the whole academic year.


21:15 · Room viewing
Seven hundred and eighty, plus bills. You can pay a holding deposit today.
I can't afford it. Even cutting everything back.
There's another viewing after you.
Then I won't take it. Thanks.
21:45
Jun... could I crash on your sofa? Until the 30th. I'm still looking.
Until the 30th. Quiet after ten. And we split expenses.
Deal. I owe you one.
You owe me some sleep. I'm up early tomorrow.

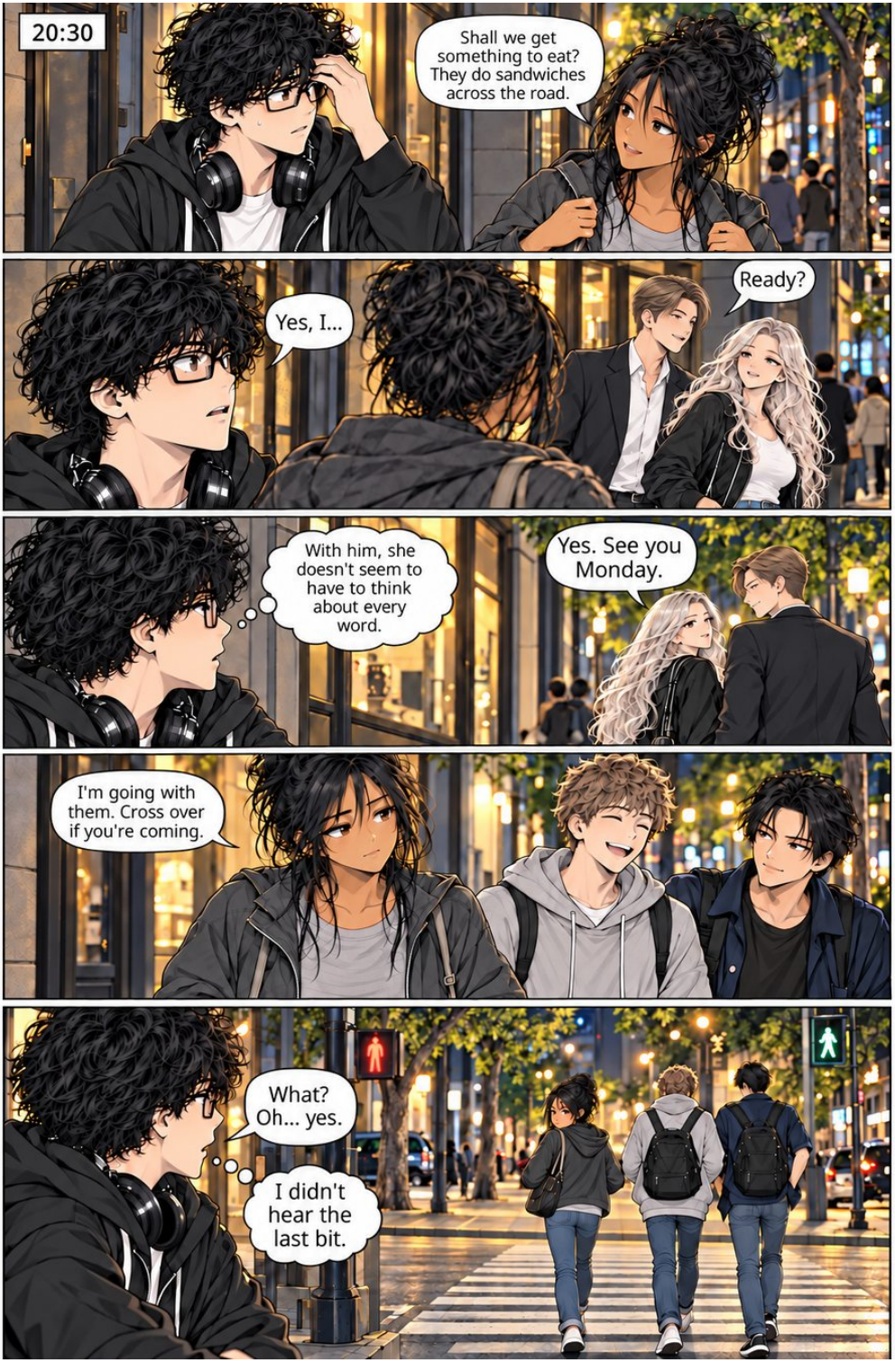

Friday, 18 September
This one's still wrong.
I know. Better isn't perfect.
I did the sums on the sofa last night. Today, I don't want to see a single number.
Then don't look at the rent.
The Lighthouse Keeper's Room, tomorrow. Six euros each. A problem that's over in an hour.
I'm in. I need to get out of here. You?
I can do an hour.
I'm leaving at half eight. I've got plans.
And I'm not asking a machine for permission to go out.


Saturday, 19 September · The Lighthouse Keeper's Room
One key and we're out. If we get stuck, we ask for a hint.
There's a torch up there.
And four different notes.
N
E
S
W
The sun rises where the key begins. Follow the signs as the clock hands turn.
East, south, west, north. We need the number for each sign.
And to know where east is in this room.
N
E
S
W
The lighthouse keeper doesn't know ARGOS. Stop rushing.
I'll get the torch.


Oof... I'm out of shape. If it shines through the backs of the cards...
THUNK!
Don't move! Are you all right?
Yes. The ceiling was badly documented.
The ceiling's been here a hundred years.
Then file a bug report.
That I can do.
Hold it. Without moving.
Wait a moment. This map's rotated.
E
N
S
W


The north arrow's on the left. We were looking at the paper the wrong way round.
N
E
S
W
The notes do go from low to high. They give us the same route.
Don't worry, it won't move.
Done. Thanks.
There are numbers under the lighthouses. This one's three.
3
Forty seconds left.
East first. Then the one Alicia's pointing at.
One... four... two.
3
1
4
2


Three, one, four, two?
00:38
Yes. Put it in.
3 1 4 2
CLICK!
I've lost a neuron, but we won.
With thirty-eight seconds to spare.
You didn't have that many to spare.
I said a problem that lasted an hour.
This one didn't come with rent.


20:30
Shall we get something to eat? They do sandwiches across the road.
Yes, I...
Ready?
With him, she doesn't seem to have to think about every word.
Yes. See you Monday.
I'm going with them. Cross over if you're coming.
What? Oh... yes.
I didn't hear the last bit.


Sunday, 20 September
I'll move my things now.
The table needs to be clear. I practise here.
I can't afford the extra €70 from October. I'll leave the room on the 30th.
Next week: Monday, Tuesday and Wednesday until 16:00. Then tutorials and study.
Understood. We'll organise the submission around those hours.
I know a clue by heart, then fail to see a beam.
I could sleep. I could go out without looking for something to solve.


The crop was Mara's idea. I wanted to keep training the same thing.
In the room, I needed more hands. Listening would have helped too.
Parts challenge · Deadline 25 September
I still want to do something of my own.
23
NEXO
23:18
Next week, I'll have to take time from somewhere. Once again, I'm taking it from sleep.

Looking at the Image the Network Receives



Images grouped by patient before splitting the data.

Mara notices the white margin. If we shrink the whole image before looking at where the tissue is, we may leave the part we care about in just a few pixels. However much network we put after it, it will work with whatever has survived that preparation.

Cropping with a case in front of us

The first attempt removes pale tissue, and the internal test drops from 0.41 to 0.39. Something that looked like background was part of the sample. Mara returns to the damaged cases, keeps a margin and crops before resizing. The order changes the detail that reaches the model. Looking at those images together is useful: it makes clear what she is correcting, rather than treating cropping as cosmetic tidying.

Keeping a recognisable baseline

They keep the model and patient split to test the new preparation. Nico repeats the run, and B goes from 0.41 to 0.53 in that internal validation: twelve hundredths of F1. Calling it twelve percentage points of patients recognised would change what the test measures. They still need Nexo's common exam to compare themselves with ARGOS. While they wait, fixing B lets them prepare a submission everyone can identify.

The last-minute idea

Leo proposes lowering the minimum similarity with seven minutes left before Mara leaves. Accepting more candidates may recover missed pairs and add false matches. Both need reviewing. Inés keeps the version fixed because it is the one they can prepare for that milestone. The other idea remains pending, and testing it will need team time as well as a change to the program.

Sources: · Preparation and leakage · Thresholds

Time with the Group



A sofa, a table and several uses for the same space.

In the lighthouse keeper's room, Leo brings the lab's ideas and urgency with him. He finds a clue, moves without looking and bumps into something. What interests me is that the game continues and he can stay part of it, even after looking a little foolish.

Changing roles during the game

For a while he interprets the route; later he holds the stool so Mara can reach the torch. Alicia gives directions, Jun reads notes, Mara gets numbers and Nico watches the time. Sharing the roles lets him participate without finding every answer. The jokes and the bump belong in the scene alongside the help he receives. They are solving a game together and can enjoy themselves while getting in each other's way a little.

No longer hearing the person nearby

Alicia leaves with Álvaro, and Leo's attention follows her. Mara is suggesting sandwiches, but he starts comparing himself with the ease he imagines between the other two. The invitation exists; to take it up, he would need to keep listening. We can see clearly here how he can want company and lose a specific opportunity while thinking about another he would prefer.

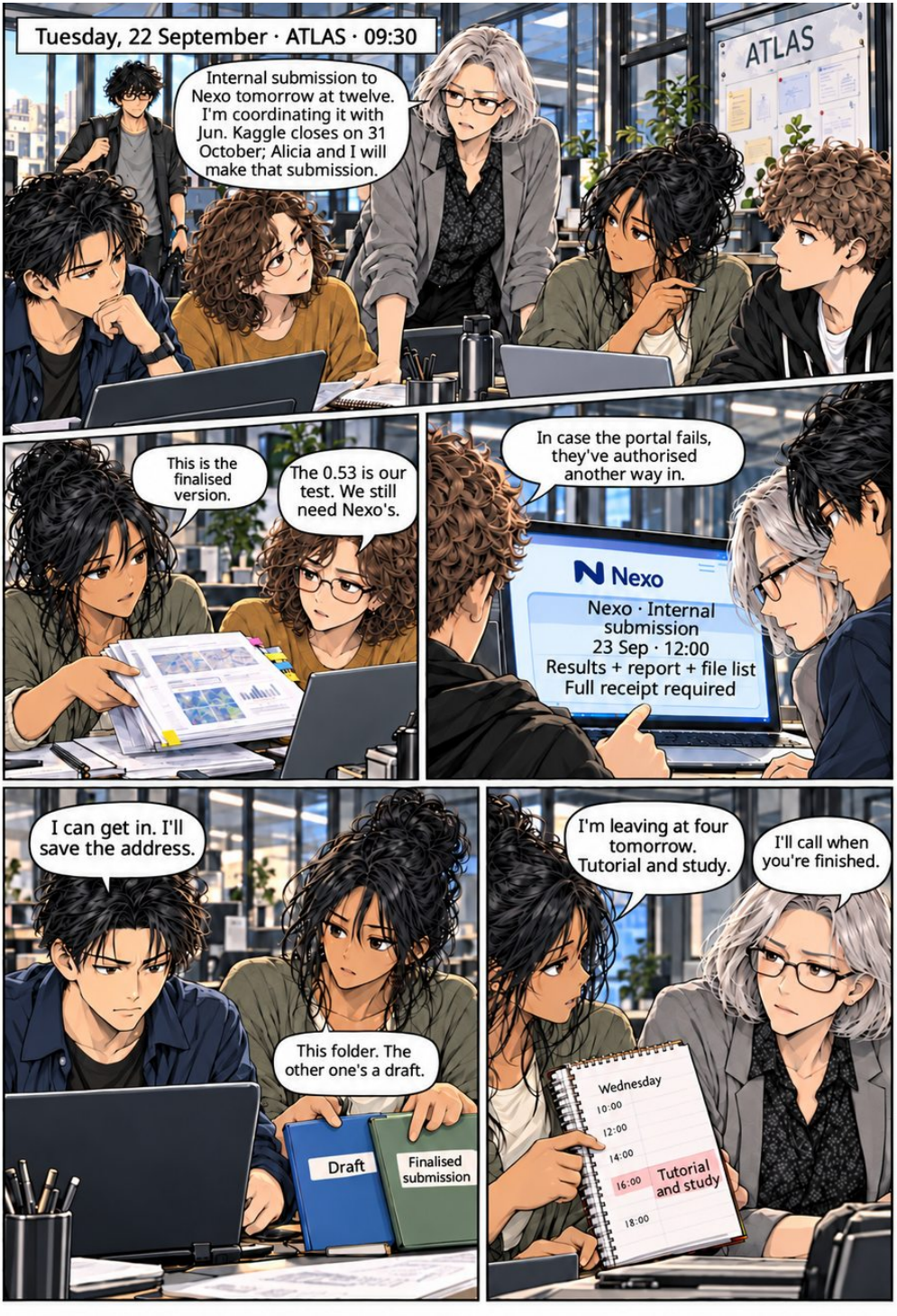
What each person goes back to

Mara negotiates shifts and works out what another year to finish the master's would mean. Nico turns down an expensive room, and Jun offers him a sofa until a set date. Those decisions still have consequences when Leo returns to his challenge that night. His reflection acknowledges the group's contributions. When it is time to close the computer, however, he still struggles to keep time for sleep.

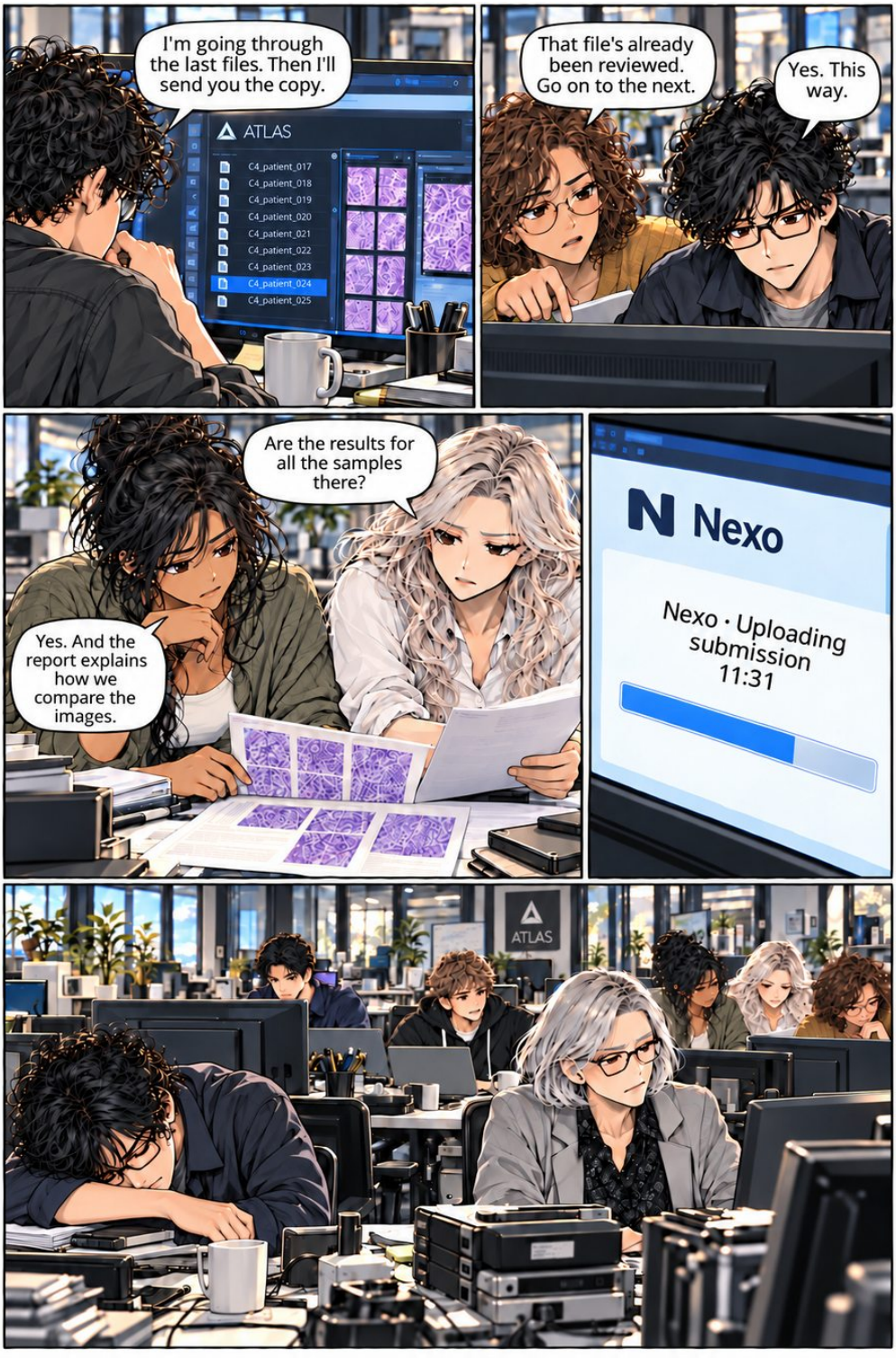

Monday, 21 September · Night
I've Finished My Studies — What Now?
Roberto Canales Mora
Contents
· Love Yourself
· Take Care of Yourself
· Think for Yourself
· Accept Diversity
Take care of myself... I sit down all day, then expect to move as if it's nothing. I could start with a class.
23:17
Booking confirmed
Tuesday 22 · 08:00
Beginners


I like what I do. But I don't want it to be my whole life.
I don't sleep enough. I promise I'll be there, then get caught up. I want company, then I want to leave.
In a few years, I'd like to be working on something that interests me. And still have a life afterwards.
If I try to change everything tomorrow, I won't do anything.
Booking confirmed
Tuesday 22 - 08:00
Beginners
For now, I'll go to that class. Even if I'm the one who gets it least.
Booking confirmed
Tuesday 22 - 08:00
Beginners
Then I'll see what it cost me. One thing at a time.


Tuesday, 22 September · 08:00
CLACK
Some pull...
Less weight. Brace yourself here and pull without hunching your shoulders.
Now it moves.
That one counts.

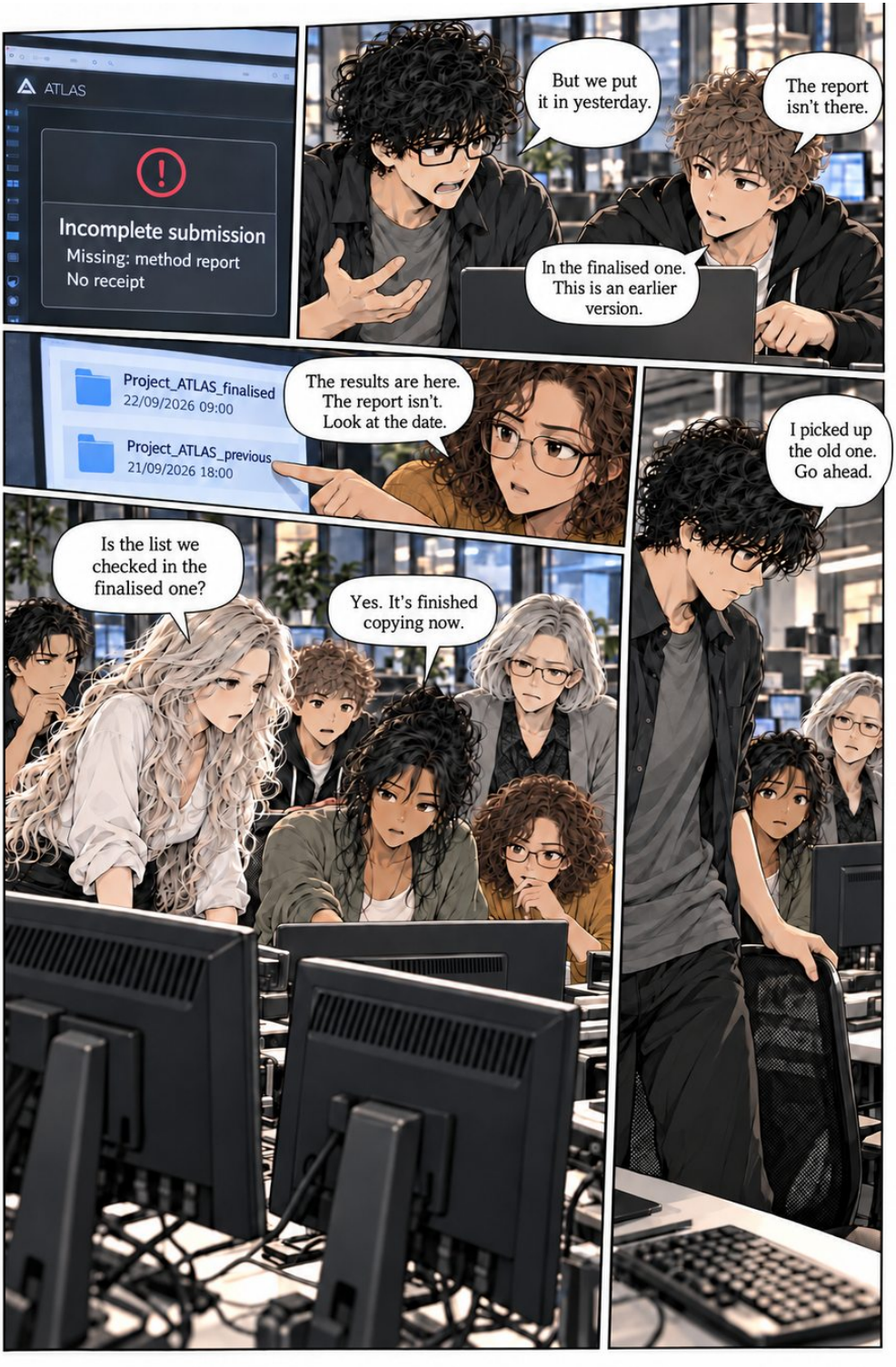

Tuesday, 22 September · ATLAS · 09:30
ATLAS
Internal submission to Nexo tomorrow at twelve. I'm coordinating it with Jun. Kaggle closes on 31 October; Alicia and I will make that submission.
This is the finalised version.
The 0.53 is our test. We still need Nexo's.
In case the portal fails, they've authorised another way in.
Nexo
Nexo · Internal submission
23 Sep · 12:00
Results + report + file list
Full receipt required
I can get in. I'll save the address.
This folder. The other one's a draft.
Draft
Finalised submission
I'm leaving at four tomorrow. Tutorial and study.
I'll call when you're finished.
Wednesday
10:00
12:00
14:00
16:00
Tutorial and study
18:00


Tuesday 22 · 23:48
Industrial parts challenge
Authorised public data
Deadline: 25 September
I've Finished My Studies— What Now?
Roberto Canales Mora
What if I use what another model has already learnt about shapes? One short test
It's time to sleep
I've been awake for a while.
Not now
Remind me later
01:26
Test in progress
Shit.
Wednesday, 23 September · 09:15
ATLAS
The door, Leo.
Yes. Sorry.


I'm going through the last files. Then I'll send you the copy.
ATLAS
C4_patient_017
C4_patient_018
C4_patient_019
C4_patient_020
C4_patient_021
C4_patient_022
C4_patient_023
C4_patient_024
C4_patient_025
That file's already been reviewed. Go on to the next.
Yes. This way.
Are the results for all the samples there?
Yes. And the report explains how we compare the images.
Nexo
Nexo · Uploading submission
11:31
ATLAS

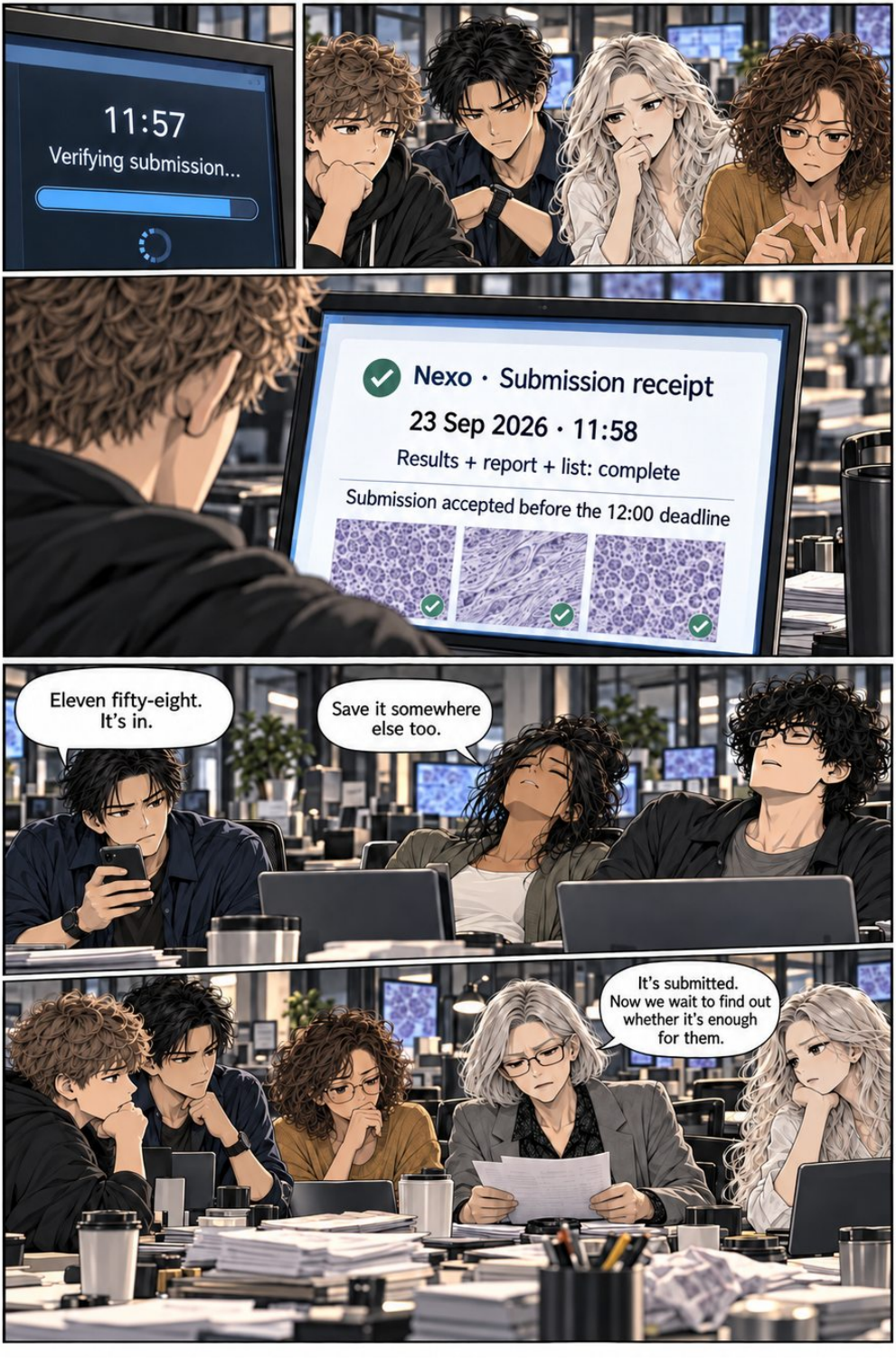

ATLAS
Projects
Data
Models
Tests
11:35
Connection interrupted
It's dropped.
It won't open here either.
ATLAS
CRACK
We've got the files. It's the way in that isn't responding.
Nexo
Operator notice
Denial of service: mass traffic is blocking the portal.
Alternative submission available.
Deadline 12:00.
Let's use the other way in. Twenty-five minutes left.

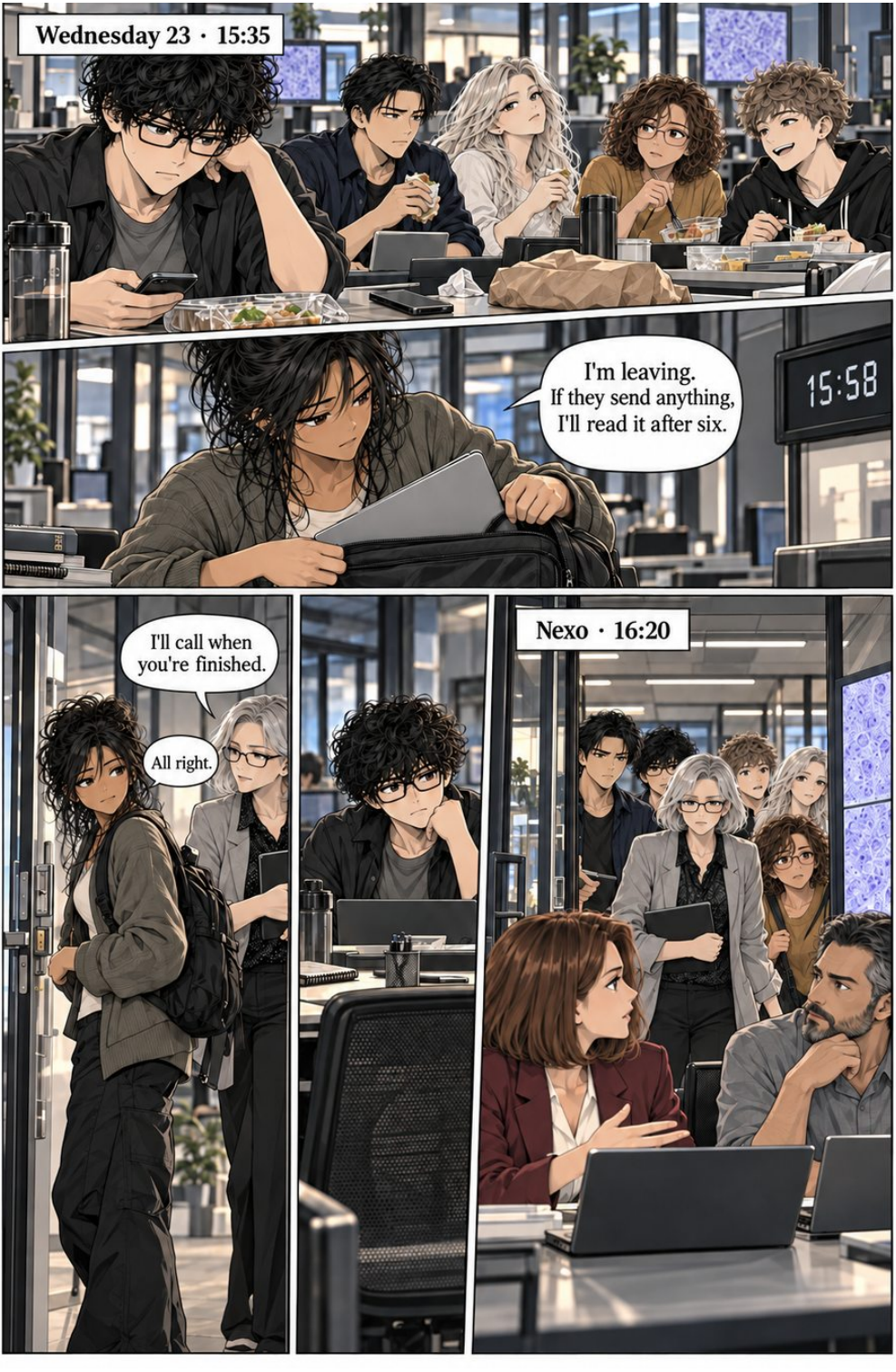

I'll watch the notice. You handle the submission.
I've got the address.
I need the report that goes with those results.
I'll cross-check it. Don't send me three folders.
I've got a copy here. I'll send it now.
I'm copying the finalised one. Wait.
This is it. I had it ready.
11:41 · Checking files

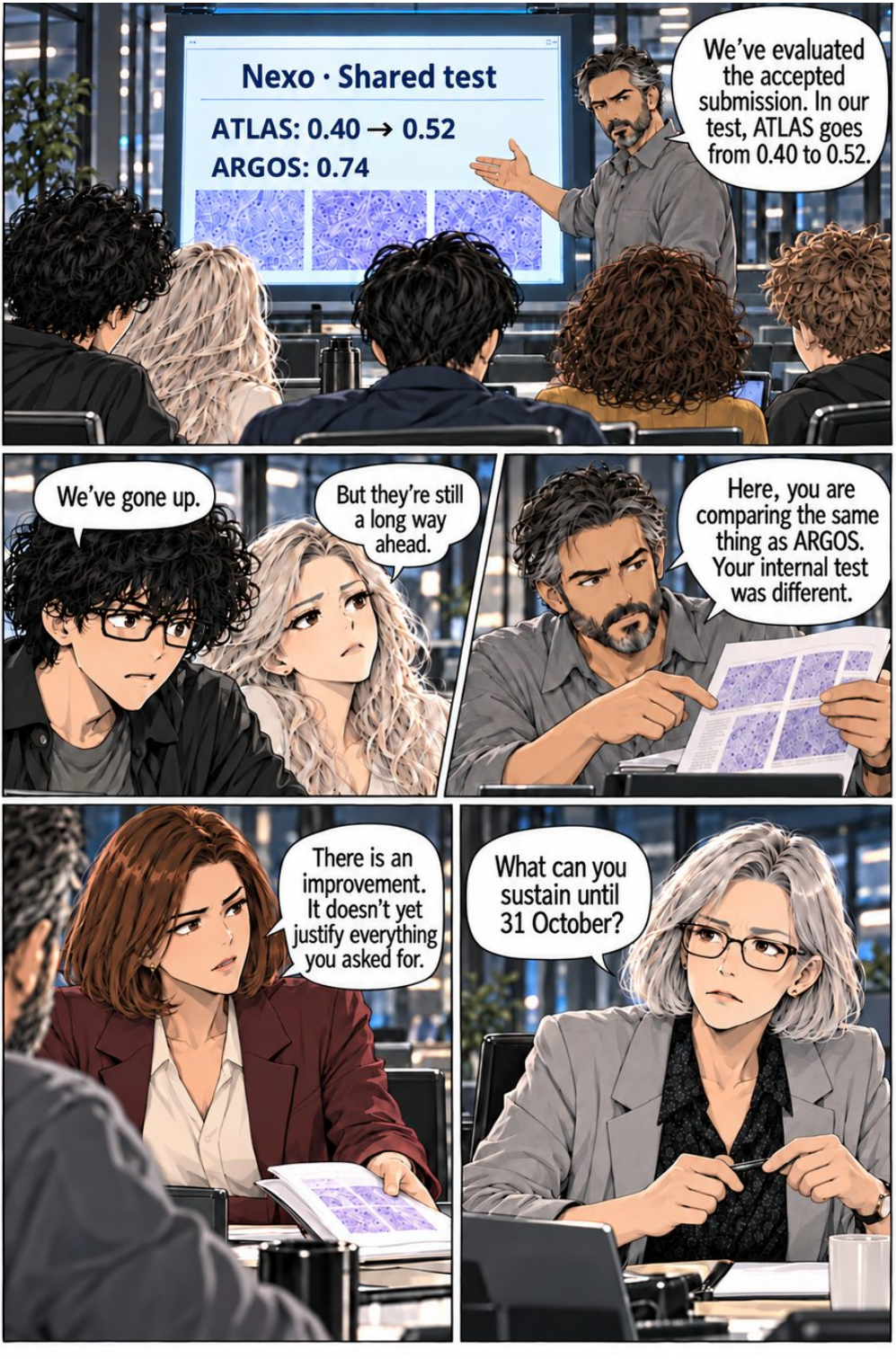

ATLAS
Incomplete submission
Missing: method report
No receipt
But we put it in yesterday.
The report isn't there.
In the finalised one. This is an earlier version.
Project_ATLAS_finalised
22/09/2026 09:00
Project_ATLAS_previous
21/09/2026 18:00
The results are here. The report isn't. Look at the date.
I picked up the old one. Go ahead.
Is the list we checked in the finalised one?
Yes. It's finished copying now.


11:46
I'll put the report in my copy and send it.
We don't know what else changed in that copy.
The finalised one's here. We need to check it.
Fourteen minutes. Do we carry on with that one or start with the complete one?
I've messed up. The complete one. Nico, you submit it.
I'll read the list. Check it with me.
Start with the results.


I've got the submission ready. Is everything there?
All the results. Report, ready. List too.
And it matches this version.
And the last results file… That one. They're all there.
They match.
11:52
Alternative submission · Uploading


11:57
Verifying submission...
Nexo · Submission receipt
23 Sep 2026 · 11:58
Results + report + list: complete
Submission accepted before the 12:00 deadline
Eleven fifty-eight. It's in.
Save it somewhere else too.
It's submitted. Now we wait to find out whether it's enough for them.


Wednesday 23 · 15:35
I'm leaving.
If they send anything,
I'll read it after six.
15:58
I'll call when
you're finished.
All right.
Nexo · 16:20


Nexo · Shared test
ATLAS: 0.40 → 0.52
ARGOS: 0.74
We've evaluated the accepted submission. In our test, ATLAS goes from 0.40 to 0.52.
We've gone up.
But they're still a long way ahead.
Here, you are comparing the same thing as ARGOS. Your internal test was different.
There is an improvement. It doesn't yet justify everything you asked for.
What can you sustain until 31 October?


Partial funding until 31 October. It doesn't cover the whole team or all the resources.
Another GPU would shorten the queue. But this won't pay for Mara's hours as well.
The queue's holding us back.
So would losing the person who prepares the data.
Those hours come first. The other GPU waits.
And my renewal?
We can't cover it now.


So my contract still ends on the 31st.
Yes. If anything changes, I'll tell you. Today, I have nothing more.
I'll send the other application. Can I put you down as a referee?
Yes.
Application sent
CV.pdf


Wednesday 23 · 18:25
I'm free now. What happened?
We're up to 0.52. There's support until 31 October. I can offer you paid hours to prepare the data.
Wait, I'm reading it.
Terms
Amount
Hours
Duration
With this, I can drop one café shift. I can do the afternoons listed here.
Café
Tutorial
I accept. But I still have to decide about my modules. I can't fit everything in.


That afternoon • Elsewhere
The contribution made it possible to arrange the additional interview you requested.
Alicia passed the test on her own. She doesn't know about this arrangement.
Wednesday 23 • 23:07
Today, it got in on time.
Mara found the package.
Nadia and Alicia checked it.
Nico and Jun got it through.
I sent the old copy.
It annoys me to still be behind. I'm glad we're carrying on. Both things.
Next class • Friday 08:00
Industrial parts challenge
Deadline: 25 September
I want to carry on with my own thing. But today, because I didn't sleep, I put the team in trouble.
Test in progress
23:19
Half an hour... I'm telling myself the same thing again.

Getting the Reviewed Version Delivered



GPU, requests to a service and people's time.

The portal goes down with twenty-five minutes left. The files are still at ATLAS and the model's score has not changed, but the submission is at risk. Leo tries to help by passing over an earlier copy and puts a second problem inside the first.

What that folder contains

The copy has results but lacks the report. They could add it quickly, but Nadia asks what else might have changed. That question makes them return to the closed folder and compare results, report and checklist. When we combine files from different versions, we can create a package nobody has checked. Under that pressure, I prefer an identified version and a list in front of us to someone remembering which was the right one.

What each step proves

The alternative channel was authorised. Using it lets them send; receiving the acknowledgement confirms that the package arrived complete and on time. Evaluation follows: Nexo reports that ATLAS has moved from 0.40 to 0.52 on its common test. The internal 0.53 remains a result from another check. They are useful data for different questions. This Nexo submission requires a research package; the competition will receive a CSV file with the matches.

Who the budget can cover

The improvement supports partial funding until 31 October. Inés has to choose: another GPU would reduce the queue, while paying Mara's hours secures data preparation. She chooses Mara and cannot renew Nadia's contract. The scene makes limited resources concrete. You can buy more computation and still lack someone to prepare images or time to review an output.

Sources: · WSI challenge · Saving the model

Packing a Bag and Going to Class



Preparing an activity and reserving an hour for it in the week.

Leo starts asking what he wants from life beyond a good result. He feels tired, stretches promises and sometimes seeks company only to want to leave straight away. Reading the book helps him put that mixture into words.

A first attempt that fits into one evening

He wants interesting work and a life afterwards. That evening, he could design an enormous change; instead he packs a bag for a class he has already booked. The next day he has to be among those with the least control over their movement and ask the instructor for help. These are small, specific actions: a T-shirt, trainers, an hour and a willingness to try something he still does awkwardly.

The rush returns with the submission

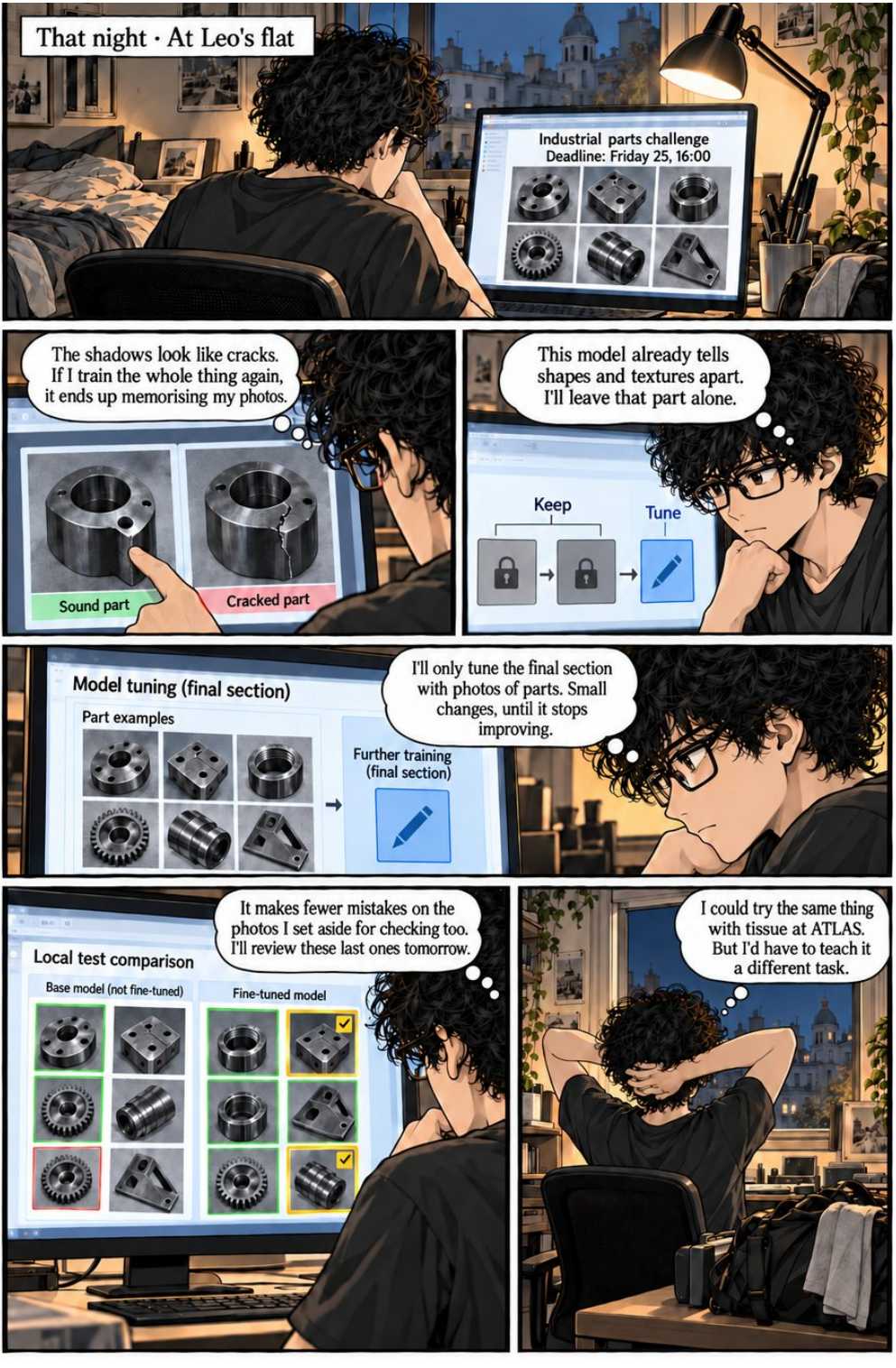
Soon afterwards, he gets ahead of Mara and offers an old copy. Again he acts before waiting for the team's checks. Now he acknowledges the mistake, leaves the keyboard to Nico and helps check the correct folder. I am interested in that response: he has messed up again, but this time he lets someone else send while he reads the list. The fear of looking bad remains, even with the bag packed and the first class done.

What makes a free afternoon possible

Inés pays Mara's hours instead of buying another machine; Nadia looks for a job because her renewal does not come through. Looking after yourself also takes income and schedules you can count on. Leo sees the tiredness he is bringing into the team and still opens the individual challenge again. He really wants to. Recognising the conflict does not mean he has managed to stop: he is still negotiating another stretch of sleep with himself.

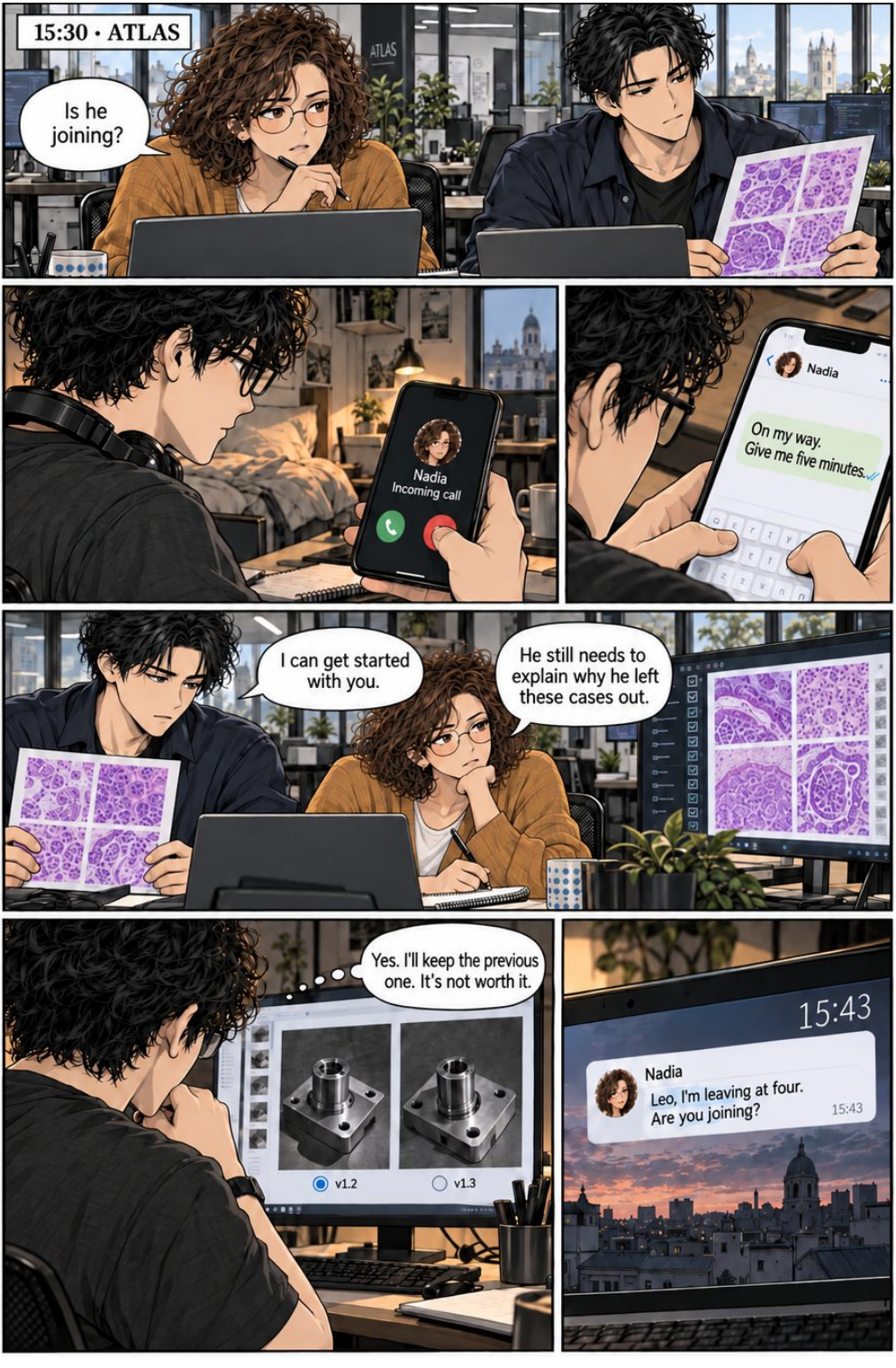
What would he need to finish the evening before to get to class without losing more sleep?

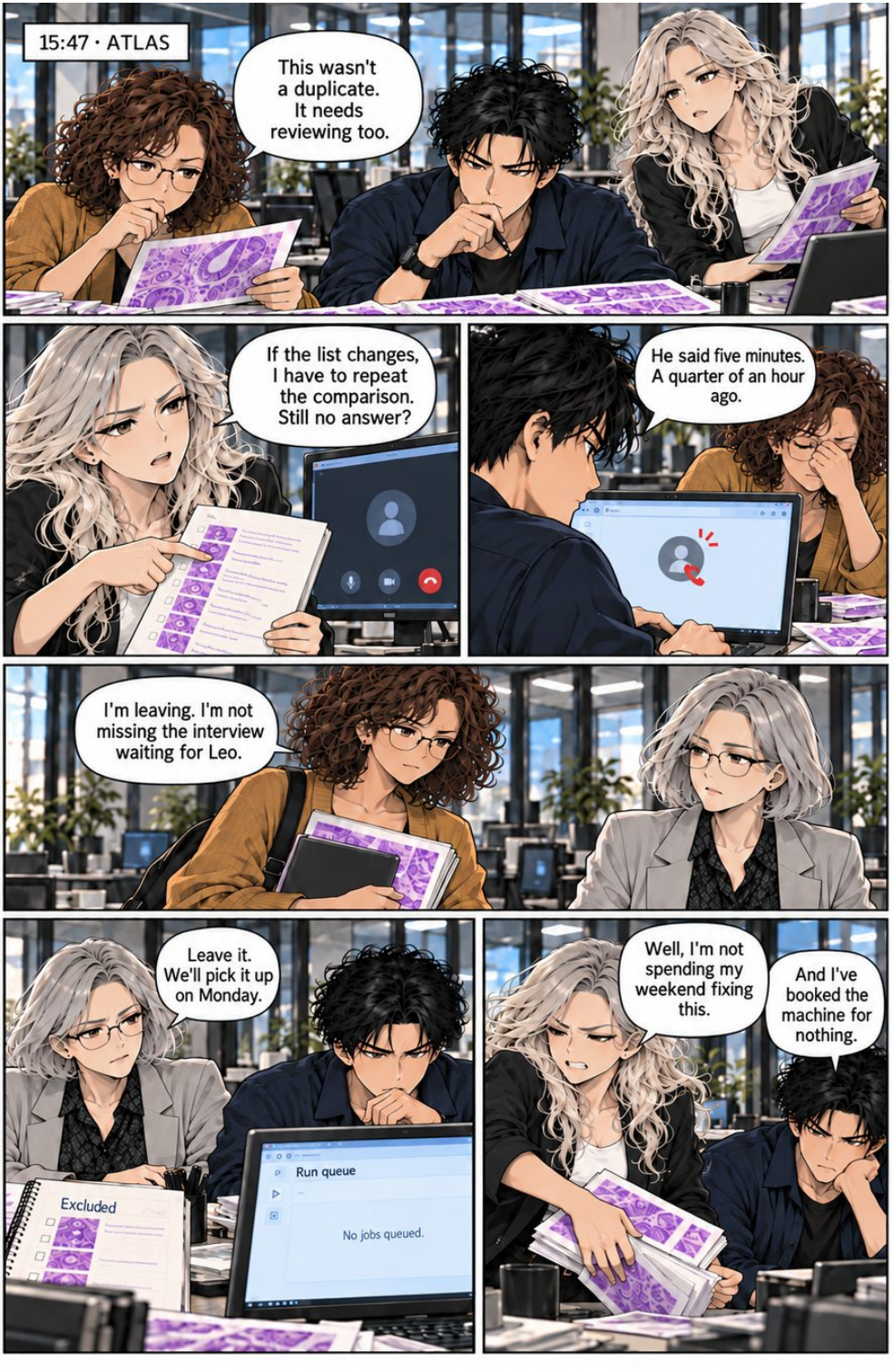

Thursday, 24 September · ATLAS
We've got until 31 October. We're still a long way behind.
Nexo test ·
ATLAS 0.52
ARGOS 0.74
Tomorrow, we review the variant's failures. If it accepts more matches, I want to see how many are false.
I'll be at home in the afternoon. I'll join online.
Friday 25
Review with Nadia
Variant analysis
Nexo report
Run tests
Half three to four. After that, I've got a job interview.
I'll prepare the run when you're finished.
Yes. Half an hour. I'll be there.
10:24
New reminder
Friday 25
15:30
Review with Nadia
Add note
Add location
Save


That night · At Leo's flat
Industrial parts challenge
Deadline: Friday 25, 16:00
The shadows look like cracks. If I train the whole thing again, it ends up memorising my photos.
Sound part
Cracked part
This model already tells shapes and textures apart. I'll leave that part alone.
Keep
Tune
Model tuning (final section)
Part examples
Further training (final section)
I'll only tune the final section with photos of parts. Small changes, until it stops improving.
Local test comparison
Base model (not fine-tuned)
Fine-tuned model
It makes fewer mistakes on the photos I set aside for checking too. I'll review these last ones tomorrow.
I could try the same thing with tissue at ATLAS. But I'd have to teach it a different task.

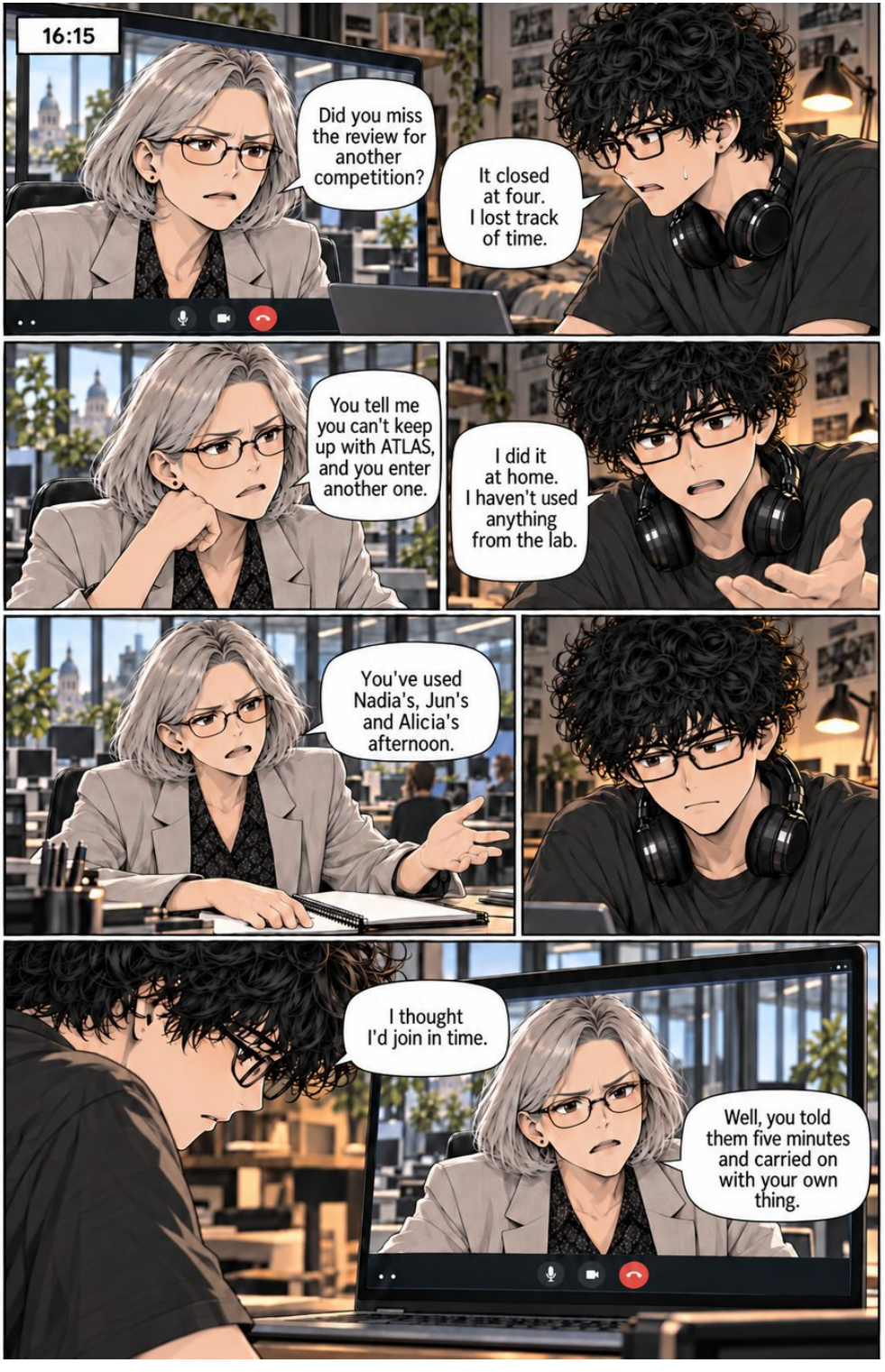

Friday 25 · 08:00
Back again?
I came down the stairs sideways.
Easy today. And don't hunch your shoulders.
5
10
15
20
25
30
35
50
65
80
All right. That one was less horrible.
14:35 · At home
I'll submit this and have plenty of time.
Review with Nadia · 15:30
Challenge submission · 16:00
Wait. Not this one again.


This one finds the cracks better. But it invents some too.
15:08 • Local comparison
Previous model
Current model
I could keep the previous one. It's ready.
Industrial challenge submission
predictions.csv
Ready
Submit
I just want to check whether it's always the shadow.
15:25 • Review with Nadia in 5 minutes
15:25
Five minutes.

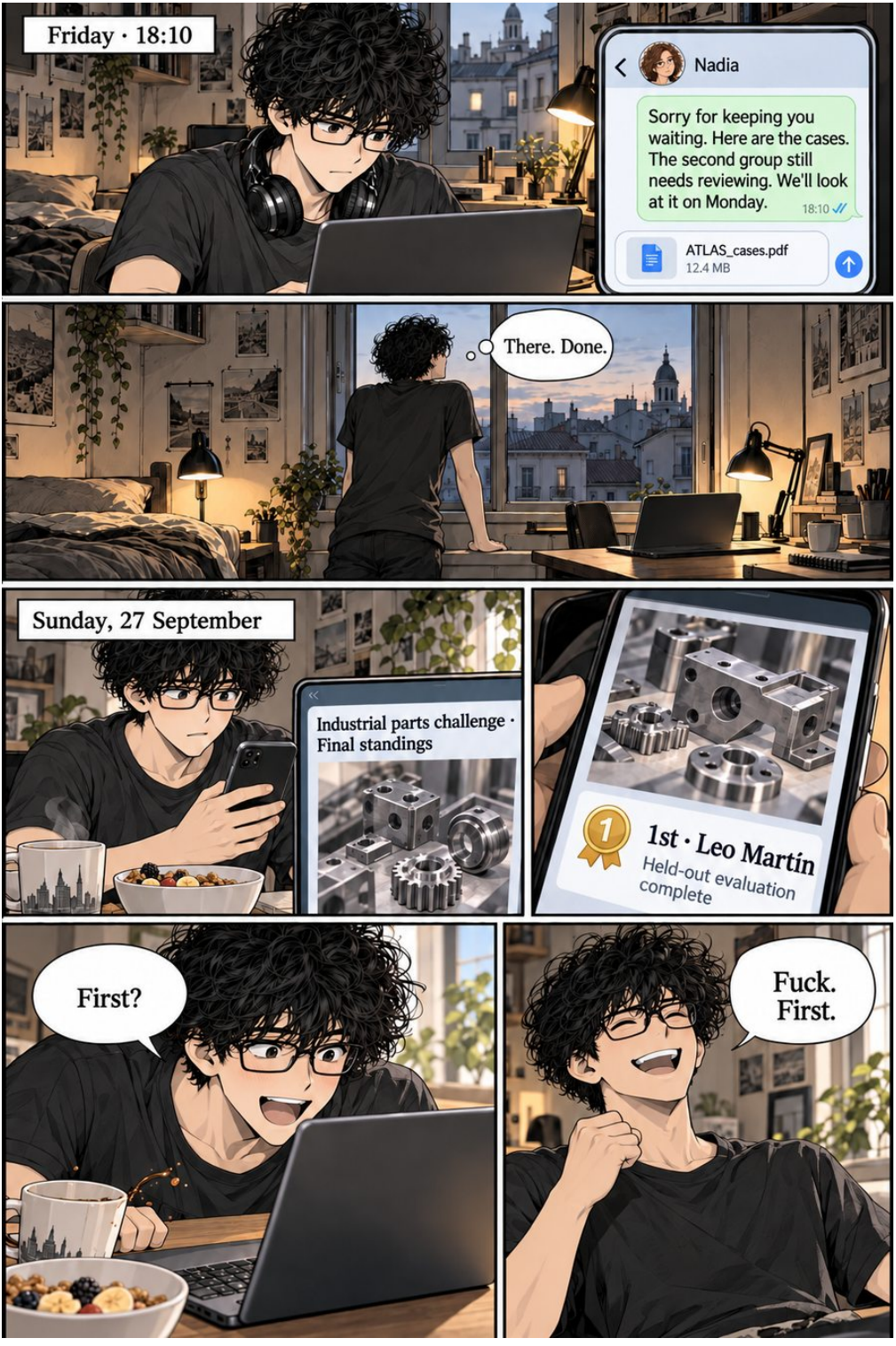

15:30 · ATLAS
ATLAS
Is he joining?
Nadia
Incoming call
Nadia
On my way.
Give me five minutes.
I can get started with you.
He still needs to explain why he left these cases out.
Yes. I'll keep the previous one. It's not worth it.
v1.2
v1.3
15:43
Nadia
Leo, I'm leaving at four.
Are you joining?
15:43


15:47 · ATLAS
This wasn't a duplicate. It needs reviewing too.
If the list changes, I have to repeat the comparison. Still no answer?
He said five minutes. A quarter of an hour ago.
I'm leaving. I'm not missing the interview waiting for Leo.
Leave it. We'll pick it up on Monday.
Excluded
Run queue
No jobs queued.
Well, I'm not spending my weekend fixing this.
And I've booked the machine for nothing.


Home, 15:54.
15:54 · Sending predictions
Submission received · 15:57
Results: Sunday 27
In.
Nadia
16:04 · Nadia · No answer
I'm here.
Are you still there?
Nadia's gone.
We told you the time.
I was finishing something else.
Right.
So were we.


16:15
Did you miss the review for another competition?
It closed at four. I lost track of time.
You tell me you can't keep up with ATLAS, and you enter another one.
I did it at home. I haven't used anything from the lab.
You've used Nadia's, Jun's and Alicia's afternoon.
I thought I'd join in time.
Well, you told them five minutes and carried on with your own thing.

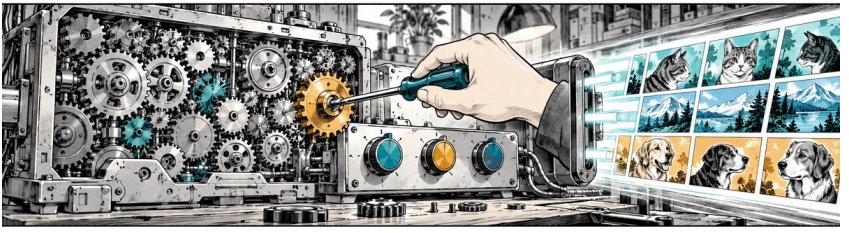

You find time for your own stuff. Ours can wait, can it?
I've been working my guts out here for weeks too.
So have we. That's why it pisses me off.
There, I can try something without asking permission for everything.
And here, you have to do things you don't feel like doing. Like everyone else.
We start on time on Monday. With or without you.
I'll be there.


Friday · 18:10
Nadia
Sorry for keeping you waiting. Here are the cases. The second group still needs reviewing. We'll look at it on Monday.
18:10
ATLAS_cases.pdf
12.4 MB
There. Done.
Sunday, 27 September
Industrial parts challenge · Final standings
1st · Leo Martin
Held-out evaluation complete
First?
Fuck. First.


ATLAS
Leo
I came first in the parts challenge.
Nico
Congratulations! Now explain how you did it without talking in code.
Nadia
Congratulations, Leo. See you Monday.
Mara
That's brilliant! Tell me later. I'm busy with orders.
Alicia
Congratulations. Let's see if we can make progress here tomorrow too.
Nico
Fancy a drink today?
Nico
I've got to move the boxes today. Monday after the lab?
Leo
All right.


Sunday · Night
I've won. And I'm still eating dinner in front of the computer.
I won. I'm absolutely thrilled.
The test I left hanging is still waiting for Monday. Winning doesn't change that.
I could enter this one too.
New challenge · Entries open
Another one, once I've done what I've already promised.

Taking an Idea to Another Task



A metaphor for adjusting part of an already trained model.

Leo finds a way to use a general model in his parts challenge: he keeps features and adjusts the final section to distinguish cracks from shadows. It is worthwhile work. The next question is which part of that idea he can test on tissue images.

What keeping one part fixed means

Training changes parameters, the internal values the network uses to respond. If we freeze part of it, those values stay while we adjust others. We can teach a new output on fixed features or allow previously learnt layers to change too. So when someone says fine-tuning, I would ask exactly what they allowed to change. The earlier base saves starting from scratch, but we need to know the procedure to repeat it.

Comparing the cracks

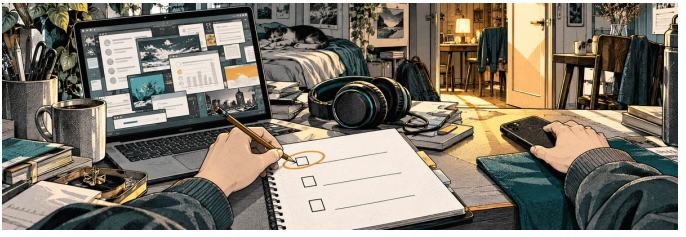
With held-out photographs, Leo sees that one variant finds more cracks and also invents some. He keeps the previous one. His success belongs to that task and its conditions. To investigate transfer to tissue, he will need to define the examples, question and expected response again, using data and models permitted for that use. He can take the adaptation strategy; the industrial weights and achieved result do not come with a test on patients.

The review he left pending

While he works at home, Nadia needs his explanation of the cases excluded from ATLAS. Jun has prepared the run, and one part is missing before they can close the comparison. Leo uses his own computer, but he had committed that half hour to them. Using his own machine avoids consuming the lab's GPU. He also needed to reserve time for the appointment he accepted.

Sources: · Transfer learning · Adjusting the network

Winning While Others Wait



Pursuing an idea of your own while agreed team work remains unfinished.

Through Leo I also wanted to show attention deficit in tasks he finds monotonous, and hyperactivity that can leave him exhausted. He has the energy to keep testing cracks, but postpones the review he accepted. That procrastination ends in rushing and unfinished work for others. Winning means an enormous amount to him. Nadia, Jun and Alicia still need him to honour his commitments.

What he had committed to

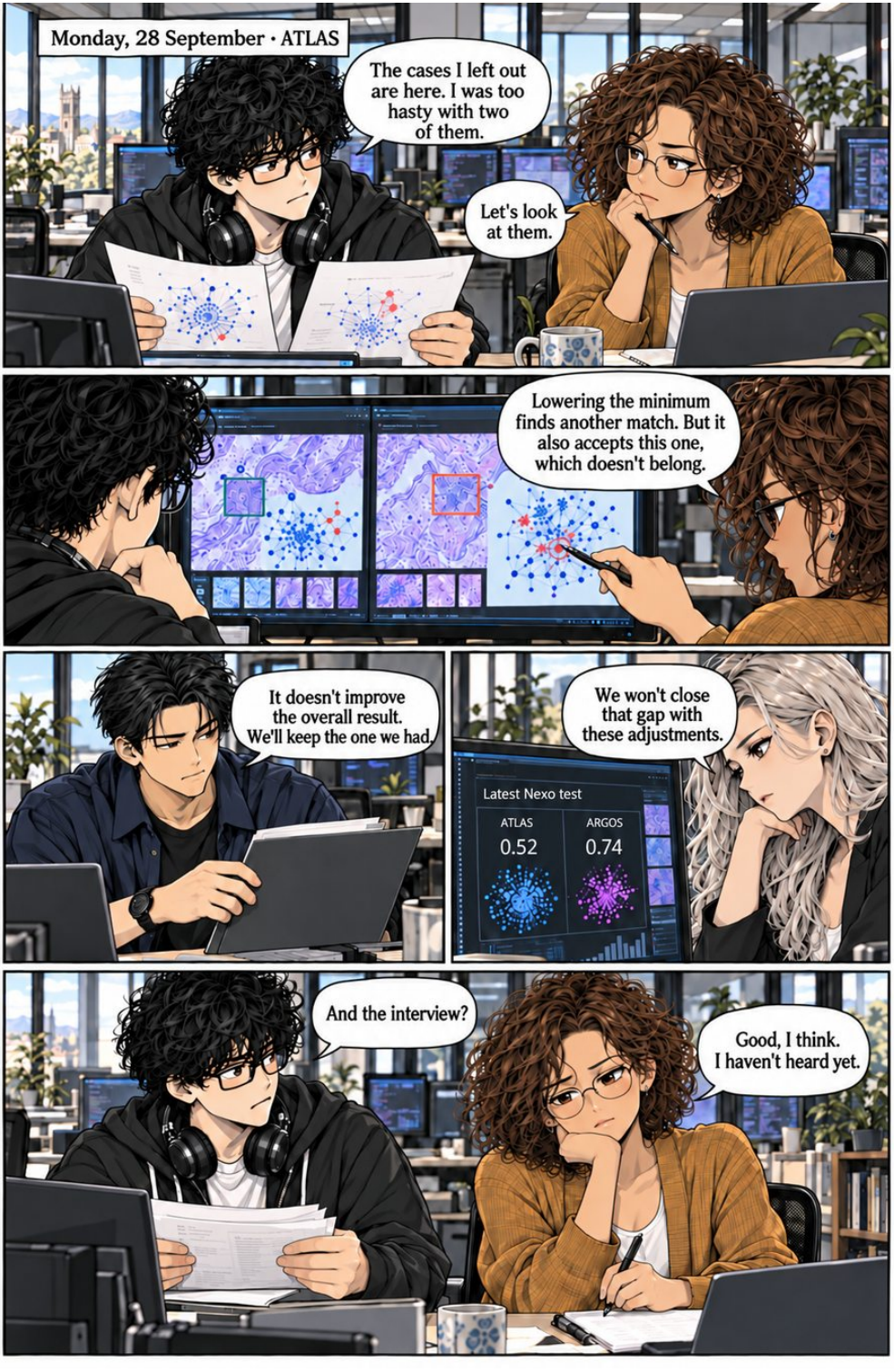
Using his computer lets him run his own test without consuming ATLAS's GPU. He has a version ready to submit, but keeps looking at the cracks: he wants to check whether the error always comes from the shadow. Reviewing the excluded cases is left for later. He had booked that half hour with Nadia, Jun and Alicia and tells them five minutes when he does not yet know how long he will take. He is doing something that interests him, yes. The others are still waiting for an explanation they need to continue.

Nadia's interview starts on time

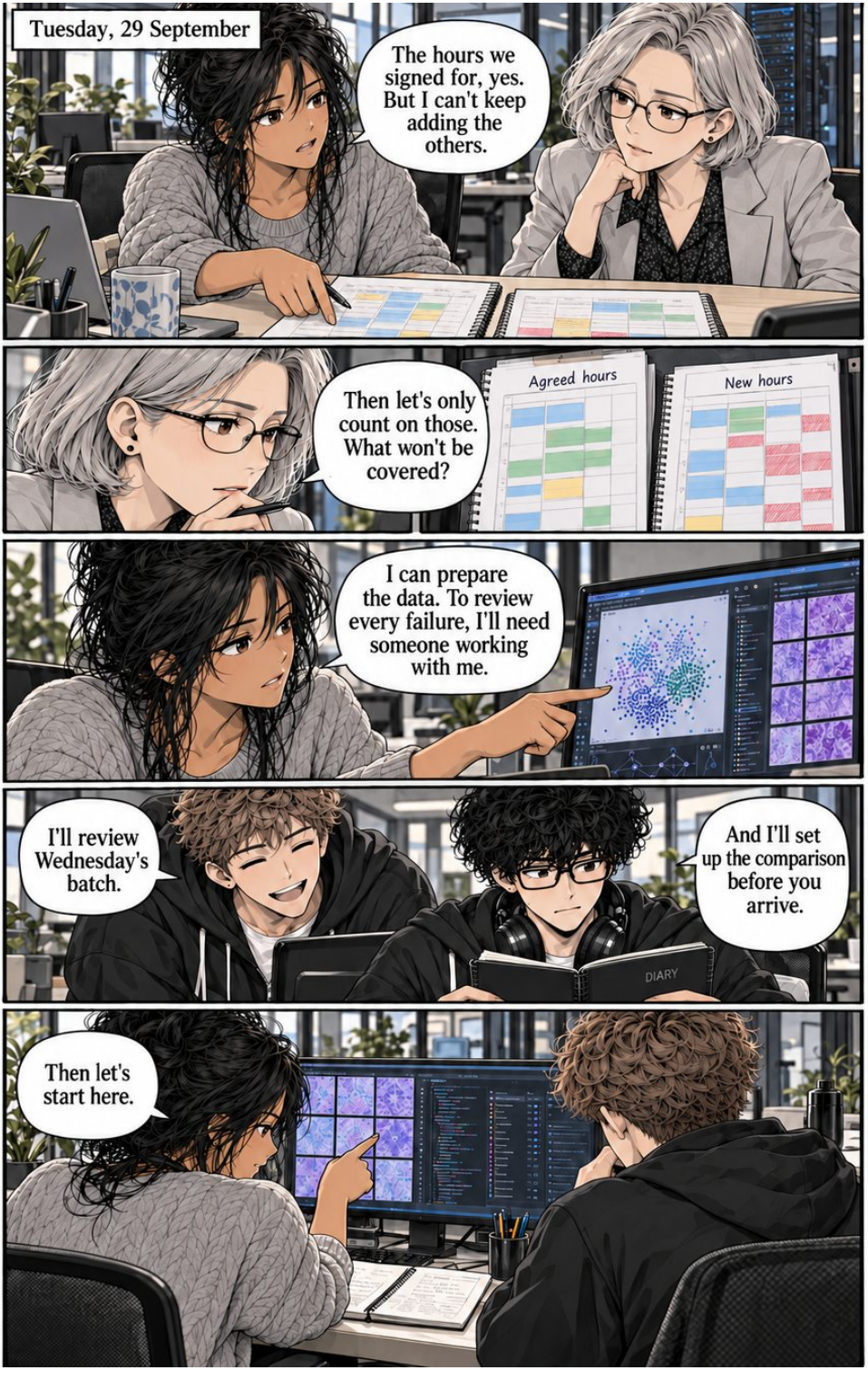
Nadia leaves and Alicia refuses to spend the weekend redoing it. They have commitments Leo knew about when he accepted the appointment. Inés makes him include those hours in his explanation. His result at home is his own and has merit; meeting his commitment to them required finishing another task or negotiating an alternative beforehand. That work remains pending after first place.

Who he wants to tell

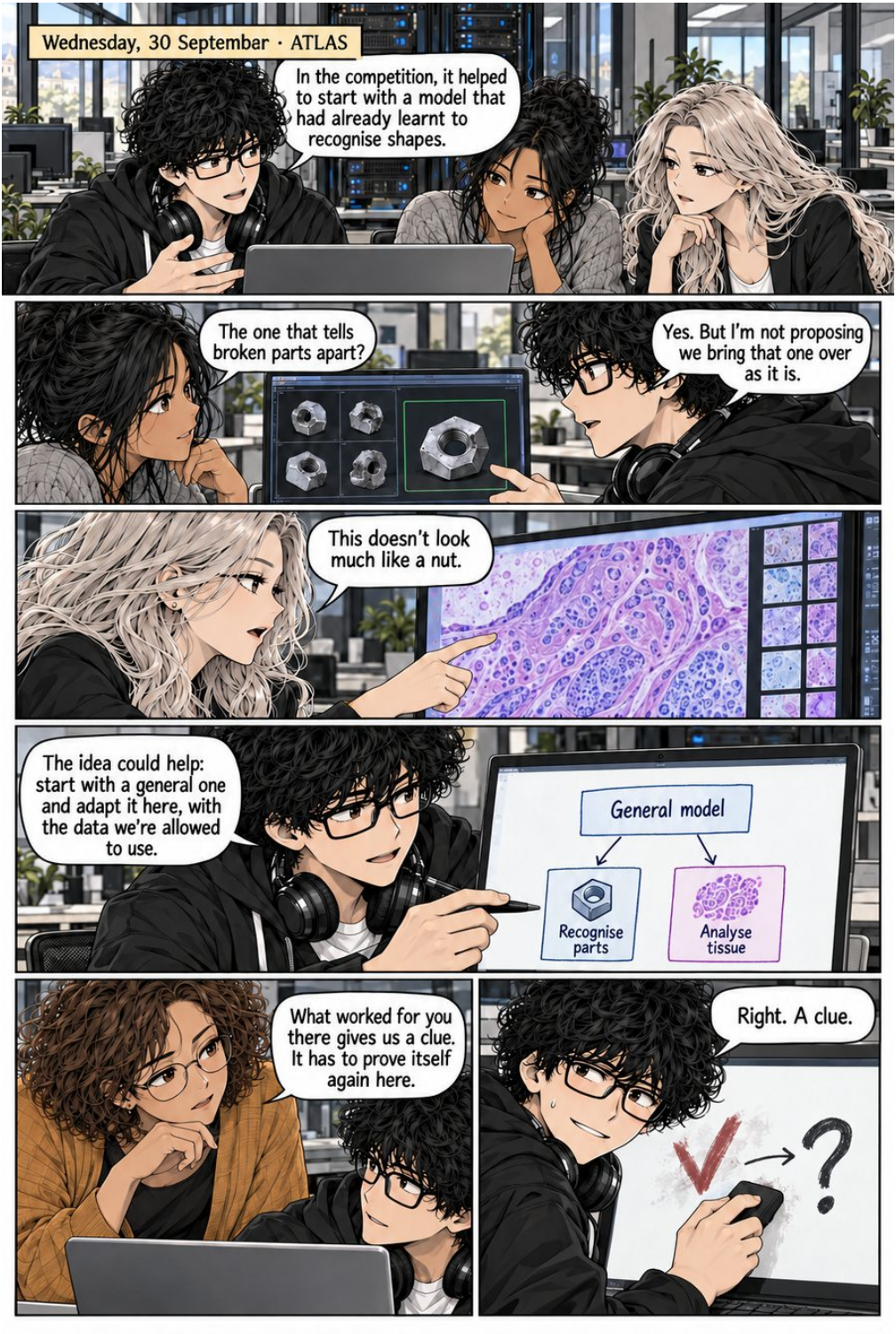
He receives congratulations and finds different schedules: Nico cannot make that Sunday and Mara is handling orders. He ends up eating dinner alone, far from the celebration he expected. From there, he can think winning achieved nothing, or look at how much he expected victory to bring him company. For now, he closes the tab for another challenge. His next free afternoon will need more organisation than waiting for everyone else to be available.

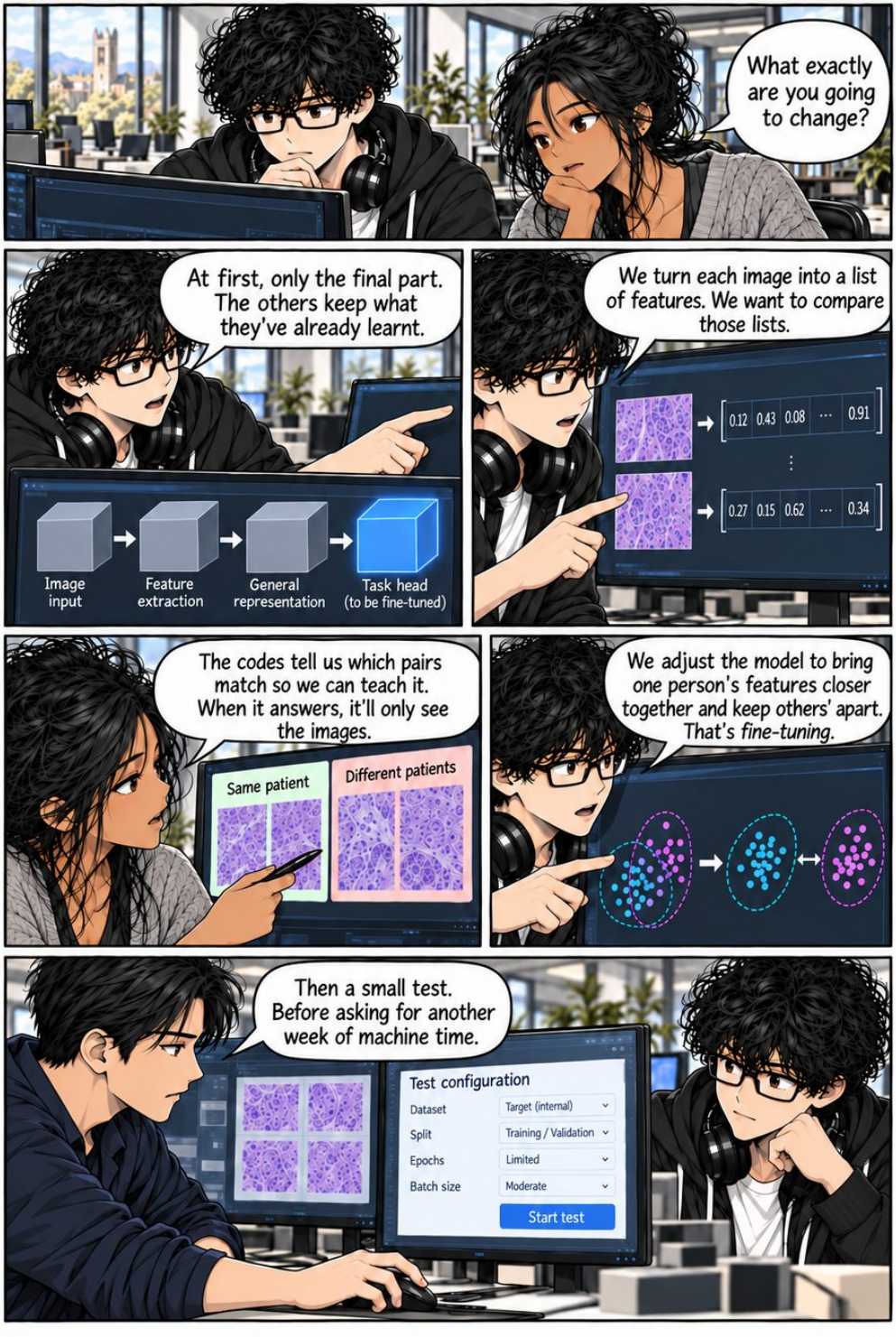

Monday, 28 September · ATLAS
The cases I left out are here. I was too hasty with two of them.
Let's look at them.
Lowering the minimum finds another match. But it also accepts this one, which doesn't belong.
It doesn't improve the overall result. We'll keep the one we had.
We won't close that gap with these adjustments.
Latest Nexo test
ATLAS
0.52
ARGOS
0.74
And the interview?
Good, I think. I haven't heard yet.

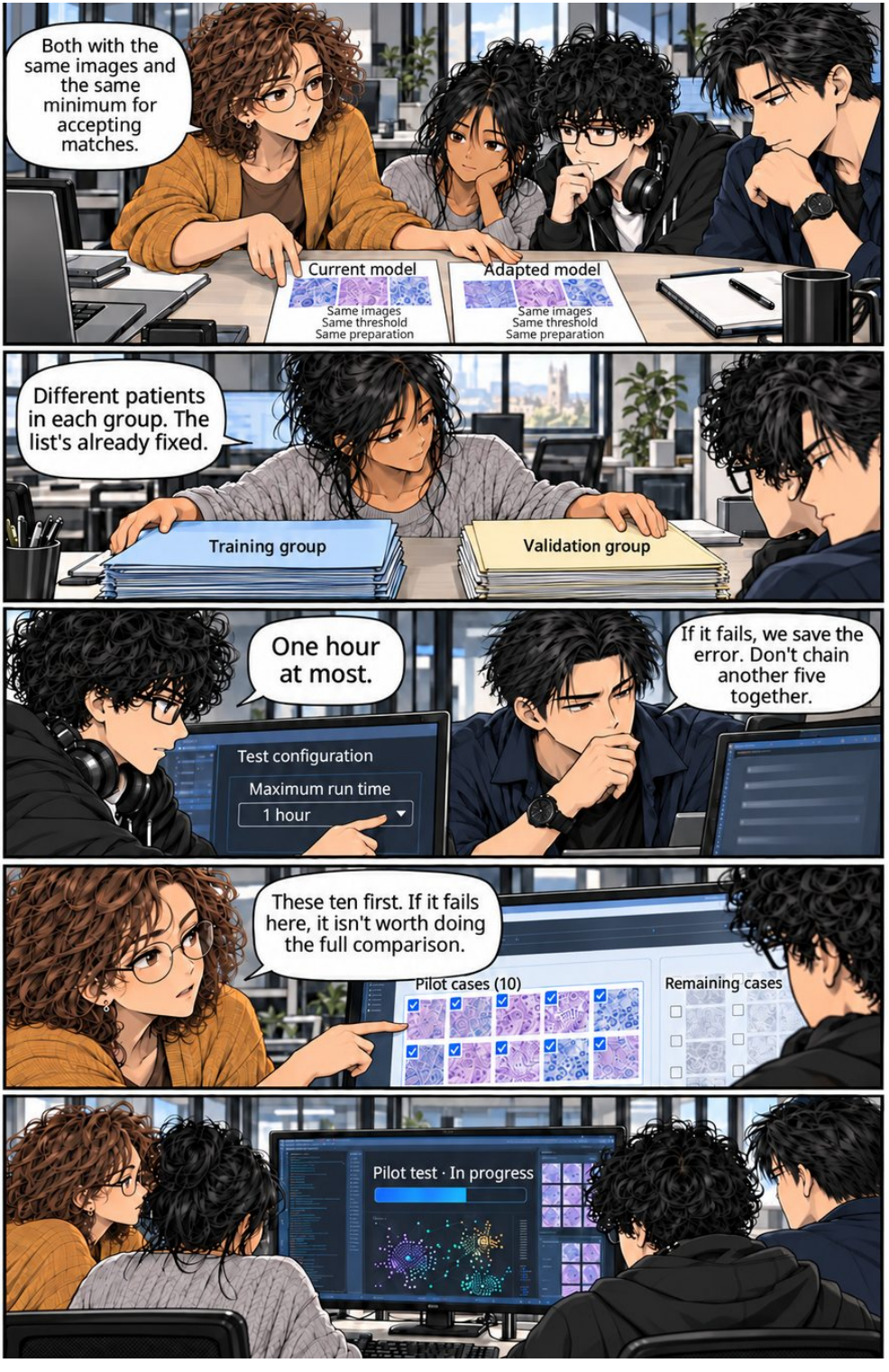

That lunchtime
Enrolment · Two modules selected
I'll take two. If I try all four, I won't manage.
Does that put you back a year?
Yes. And I'll have to pay for another academic year.
But I can't keep doing everything by halves.
Enrolment · Two modules selected
Confirm enrolment
La Segunda Taza
We'll cover Tuesday, as we said.
The other shifts stay. I know.

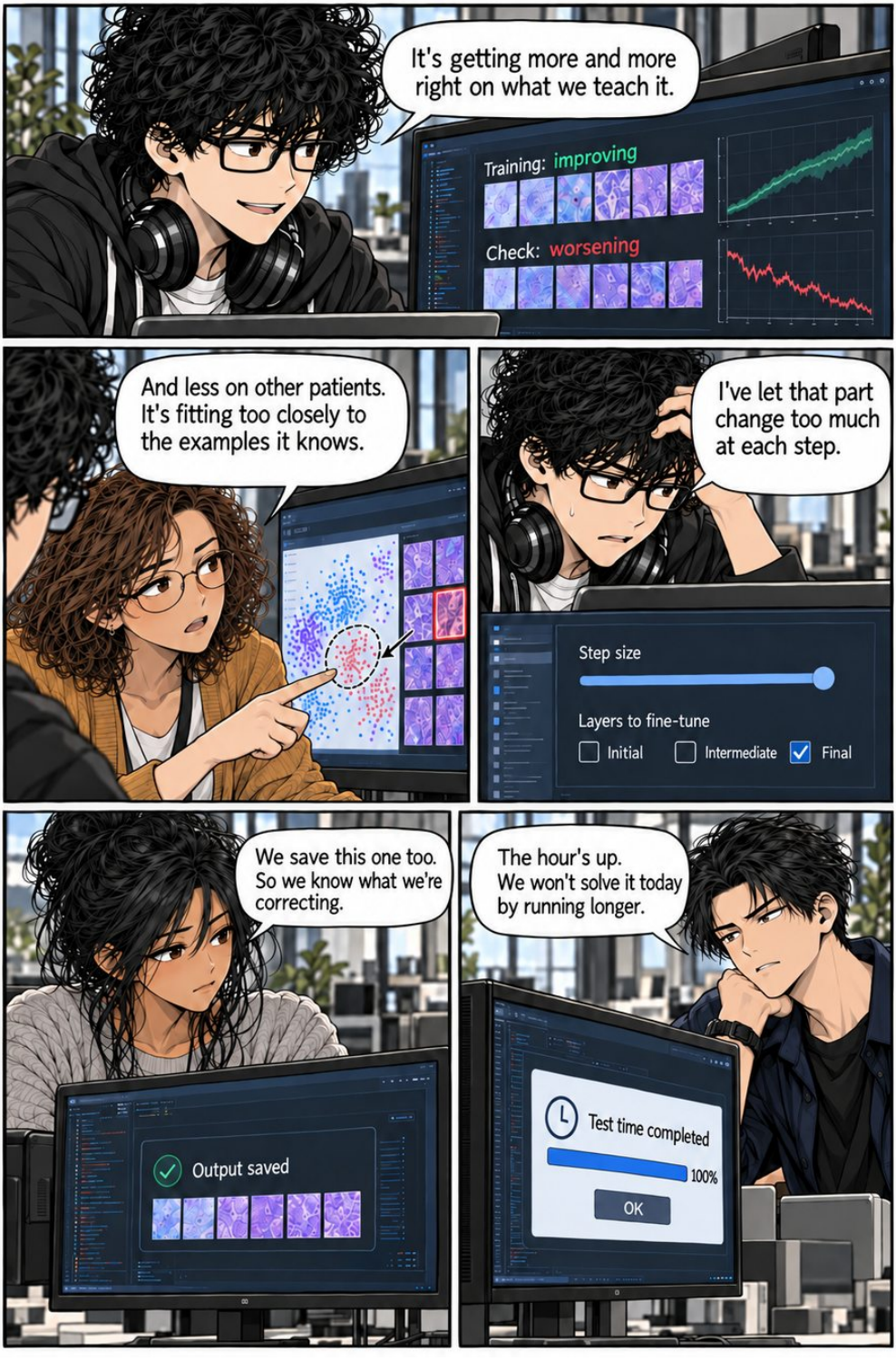

Tuesday, 29 September
The hours we signed for, yes. But I can't keep adding the others.
Then let's only count on those. What won't be covered?
Agreed hours
New hours
I can prepare the data. To review every failure, I'll need someone working with me.
I'll review Wednesday's batch.
And I'll set up the comparison before you arrive.
DIARY
Then let's start here.

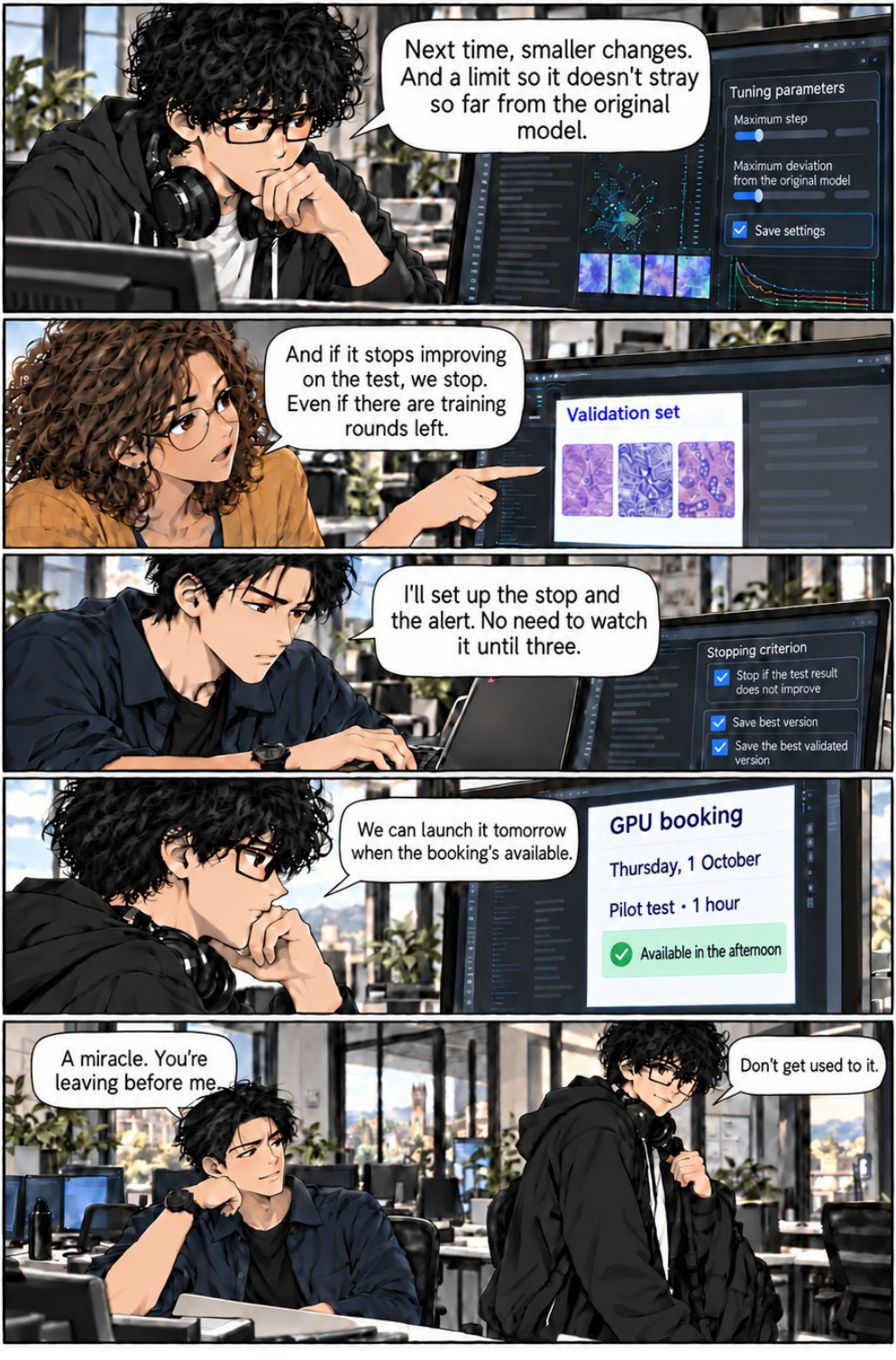

Tuesday · Night
I need the table clear. Again.
I'm looking at a room. They'll tell me tomorrow.
I meant it about the thirtieth. I need my home back.
I know. This one's cheaper. It's miles away.
If you take it, I'll help you with this.
Thanks. Leave that one to me, it's heavy.

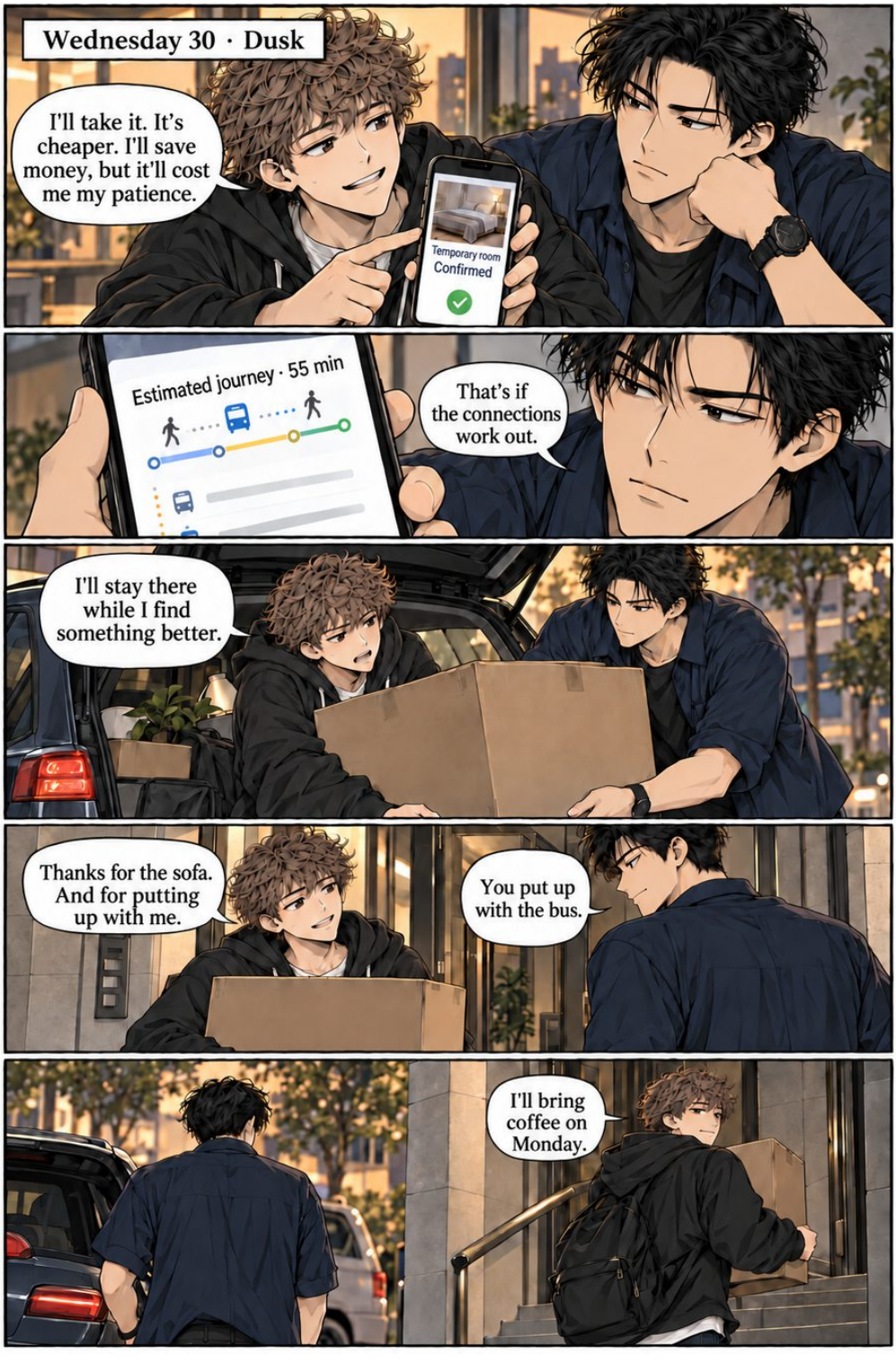

Wednesday, 30 September · ATLAS
In the competition, it helped to start with a model that had already learnt to recognise shapes.
The one that tells broken parts apart?
Yes. But I'm not proposing we bring that one over as it is.
This doesn't look much like a nut.
The idea could help: start with a general one and adapt it here, with the data we're allowed to use.
General model
Recognise parts
Analyse tissue
What worked for you there gives us a clue. It has to prove itself again here.
Right. A clue.

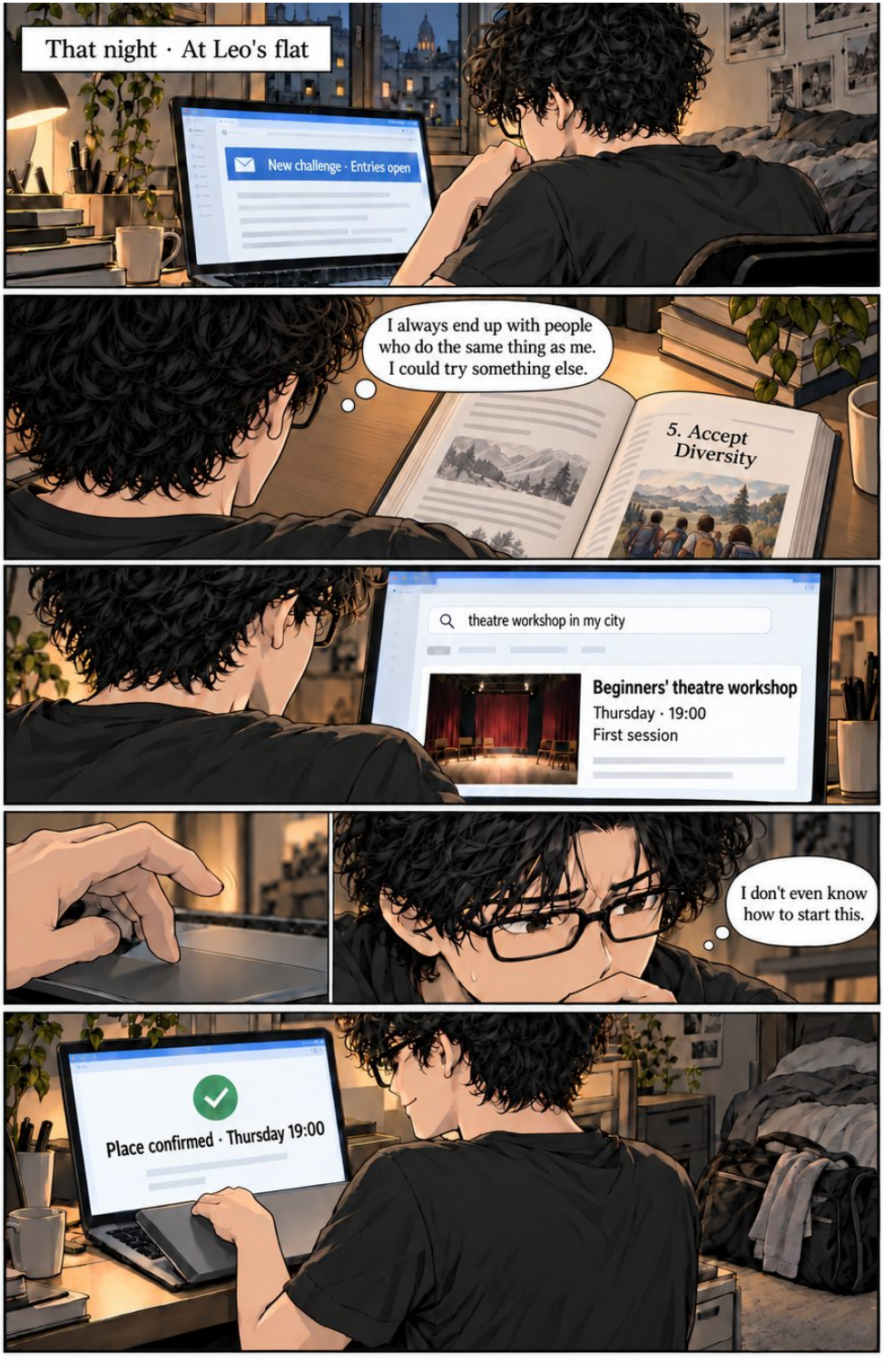

What exactly are you going to change?
At first, only the final part. The others keep what they've already learnt.
Image input
Feature extraction
General representation
Task head (to be fine-tuned)
We turn each image into a list of features. We want to compare those lists.
0.12 0.43 0.08 ··· 0.91
0.27 0.15 0.62 ··· 0.34
The codes tell us which pairs match so we can teach it. When it answers, it'll only see the images.
Same patient
Different patients
We adjust the model to bring one person's features closer together and keep others' apart. That's fine-tuning.
Then a small test. Before asking for another week of machine time.
Test configuration
Dataset
Target (internal)
Split
Training / Validation
Epochs
Limited
Batch size
Moderate
Start test


Both with the same images and the same minimum for accepting matches.
Current model
Same images
Same threshold
Same preparation
Adapted model
Same images
Same threshold
Same preparation
Different patients in each group. The list's already fixed.
Training group
Validation group
One hour at most.
Test configuration
Maximum run time
1 hour
If it fails, we save the error. Don't chain another five together.
These ten first. If it fails here, it isn't worth doing the full comparison.
Pilot cases (10)
Remaining cases
Pilot test · In progress


It's getting more and more right on what we teach it.
Training: improving
Check: worsening
And less on other patients. It's fitting too closely to the examples it knows.
I've let that part change too much at each step.
Step size
Layers to fine-tune
Initial
Intermediate
Final
We save this one too. So we know what we're correcting.
Output saved
The hour's up. We won't solve it today by running longer.
Test time completed
100%
OK


Next time, smaller changes. And a limit so it doesn't stray so far from the original model.
Tuning parameters
Maximum step
Maximum deviation from the original model
Save settings
And if it stops improving on the test, we stop. Even if there are training rounds left.
Validation set
I'll set up the stop and the alert. No need to watch it until three.
Stopping criterion
Stop if the test result does not improve
Save best version
Save the best validated version
We can launch it tomorrow when the booking's available.
GPU booking
Thursday, 1 October
Pilot test · 1 hour
Available in the afternoon
A miracle. You're leaving before me.
Don't get used to it.


Wednesday 30 · Dusk
I'll take it. It's cheaper. I'll save money, but it'll cost me my patience.
Temporary room
Confirmed
Estimated journey · 55 min
That's if the connections work out.
I'll stay there while I find something better.
Thanks for the sofa. And for putting up with me.
You put up with the bus.
I'll bring coffee on Monday.


That night · At Leo's flat
New challenge · Entries open
I always end up with people who do the same thing as me. I could try something else.
5. Accept Diversity
theatre workshop in my city
Beginners' theatre workshop
Thursday · 19:00
First session
I don't even know how to start this.
Place confirmed · Thursday 19:00


Thursday, 1 October · ATLAS
Same ten cases. That way, we can compare with yesterday.
I'm leaving at six. We'll review this pilot tomorrow.
Pilot test · Limit: 1 hour
1 hour
Got plans?
Theatre. A class.
Yes. Don't give me that look.
I haven't said anything.
We'll look at the results tomorrow. Go on, or you'll be late for class.

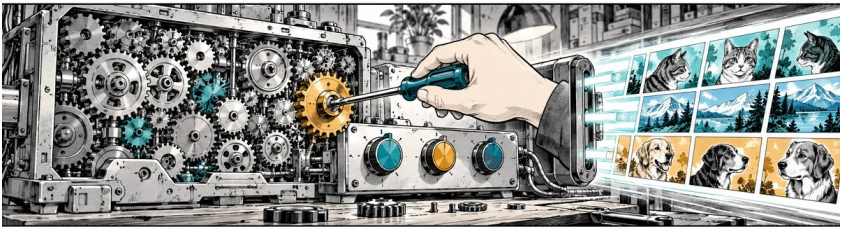

19:00 · Theatre workshop
Leave your bags at the side. We'll work in pairs.
I'm Vera. Is this your first time?
Is it that obvious?
A bit. I started with my bag still on too.
It's in case I need to run for it.
We'll work on this scene. Before touching anyone, we agree how.
Romeo and Juliet
First meeting scene
Romeo and Juliet
First meeting scene


There's a kiss here. Do you want to try it, or shall we mark it without touching?
We can try. But tell me when.
At the end of your line. A short one, then we step apart.
All right.
When you're ready.
Then will you let me come closer?
You blocked me with your shoulder. Let's turn a little.
She's already thinking about the next take.


Romeo's run out of lines.
It's his first day. Didn't you ever go blank?
Just in case, next time I'll bring a prompter.
That would be me. Better not.
Let's try again. Look at me before you start.
Here goes.
I'm preparing a neighbourhood show. I need proper practice.
If you come back, read my lines too.
I'll read the whole thing.


Friday, 2 October · ATLAS
Pilot · 10 cases
Current model: 6 correct
Fine-tuned model: 8 correct
Two more.
Here. Now we still need the full group.
I'll get it ready for Monday. With the same patients we'd kept separate.
Sunday, 4 October
Don't just memorise your own line. Listen to the other one.
The idea from the parts challenge helped us try something else. We still don't know if it'll hold up.
At theatre, I didn't have an answer ready. I want to go back.

Teaching the Model What We Want to Compare



A metaphor for adjusting part of an already trained model.

Mara asks what will change. Leo proposes starting from an authorised general model and adapting the final part. Now he must explain how we move from features useful for recognising objects to comparing samples from the same patient.

From an image to numbers

The network turns each image into a representation, or embedding: a list of numbers summarising features. We can compare that list with those of the hundred candidates and calculate similarities. Then we must decide which to accept, including the possibility of returning none. If the representation keeps differences useful for recognising objects but loses those that distinguish these samples, the comparison will have little to work with.

What the examples teach

Patient codes let us form genuine pairs and pairs from different people. During adjustment, we aim to bring the first representations closer and separate the others. That specifies the relationship we want it to learn. When the program responds, it will receive images without those codes we know for marking it. Leo keeps some general features and adapts the final section to the new task. Testing it will require patients kept out of training.

The first attempt is worse

The run improves on the examples used for learning and gets worse on other patients. Mara keeps that output and Jun respects the maximum time. The failure tells us something about that configuration: they must review how they adjusted it before spending another slot. Winning with the parts may have suggested an interesting route, but here we already have our own images, errors and conditions to examine.

Sources: · Learnt similarity · Adjusting the network

What We Do with the Pilot's Ten Cases



The pilot cases guide adjustment; final evaluation needs held-out patients.

Training goes well and the check gets worse. Leo begins to see overfitting: the model is specialising in familiar examples. Nadia proposes using validation to decide when to stop, and Jun has to turn that rule into a concrete run.

Three decisions worth distinguishing

A lower learning rate reduces the size of each training change. A penalty limiting distance from the initial parameters restricts how far the network adapts. Early stopping watches validation and ends the passes when they no longer improve it. Each decision acts on a part of the process. Jun also sets a maximum duration to protect the slot: reaching the one-hour limit may stop a test even if it was still improving.

What eight against six means

They repeat the same ten cases, and the adjusted model gets eight right against the current model's six. There is a reason to continue. But those cases have already helped correct the recipe, so they belong to development validation. Turning eight out of ten into an overall F1 of 0.80 would even mix the units: counting correctly solved cases may differ from counting correct matches.

The test they leave ready

Nadia asks for the whole group and Mara prepares it for Monday. We do not yet know its result. Before opening it, we should make clear which patients were used for learning, which helped choose adjustments and which were held out to check the chosen version. If we move those roles after seeing each score, we will end up examining ourselves with questions we already used to study.

Sources: · Regularisation · Early stopping · Validation

What a Kiss on Stage Means



Listening and waiting for the next line are part of rehearsal too.

Leo arrives at theatre wanting to meet people and finds himself preparing Romeo and Juliet. He must look at another person, wait for her line and accept that the body is part of the scene too.

Vera is rehearsing

For Vera, preparing the approach and agreeing a kiss are part of the group's work. Leo may feel embarrassment or attraction without that revealing what she feels outside rehearsal. His bewilderment makes sense: he has felt out of place for a long time and any sign of closeness matters a great deal.

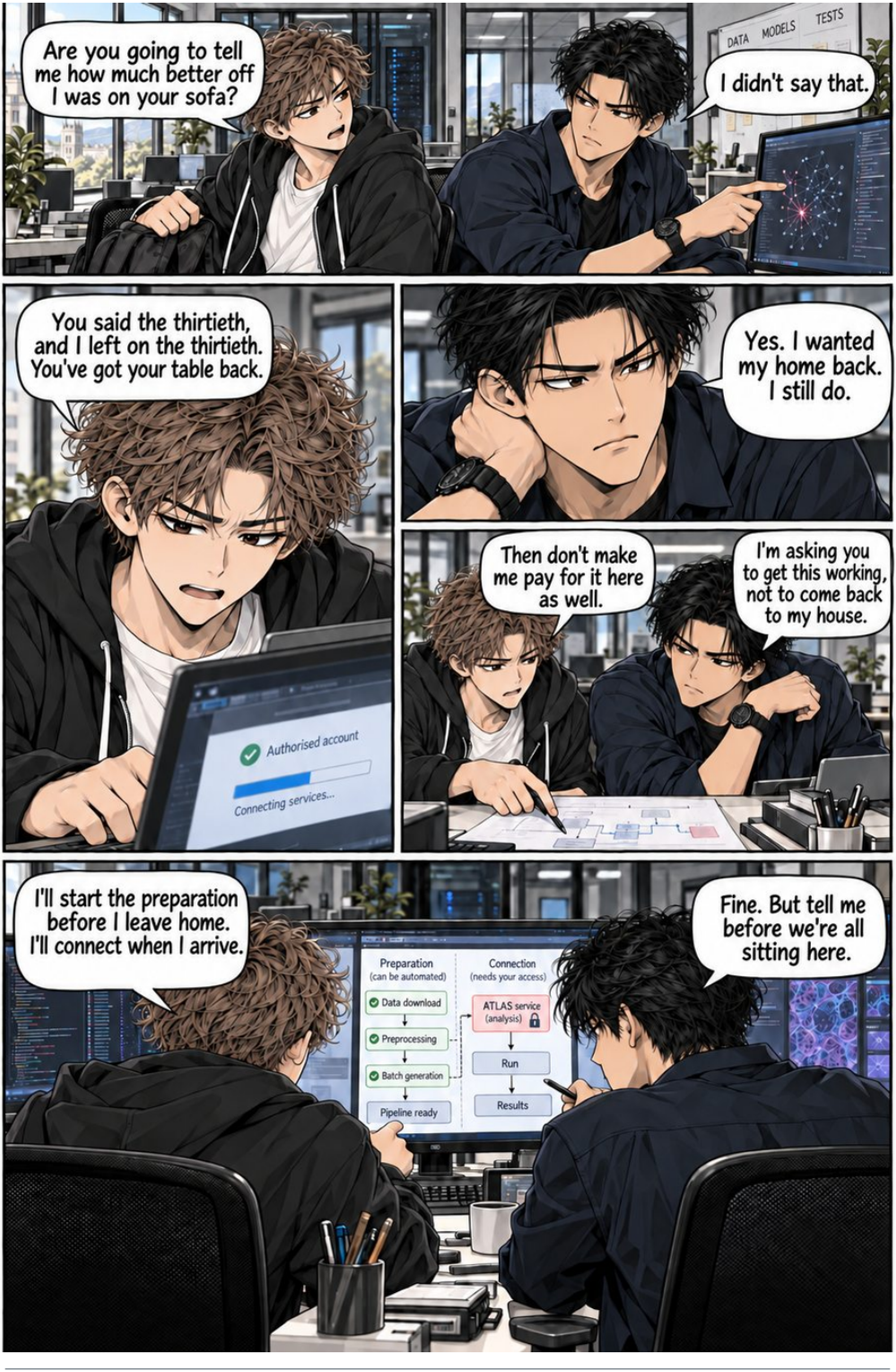
He needs to return prepared

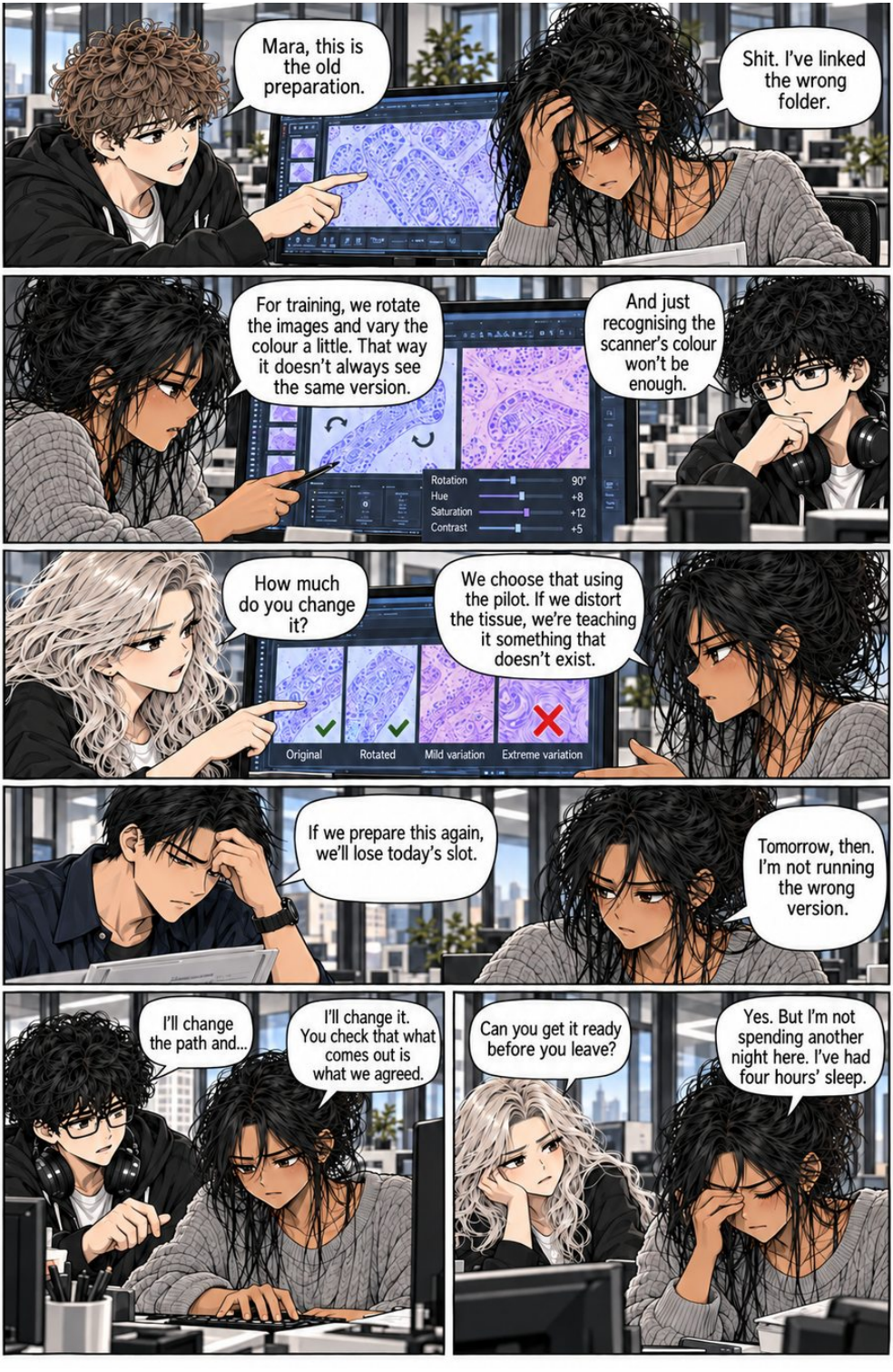
Vera needs a partner who learns the scene and listens. She asks him to read her lines too, and Leo agrees to read the whole thing. Now someone else is counting on his preparation for Thursday. He must fit it between ATLAS's work and the sleep he has already been cutting back.

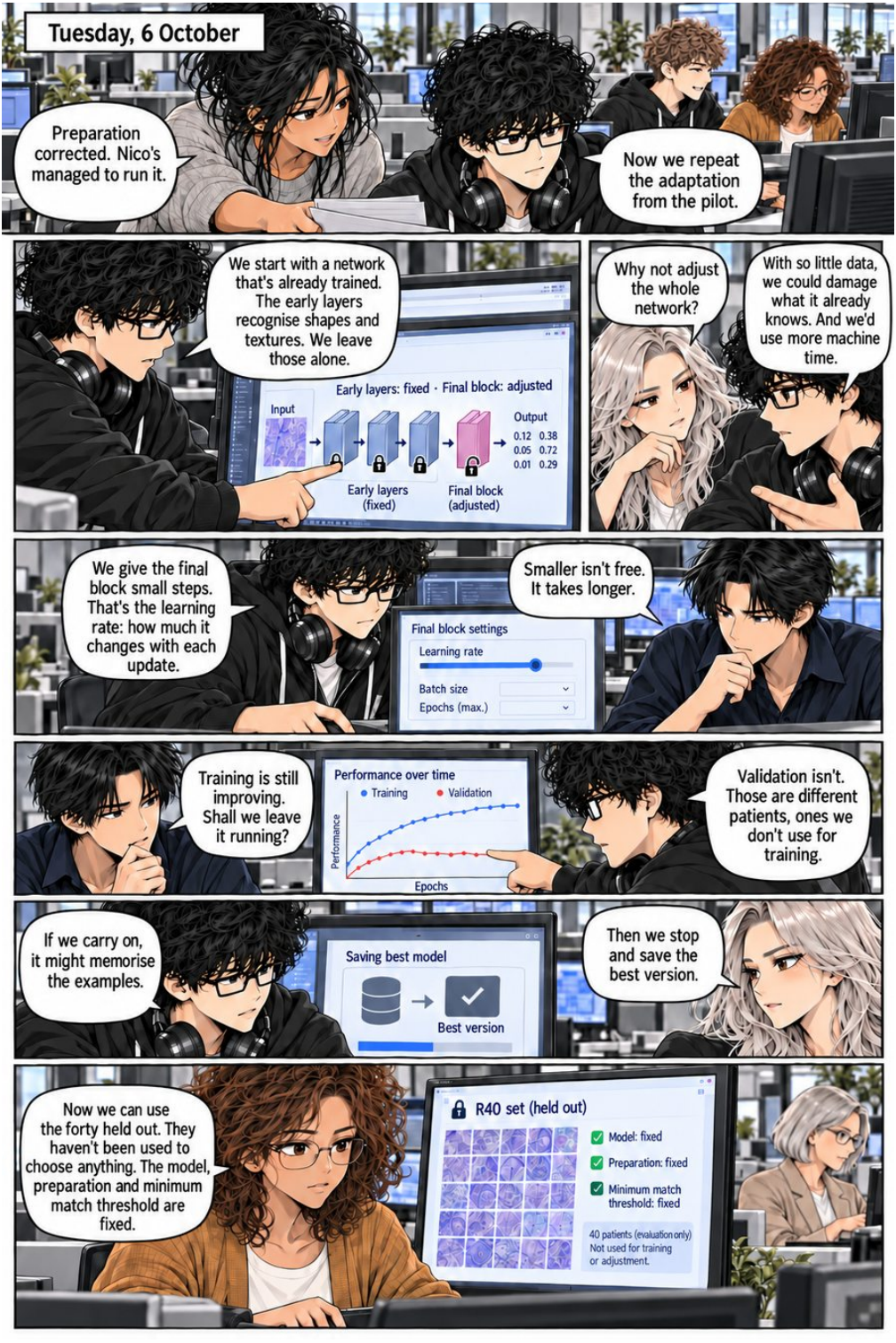
A week with more in it

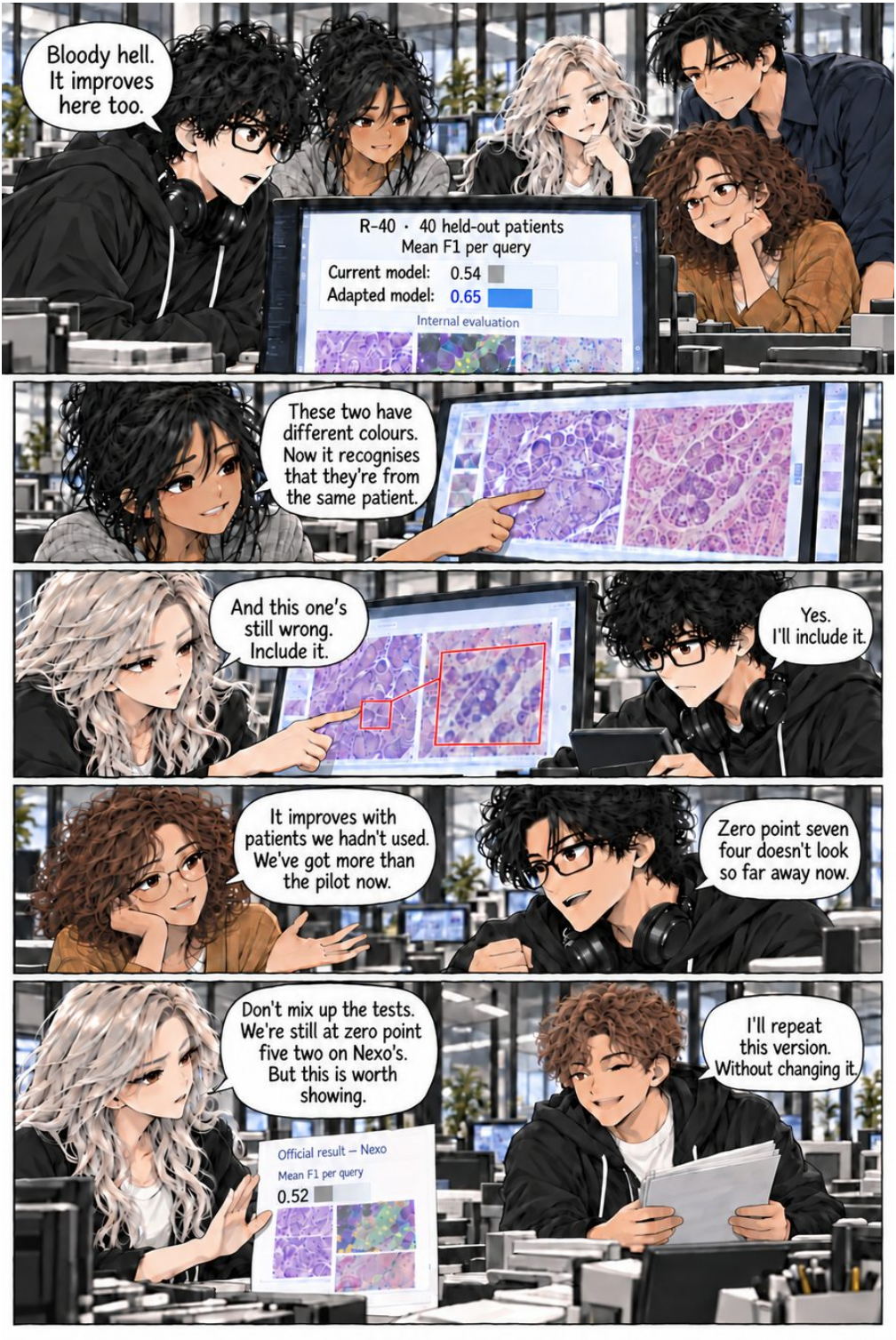
Stopping training lets them stop watching a screen until the early hours. Mara protects her time and Nico accepts living further away. Each is looking for room for something different. Leo starts doing that too, awkwardly and without having sorted out sleep yet.

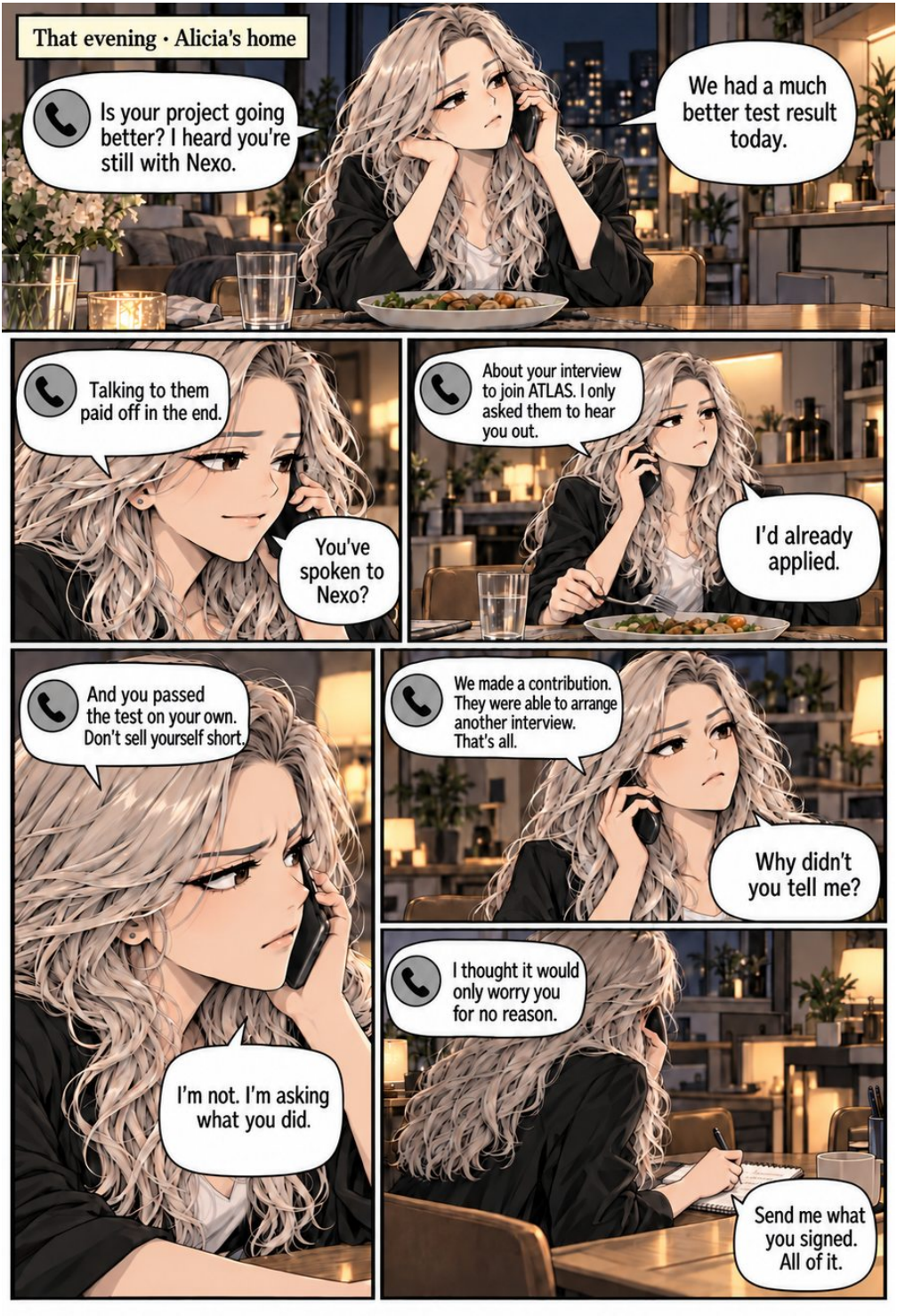

Monday, 5 October
Come on...
We said nine. Have you seen him?
No.
ATLAS
Jun
We're waiting to start.
Now I've missed the other bus.
Sorry. I missed the second bus.
ATLAS
We've been here twenty minutes.
I know. I didn't choose to be late.
Then tell me what I can get ready without you. Waiting doesn't solve anything.

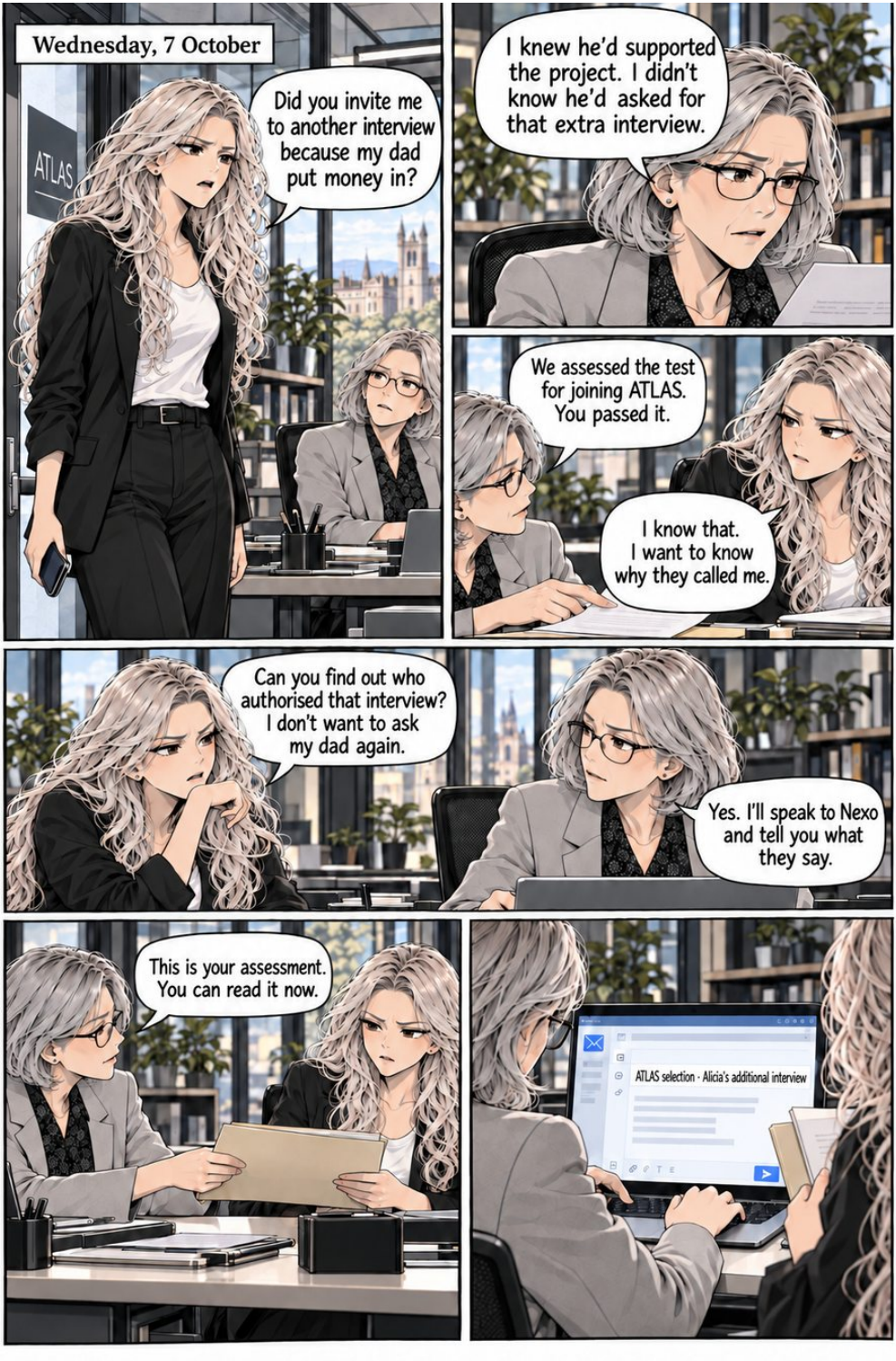

Are you going to tell me how much better off I was on your sofa?
DATA
MODELS
TESTS
I didn't say that.
You said the thirtieth, and I left on the thirtieth. You've got your table back.
Authorised account
Connecting services...
Yes. I wanted my home back. I still do.
Then don't make me pay for it here as well.
I'm asking you to get this working, not to come back to my house.
I'll start the preparation before I leave home. I'll connect when I arrive.
Preparation
(can be automated)
Data download
Preprocessing
Batch generation
Pipeline ready
Connection
(needs your access)
ATLAS service
(analysis)
Run
Results
Fine. But tell me before we're all sitting here.

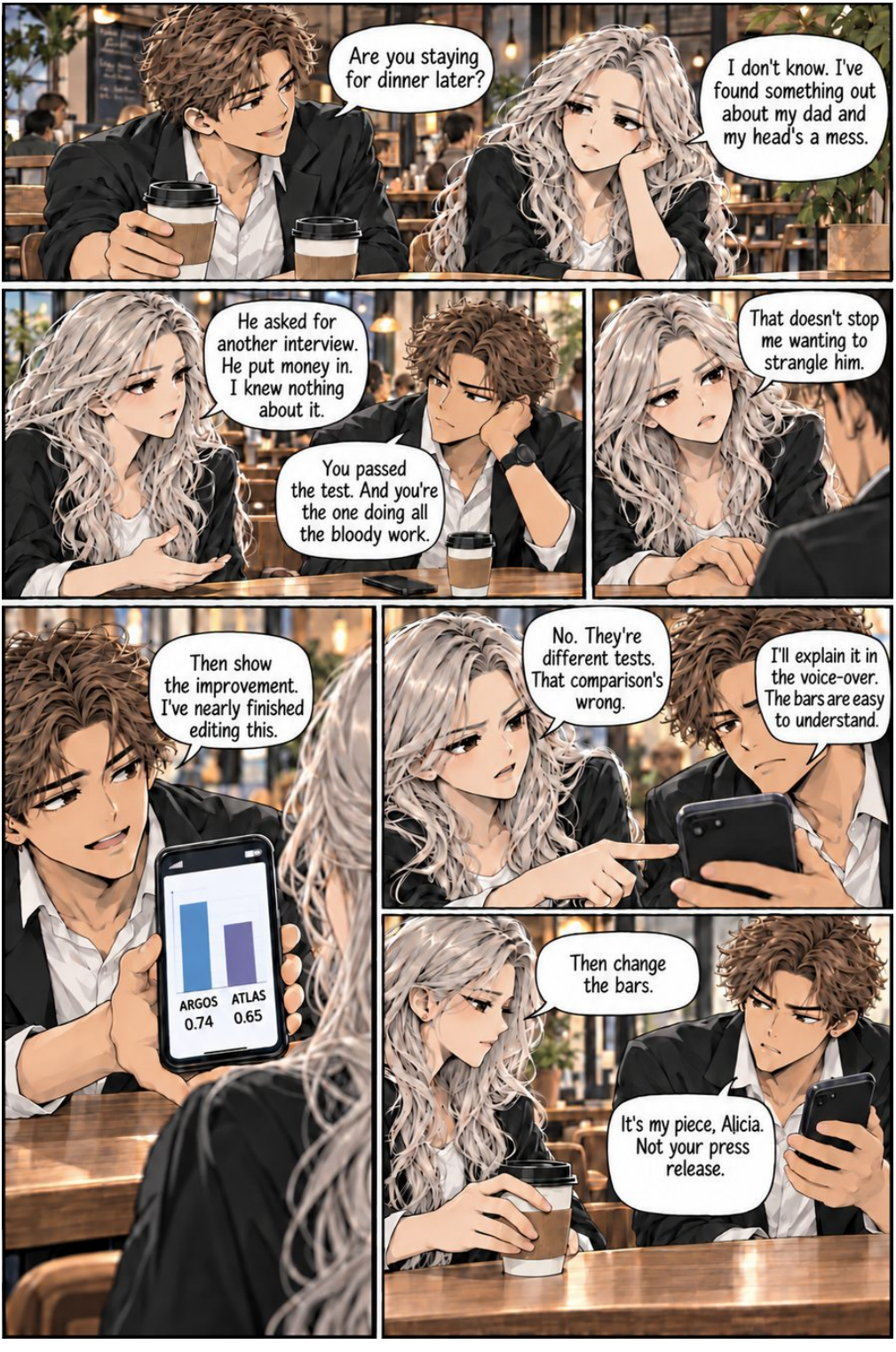

Mara, this is the old preparation.
Shit. I've linked the wrong folder.
For training, we rotate the images and vary the colour a little. That way it doesn't always see the same version.
And just recognising the scanner's colour won't be enough.
Rotation 90°
Hue +8
Saturation +12
Contrast +5
How much do you change it?
We choose that using the pilot. If we distort the tissue, we're teaching it something that doesn't exist.
Original
Rotated
Mild variation
Extreme variation
If we prepare this again, we'll lose today's slot.
Tomorrow, then. I'm not running the wrong version.
I'll change the path and...
I'll change it. You check that what comes out is what we agreed.
Can you get it ready before you leave?
Yes. But I'm not spending another night here. I've had four hours' sleep.

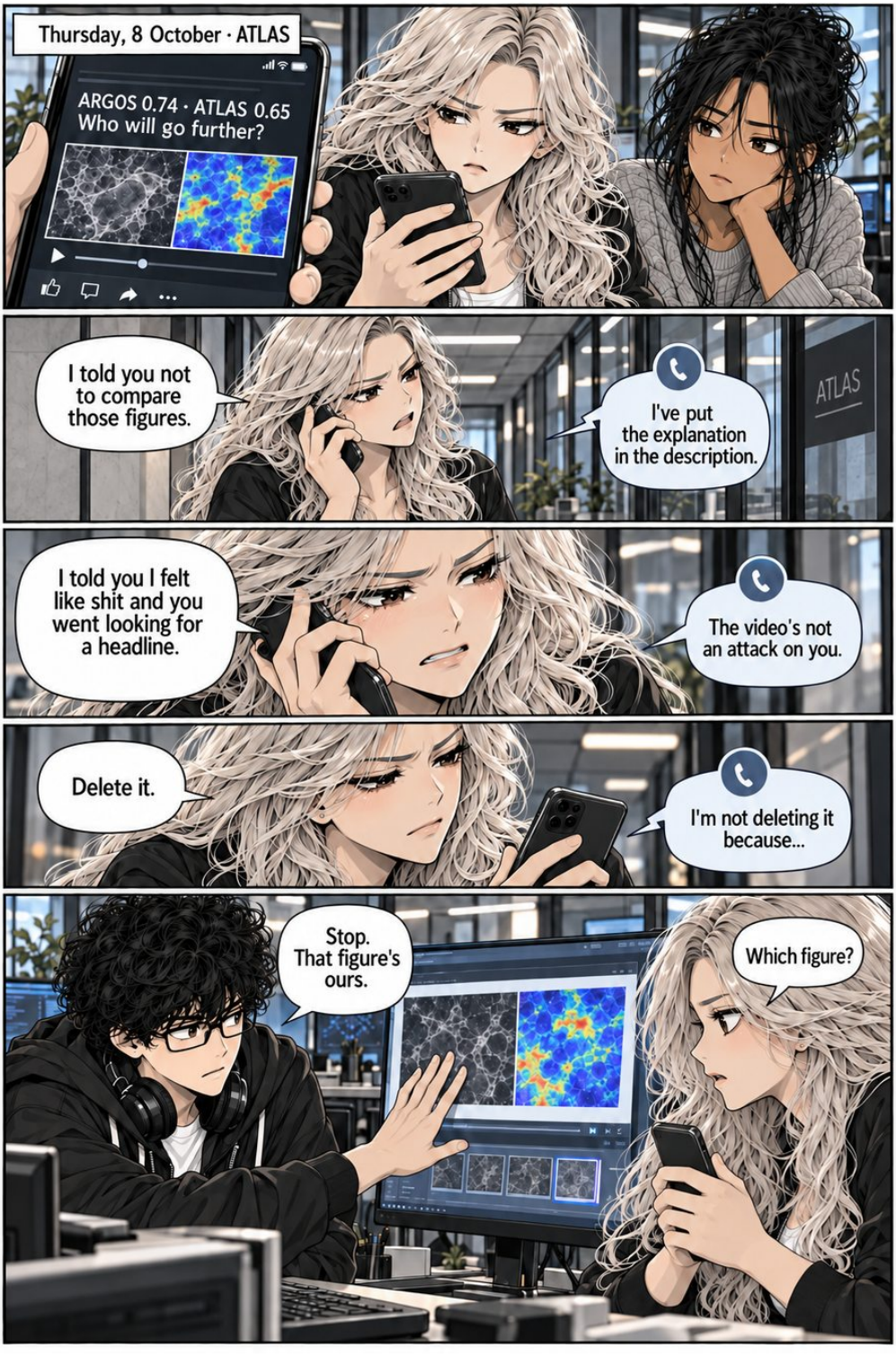

Tuesday, 6 October
Preparation corrected. Nico's managed to run it.
Now we repeat the adaptation from the pilot.
We start with a network that's already trained. The early layers recognise shapes and textures. We leave those alone.
Early layers: fixed · Final block: adjusted
Input
Output
0.12 0.38
0.05 0.72
0.01 0.29
Early layers (fixed)
Final block (adjusted)
Why not adjust the whole network?
With so little data, we could damage what it already knows. And we'd use more machine time.
We give the final block small steps. That's the learning rate: how much it changes with each update.
Smaller isn't free. It takes longer.
Final block settings
Learning rate
Batch size
Epochs (max.)
Training is still improving. Shall we leave it running?
Performance over time
Training
Validation
Performance
Epochs
Validation isn't. Those are different patients, ones we don't use for training.
If we carry on, it might memorise the examples.
Saving best model
Best version
Then we stop and save the best version.
Now we can use the forty held out. They haven't been used to choose anything. The model, preparation and minimum match threshold are fixed.
R40 set (held out)
Model: fixed
Preparation: fixed
Minimum match threshold: fixed
40 patients (evaluation only)
Not used for training or adjustment.

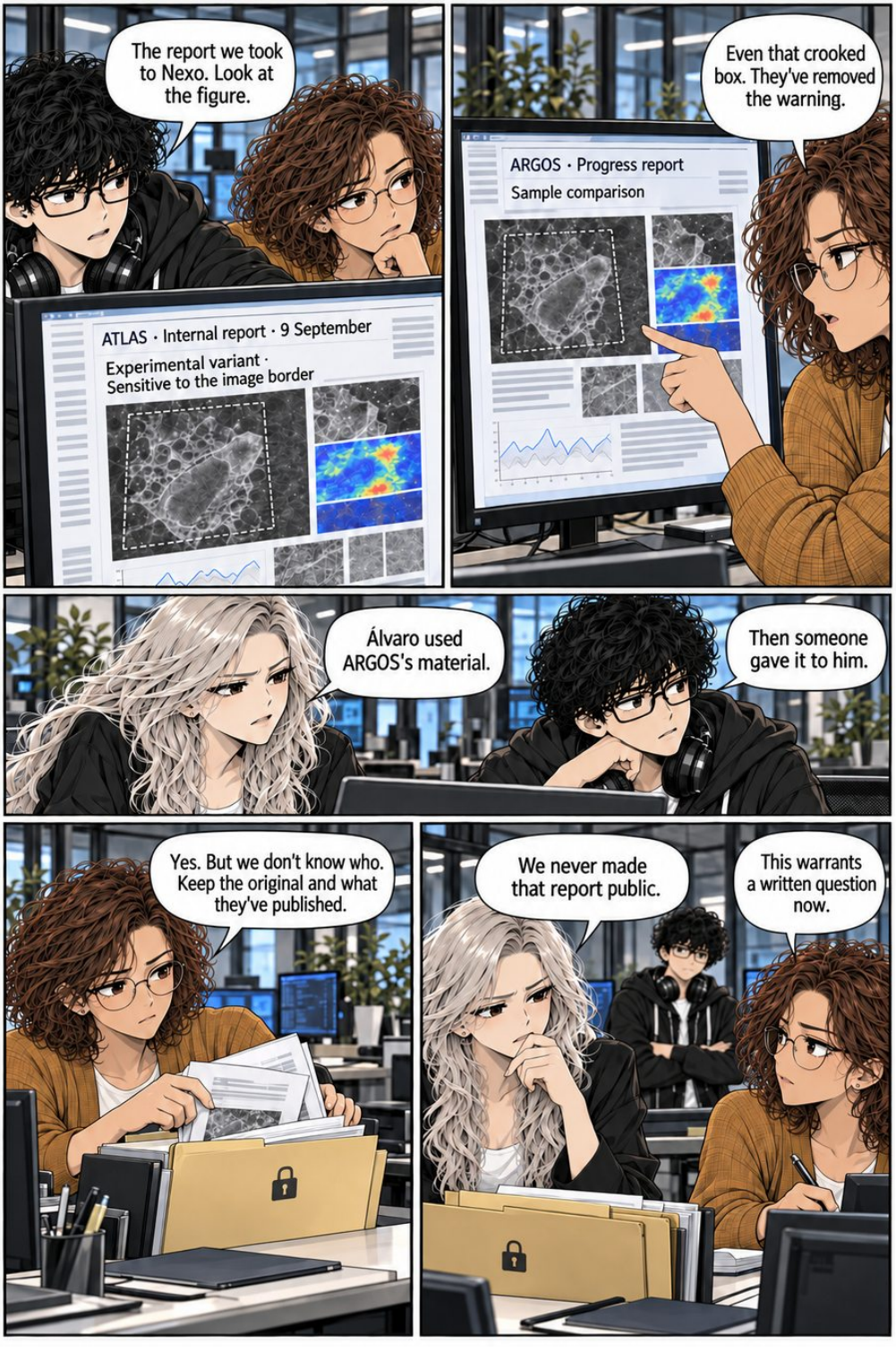

Bloody hell. It improves here too.
R-40 · 40 held-out patients
Mean F1 per query
Current model: 0.54
Adapted model: 0.65
Internal evaluation
These two have different colours. Now it recognises that they're from the same patient.
And this one's still wrong. Include it.
Yes. I'll include it.
It improves with patients we hadn't used. We've got more than the pilot now.
Zero point seven four doesn't look so far away now.
Don't mix up the tests. We're still at zero point five two on Nexo's. But this is worth showing.
Official result — Nexo
Mean F1 per query
0.52
I'll repeat this version. Without changing it.

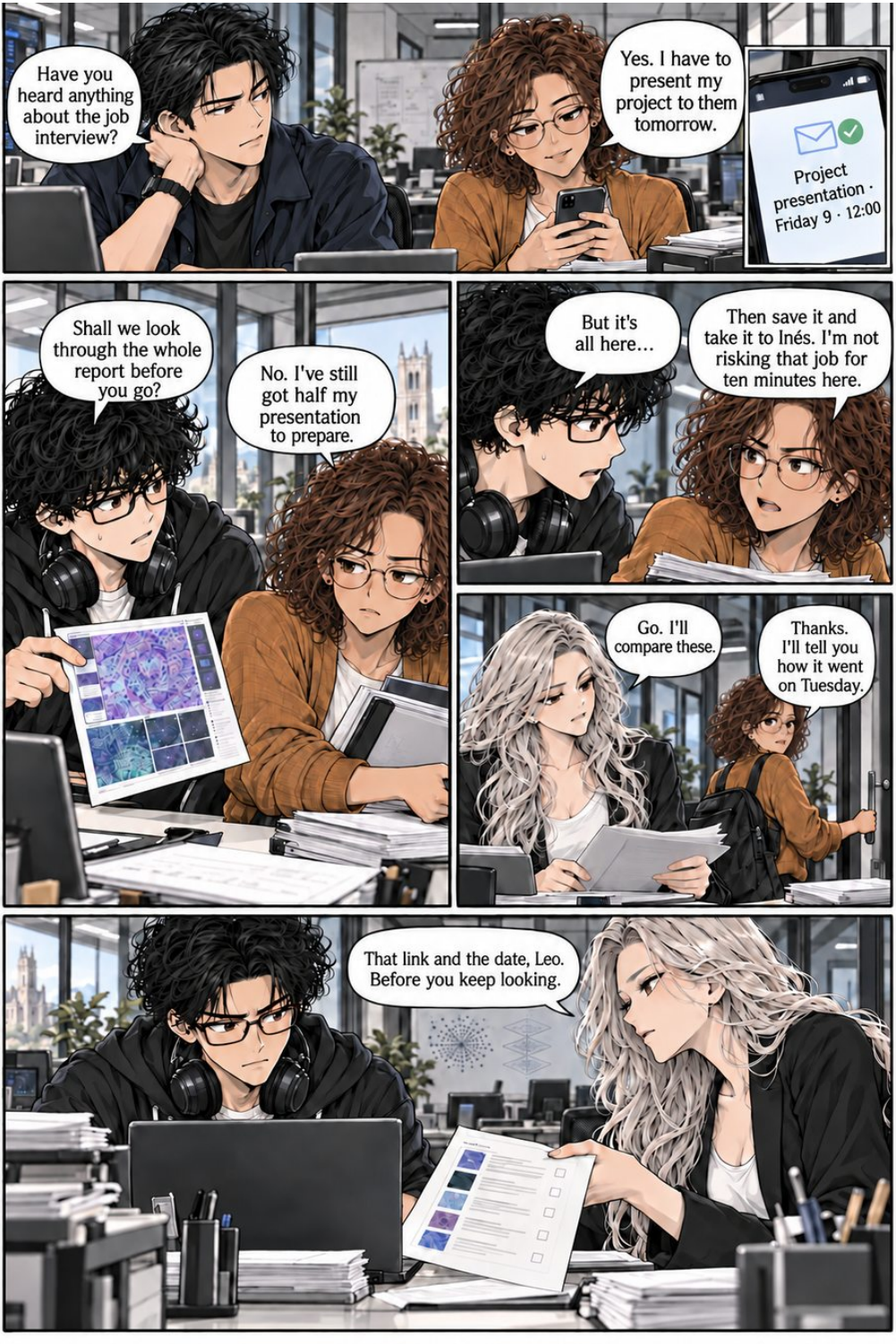

That evening · Alicia's home
Is your project going better? I heard you're still with Nexo.
We had a much better test result today.
Talking to them paid off in the end.
You've spoken to Nexo?
About your interview to join ATLAS. I only asked them to hear you out.
I'd already applied.
And you passed the test on your own. Don't sell yourself short.
I'm not. I'm asking what you did.
We made a contribution. They were able to arrange another interview. That's all.
Why didn't you tell me?
I thought it would only worry you for no reason.
Send me what you signed. All of it.

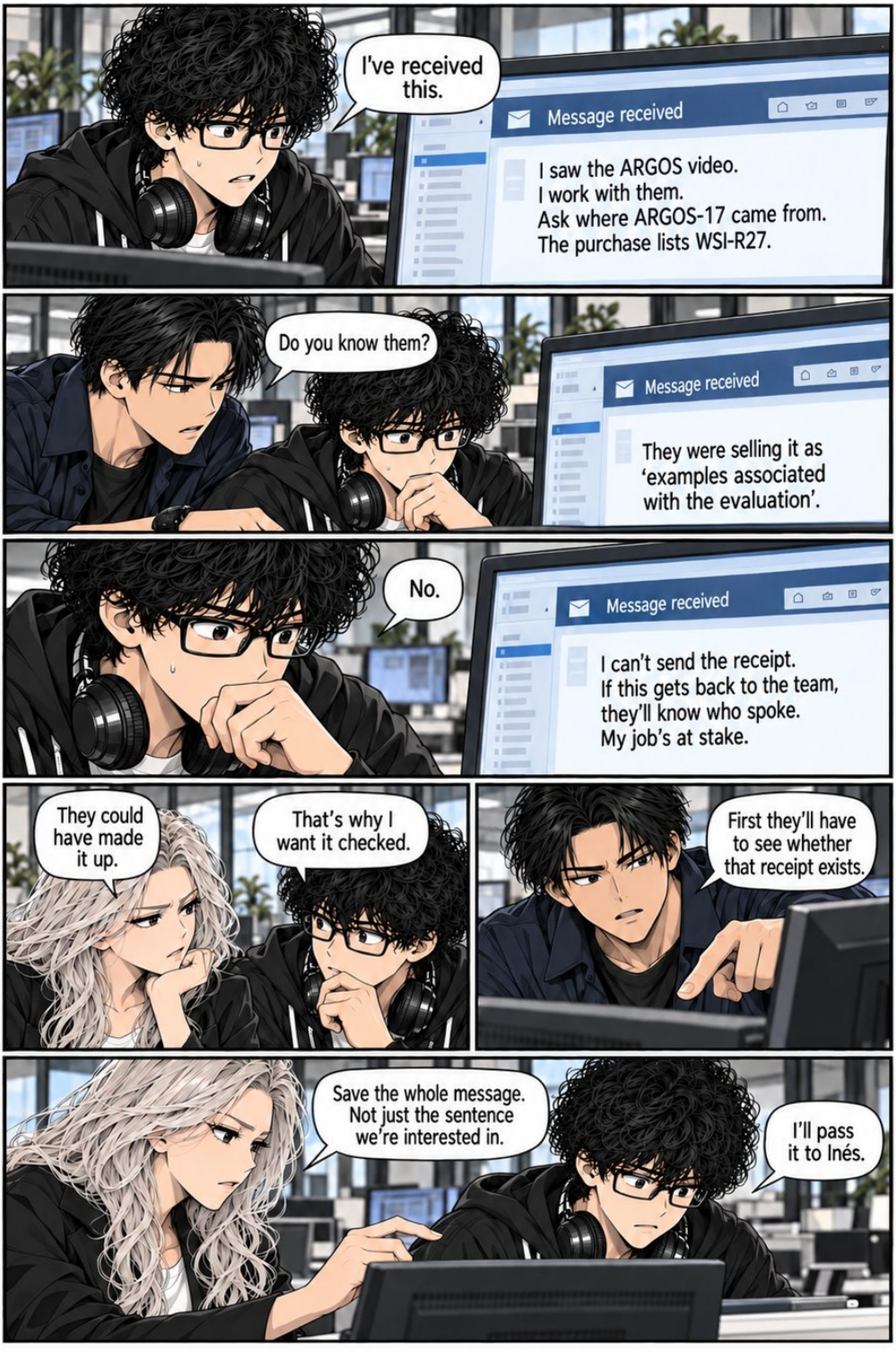

Wednesday, 7 October
ATLAS
Did you invite me to another interview because my dad put money in?
I knew he'd supported the project. I didn't know he'd asked for that extra interview.
We assessed the test for joining ATLAS. You passed it.
I know that. I want to know why they called me.
Can you find out who authorised that interview? I don't want to ask my dad again.
Yes. I'll speak to Nexo and tell you what they say.
This is your assessment. You can read it now.
ATLAS selection · Alicia's additional interview


Are you staying for dinner later?
I don't know. I've found something out about my dad and my head's a mess.
He asked for another interview. He put money in. I knew nothing about it.
You passed the test. And you're the one doing all the bloody work.
That doesn't stop me wanting to strangle him.
Then show the improvement. I've nearly finished editing this.
ARGOS 0.74
ATLAS 0.65
No. They're different tests. That comparison's wrong.
I'll explain it in the voice-over. The bars are easy to understand.
Then change the bars.
It's my piece, Alicia. Not your press release.


Thursday, 8 October · ATLAS
ARGOS 0.74 · ATLAS 0.65
Who will go further?
I told you not to compare those figures.
I've put the explanation in the description.
ATLAS
I told you I felt like shit and you went looking for a headline.
The video's not an attack on you.
Delete it.
I'm not deleting it because...
Stop. That figure's ours.
Which figure?

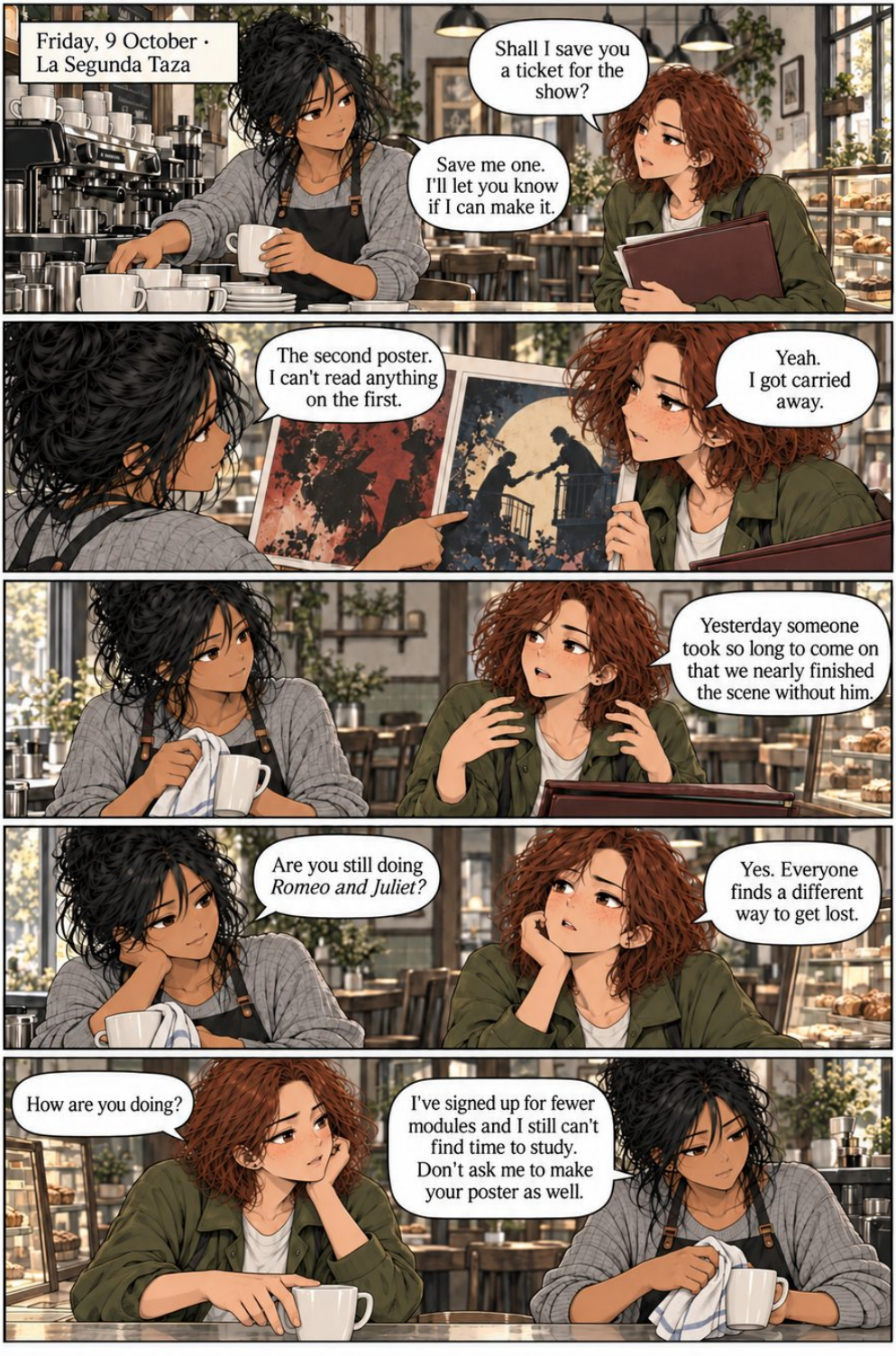

The report we took to Nexo. Look at the figure.
ATLAS · Internal report · 9 September
Experimental variant · Sensitive to the image border
Even that crooked box. They've removed the warning.
ARGOS · Progress report
Sample comparison
Álvaro used ARGOS's material.
Then someone gave it to him.
Yes. But we don't know who. Keep the original and what they've published.
We never made that report public.
This warrants a written question now.


Have you heard anything about the job interview?
Yes. I have to present my project to them tomorrow.
Project presentation · Friday 9 · 12:00
Shall we look through the whole report before you go?
No. I've still got half my presentation to prepare.
But it's all here…
Then save it and take it to Inés. I'm not risking that job for ten minutes here.
Go. I'll compare these.
Thanks. I'll tell you how it went on Tuesday.
That link and the date, Leo. Before you keep looking.

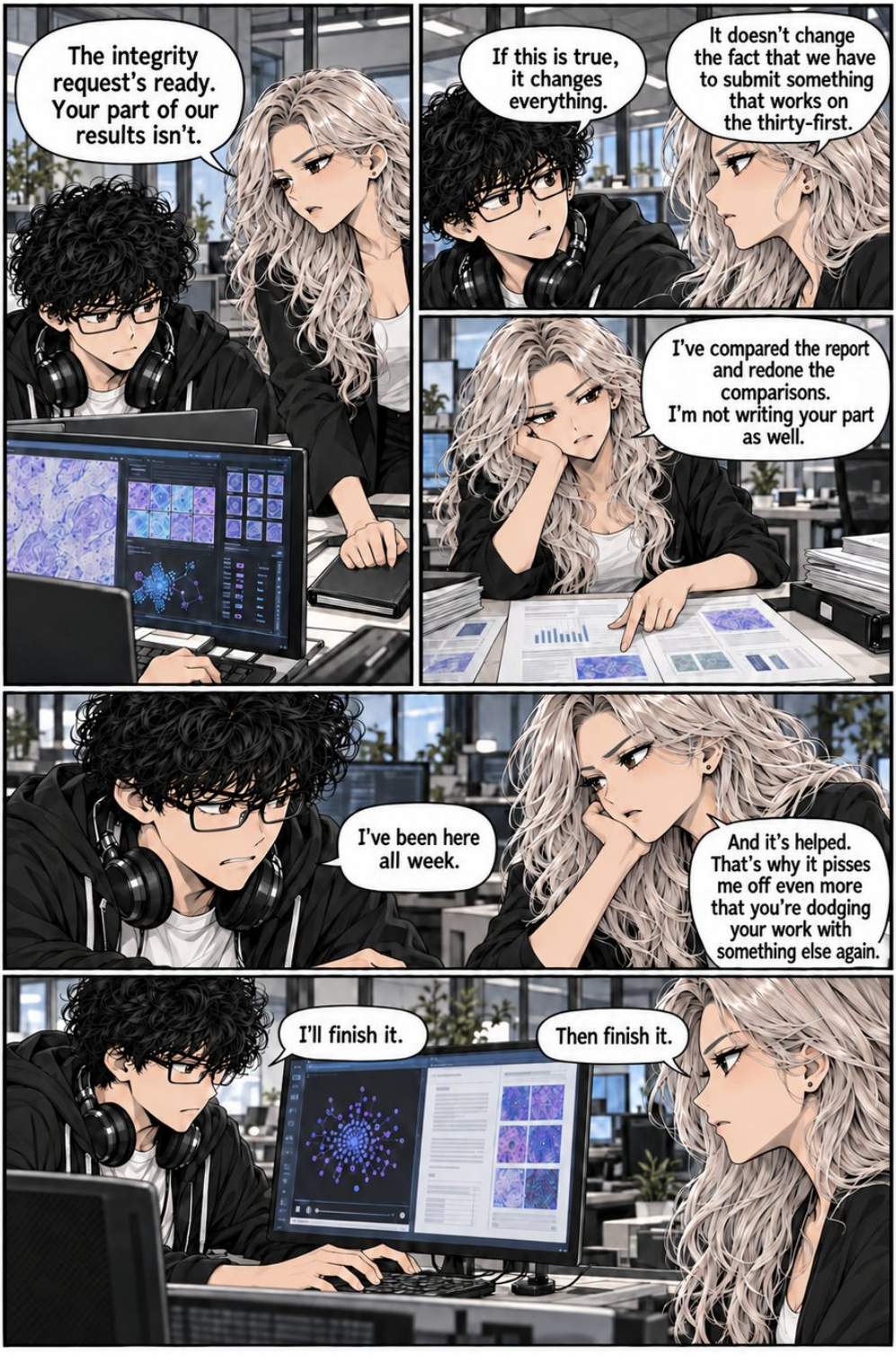

I've received this.
Message received
I saw the ARGOS video.
I work with them.
Ask where ARGOS-17 came from.
The purchase lists WSI-R27.
Do you know them?
Message received
They were selling it as 'examples associated with the evaluation'.
No.
Message received
I can't send the receipt.
If this gets back to the team, they'll know who spoke.
My job's at stake.
They could have made it up.
That's why I want it checked.
First they'll have to see whether that receipt exists.
Save the whole message. Not just the sentence we're interested in.
I'll pass it to Inés.

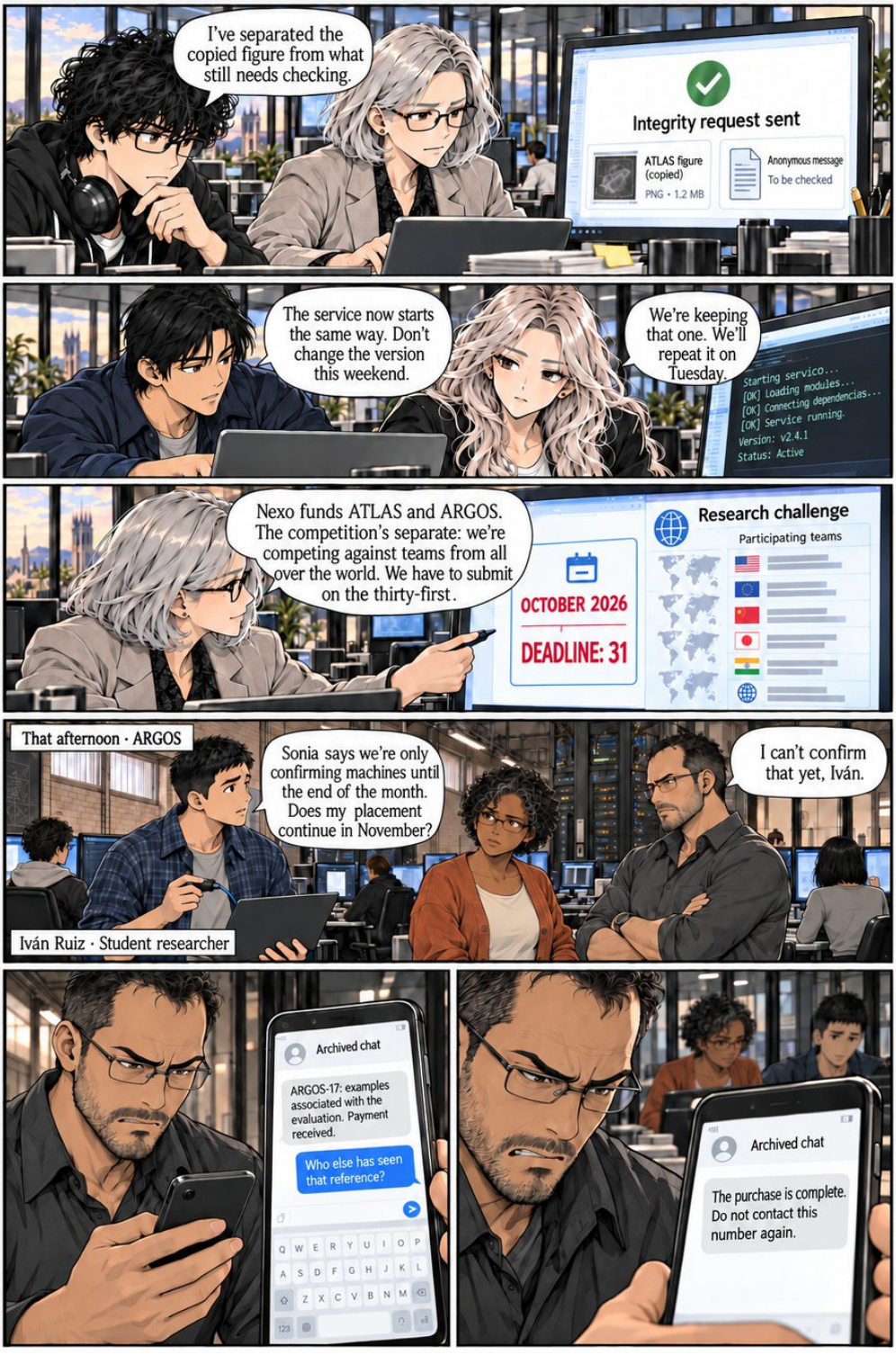

The figure's ours. The message still needs checking. I'll send it to the integrity team tomorrow.
And we carry on as if nothing's happened?
We keep preparing the submission for the thirty-first. And we don't accuse anyone on the strength of an anonymous message.
Then I'll schedule the full repeat for Tuesday.
That Thursday · ARGOS
Héctor, my contract ends in November. Without Nexo's renewal, what do I do about the machine bookings?
Keep this month's.
Sonia Vidal · Infrastructure
I should have removed that ATLAS figure.
Shall I confirm we're going ahead?
Yes. Leave it to me.


Thursday, 8 October · 19:00
From your entrance, Leo.
So, will you let me come closer?
Wait... what was my line?
Only this far.
VERA:
Only this far.
And his cue?
Who's there?
That's it.
Have you learnt everyone's lines?
They stick in my head.
That's brilliant for rehearsing.
Leo, now.
You've got the words. Now look at me.
All right. Again.


Friday, 9 October · La Segunda Taza
Shall I save you a ticket for the show?
Save me one. I'll let you know if I can make it.
The second poster. I can't read anything on the first.
Yeah. I got carried away.
Yesterday someone took so long to come on that we nearly finished the scene without him.
Are you still doing *Romeo and Juliet?*
Yes. Everyone finds a different way to get lost.
How are you doing?
I've signed up for fewer modules and I still can't find time to study. Don't ask me to make your poster as well.


The preparation's here. The version that scored zero point six five, unchanged.
If we finish today, we won't lose Tuesday morning.
I have to catch the bus. The next one's in an hour.
We've spent two days fitting everything round your timetable.
I fitted round your timetable at your place too. Are you always going to throw the sofa back in my face?
I'm leaving too. I'll leave you the steps here, but I'm not staying another night.
I still need to check this. I'm not saying it's finished.
I'll do it. On Tuesday we'll look at how we divide the work.


The integrity request's ready. Your part of our results isn't.
If this is true, it changes everything.
It doesn't change the fact that we have to submit something that works on the thirty-first.
I've compared the report and redone the comparisons. I'm not writing your part as well.
I've been here all week.
And it's helped. That's why it pisses me off even more that you're dodging your work with something else again.
I'll finish it.
Then finish it.

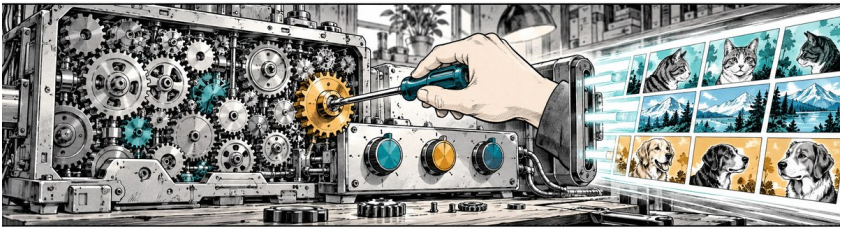

I've separated the copied figure from what still needs checking.
Integrity request sent
ATLAS figure (copied)
PNG · 1.2 MB
Anonymous message
To be checked
The service now starts the same way. Don't change the version this weekend.
We're keeping that one. We'll repeat it on Tuesday.
Starting servico...
[OK] Loading modules...
[OK] Connecting dependencias...
[OK] Service running.
Version: v2.4.1
Status: Active
Nexo funds ATLAS and ARGOS. The competition's separate: we're competing against teams from all over the world. We have to submit on the thirty-first.
OCTOBER 2026
DEADLINE: 31
Research challenge
Participating teams
That afternoon · ARGOS
Sonia says we're only confirming machines until the end of the month. Does my placement continue in November?
I can't confirm that yet, Iván.
Iván Ruiz · Student researcher
Archived chat
ARGOS-17: examples associated with the evaluation. Payment received.
Who else has seen that reference?
Archived chat
The purchase is complete. Do not contact this number again.


Saturday, 10 October
I'm off. We're playing tonight.
Can't you finish this first?
They've been rehearsing without me all week. I'm not being late again.
Jun, soundcheck at seven. It's our gig too.
I'm going. Are you coming?
All right. We'll leave it as it is.
About time! Let's try the opening.
I'm here. From the second bar.
Here he knows what to do with his hands.


Are you staying here all night?
I can't dance.
Neither can I. Come on.
I think I'm out of time.
Keep going. Then there'll be two of us.
We want you at rehearsal on Tuesday. Without the lab laptop.
I can bring the mixing one.
That one, yes.
There'll still be work left tomorrow. I'm glad I came.


Sunday, 11 October
The same video again.
ARGOS 0.74 · ATLAS 0.65
The model improves with new patients. And I'm still waiting for ARGOS to fall.
I'd like to be right. I'd also like Alicia to stop being so pissed off with me.
This still doesn't make sense.
Right. What do we do to the image before we compare it?

What We Change When Adapting the Network



A metaphor for adjusting part of an already trained model.

Before training, Mara spots that they are using old preparation and decides to lose the slot. That seems a good place to start: the model will receive those files, however much the team had another version in mind.

Which part they let learn

A network transforms the image into a representation through operations whose values were learnt from examples. Those values are its parameters. Leo keeps the initial blocks fixed and allows the last one to change. That concentrates the adjustment in a smaller part of the system. He will then need to check whether the features he kept serve this collection; starting from a trained network saves work but leaves that question open.

Steps and examples

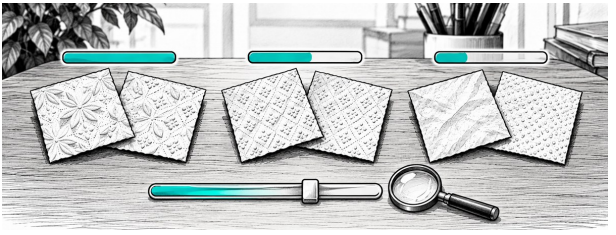
The learning rate controls the size of changes. Lowering it may require more steps; the final time will also depend on the data, machine and stopping rule. Mara makes another decision by preparing rotations and moderate colour variations. She changes the examples the network sees while keeping the information it needs to recognise. If a distortion destroys that signal, adding it to training would teach something else. The transformations still come from the same patients.

When to save the result

The training curve keeps improving while validation levels off. That is the moment to see whether further passes only specialise the network more in familiar material. They save the state chosen using patients different from the training ones. Preparation, adjustable blocks and stopping stay in the same recipe, because they must be kept together to interpret the test that follows.

Sources: · Adjusting the network · Data augmentation

Comparing Scores from the Same Exam



Illustrated textures: similarity scores and an acceptance selector.

ATLAS scores 0.65 on R-40 against its baseline's 0.54. Leo looks at ARGOS's 0.74 and already has a comparison in mind. Alicia stops him: first they must establish which exam each figure comes from.

When they open R-40

They used training and validation to adjust the model, choose the saved state and set the minimum match threshold. R-40's forty patients are held out until that recipe is fixed. That lets them check it on examples that played no part in the choice. If they change it again by looking at that result, R-40 will guide the new version, and they will need to acknowledge that use when evaluating it.

What F1 captures

Imagine the program proposes several matches. Precision asks how many proposals are correct; recall asks how many genuine pairs it recovered. F1 combines both measures, and this test averages the queries' scores. A score of 0.65 does not let us say 65% of patients were identified or assign that probability to each pair. To understand it, we must keep track of what was counted.

What the two bars say

R-40 compares two ATLAS versions with the same patients and procedure. ARGOS's 0.74 comes from another evaluation. Drawing both values side by side, as the video does, conceals that difference without resolving it. Even within R-40, a better mean can accompany cases that worsen. That is why Alicia keeps the visible error: it is part of what they will need to explain too.

Sources: · Validation · F1 metric

Getting to Know Jun Outside the Lab



Jun has friends and commitments outside the lab too.

Leo has spent several days watching ARGOS. He has reasons to investigate, but is also drawn to proving the rival wrong. Meanwhile, Alicia is still waiting for him to finish explaining their own work.

The friend whose band is missing him

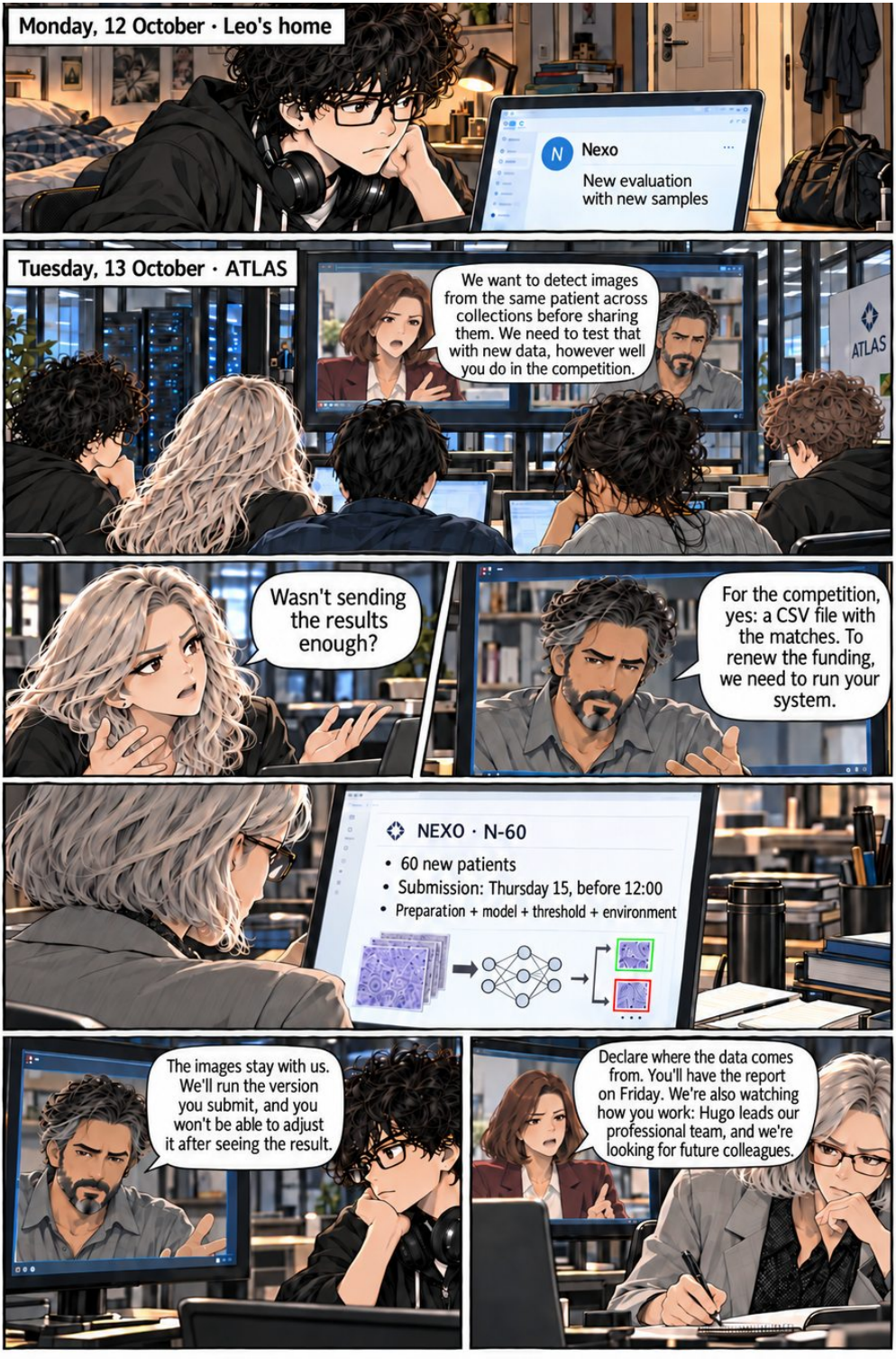
At the gig, Jun has bandmates who know him through music. They have missed him at rehearsals and tell him so. ATLAS may also be taking up too much space in the life of someone who usually seems confident and resourceful.

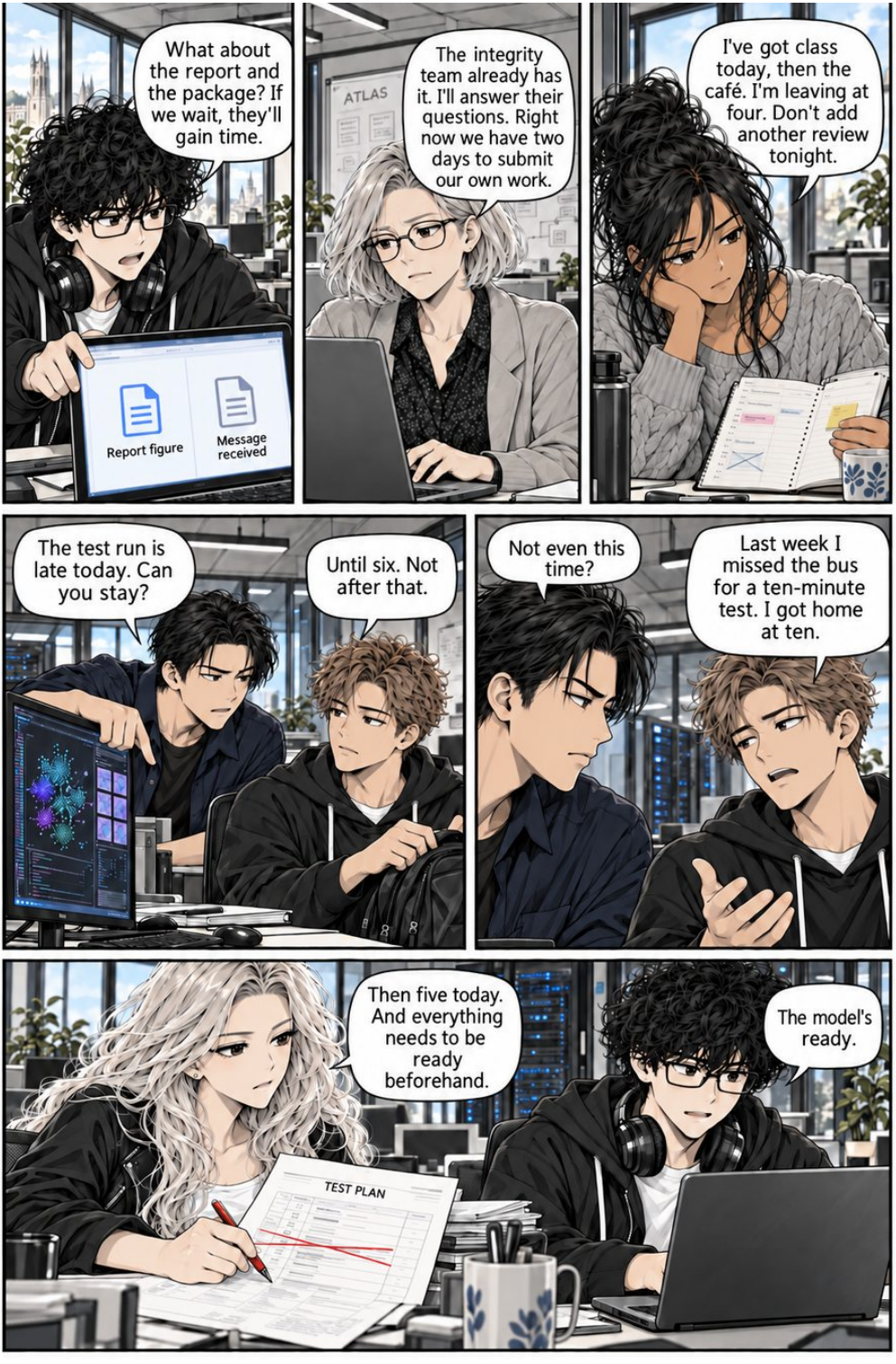
Learning the Lines and Finding His Place

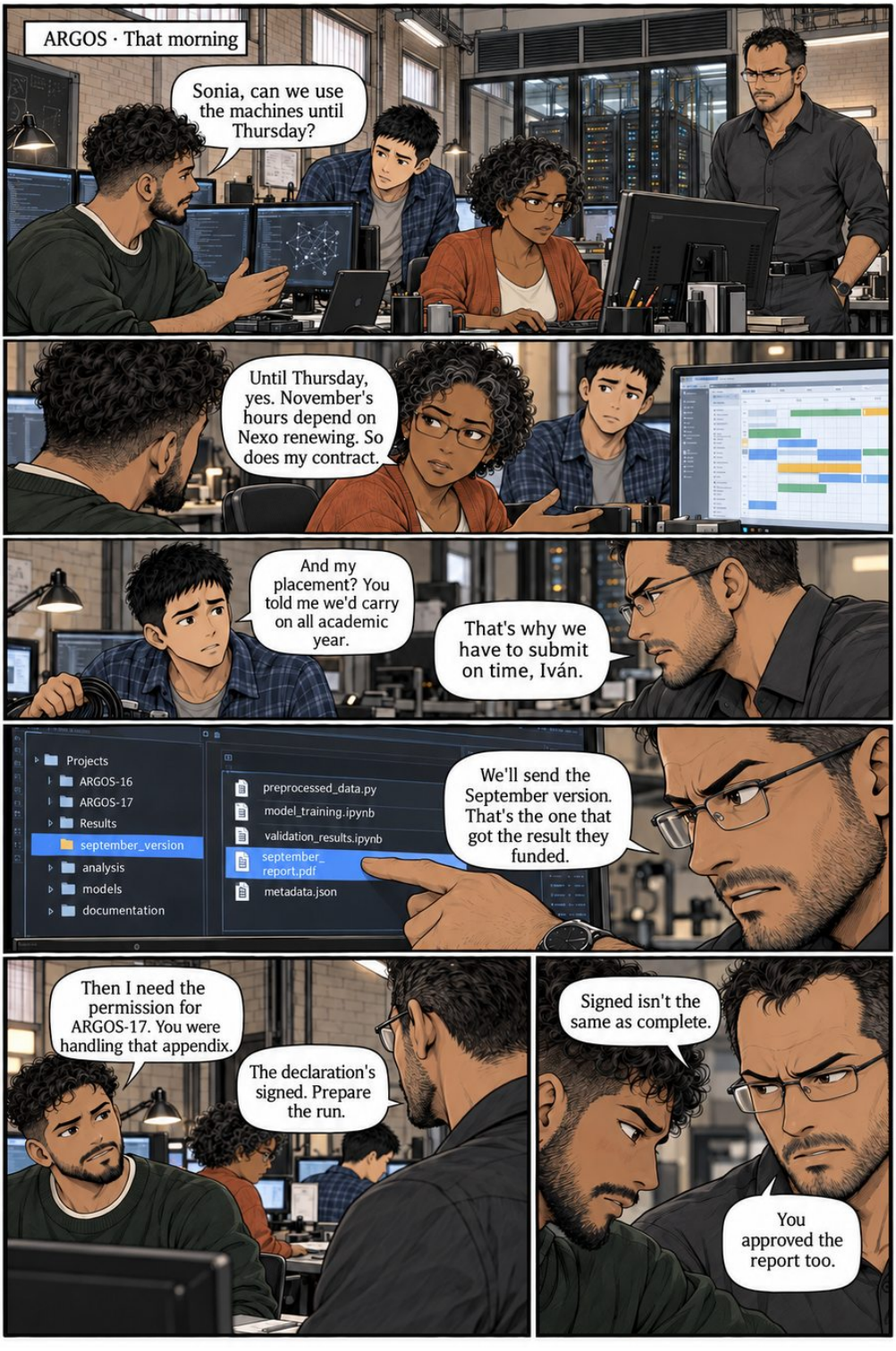
At theatre rehearsal, Leo knows his lines and everyone else's too. That is useful to Vera for practising. When it is time to enter, he crosses his feet and gets caught up watching the floor mark. Knowing what comes next does not tell him when to move or how to look at the person in front of him. At the gig he joins in too: he has had a drink, dances and enjoys himself with the group. That week is offering him things to do besides watching ARGOS's video again.

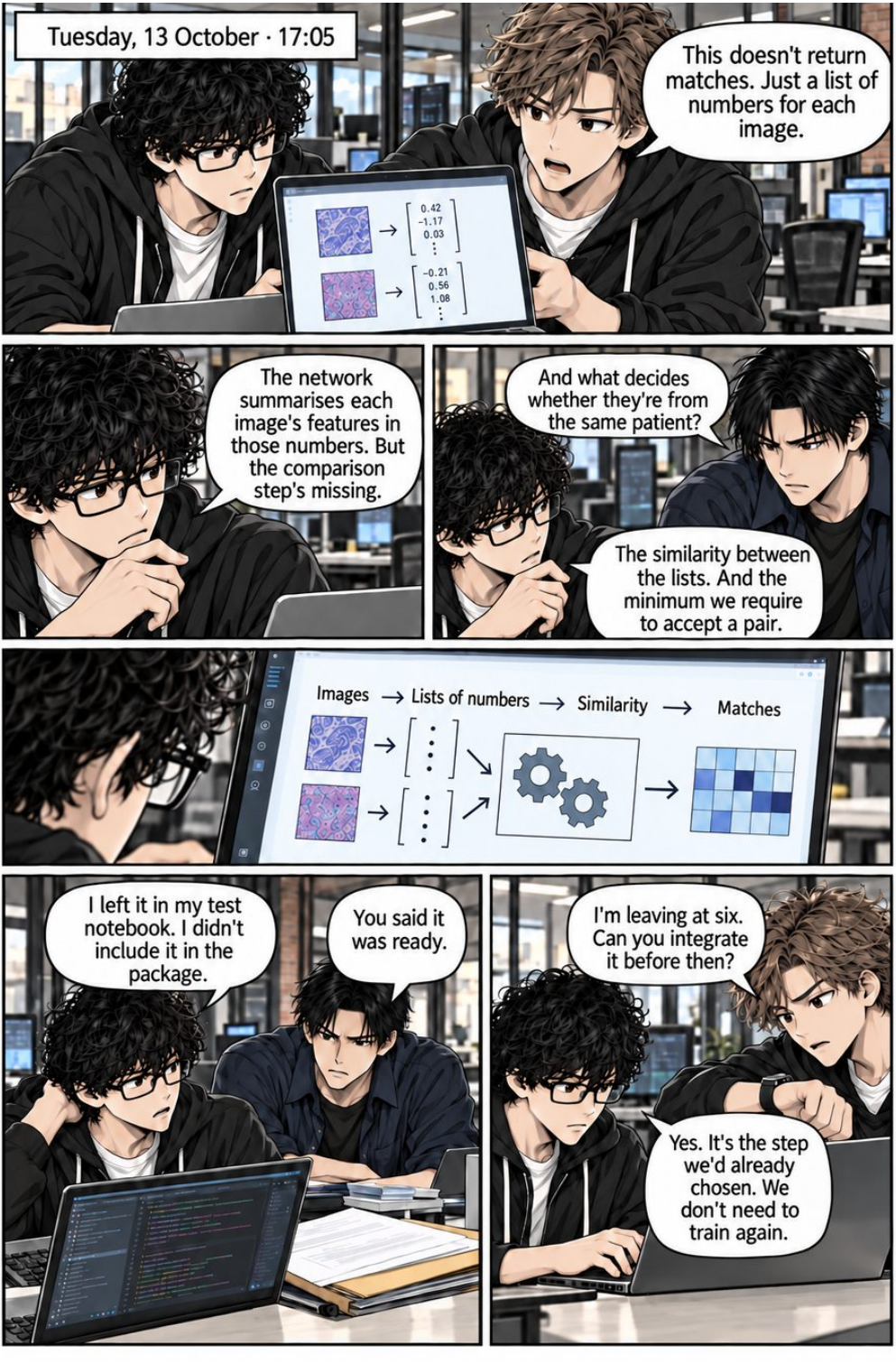
Alicia is not telling everything

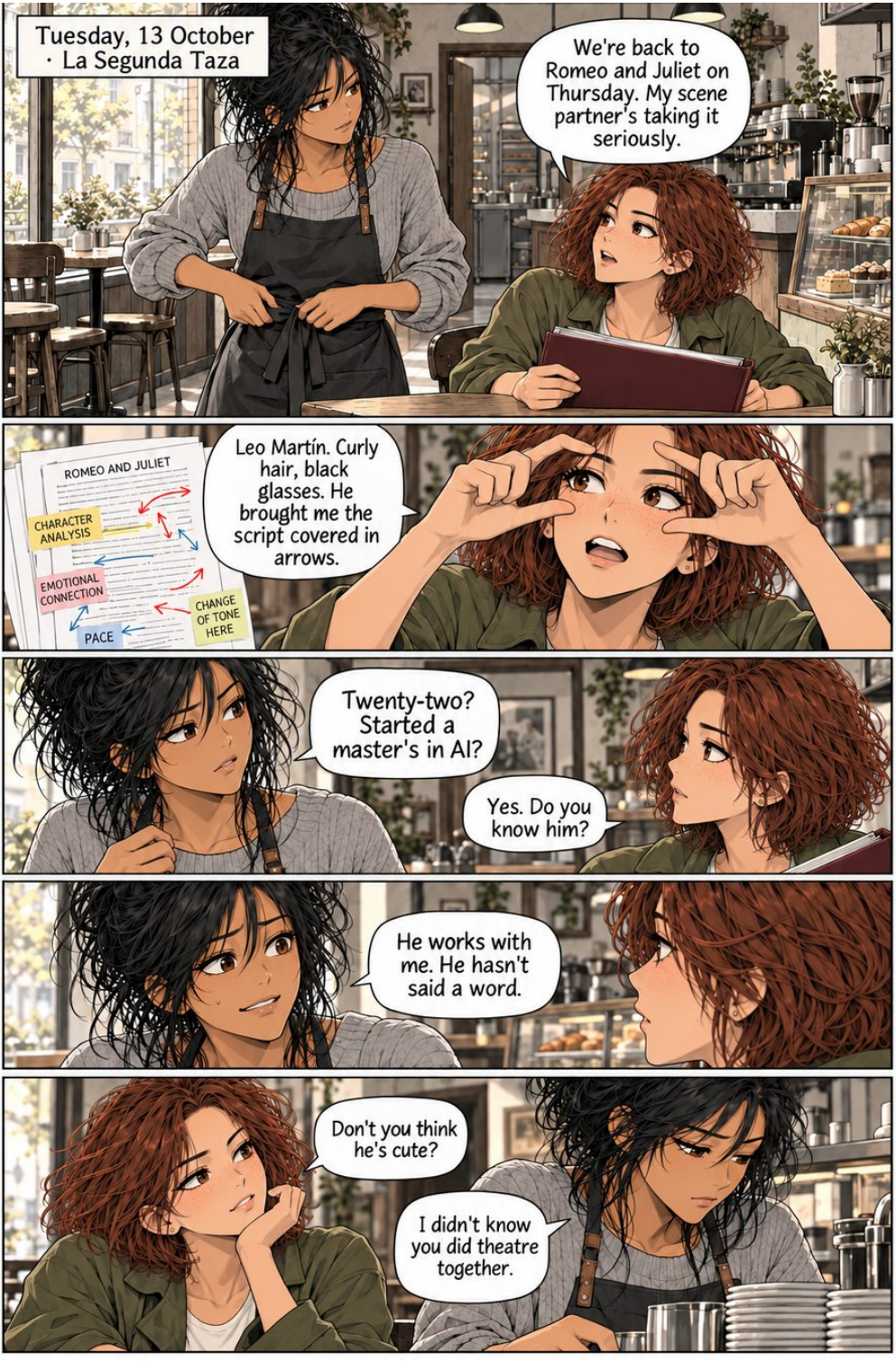
The reader knows about her father's intervention and how Álvaro's published comparison hurts her. Leo does not have that information. He can be there for her without guessing a complete explanation, just as he needs to hear Jun before assuming he will always be available.

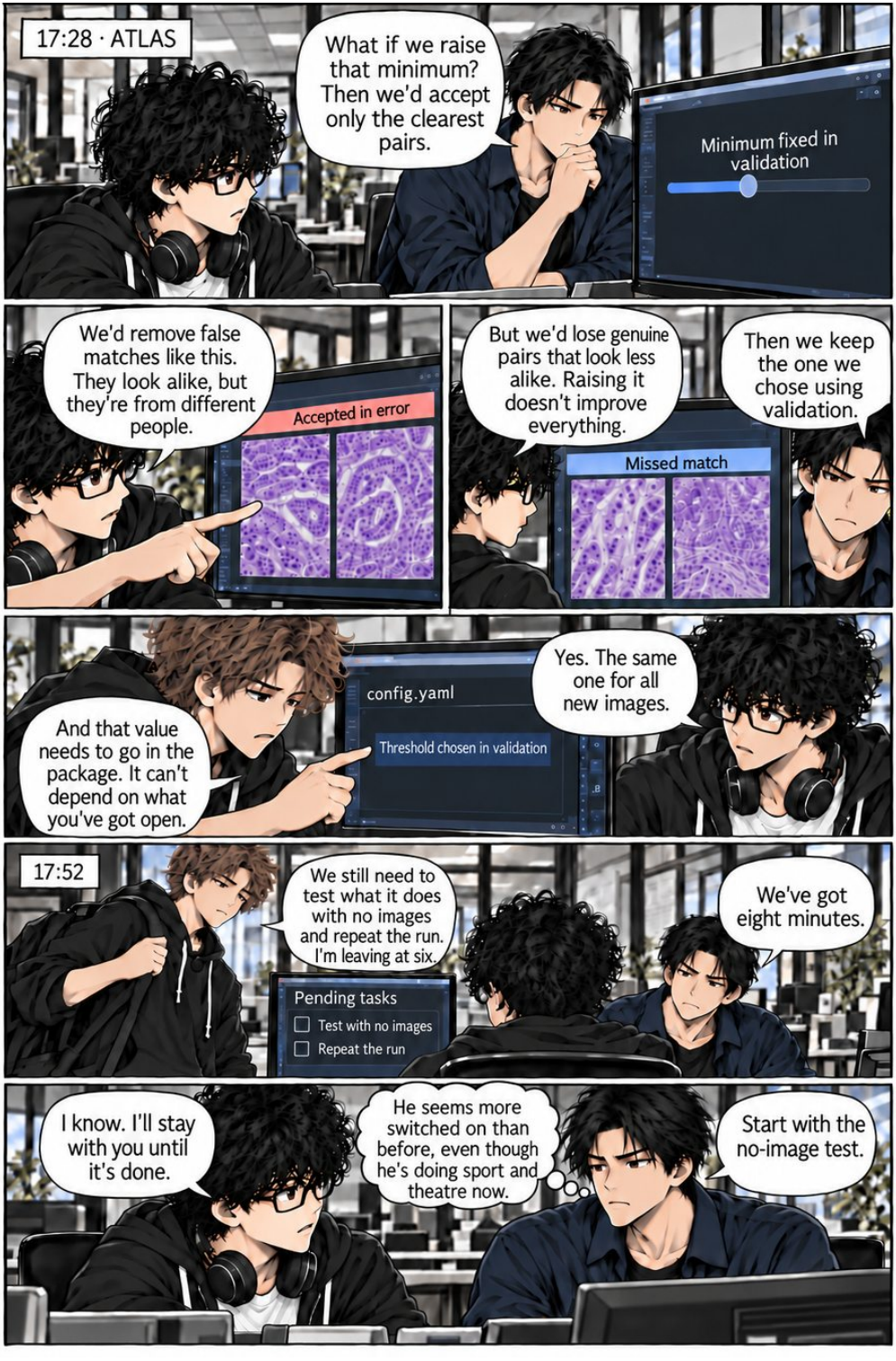

Monday, 12 October · Leo's home
Nexo
New evaluation with new samples
Tuesday, 13 October · ATLAS
We want to detect images from the same patient across collections before sharing them. We need to test that with new data, however well you do in the competition.
ATLAS
Wasn't sending the results enough?
For the competition, yes: a CSV file with the matches. To renew the funding, we need to run your system.
NEXO · N-60
• 60 new patients
• Submission: Thursday 15, before 12:00
• Preparation + model + threshold + environment
The images stay with us. We'll run the version you submit, and you won't be able to adjust it after seeing the result.
Declare where the data comes from. You'll have the report on Friday. We're also watching how you work: Hugo leads our professional team, and we're looking for future colleagues.

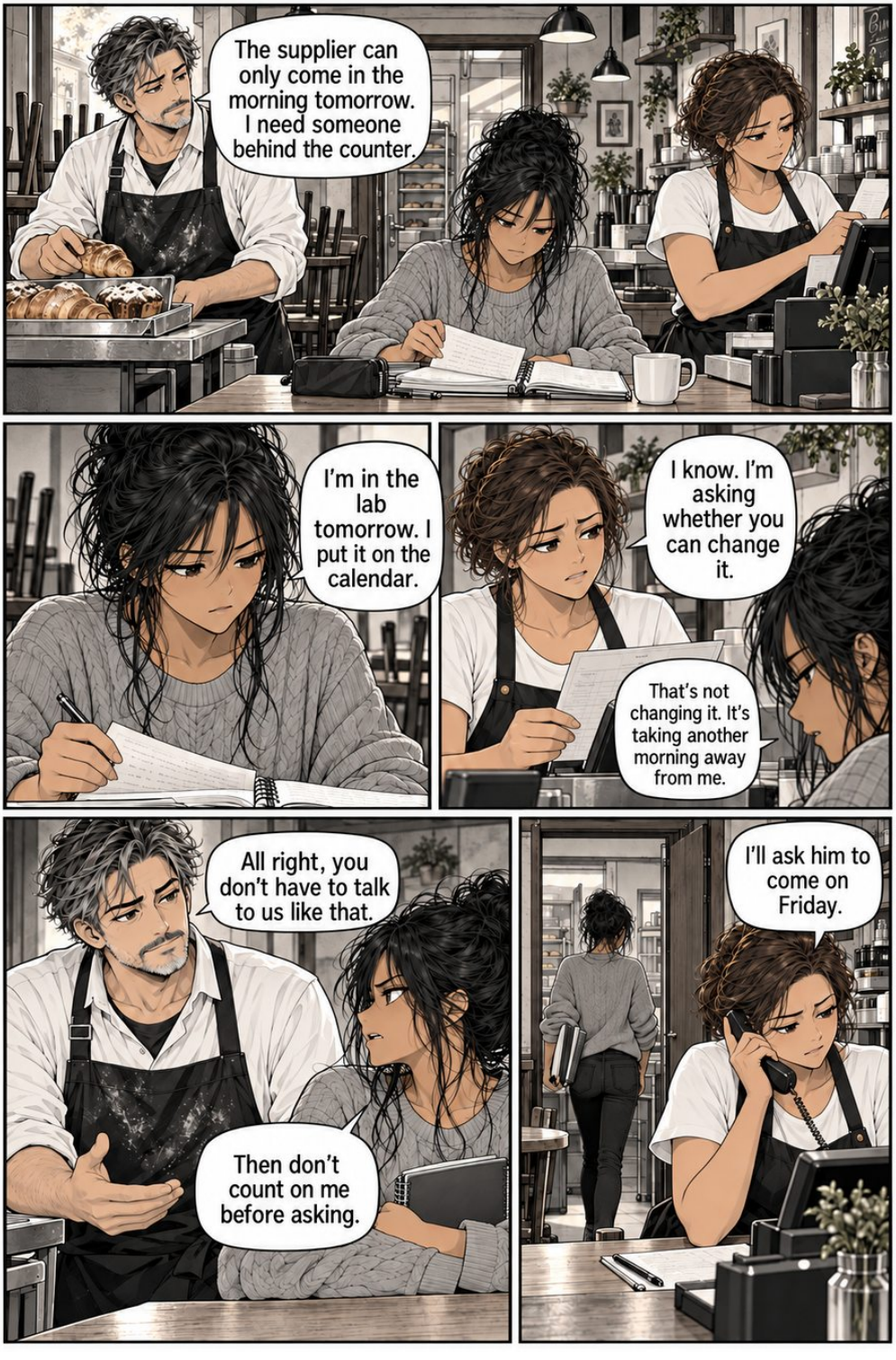

What about the report and the package? If we wait, they'll gain time.
Report figure
Message received
ATLAS
The integrity team already has it. I'll answer their questions. Right now we have two days to submit our own work.
I've got class today, then the café. I'm leaving at four. Don't add another review tonight.
The test run is late today. Can you stay?
Until six. Not after that.
Not even this time?
Last week I missed the bus for a ten-minute test. I got home at ten.
Then five today. And everything needs to be ready beforehand.
TEST PLAN
The model's ready.

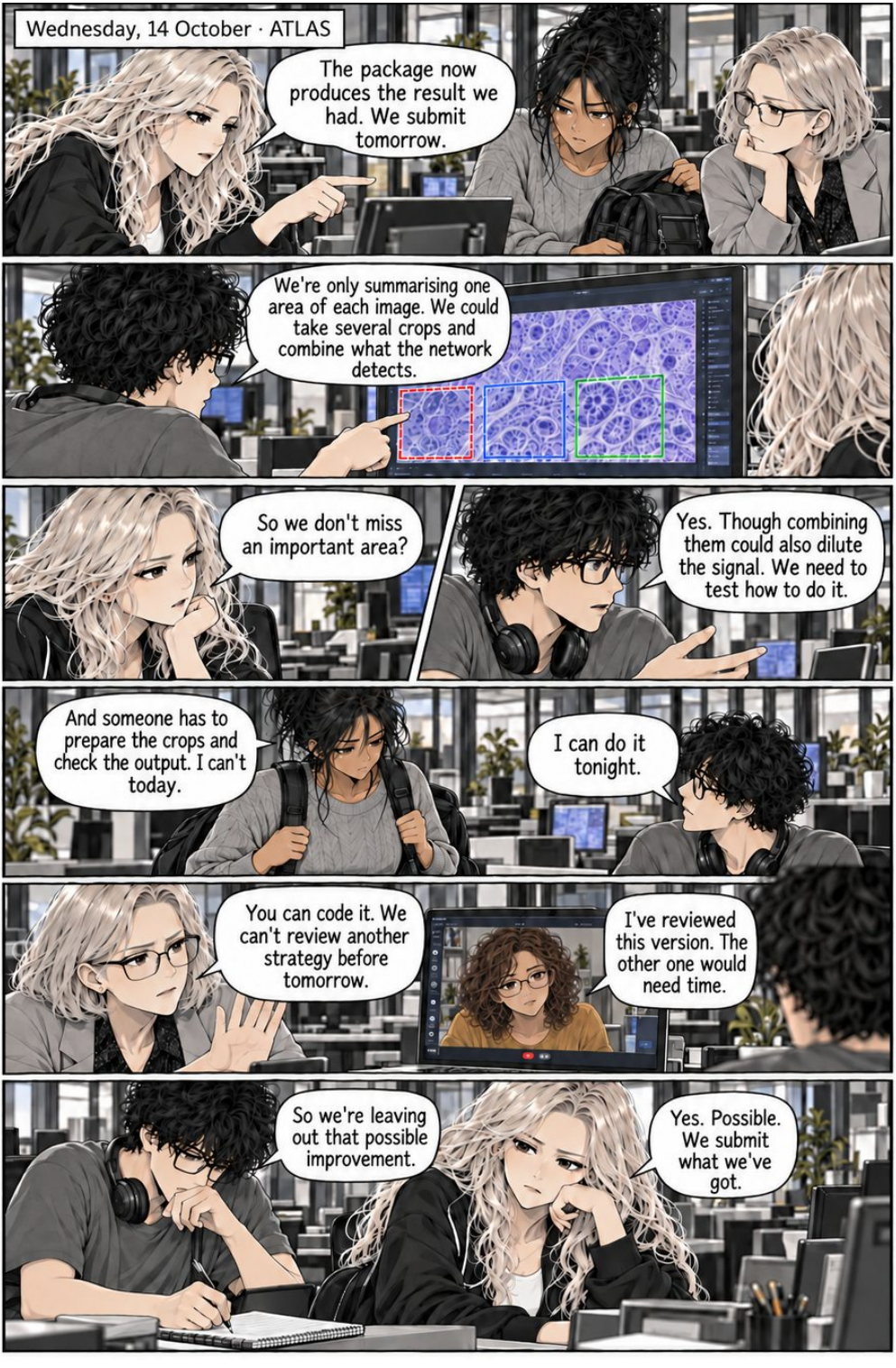

ARGOS · That morning
Sonia, can we use the machines until Thursday?
Until Thursday, yes. November's hours depend on Nexo renewing. So does my contract.
And my placement? You told me we'd carry on all academic year.
That's why we have to submit on time, Iván.
Projects
ARGOS-16
ARGOS-17
Results
september_version
analysis
models
documentation
preprocessed_data.py
model_training.ipynb
validation_results.ipynb
september_report.pdf
metadata.json
We'll send the September version. That's the one that got the result they funded.
Then I need the permission for ARGOS-17. You were handling that appendix.
The declaration's signed. Prepare the run.
Signed isn't the same as complete.
You approved the report too.

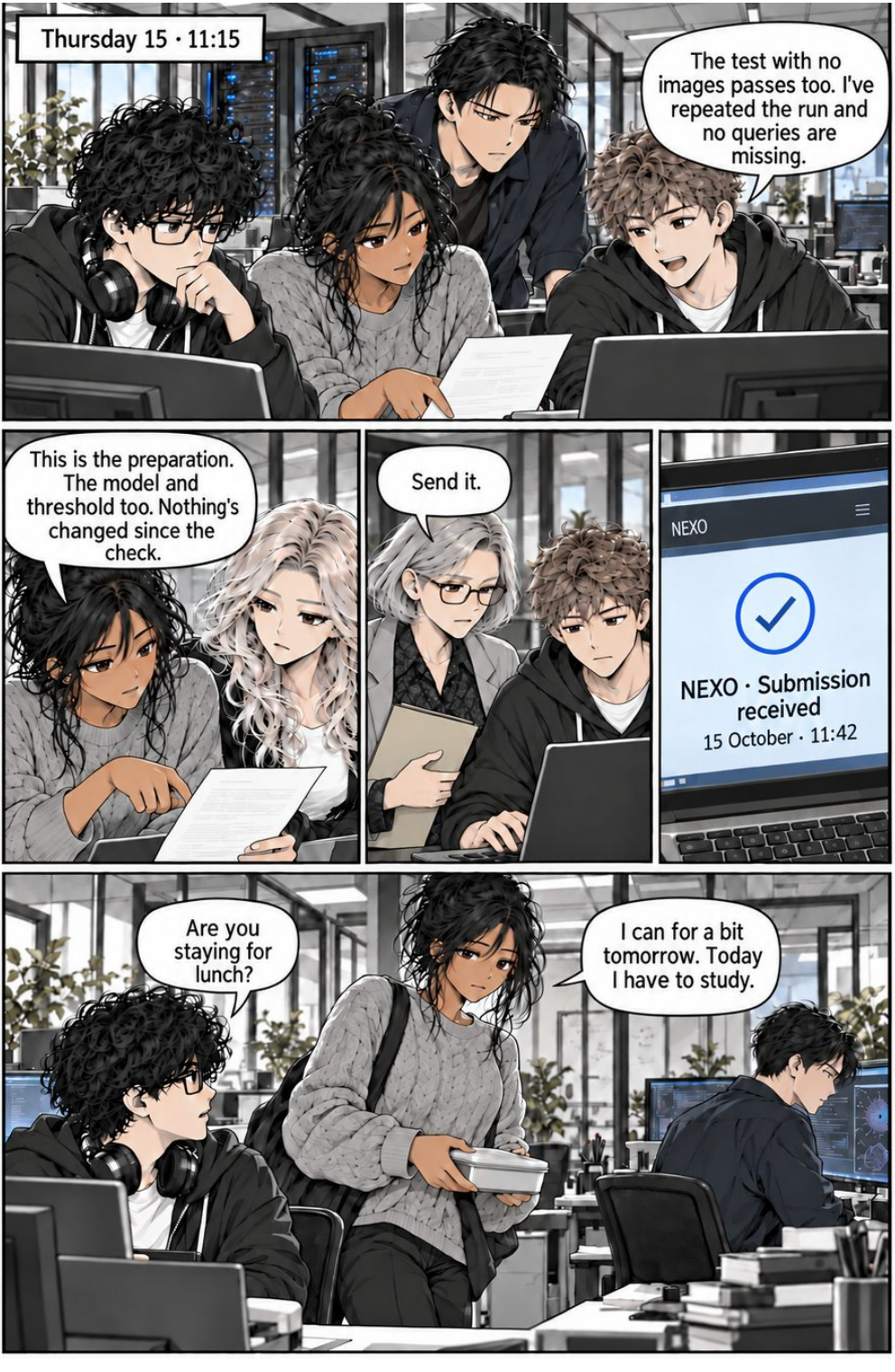

Tuesday, 13 October · 17:05
This doesn't return matches. Just a list of numbers for each image.
0.42
-1.17
0.03
-0.21
0.56
1.08
The network summarises each image's features in those numbers. But the comparison step's missing.
And what decides whether they're from the same patient?
The similarity between the lists. And the minimum we require to accept a pair.
Images
Lists of numbers
Similarity
Matches
I left it in my test notebook. I didn't include it in the package.
You said it was ready.
I'm leaving at six. Can you integrate it before then?
Yes. It's the step we'd already chosen. We don't need to train again.

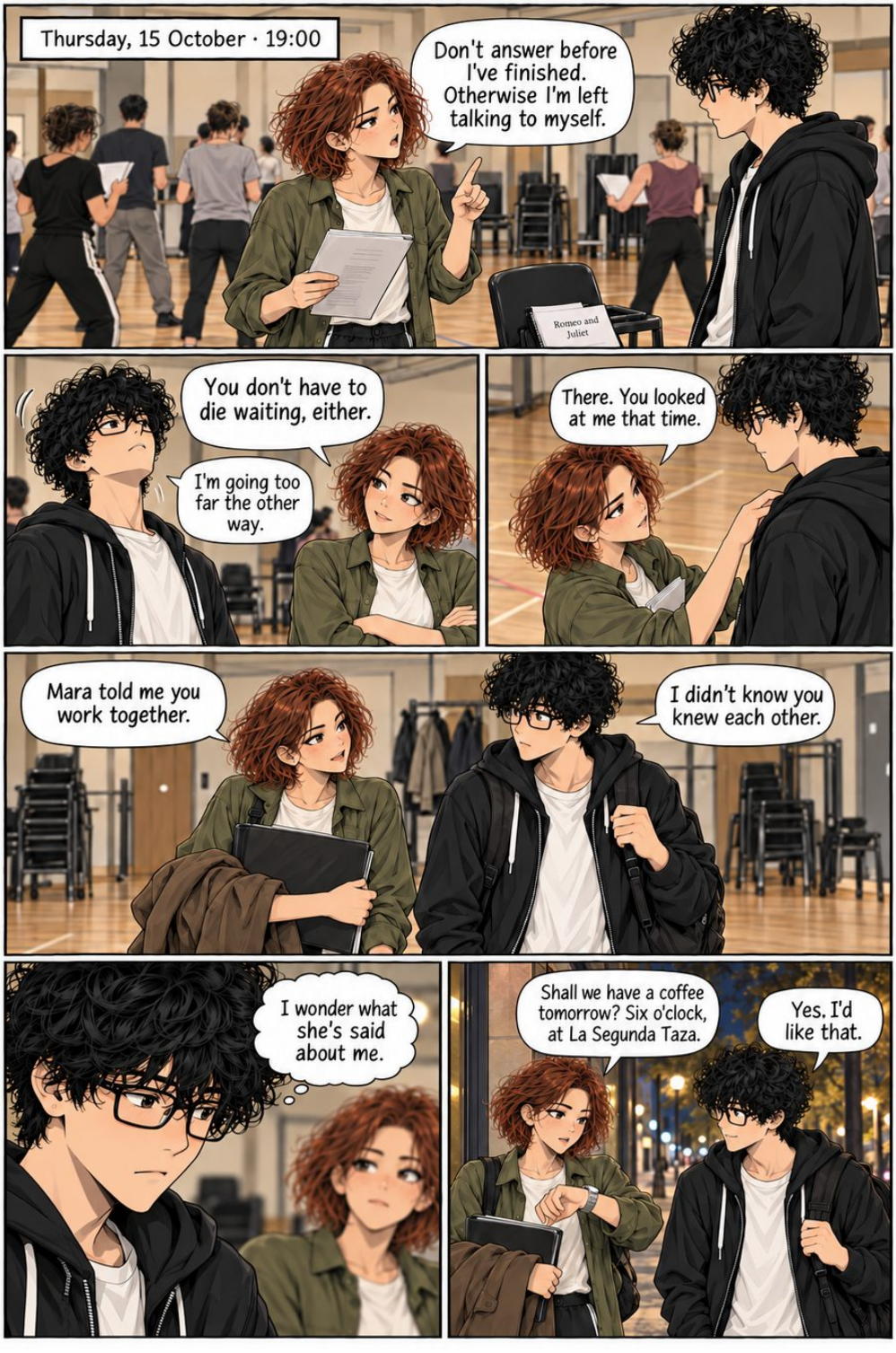

Tuesday, 13 October · La Segunda Taza
We're back to Romeo and Juliet on Thursday. My scene partner's taking it seriously.
ROMEO AND JULIET
CHARACTER ANALYSIS
EMOTIONAL CONNECTION
PACE
CHANGE OF TONE HERE
Leo Martín. Curly hair, black glasses. He brought me the script covered in arrows.
Twenty-two? Started a master's in AI?
Yes. Do you know him?
He works with me. He hasn't said a word.
Don't you think he's cute?
I didn't know you did theatre together.

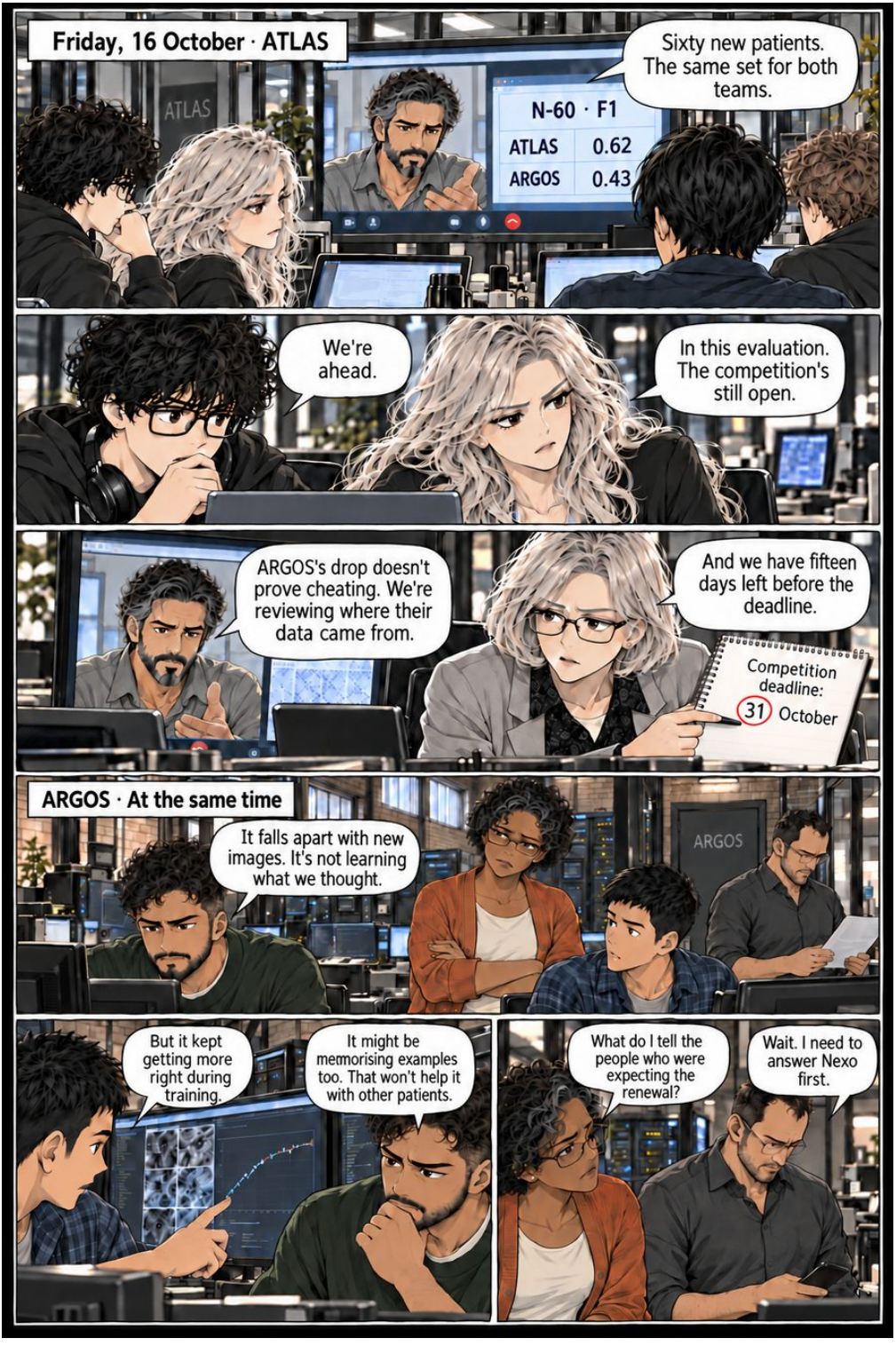

17:28 · ATLAS
What if we raise that minimum? Then we'd accept only the clearest pairs.
Minimum fixed in validation
We'd remove false matches like this. They look alike, but they're from different people.
Accepted in error
But we'd lose genuine pairs that look less alike. Raising it doesn't improve everything.
Missed match
Then we keep the one we chose using validation.
And that value needs to go in the package. It can't depend on what you've got open.
config.yaml
Threshold chosen in validation
Yes. The same one for all new images.
17:52
We still need to test what it does with no images and repeat the run. I'm leaving at six.
Pending tasks
Test with no images
Repeat the run
We've got eight minutes.
I know. I'll stay with you until it's done.
He seems more switched on than before, even though he's doing sport and theatre now.
Start with the no-image test.

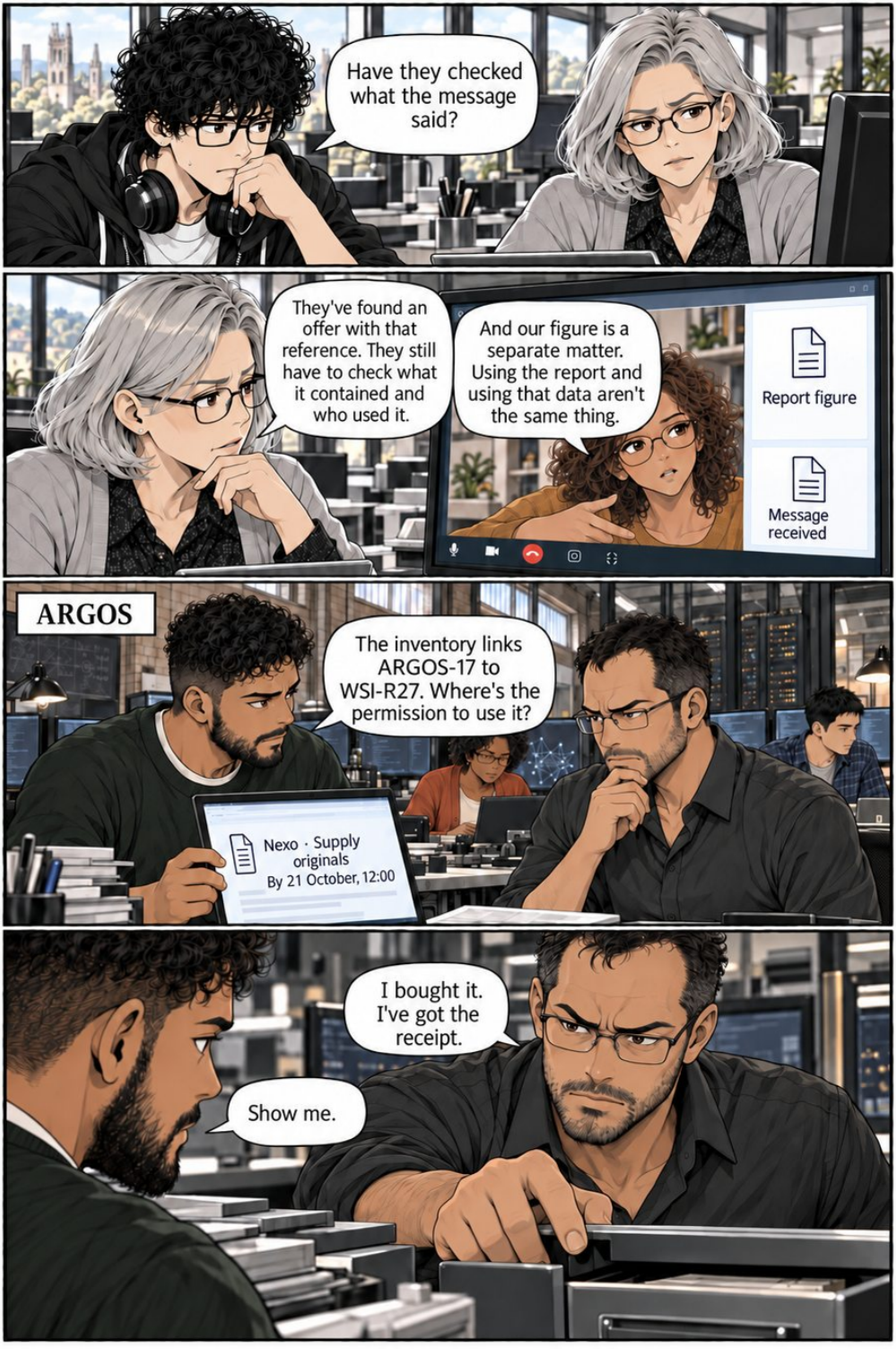

The supplier can only come in the morning tomorrow. I need someone behind the counter.
I'm in the lab tomorrow. I put it on the calendar.
I know. I'm asking whether you can change it.
That's not changing it. It's taking another morning away from me.
All right, you don't have to talk to us like that.
Then don't count on me before asking.
I'll ask him to come on Friday.


Wednesday, 14 October · ATLAS
The package now produces the result we had. We submit tomorrow.
We're only summarising one area of each image. We could take several crops and combine what the network detects.
So we don't miss an important area?
Yes. Though combining them could also dilute the signal. We need to test how to do it.
And someone has to prepare the crops and check the output. I can't today.
I can do it tonight.
You can code it. We can't review another strategy before tomorrow.
I've reviewed this version. The other one would need time.
So we're leaving out that possible improvement.
Yes. Possible. We submit what we've got.


Thursday 15 · 11:15
The test with no images passes too. I've repeated the run and no queries are missing.
This is the preparation. The model and threshold too. Nothing's changed since the check.
Send it.
NEXO
NEXO · Submission received
15 October · 11:42
Are you staying for lunch?
I can for a bit tomorrow. Today I have to study.


Thursday, 15 October · 19:00
Don't answer before I've finished. Otherwise I'm left talking to myself.
Romeo and Juliet
You don't have to die waiting, either.
I'm going too far the other way.
There. You looked at me that time.
Mara told me you work together.
I didn't know you knew each other.
I wonder what she's said about me.
Shall we have a coffee tomorrow? Six o'clock, at La Segunda Taza.
Yes. I'd like that.

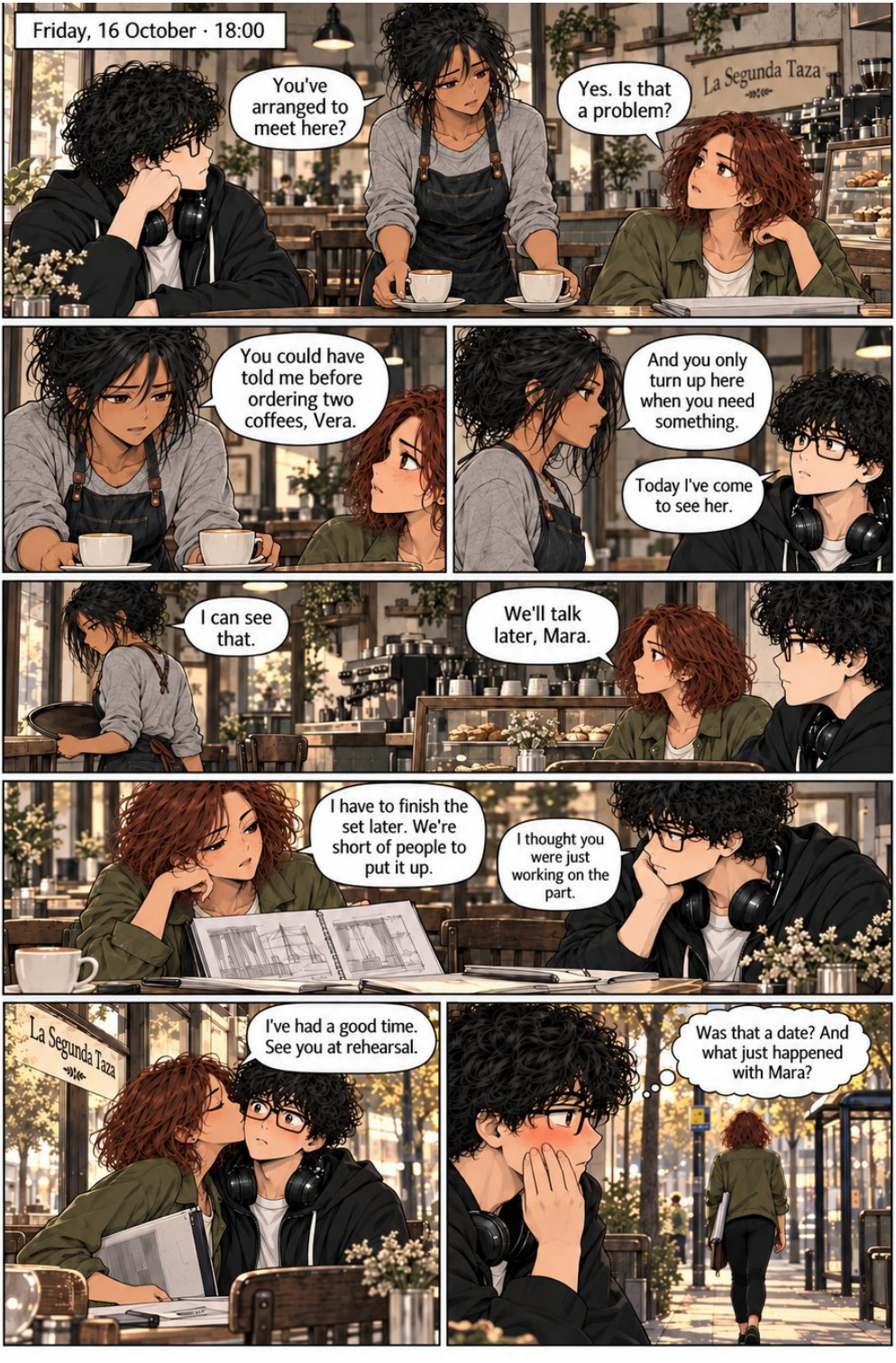

Friday, 16 October · ATLAS
ATLAS
N-60 · F1
ATLAS 0.62
ARGOS 0.43
Sixty new patients. The same set for both teams.
We're ahead.
In this evaluation. The competition's still open.
ARGOS's drop doesn't prove cheating. We're reviewing where their data came from.
And we have fifteen days left before the deadline.
Competition deadline: 31 October
ARGOS · At the same time
ARGOS
It falls apart with new images. It's not learning what we thought.
But it kept getting more right during training.
It might be memorising examples too. That won't help it with other patients.
What do I tell the people who were expecting the renewal?
Wait. I need to answer Nexo first.

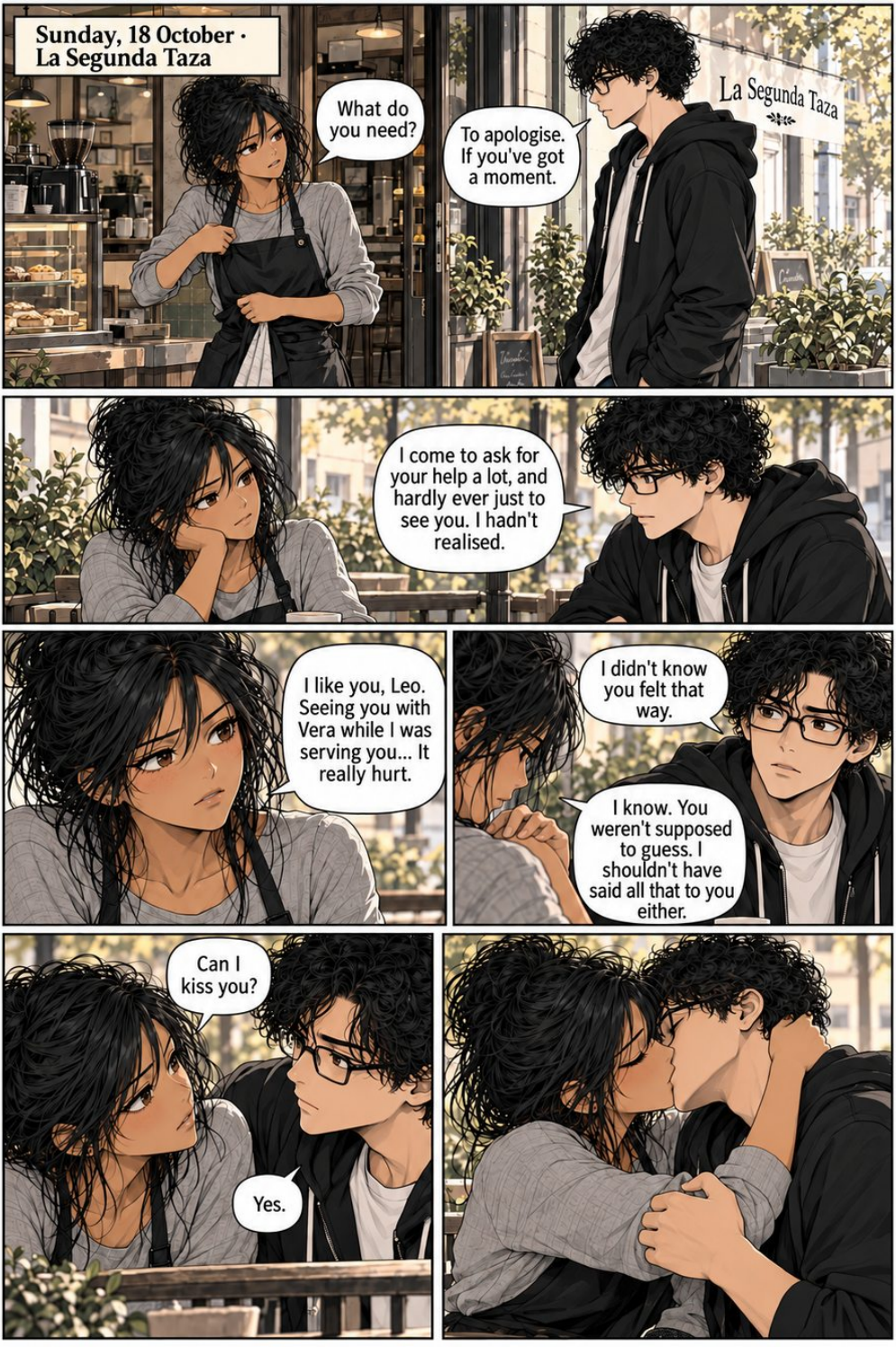

Have they checked what the message said?
They've found an offer with that reference. They still have to check what it contained and who used it.
And our figure is a separate matter. Using the report and using that data aren't the same thing.
Report figure
Message received
ARGOS
The inventory links ARGOS-17 to WSI-R27. Where's the permission to use it?
Nexo · Supply originals
By 21 October, 12:00
I bought it. I've got the receipt.
Show me.


18 September
WSI-R27
Examples associated with evaluation
You gave me examples linked to the exam to prepare the model?
The seller said 'associated'. I didn't know how much they overlapped.
If the model's already seen the exam images, the score doesn't tell us whether it'll work with others.
You told me the data was authorised. Have you got anything besides this receipt?
No.


We already had good results before that purchase. We could have carried on with our work.
That wasn't enough to guarantee the funding.
What have you bought?
Examples linked to the evaluation. I loaded them because Héctor said we could use them.
If you send that, we lose the funding. So does Sonia. So does Iván.
Don't use me as an excuse. I've spent months keeping this lab going.
Did my placement depend on nobody asking where the data came from?
And you approved the report without reading the appendix.
I didn't read it. What a fuck-up. But I'm not signing again to say this is authorised.
I'll speak to management. Don't send anything yourselves.
They've asked for the originals.

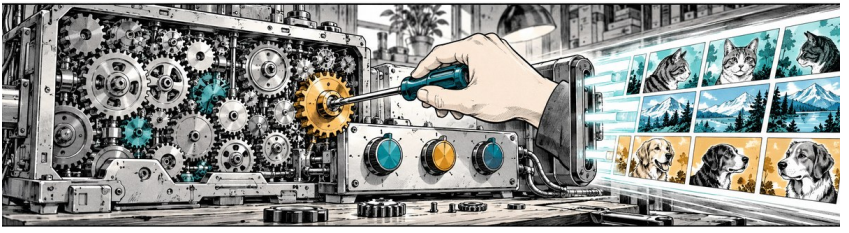

From: Héctor
Include ARGOS-17 in the Nexo test. The provenance is covered.
I'm not changing the server log.
I don't want you to. I need to keep it exactly as it is.
Originals
Receipt
Héctor's email
Import log
Provenance declaration
Report approved by Damián
Do we carry on working?
With the data we were authorised to use. The rest stays out.
From: Héctor
I'll handle the meeting with management.
My sign-off stays in there too.


Friday, 16 October · 18:00
You've arranged to meet here?
Yes. Is that a problem?
La Segunda Taza
You could have told me before ordering two coffees, Vera.
And you only turn up here when you need something.
Today I've come to see her.
I can see that.
We'll talk later, Mara.
I have to finish the set later. We're short of people to put it up.
I thought you were just working on the part.
La Segunda Taza
I've had a good time. See you at rehearsal.
Was that a date? And what just happened with Mara?


Sunday, 18 October · La Segunda Taza
What do you need?
To apologise. If you've got a moment.
La Segunda Taza
I come to ask for your help a lot, and hardly ever just to see you. I hadn't realised.
I like you, Leo. Seeing you with Vera while I was serving you... It really hurt.
I didn't know you felt that way.
I know. You weren't supposed to guess. I shouldn't have said all that to you either.
Can I kiss you?
Yes.


Leo and Mara? Maybe I've spent too long looking straight past him.
La Segunda Taza
And I still look at the first thing that catches my eye.
I have to go. I'm not really sure what we do with this.
Neither am I.
That night · Damián's home
To: Nexo Integrity
Draft · Not sent
Leo's home
We submitted what we'd tested. Everything else stayed out.
A rehearsed kiss, one on the cheek, and this one. I don't know what I want. I'll have to talk to them.

What Nico Needs to Run



A metaphor for adjusting part of an already trained model.

Leo says the model is ready. Nico starts it and receives lists of numbers when he expected matches. It is a familiar project conversation: each person has considered a different part finished, and now they must join them up.

What comes out of the network

The representation, also called a vector, summarises an image's features. The program still has to compare it with the candidates and score their similarity. Two images from different people can look quite alike, so it then needs a rule to accept or reject pairs. The representation enables that comparison; on its own, it still does not give Nico the answer he must submit.

Where the threshold comes in

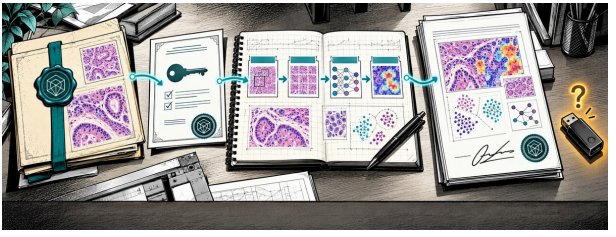
With the same scores, requiring greater similarity rejects proposals: it removes some false ones and may lose genuine pairs. Lowering the minimum admits more of both kinds. We are changing a decision rule without changing the network's parameters. ATLAS includes the value it had already chosen using validation. The comparator existed in Leo's notebook; it had yet to reach the program Nico was running.

Who we are sending it to

The competition asks for a CSV: text organised into rows and columns containing predicted matches. Nexo wants to run the system on its images and needs the whole chain: preparation, model, decision and software environment. Nico must be able to follow it without asking about every step of the notebook. The no-image test checks what the program does with no input. To check the full chain, we also need to take one example through and review its output. Checking match quality requires examples whose answers we can verify.

Sources: · Thresholds · WSI challenge

Checking the Result and the Provenance



Origin, permissions and steps kept alongside the result.

Nexo evaluates both teams with the same sixty new patients. ARGOS drops, and it is easy to complete the story before having the evidence. Hugo asks them to separate two jobs: measuring what the system does and finding out what material was used to build it.

What may have entered training

A program can achieve a high score because its preparation exposed it to images related to the exam. We call it data leakage when information enters that should have remained held out for evaluation. It can happen without copying a literal answer: the examples need only stop letting us test what it will do with new patients. A falling score requires investigation, but does not by itself reconstruct that entry of information.

What permission comes with the purchase

Paying for an image package gives us a receipt. We still have to check whether the seller could supply it and whether our intended use was authorised. The story shows reference WSI-R27, its internal name ARGOS-17, an email and logs that let us trace decisions. They are being gathered and compared. Héctor's evasions may raise suspicion; the documents must support what is claimed.

What each trail explains

ATLAS's figure republished without its warning raises another investigation. Copying that figure does not automatically explain a model's performance, just as a poor score does not identify where its images came from. Nadia keeps the questions separate. Damián also keeps his own sign-off: to reconstruct what happened, the file must include decisions that implicate the person supplying the evidence too.

Sources: · Preparation and leakage · WSI challenge

The Same Closeness Can Mean Different Things



Listening and waiting for the next line are part of rehearsal too.

A rehearsed kiss, a date and an apology are different situations. Leo is eager to know where he stands with Vera and Mara, but tries to draw a final answer from every gesture.

What happens in front of Mara

Vera and Leo go to the café and Mara has to serve them. Vera had not told her they were meeting; Leo did not know Mara liked him. She gets angry with both. When they talk again, she acknowledges he was not supposed to guess. Her disappointment does not prove Leo had accepted a commitment.

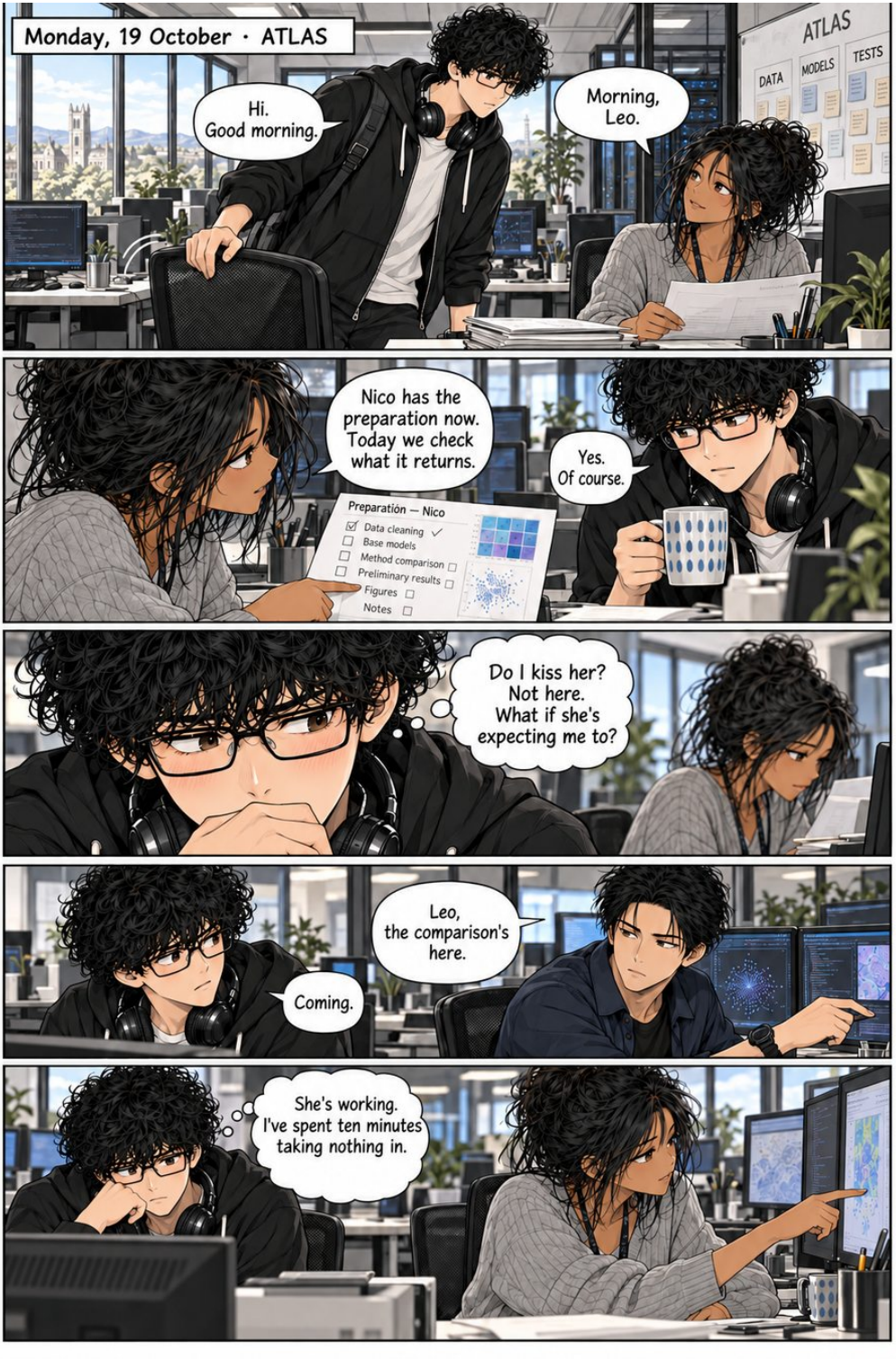
An apology does not decide a relationship

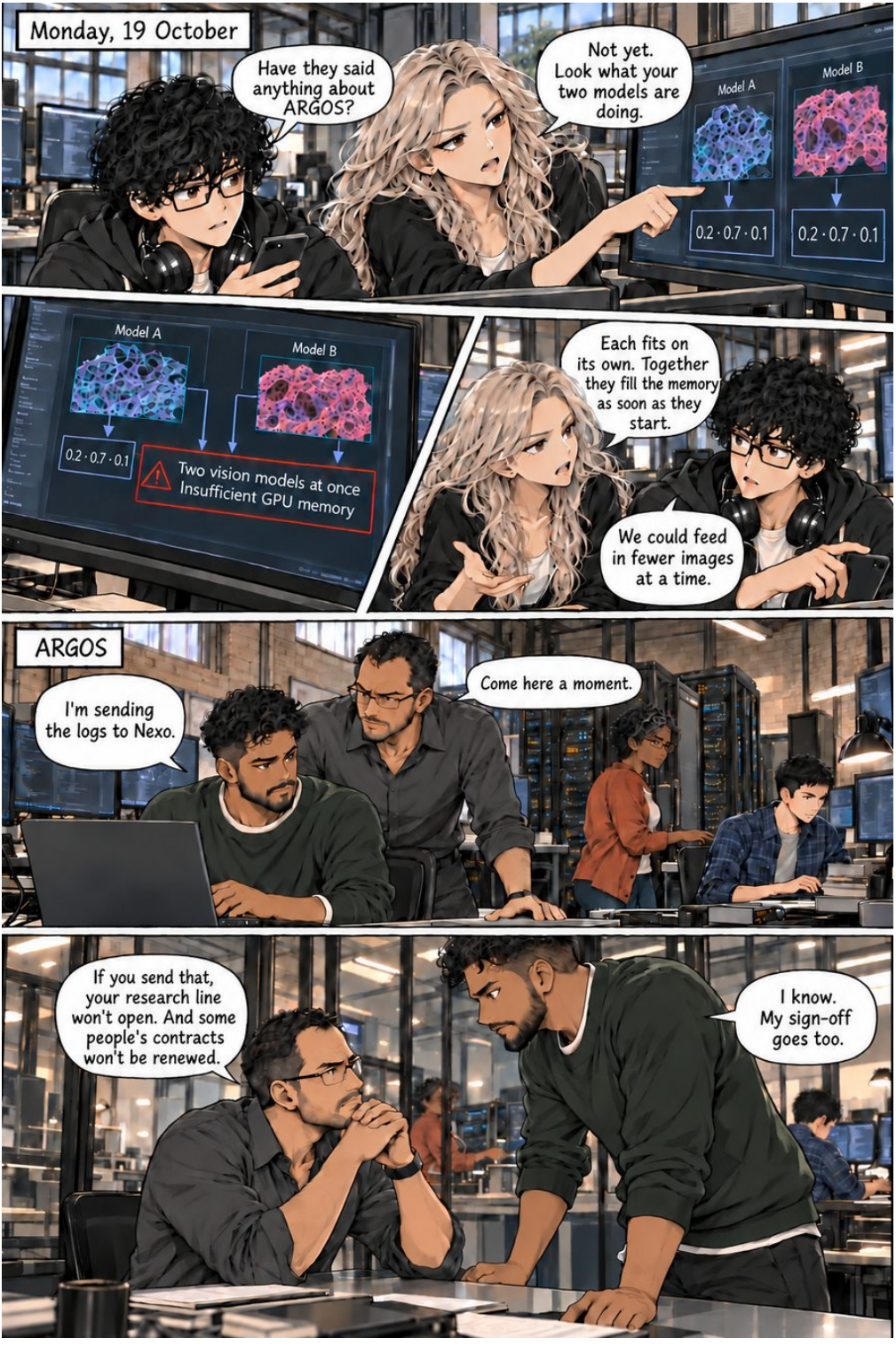
Leo returns to talk to Mara and the meeting ends in a mutual kiss. He can like it and still not know what he wants to do next. It is not Mara's job to resolve all that confusion either. They will have to discuss what each wants, even if it feels less comfortable than looking for clues.

And Alicia is still Alicia

Seeing them, Alicia wonders whether she has spent too long looking straight past Leo. Soon afterwards, she is attracted to someone else again. That hesitation occupies a moment of her afternoon. Leo does not even know she saw them: he has enough to do understanding what he wants after kissing Mara.

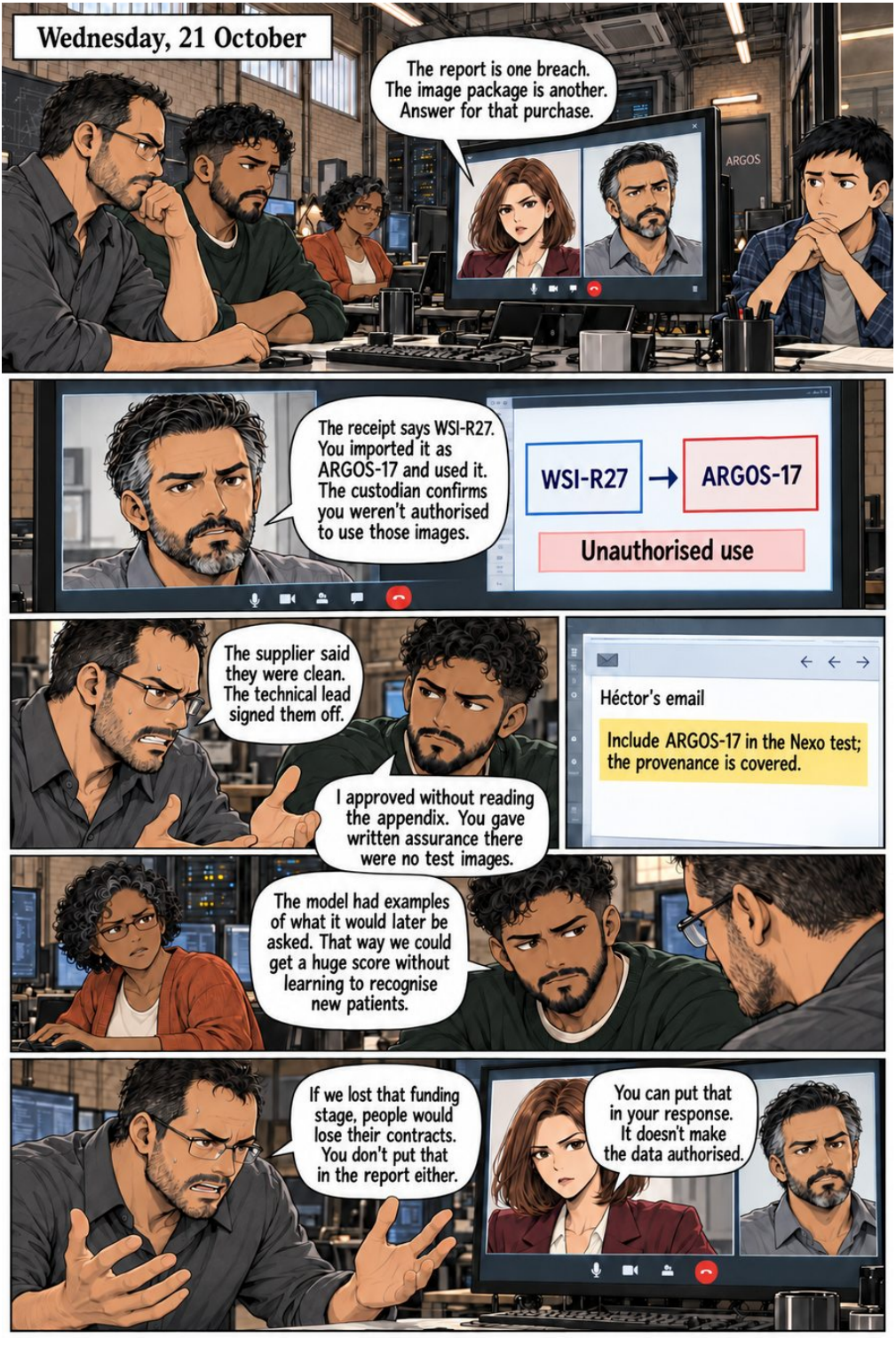
What could Leo ask without assuming a kiss means the same thing to both of them?

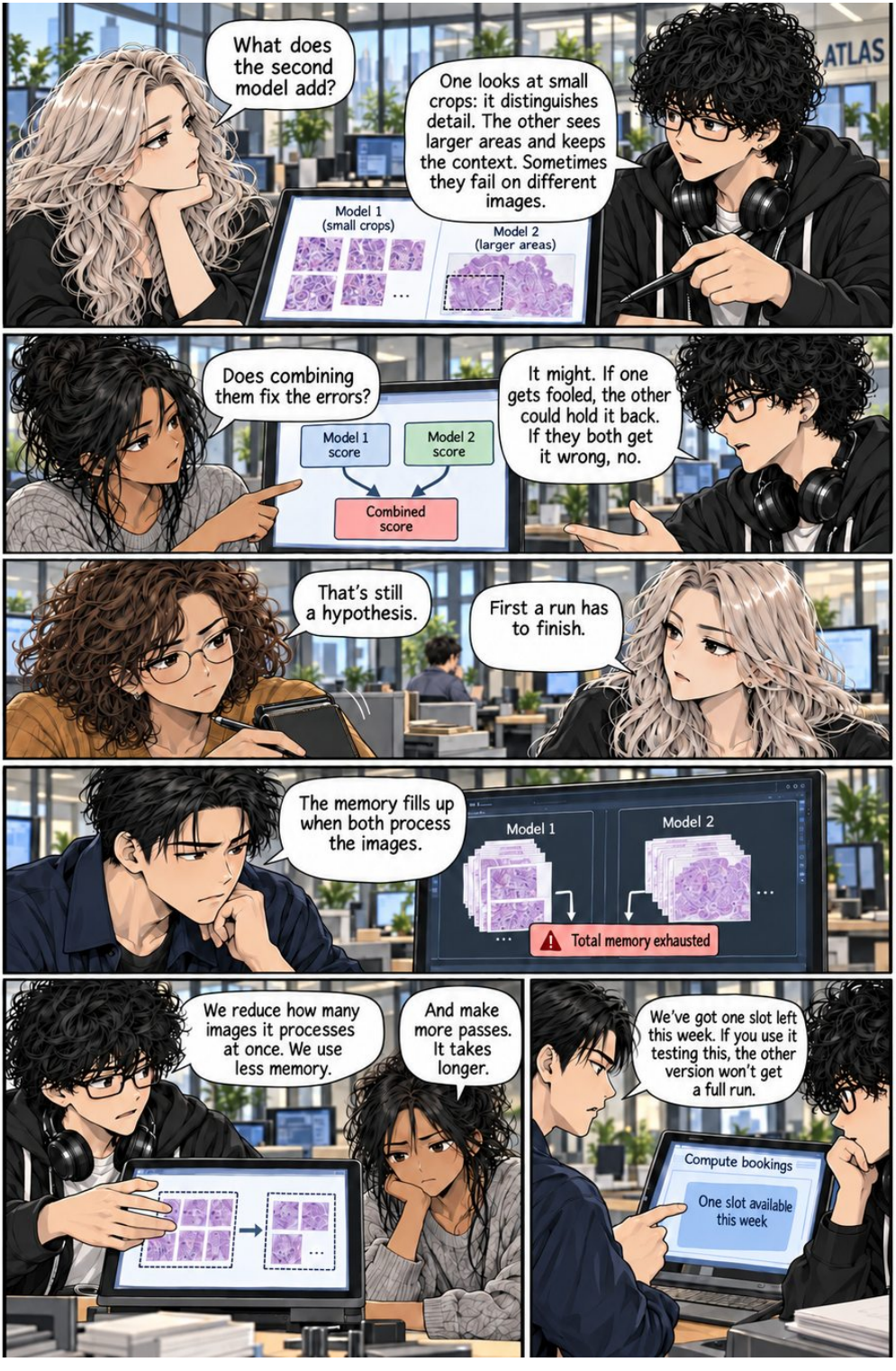

Monday, 19 October · ATLAS
ATLAS
DATA
MODELS
TESTS
Hi.
Good morning.
Morning,
Leo.
Nico has the preparation now. Today we check what it returns.
Yes.
Of course.
Preparation — Nico
Data cleaning
Base models
Method comparison
Preliminary results
Figures
Notes
Do I kiss her? Not here. What if she's expecting me to?
Leo,
the comparison's here.
Coming.
She's working.
I've spent ten minutes taking nothing in.

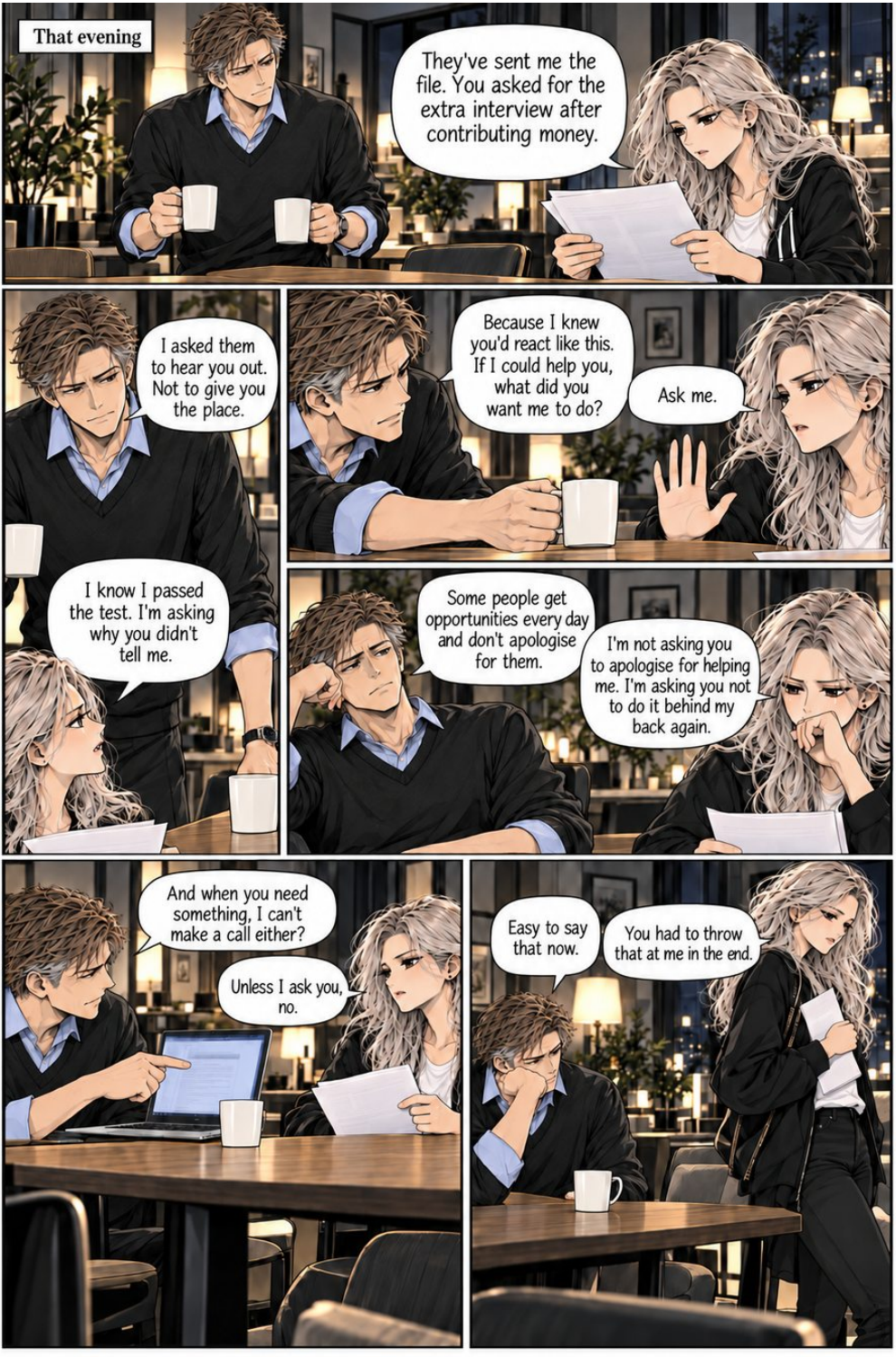

Monday, 19 October
Have they said anything about ARGOS?
Not yet. Look what your two models are doing.
Model A
Model B
0.2 · 0.7 · 0.1
0.2 · 0.7 · 0.1
Model A
Model B
0.2 · 0.7 · 0.1
Two vision models at once
Insufficient GPU memory
Each fits on its own. Together they fill the memory as soon as they start.
We could feed in fewer images at a time.
ARGOS
I'm sending the logs to Nexo.
Come here a moment.
If you send that, your research line won't open. And some people's contracts won't be renewed.
I know. My sign-off goes too.


What's happening with the renewals, Héctor?
Sonia, I'm trying to keep the project from collapsing.
Then tell us. Don't use our contracts to keep him quiet.
Is everything we've trained useless?
Not all of it, Iván. But unauthorised data was used.
Nexo Integrity
I'm attaching the originals and my approval.
Send
I should have checked it before.
Sent
Héctor
Do not represent the team until your involvement has been clarified.
And now you tell us what you signed. Us.

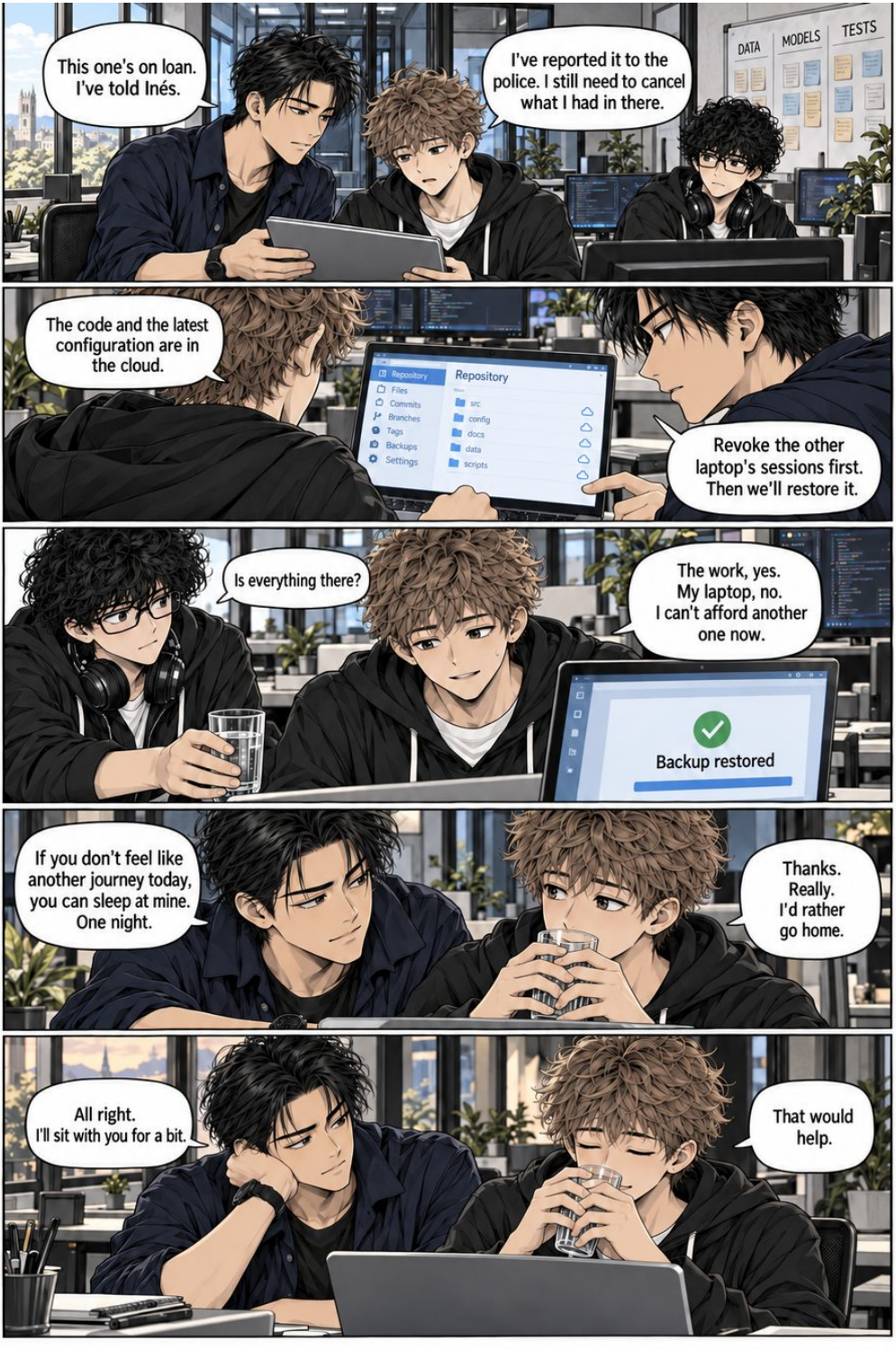

Tuesday 20 · Nexo
Óscar Molina · Evaluation analyst
Héctor asked you for a tip. Is that why you sent him ATLAS's report?
Compose
Inbox
Sent
Drafts
Archived
Óscar → Héctor
Attachment:
ATLAS report
'Don't ask. See if it's any use to us.'
He said he'd keep me in mind for another project. It was only a report.
They presented it as their own progress. And removed the warning that the model failed.
Until we finish the review, you won't have access to the teams' reports.
You review my renewal too.
NEXO
EVALUATION
That's precisely why you had to look after that access.

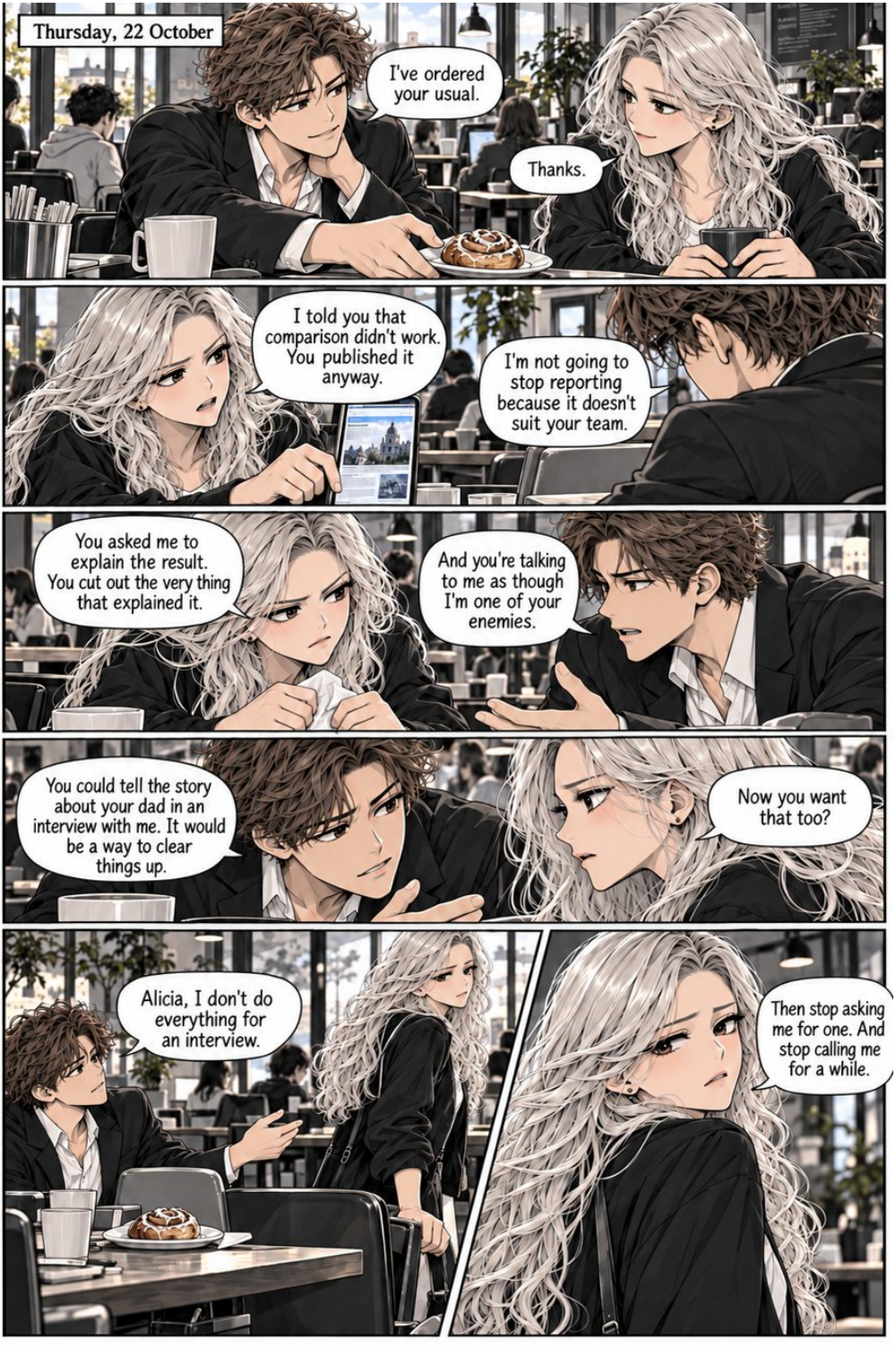

Wednesday, 21 October
The report is one breach. The image package is another. Answer for that purchase.
ARGOS
The receipt says WSI-R27. You imported it as ARGOS-17 and used it. The custodian confirms you weren't authorised to use those images.
WSI-R27
ARGOS-17
Unauthorised use
The supplier said they were clean. The technical lead signed them off.
I approved without reading the appendix. You gave written assurance there were no test images.
Héctor's email
Include ARGOS-17 in the Nexo test; the provenance is covered.
The model had examples of what it would later be asked. That way we could get a huge score without learning to recognise new patients.
If we lost that funding stage, people would lose their contracts. You don't put that in the report either.
You can put that in your response. It doesn't make the data authorised.

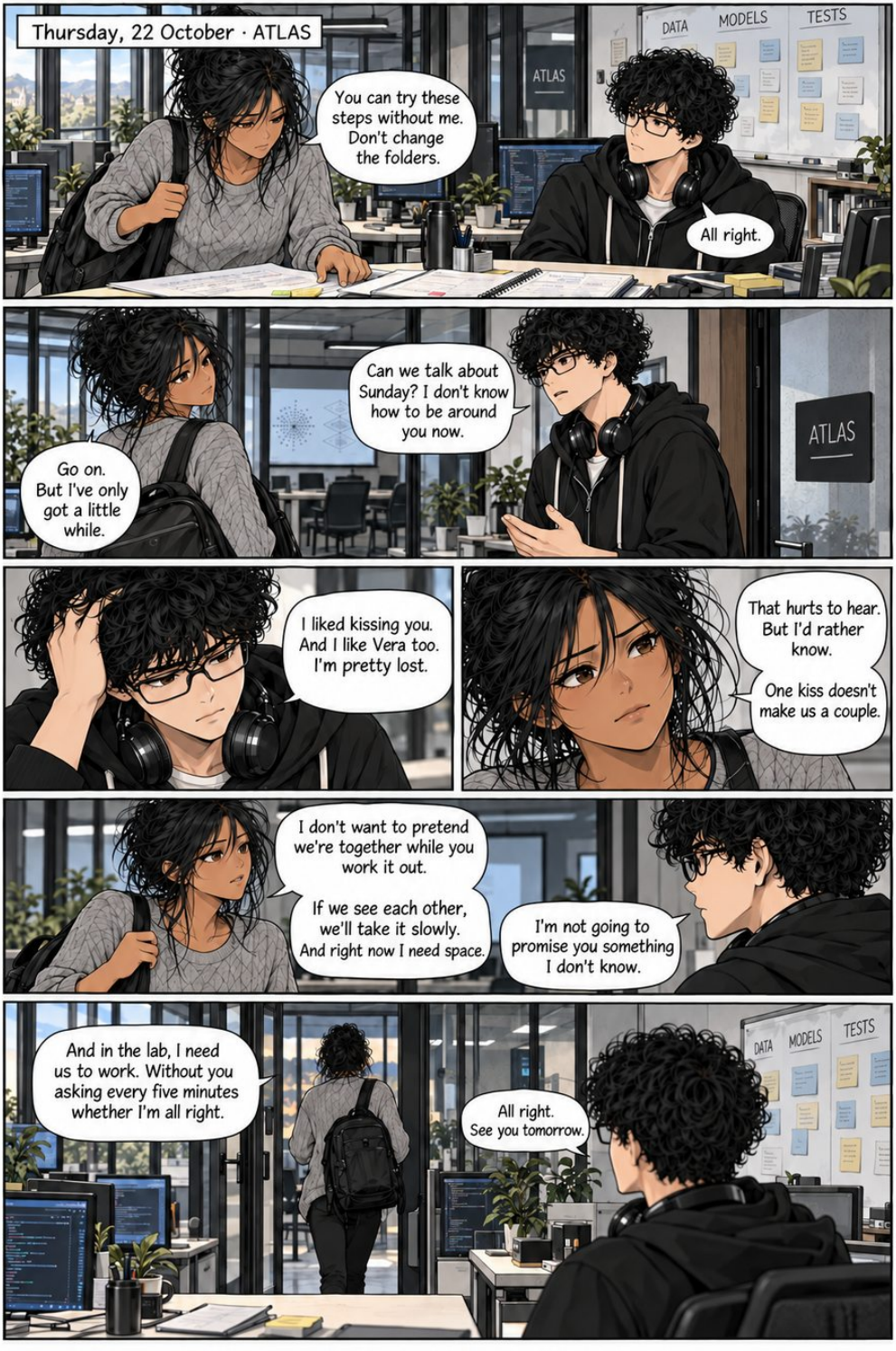

What does the second model add?
One looks at small crops: it distinguishes detail. The other sees larger areas and keeps the context. Sometimes they fail on different images.
ATLAS
Model 1 (small crops)
Model 2 (larger areas)
Does combining them fix the errors?
Model 1 score
Model 2 score
Combined score
It might. If one gets fooled, the other could hold it back. If they both get it wrong, no.
That's still a hypothesis.
First a run has to finish.
The memory fills up when both process the images.
Model 1
Model 2
Total memory exhausted
We reduce how many images it processes at once. We use less memory.
And make more passes. It takes longer.
We've got one slot left this week. If you use it testing this, the other version won't get a full run.
Compute bookings
One slot available this week


That evening
They've sent me the file. You asked for the extra interview after contributing money.
I asked them to hear you out. Not to give you the place.
I know I passed the test. I'm asking why you didn't tell me.
Because I knew you'd react like this. If I could help you, what did you want me to do?
Ask me.
Some people get opportunities every day and don't apologise for them.
I'm not asking you to apologise for helping me. I'm asking you not to do it behind my back again.
And when you need something, I can't make a call either?
Unless I ask you, no.
Easy to say that now.
You had to throw that at me in the end.

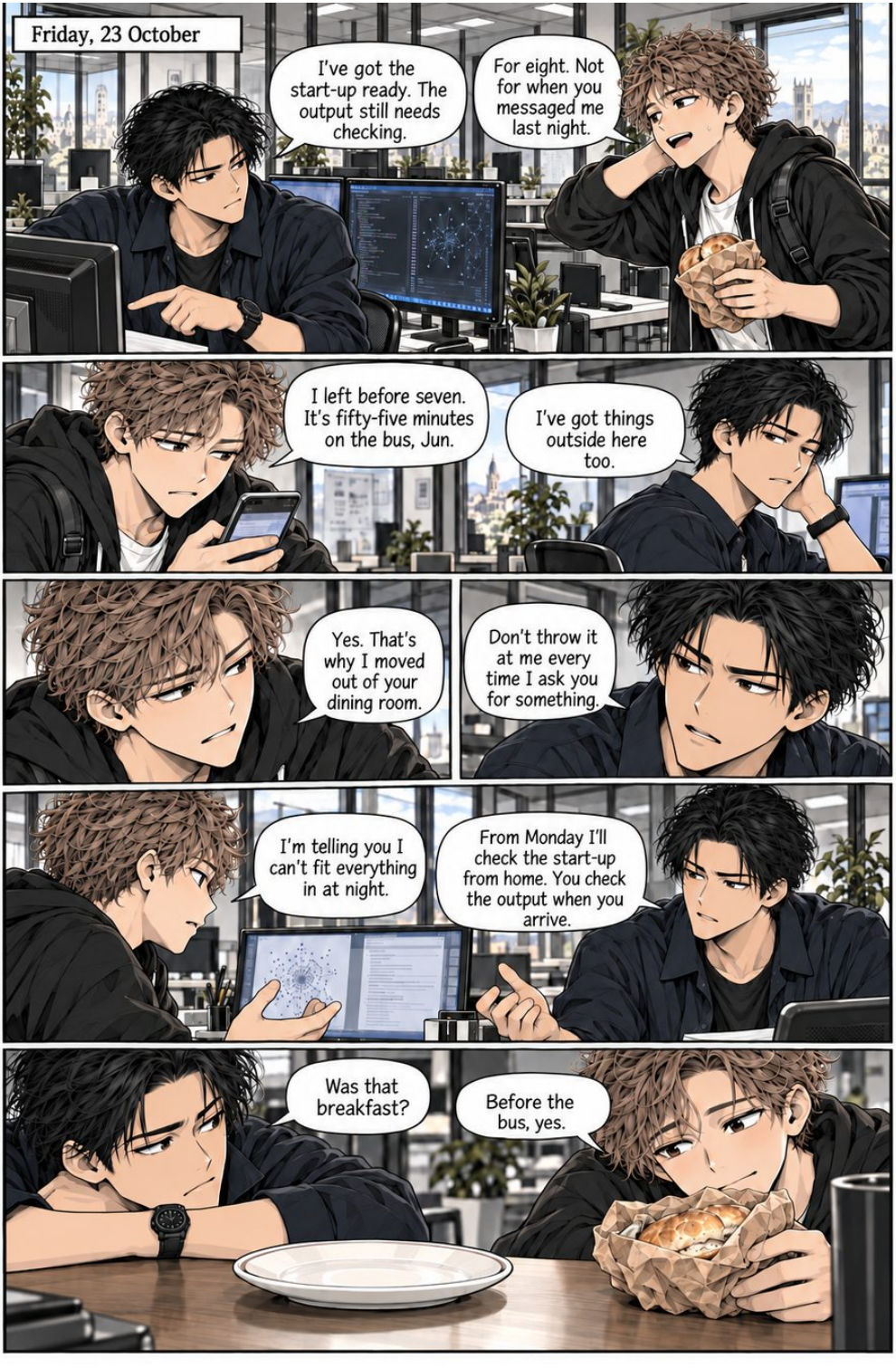

Wednesday, 21 October · On the way home
Shit. My stop.
Now I've got to walk back.
The next day
Hey!
The laptop...
Jun. Someone's taken my bag. No, I'm all right... I don't know. I'm on my way.

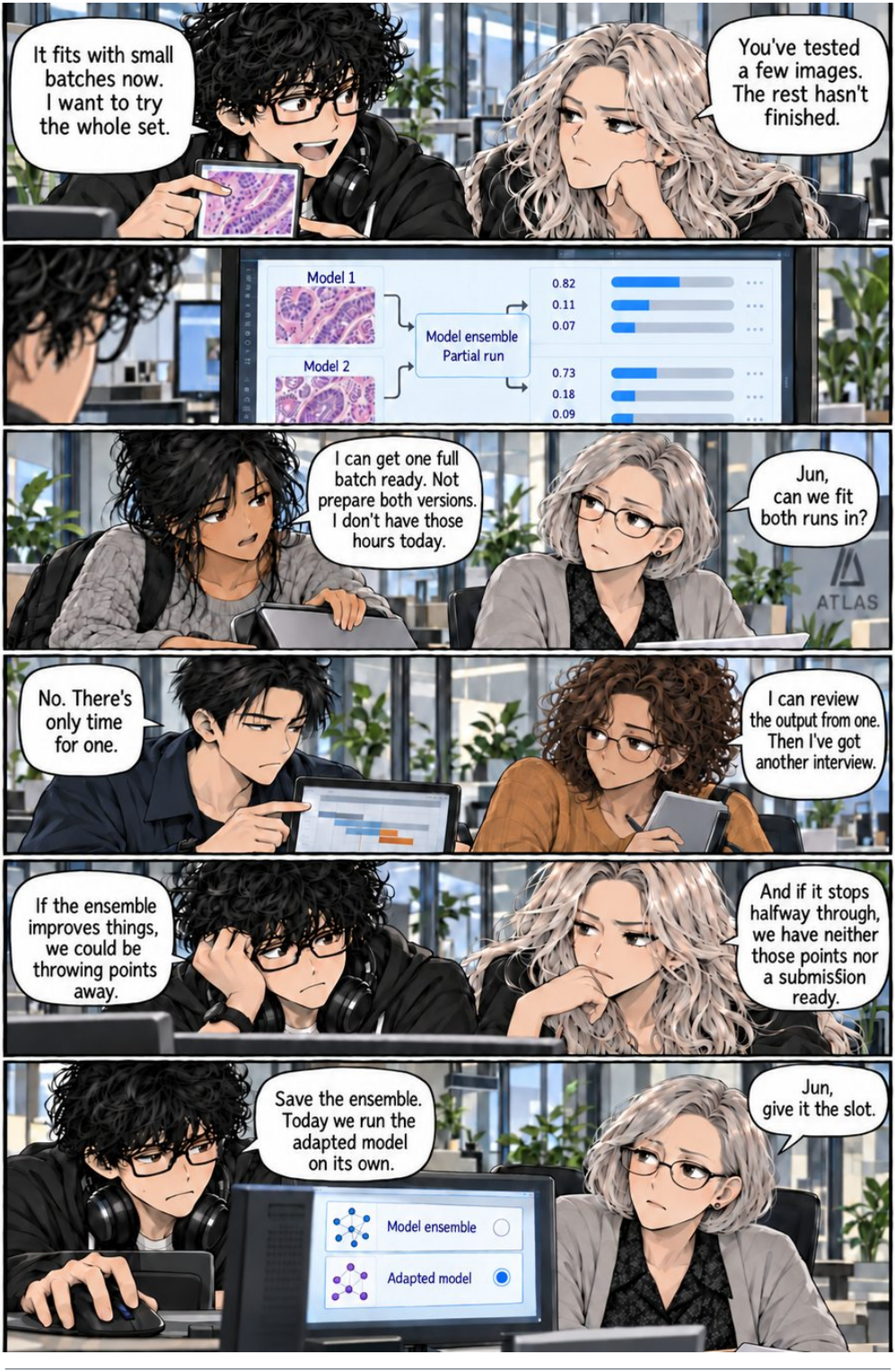

This one's on loan. I've told Inés.
I've reported it to the police. I still need to cancel what I had in there.
DATA
MODELS
TESTS
The code and the latest configuration are in the cloud.
Repository
Files
Commits
Branches
Tags
Backups
Settings
src
config
docs
data
scripts
Revoke the other laptop's sessions first. Then we'll restore it.
Is everything there?
The work, yes. My laptop, no. I can't afford another one now.
Backup restored
If you don't feel like another journey today, you can sleep at mine. One night.
Thanks. Really. I'd rather go home.
All right. I'll sit with you for a bit.
That would help.


Thursday, 22 October
I've ordered your usual.
Thanks.
I told you that comparison didn't work. You published it anyway.
I'm not going to stop reporting because it doesn't suit your team.
You asked me to explain the result. You cut out the very thing that explained it.
And you're talking to me as though I'm one of your enemies.
You could tell the story about your dad in an interview with me. It would be a way to clear things up.
Now you want that too?
Alicia, I don't do everything for an interview.
Then stop asking me for one. And stop calling me for a while.

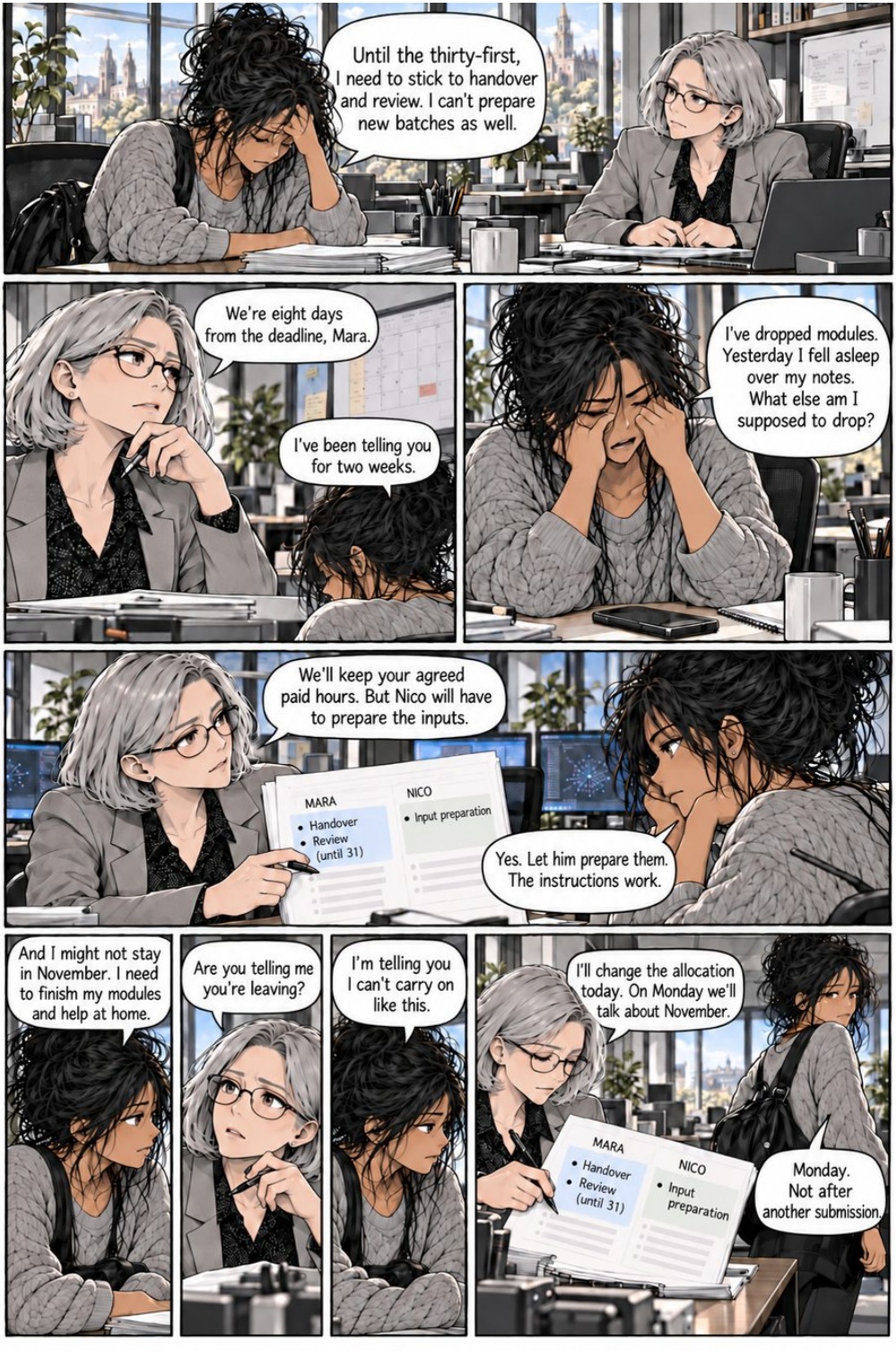

Thursday, 22 October · ATLAS
ATLAS
DATA
MODELS
TESTS
You can try these steps without me. Don't change the folders.
All right.
Can we talk about Sunday? I don't know how to be around you now.
Go on. But I've only got a little while.
ATLAS
I liked kissing you. And I like Vera too. I'm pretty lost.
That hurts to hear. But I'd rather know.
One kiss doesn't make us a couple.
I don't want to pretend we're together while you work it out.
If we see each other, we'll take it slowly. And right now I need space.
I'm not going to promise you something I don't know.
And in the lab, I need us to work. Without you asking every five minutes whether I'm all right.
All right. See you tomorrow.
DATA
MODELS
TESTS

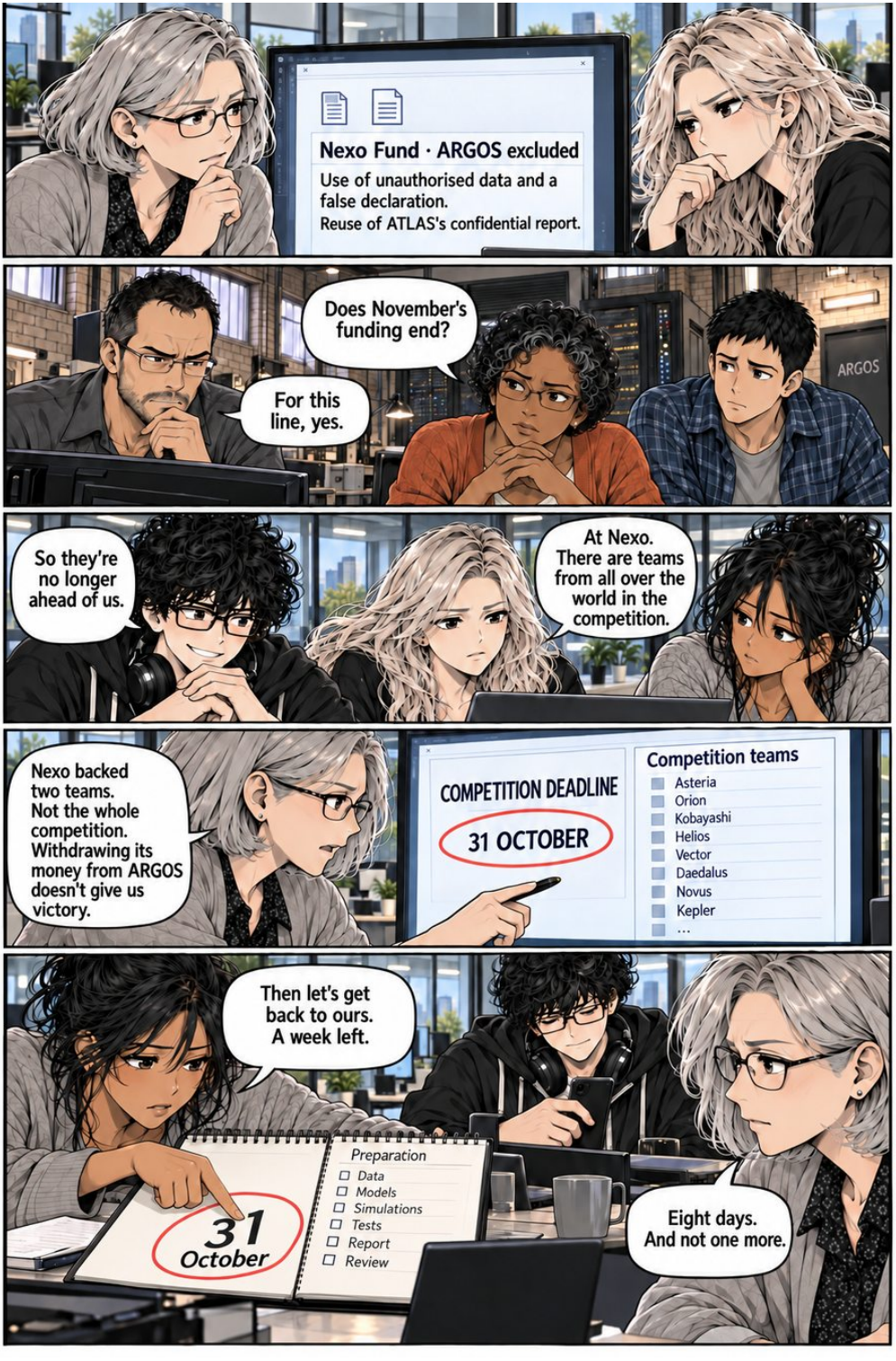

La Segunda Taza
I have to see the supplier on Tuesday afternoon. I need you to cover here.
I have class on Tuesday.
I thought, since you'd dropped some modules...
I've kept two, Mum. So I can pass them.
Module 1
Module 2
Then tell me what we do, because I can't be in two places either.
Not Tuesday. I can change my shift on Wednesday.
I'll ask if they can see me on Wednesday morning. I'm not promising anything.
I can't keep promising every afternoon either.
Inés
I need to reduce my afternoons. Tomorrow we'll hand over the preparation. We need to talk about next month.


Friday, 23 October
I've got the start-up ready. The output still needs checking.
For eight. Not for when you messaged me last night.
I left before seven. It's fifty-five minutes on the bus, Jun.
I've got things outside here too.
Yes. That's why I moved out of your dining room.
Don't throw it at me every time I ask you for something.
I'm telling you I can't fit everything in at night.
From Monday I'll check the start-up from home. You check the output when you arrive.
Was that breakfast?
Before the bus, yes.

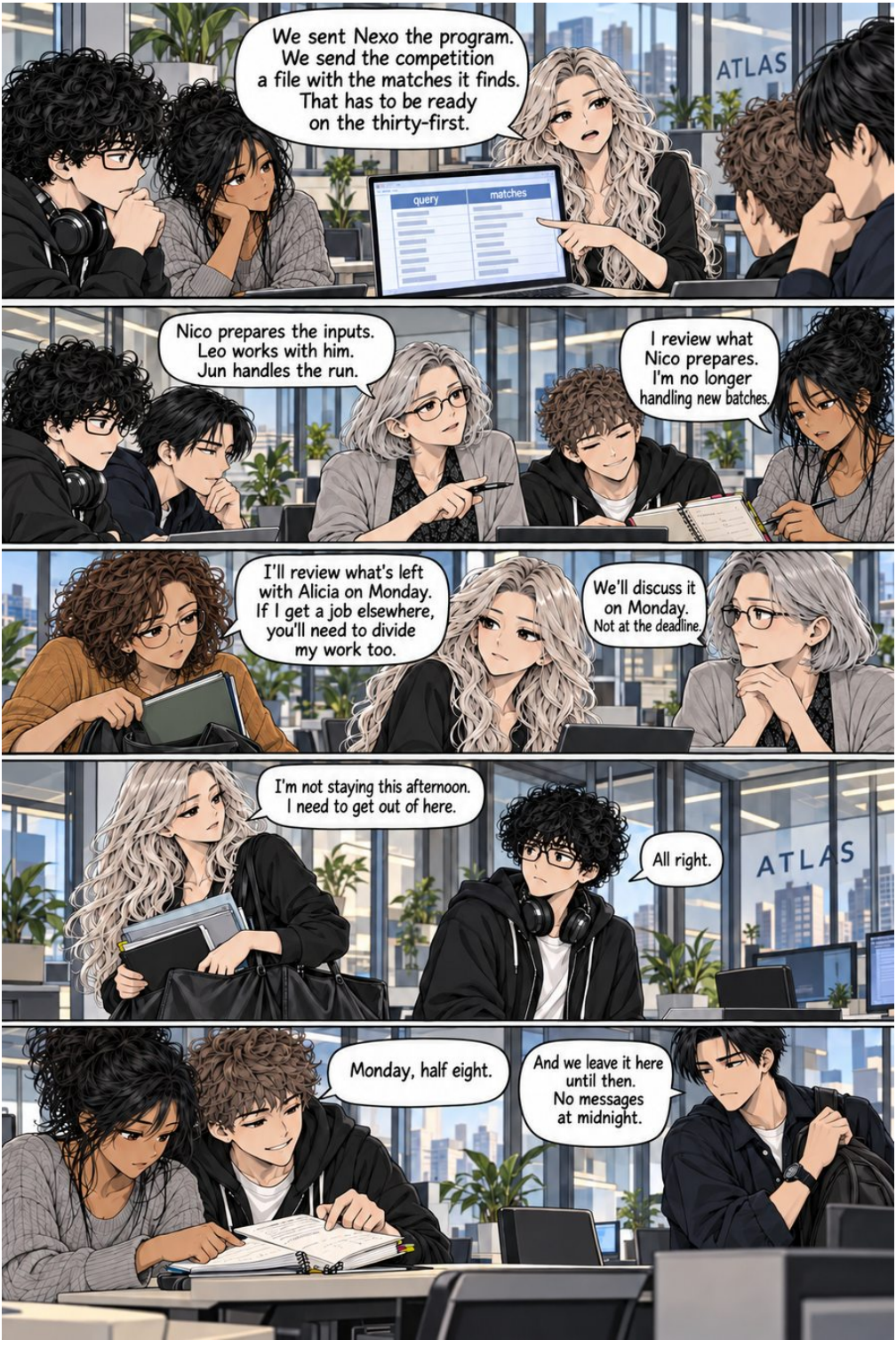

It fits with small batches now. I want to try the whole set.
You've tested a few images. The rest hasn't finished.
Model 1
Model 2
Model ensemble
Partial run
0.82
0.11
0.07
0.73
0.18
0.09
I can get one full batch ready. Not prepare both versions. I don't have those hours today.
Jun, can we fit both runs in?
ATLAS
No. There's only time for one.
I can review the output from one. Then I've got another interview.
If the ensemble improves things, we could be throwing points away.
And if it stops halfway through, we have neither those points nor a submission ready.
Save the ensemble. Today we run the adapted model on its own.
Jun, give it the slot.
Model ensemble
Adapted model


I'll start from here.
Move, I'll do it. I built this.
It's not that difficult. We've spent half the morning on one folder.
Then stop talking to me like I'm an idiot.
You've got two options: you always do it yourself, or you let me learn. But not like this.
You're the one who didn't write down the path, Leo.
All right. Sorry. I went too far.
Now tell me which folder it is. Without taking the keyboard off me.
Your "usual path" means nothing to me.
I'll replace it with the full name.
Usual path
Folder: input
It's finished. I'll let them in.
The output's complete.


Until the thirty-first, I need to stick to handover and review. I can't prepare new batches as well.
We're eight days from the deadline, Mara.
I've been telling you for two weeks.
I've dropped modules. Yesterday I fell asleep over my notes. What else am I supposed to drop?
We'll keep your agreed paid hours. But Nico will have to prepare the inputs.
MARA
• Handover
• Review (until 31)
NICO
• Input preparation
Yes. Let him prepare them. The instructions work.
And I might not stay in November. I need to finish my modules and help at home.
Are you telling me you're leaving?
I'm telling you I can't carry on like this.
I'll change the allocation today. On Monday we'll talk about November.
MARA
• Handover
• Review (until 31)
NICO
• Input preparation
Monday. Not after another submission.


Nexo Fund · ARGOS excluded
Use of unauthorised data and a false declaration.
Reuse of ATLAS's confidential report.
Does November's funding end?
For this line, yes.
ARGOS
So they're no longer ahead of us.
At Nexo. There are teams from all over the world in the competition.
Nexo backed two teams. Not the whole competition. Withdrawing its money from ARGOS doesn't give us victory.
COMPETITION DEADLINE
31 OCTOBER
Competition teams
Asteria
Orion
Kobayashi
Helios
Vector
Daedalus
Novus
Kepler
...
Then let's get back to ours. A week left.
31 October
Preparation
Data
Models
Simulations
Tests
Report
Review
Eight days. And not one more.


My renewal was for November. You've spent weeks telling me it was on track.
That's why I tried to secure that funding stage.
With other people's work and data we couldn't use. Secure it, you say.
I stayed to work on the new project. Now what?
It won't open with that money. I'll speak to the department with you. I can't promise you a place.
You approved it too.
Yes. And I've admitted it. That doesn't give you anything back.
Do we lose this too?
You made that model. Let's separate what's ours from what should never have gone in.
And this time, those of us who'll be left out find out before anything's signed.


We sent Nexo the program. We send the competition a file with the matches it finds. That has to be ready on the thirty-first.
ATLAS
query
matches
Nico prepares the inputs. Leo works with him. Jun handles the run.
I review what Nico prepares. I'm no longer handling new batches.
I'll review what's left with Alicia on Monday. If I get a job elsewhere, you'll need to divide my work too.
We'll discuss it on Monday. Not at the deadline.
I'm not staying this afternoon. I need to get out of here.
All right.
ATLAS
Monday, half eight.
And we leave it here until then. No messages at midnight.


Can you handle the last output? I'll miss the bus.
Yes. I'll mark it up before I leave.
1
Coming to hear the mix on Saturday? One hour.
If you don't make me carry another box.
Output review · Sent to Alicia
Results
Comparisons
ARGOS
Variants
Personal_test (postponed)
3. Take Care of Yourself
They've caught ARGOS cheating. I'm happier about it than I want to admit. But we've still got eight days and errors to review.
Today I saved an idea I liked for later.
It annoyed me.
Nico finished without me taking the keyboard away.
I'm not entering another challenge before the deadline.

What We Can Finish with the Remaining Slot



GPU, requests to a service and people's time.

Leo's ensemble has a reason behind it: one model looks at detail and the other at wider context, and some errors differ. But to decide whether it helps, they must run it, compare it and have time to review the output.

How we would test the ensemble

They would combine both models' scores. They might use different signals or count the same error twice; they could even spoil cases one model already solved. These pages show only a partial run. Claiming an improvement would require comparing it with the baseline on the same examples. For now, keeping what they intended to test matters, because setting the idea aside does not prove it was bad either.

Where it runs out of memory

Each model fits separately. Processing the planned batch together, the models and intermediate data exhaust available memory. Reducing the number of simultaneous images lets the run fit, but requires more batches and takes longer here. The test with a few images establishes a technical possibility; they still need to complete the route with all the input the team needs.

Who reviews the output

The slot allows one full run. Mara can prepare one batch and Nadia review one output before her interview. More memory would help the machine, but their hours would stay the same. They choose the adapted version they can run and review before the deadline. I would consider that decision with all those conditions in view: they have finished a feasible piece of work and kept another experiment, still without a result.

Sources: · GPU memory · Transfer learning

Being Right Does Not Give Permission to Treat People Badly



Listening and waiting for the next line are part of rehearsal too.

Leo knows how to do a difficult part of the work and grows impatient when Nico needs to run the process himself. This time the help comes with a remark that belittles him.

Nico does not need replacing

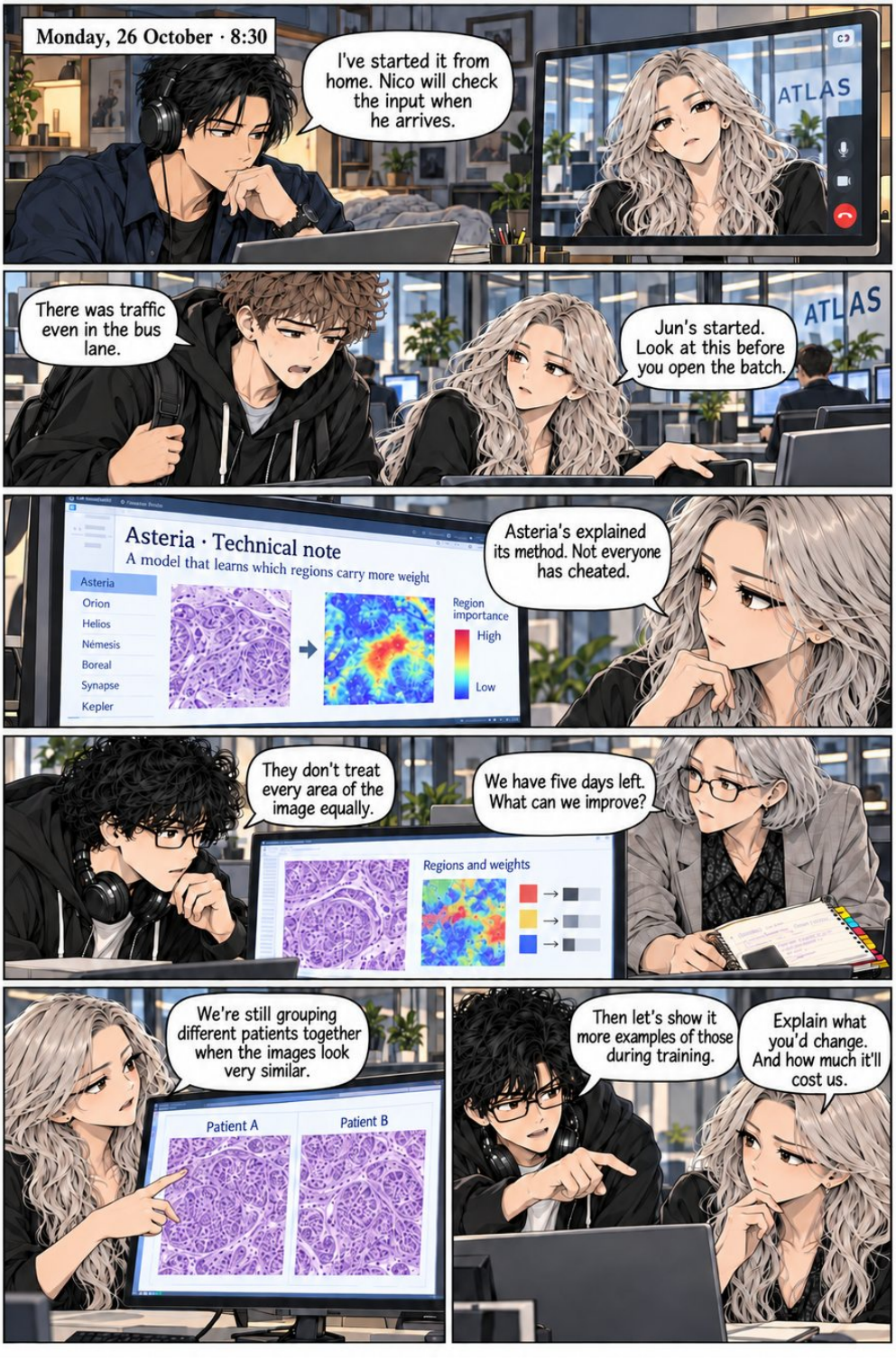
If Leo takes the keyboard whenever something takes time, nobody else will be able to take over. Nico's response sets a limit. To repair the way he treated him, Leo must let him work and attend to what he discovers, as well as correcting an instruction.

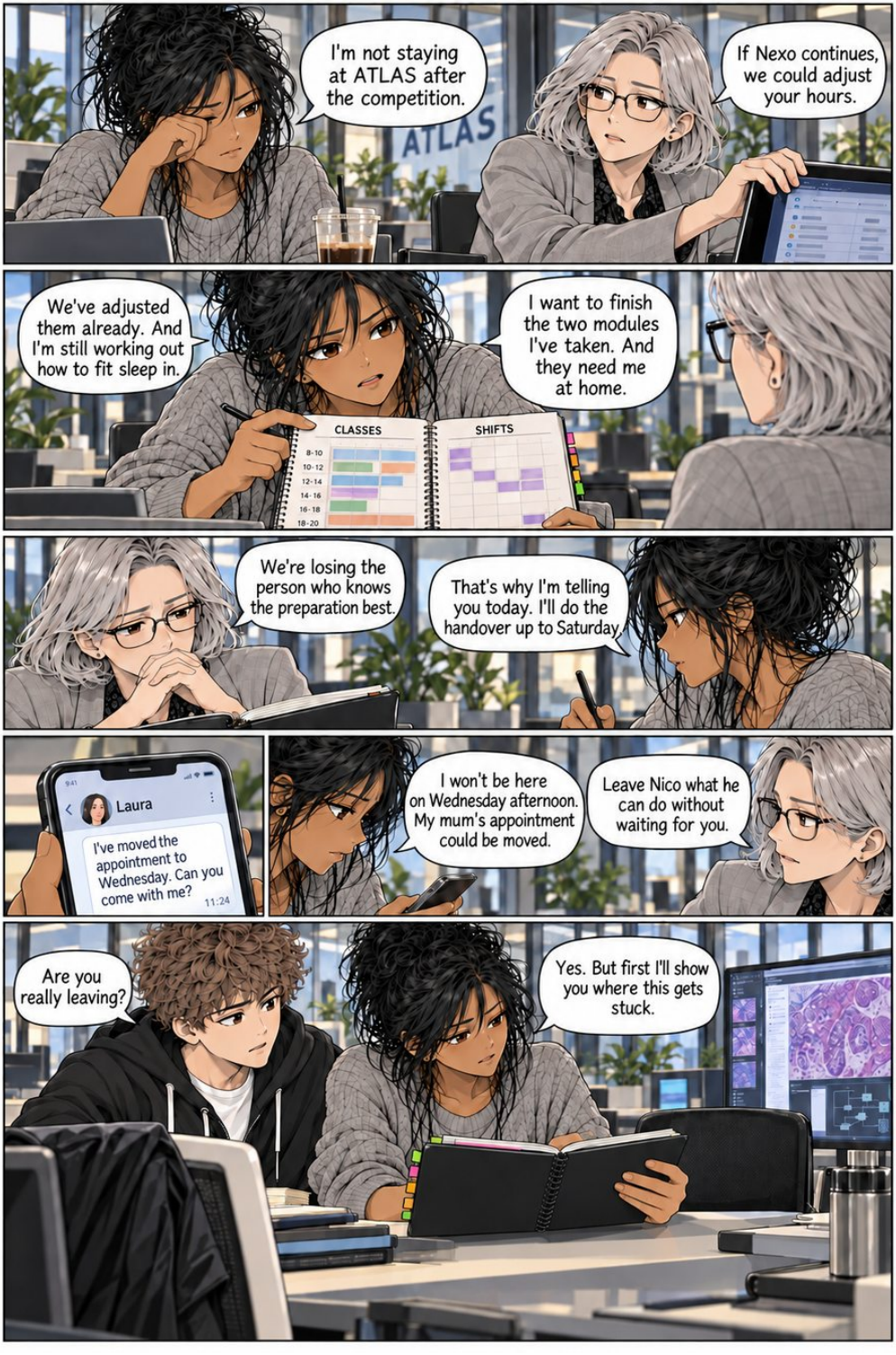
Recovering files does not undo the fright

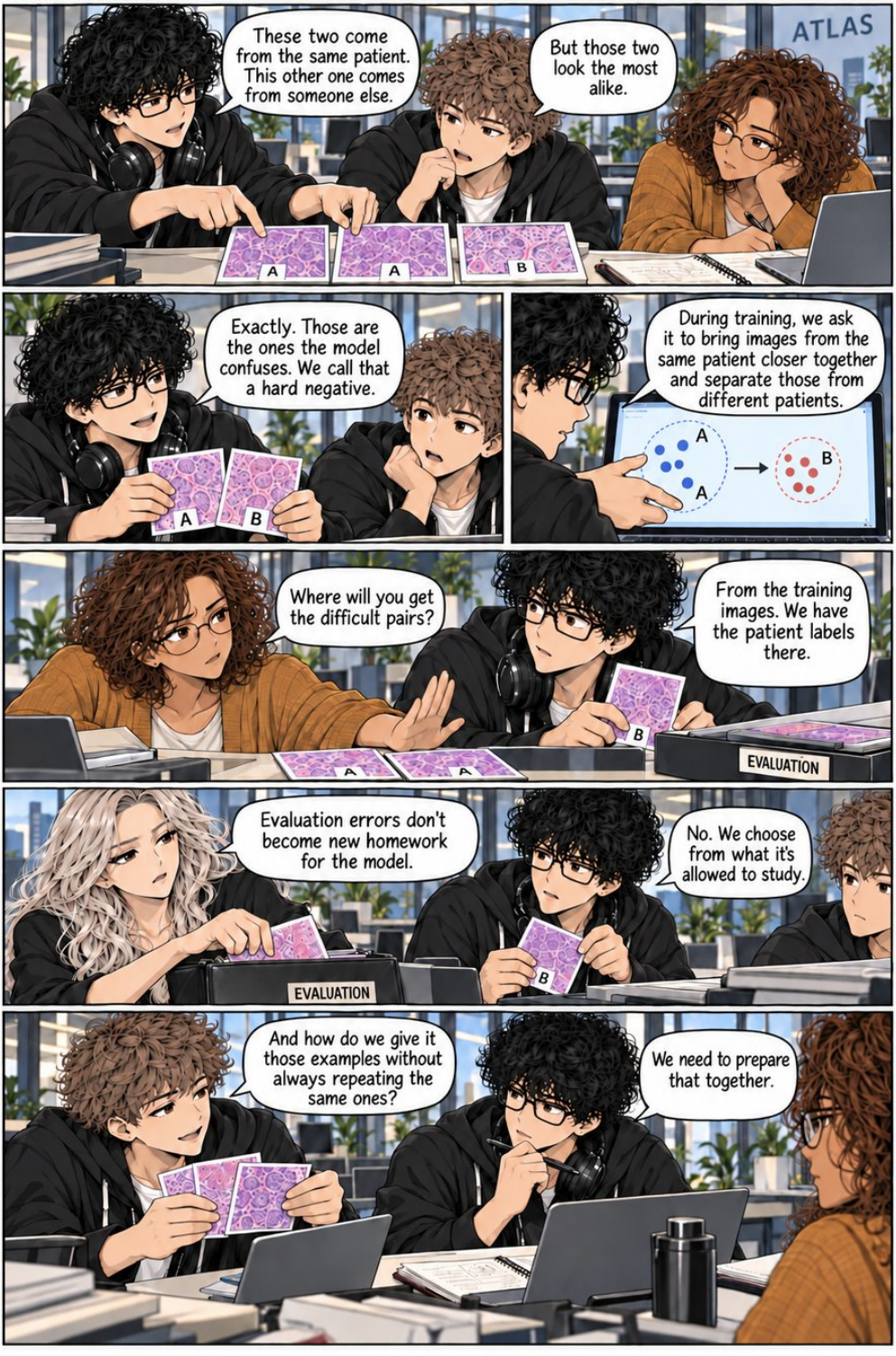
Nico recovers the work after the theft because it was stored away from the laptop. He has his files back, but still lacks his tool and has had a real fright. Jun offers him a night at his place; Nico is grateful and prefers to return to his room. The offer counts even though he does not accept it this time.

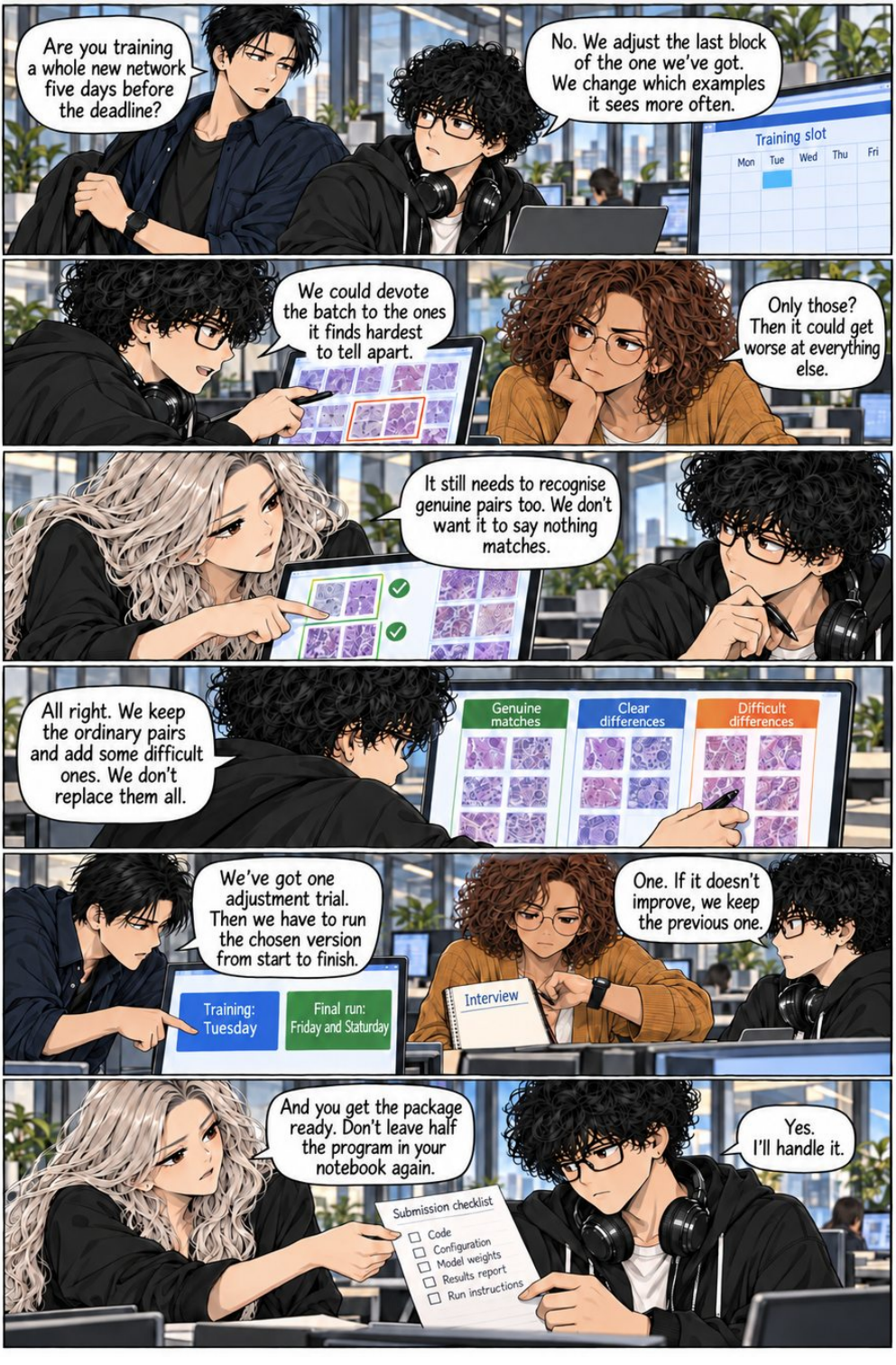
Mara still works there

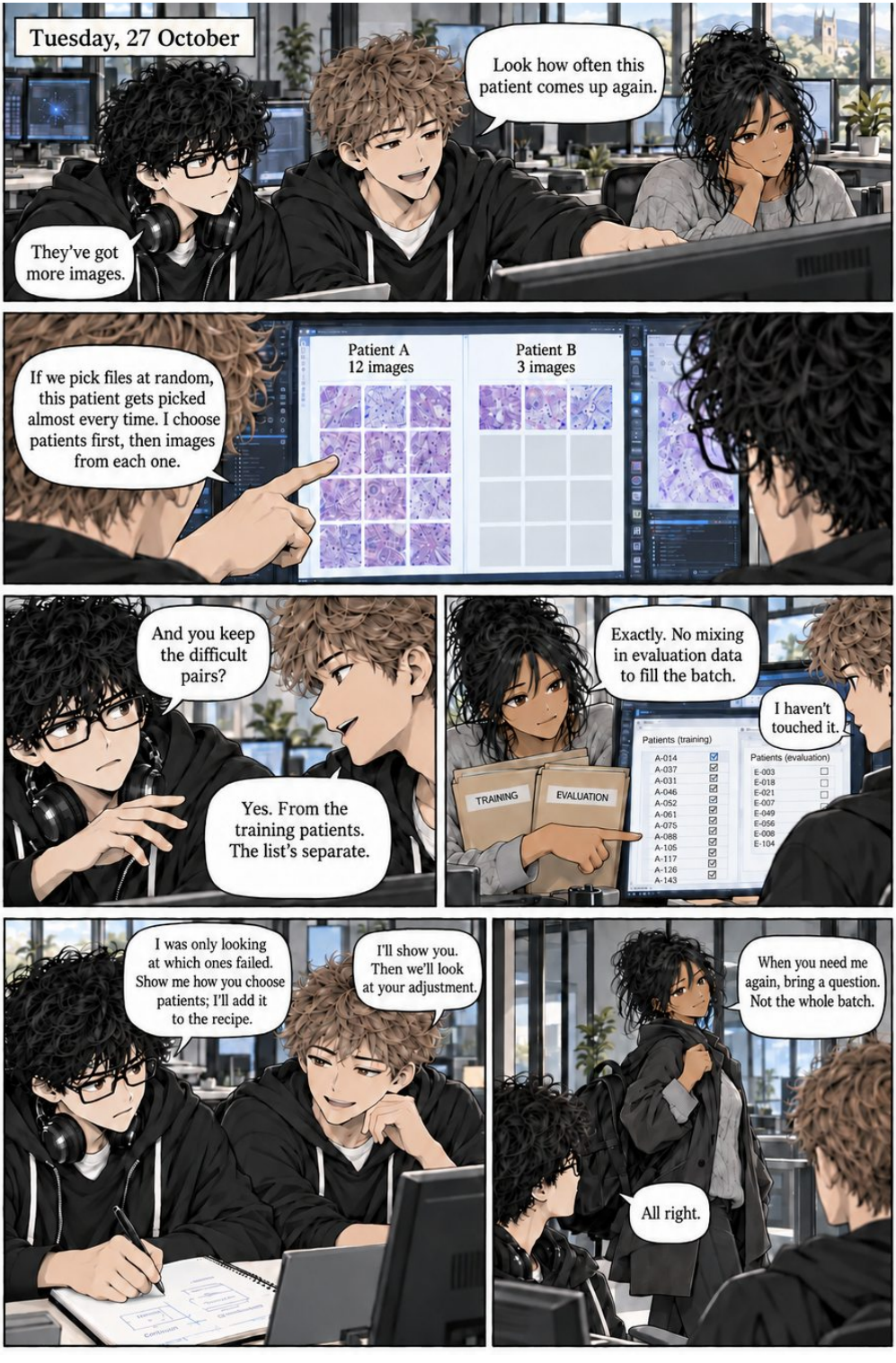
Leo feels watched after the kiss. Mara carries on with her work. When they talk, he admits he likes Vera too, and she makes clear that one kiss does not make them a couple. She wants to go slowly and needs space. Her family, studies and the limit on hours she had already explained still fill her life.


Monday, 26 October · 8:30
I've started it from home. Nico will check the input when he arrives.
ATLAS
There was traffic even in the bus lane.
Jun's started. Look at this before you open the batch.
ATLAS
Asteria · Technical note
A model that learns which regions carry more weight
Asteria
Orion
Helios
Némesis
Boreal
Synapse
Kepler
Region importance
High
Low
Asteria's explained its method. Not everyone has cheated.
They don't treat every area of the image equally.
We have five days left. What can we improve?
Regions and weights
We're still grouping different patients together when the images look very similar.
Patient A
Patient B
Then let's show it more examples of those during training.
Explain what you'd change. And how much it'll cost us.

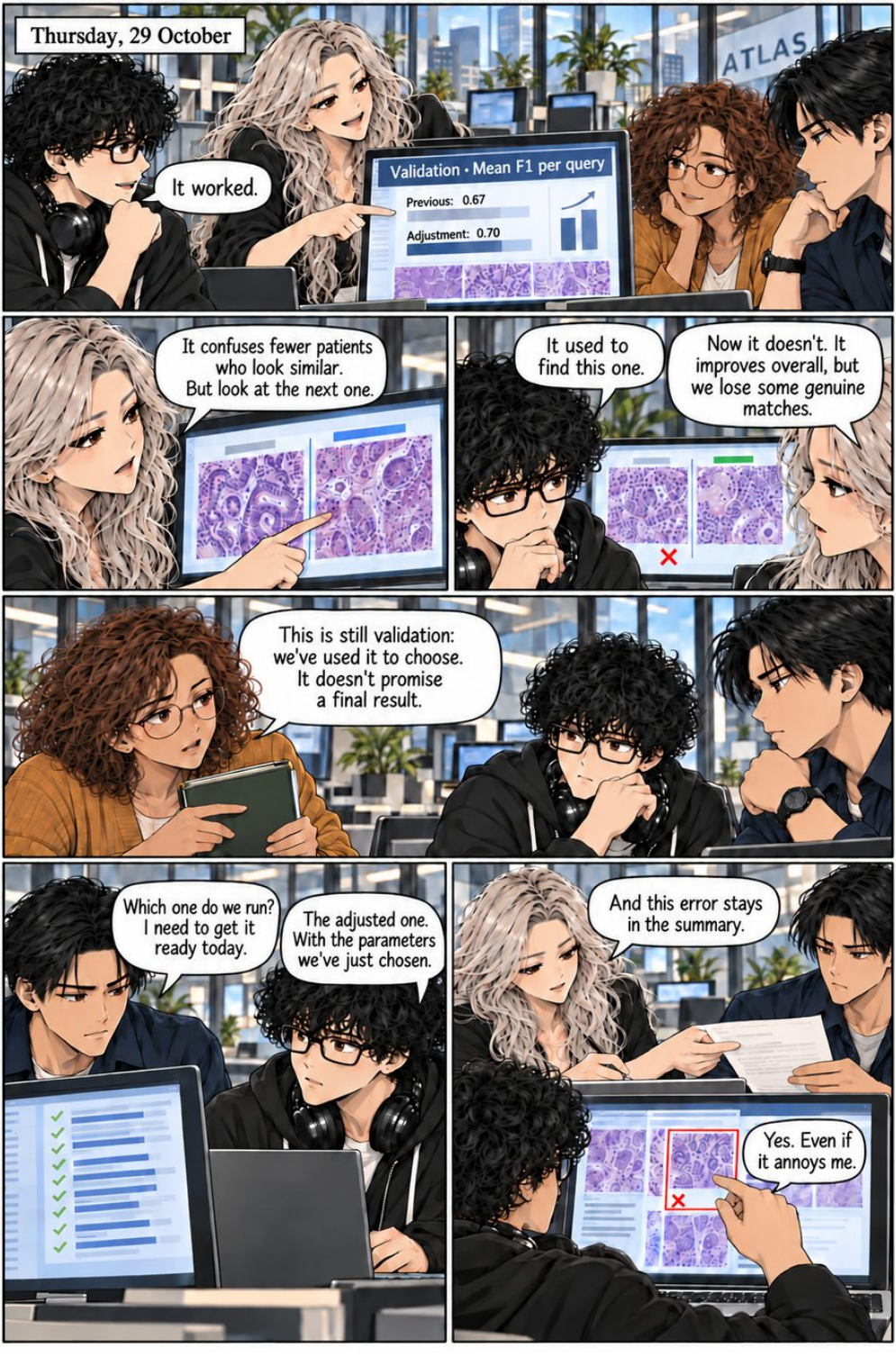

I'm not staying at ATLAS after the competition.
If Nexo continues, we could adjust your hours.
ATLAS
We've adjusted them already. And I'm still working out how to fit sleep in.
I want to finish the two modules I've taken. And they need me at home.
CLASSES
SHIFTS
8-10
10-12
12-14
14-16
16-18
18-20
We're losing the person who knows the preparation best.
That's why I'm telling you today. I'll do the handover up to Saturday.
9:41
Laura
I've moved the appointment to Wednesday. Can you come with me?
11:24
I won't be here on Wednesday afternoon. My mum's appointment could be moved.
Leave Nico what he can do without waiting for you.
Are you really leaving?
Yes. But first I'll show you where this gets stuck.


These two come from the same patient. This other one comes from someone else.
But those two look the most alike.
ATLAS
A
A
B
Exactly. Those are the ones the model confuses. We call that a hard negative.
A
B
During training, we ask it to bring images from the same patient closer together and separate those from different patients.
A
A
B
Where will you get the difficult pairs?
From the training images. We have the patient labels there.
A
A
B
EVALUATION
Evaluation errors don't become new homework for the model.
No. We choose from what it's allowed to study.
EVALUATION
B
And how do we give it those examples without always repeating the same ones?
We need to prepare that together.


Are you training a whole new network five days before the deadline?
No. We adjust the last block of the one we've got. We change which examples it sees more often.
Training slot
Mon
Tue
Wed
Thu
Fri
We could devote the batch to the ones it finds hardest to tell apart.
Only those? Then it could get worse at everything else.
It still needs to recognise genuine pairs too. We don't want it to say nothing matches.
All right. We keep the ordinary pairs and add some difficult ones. We don't replace them all.
Genuine matches
Clear differences
Difficult differences
We've got one adjustment trial. Then we have to run the chosen version from start to finish.
One. If it doesn't improve, we keep the previous one.
Training: Tuesday
Final run: Friday and Staturday
Interview
And you get the package ready. Don't leave half the program in your notebook again.
Yes. I'll handle it.
Submission checklist
Code
Configuration
Model weights
Results report
Run instructions


Tuesday, 27 October
Look how often this patient comes up again.
They've got more images.
If we pick files at random, this patient gets picked almost every time. I choose patients first, then images from each one.
Patient A
12 images
Patient B
3 images
And you keep the difficult pairs?
Yes. From the training patients. The list's separate.
Exactly. No mixing in evaluation data to fill the batch.
I haven't touched it.
TRAINING
EVALUATION
Patients (training)
A-014
A-037
A-031
A-046
A-052
A-061
A-075
A-088
A-105
A-117
A-126
A-143
Patients (evaluation)
E-003
E-018
E-021
E-007
E-049
E-056
E-008
E-104
I was only looking at which ones failed. Show me how you choose patients; I'll add it to the recipe.
I'll show you. Then we'll look at your adjustment.
When you need me again, bring a question. Not the whole batch.
All right.


Wednesday 28 · Waiting room
If you need to go...
I said I'd come with you today.
You haven't slept much again.
And I'm trying to stop it being like that all the time.
Where's the configuration we're going to use?
I'm just finishing checking something about the adjustment.
The configuration, Leo. You took on that part.
I need ten minutes.
You said that an hour ago.
All right. I'll finish it and let you know.
When it's ready. Not when you think it is.

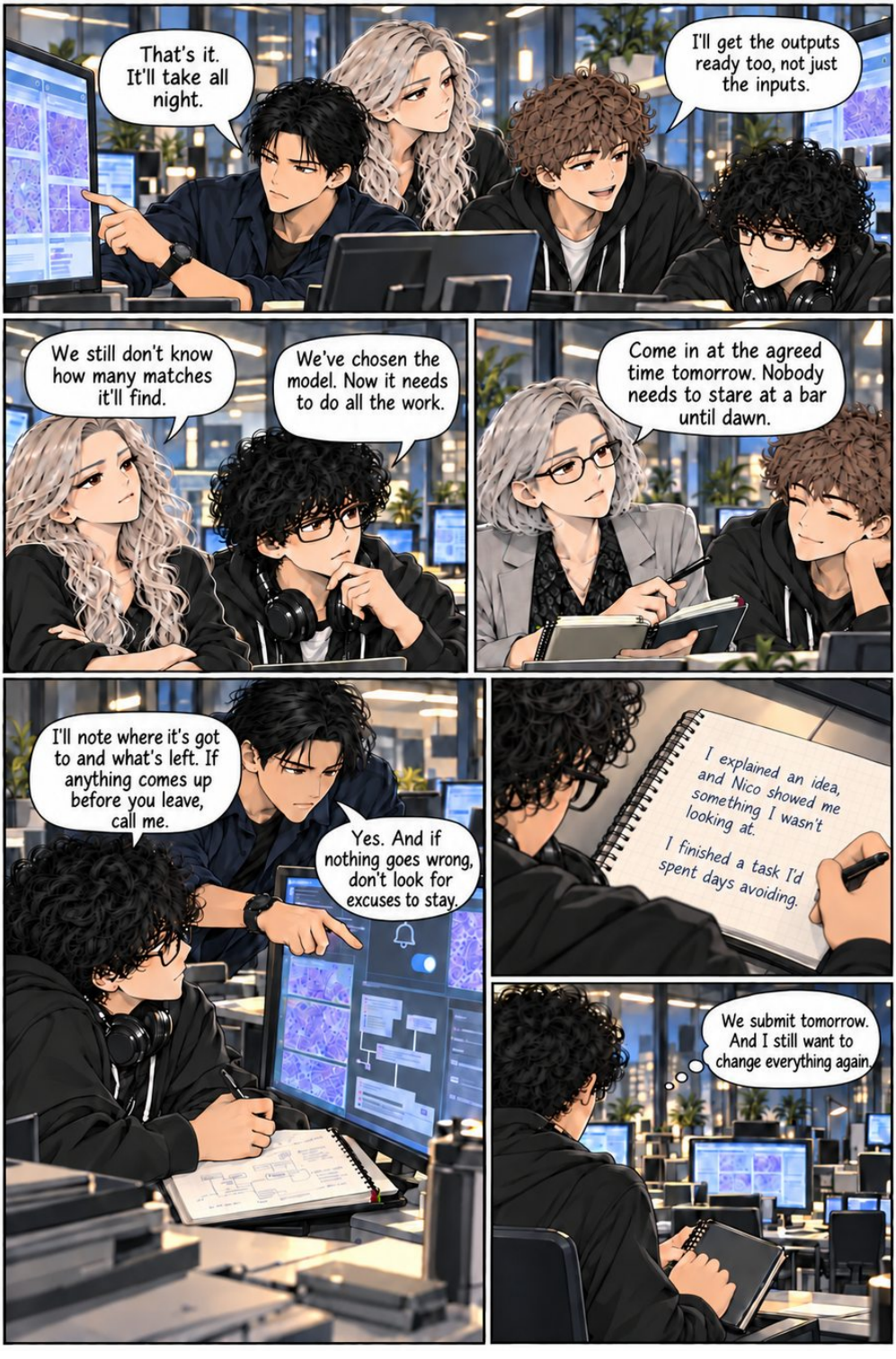

Thursday, 29 October
ATLAS
It worked.
Validation · Mean F1 per query
Previous: 0.67
Adjustment: 0.70
It confuses fewer patients who look similar. But look at the next one.
It used to find this one.
Now it doesn't. It improves overall, but we lose some genuine matches.
This is still validation: we've used it to choose. It doesn't promise a final result.
Which one do we run? I need to get it ready today.
The adjusted one. With the parameters we've just chosen.
And this error stays in the summary.
Yes. Even if it annoys me.


I got it.
I start on the sixteenth of November.
In the other group?
Yes. I've wanted to work in evaluation since I started the master's.
I'm glad.
Who takes over your work here?
We'll get the deadline work ready together. After that, Inés has to decide.
Your contract extension runs until the thirteenth of November. We'll have the handover ready by then.
Then explain that bit you always review yourself today.
One part today.
You won't learn it all in an afternoon.
She's happy to be leaving.
I was only thinking about what we'd be missing.


That afternoon · Rehearsal
Leo, you haven't heard the question yet.
I knew what was coming.
So did we. But let it happen.
All right. I'll wait until he finishes.
I can't make Saturday. It's the competition deadline, and I need to be with the team.
I thought I mattered more to you.
I thought we were talking about rehearsal. Now I don't know what to say.
You had a scene with the others.
I know. I should have said sooner.
We'll change the cast for that rehearsal. You'll have to catch up on Thursday.
Send me what changes. I'll prepare it.


What if I try the regions idea tonight?
Asteria · Technical note
4. Think for Yourself
Asteria · Technical note
Just because it works for them doesn't mean I can fit it in tonight.
Friday, 30 October
ATLAS
Now I can start it without asking you for three files.
That's what I stuck to last night.
All night?
No. And I left an idea untested. It's still nagging at me.


If the source changes, don't assume these crops will work. Check before you run it.
And if I don't know what to change, I say so.
I've heard you're not staying.
ATLAS
I would have told you. We've had a rough week.
Is it to do with...?
With sleeping, studying and my family. Not everything I do is about you.
Right. Sorry.
And I still need space between us.
Mara, before you go: do I save this here?
Coming.
ATLAS


That's it. It'll take all night.
I'll get the outputs ready too, not just the inputs.
We still don't know how many matches it'll find.
We've chosen the model. Now it needs to do all the work.
Come in at the agreed time tomorrow. Nobody needs to stare at a bar until dawn.
I'll note where it's got to and what's left. If anything comes up before you leave, call me.
Yes. And if nothing goes wrong, don't look for excuses to stay.
I explained an idea, and Nico showed me something I wasn't looking at.
I finished a task I'd spent days avoiding.
We submit tomorrow. And I still want to change everything again.

Which Examples It Studies More Often



Images grouped by patient before splitting the data.

The system confuses very similar images from different patients. Leo wants to show it more examples of that difficulty. Nico looks at how they are chosen and finds another important detail: some patients have far more files than others.

What makes a pair difficult

Two images from the same patient form a positive pair. If they come from different people, it is negative; when they look very similar and the model confuses them, we have a hard negative. Training tries to bring representations of genuine pairs closer and separate those from different people. What changes is the numerical values used to compare them. The original images stay where they are.

How they prepare the mix

If we devote every attempt to very difficult differences, we may worsen recognition of other pairs. ATLAS keeps ordinary examples and includes some difficult ones, taken from training patients with known labels. They keep the exam's failures for checking the chosen solution. Using them as new lessons would give useful information for adjustment, but would consume precisely the separation they were preserving.

What Nico proposes

Picking files at random favours the patient with more images. Nico chooses patients first, then images, even though the folder is the same. That changes who the model sees more often. The adjustment moves from 0.67 to 0.70 in validation and loses some matches it previously found. They select that version and record the loss. It is a comparison for choosing how to continue; the competition result remains unknown.

Sources: · Difficult examples · Sampling

Listening to an Idea That Was Not His



A sofa, a table and several uses for the same space.

Nico sees that some patients appear too often in training. Leo was looking at the model's errors and had overlooked how the examples were chosen.

Making amends in a way that changes the work

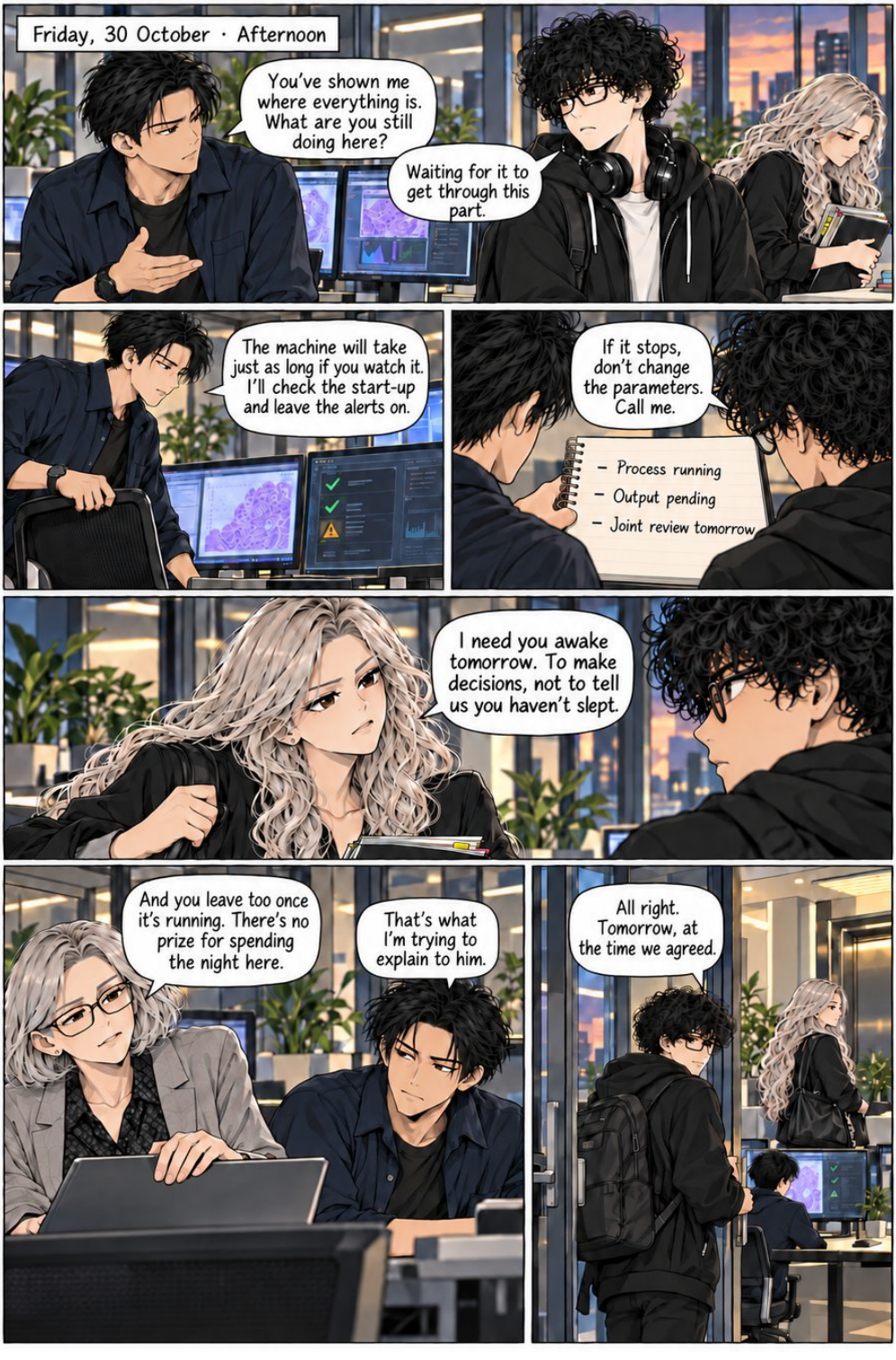
After speaking badly to him, Leo has a concrete chance to listen. Nico proposes choosing patients first, then images. Leo asks how he set it up and asks him to show him. This time the keyboard can stay where it is: there is something he had not seen and his colleague has already solved.

People can leave

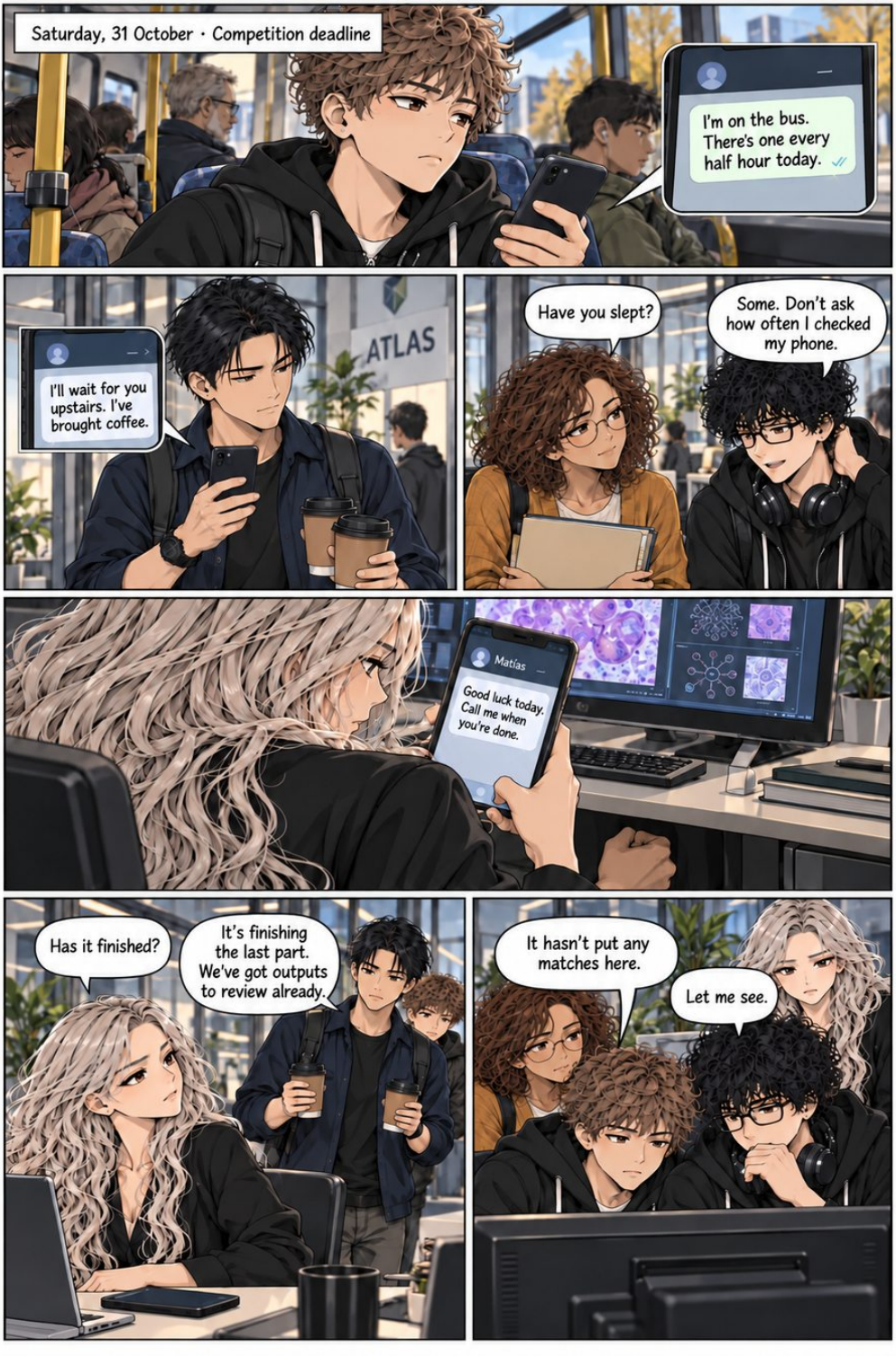
Mara needs to study, sleep and help her family. Nadia has found another job. Leo feels the loss and tends to look for which part is about him. Listening to their reasons lets them prepare the handover without turning departure into disloyalty.

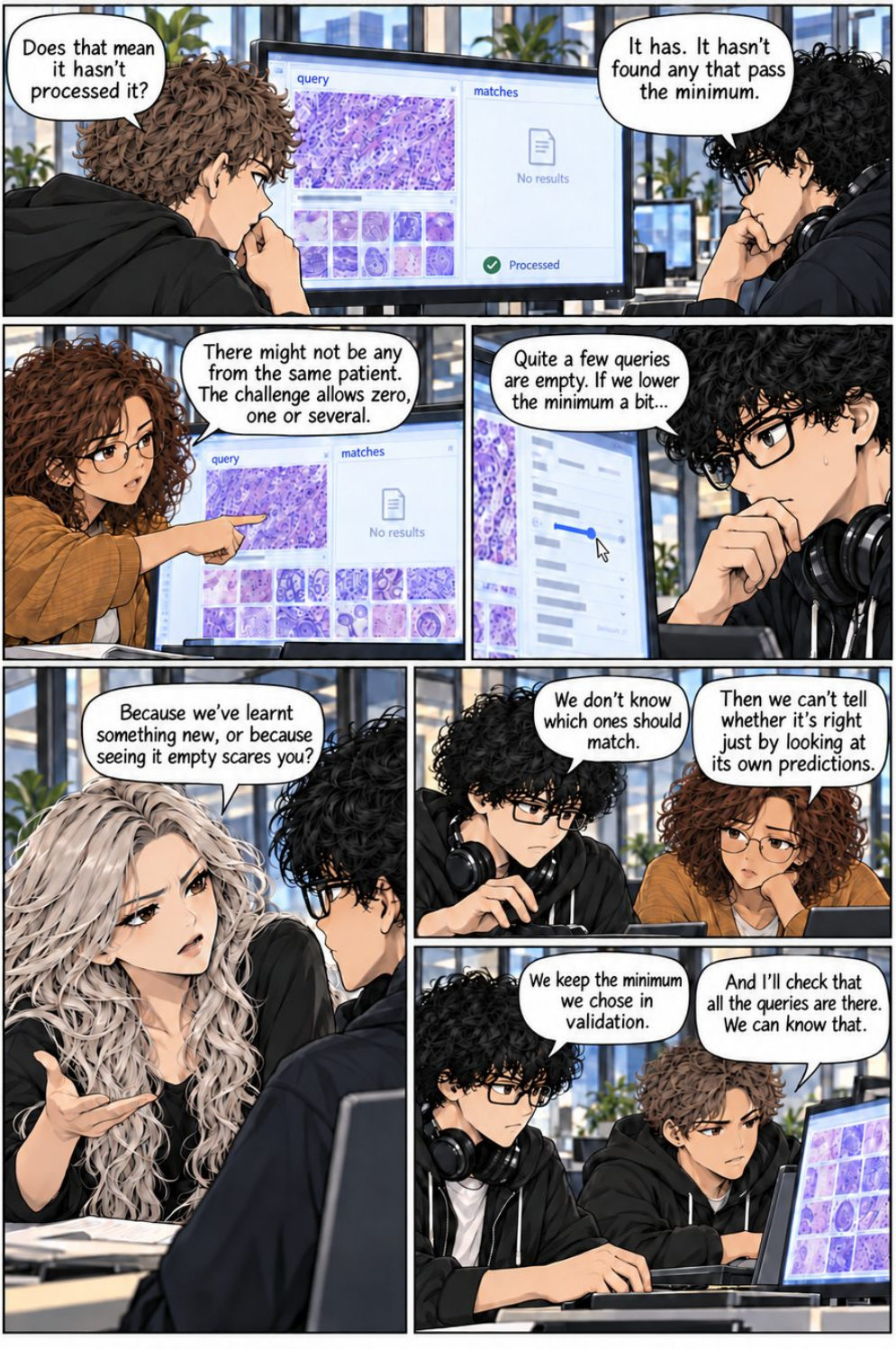
Vera expected something more

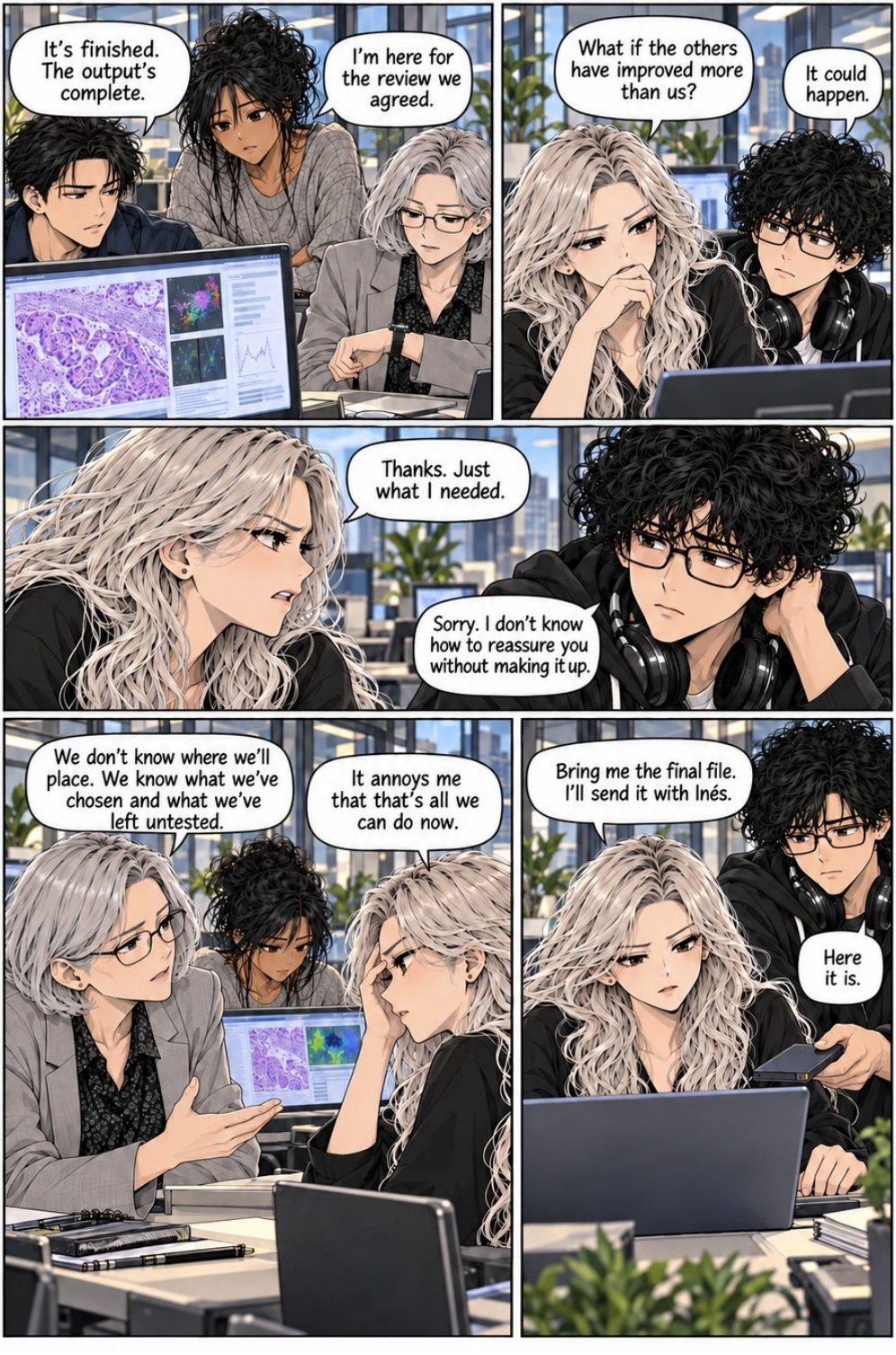
Vera says she thought she mattered more to him, and Leo is bewildered. He thought he was giving notice of an absence; she is counting on him for rehearsal and also seems to expect another kind of attention. They have a conversation pending there. Meanwhile, the others lack a partner for a scene. Leo acknowledges he gave late notice and asks for the changes so he can prepare them. Having to be with ATLAS that Saturday did not stop him giving notice sooner.

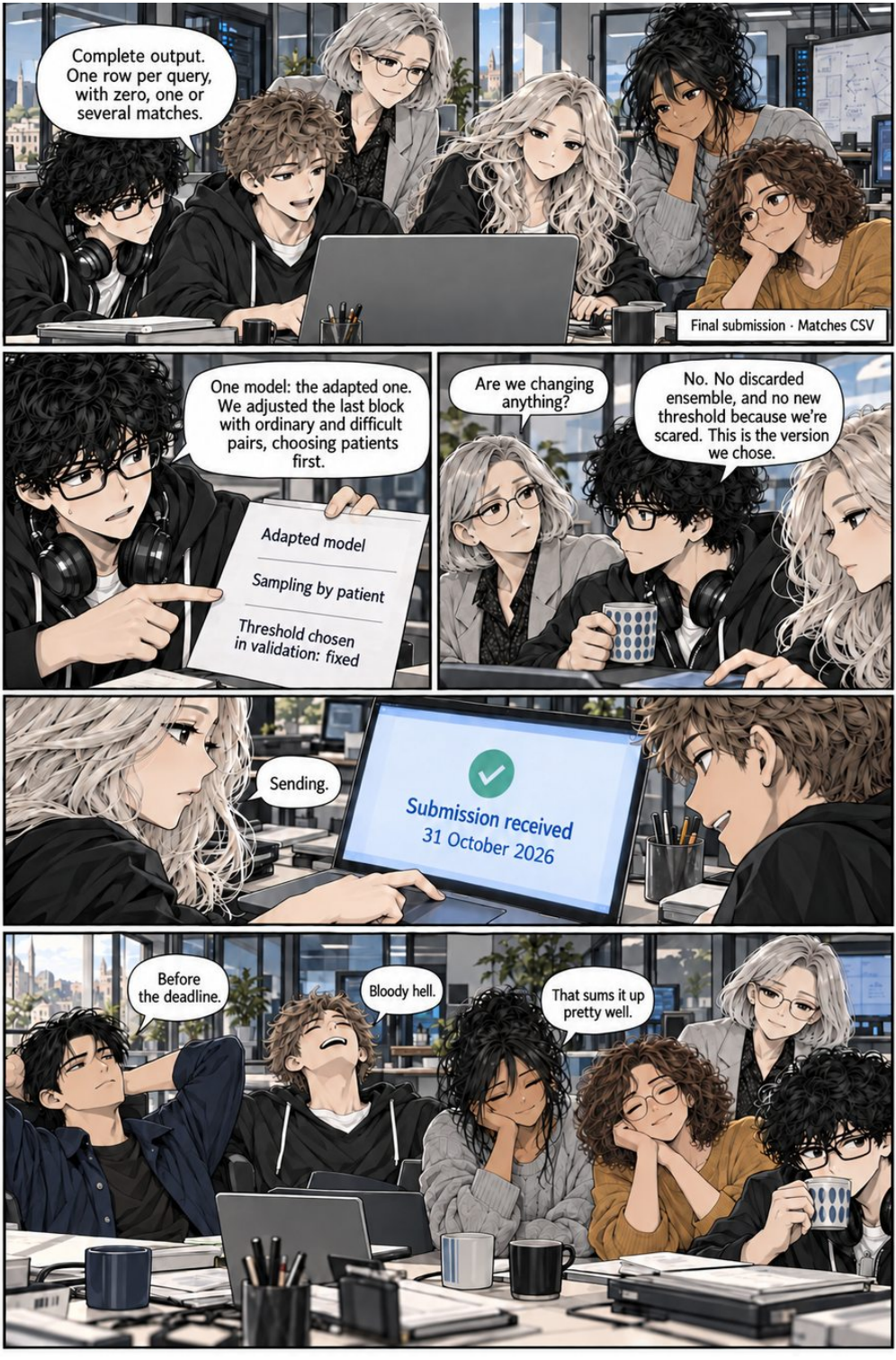

Friday, 30 October · Afternoon
You've shown me where everything is. What are you still doing here?
Waiting for it to get through this part.
The machine will take just as long if you watch it. I'll check the start-up and leave the alerts on.
If it stops, don't change the parameters. Call me.
– Process running
– Output pending
– Joint review tomorrow
I need you awake tomorrow. To make decisions, not to tell us you haven't slept.
And you leave too once it's running. There's no prize for spending the night here.
That's what I'm trying to explain to him.
All right. Tomorrow, at the time we agreed.


Jun
Start-up fine. I'm leaving. If anything fails, the alert will go off.
I can check that. I don't need to wake him up to tell me the same thing.
How far has it got now?
Now, this.
I've done the handover. I'm not running away.
3. Take Care of Yourself
Now let's see if I can get to sleep.

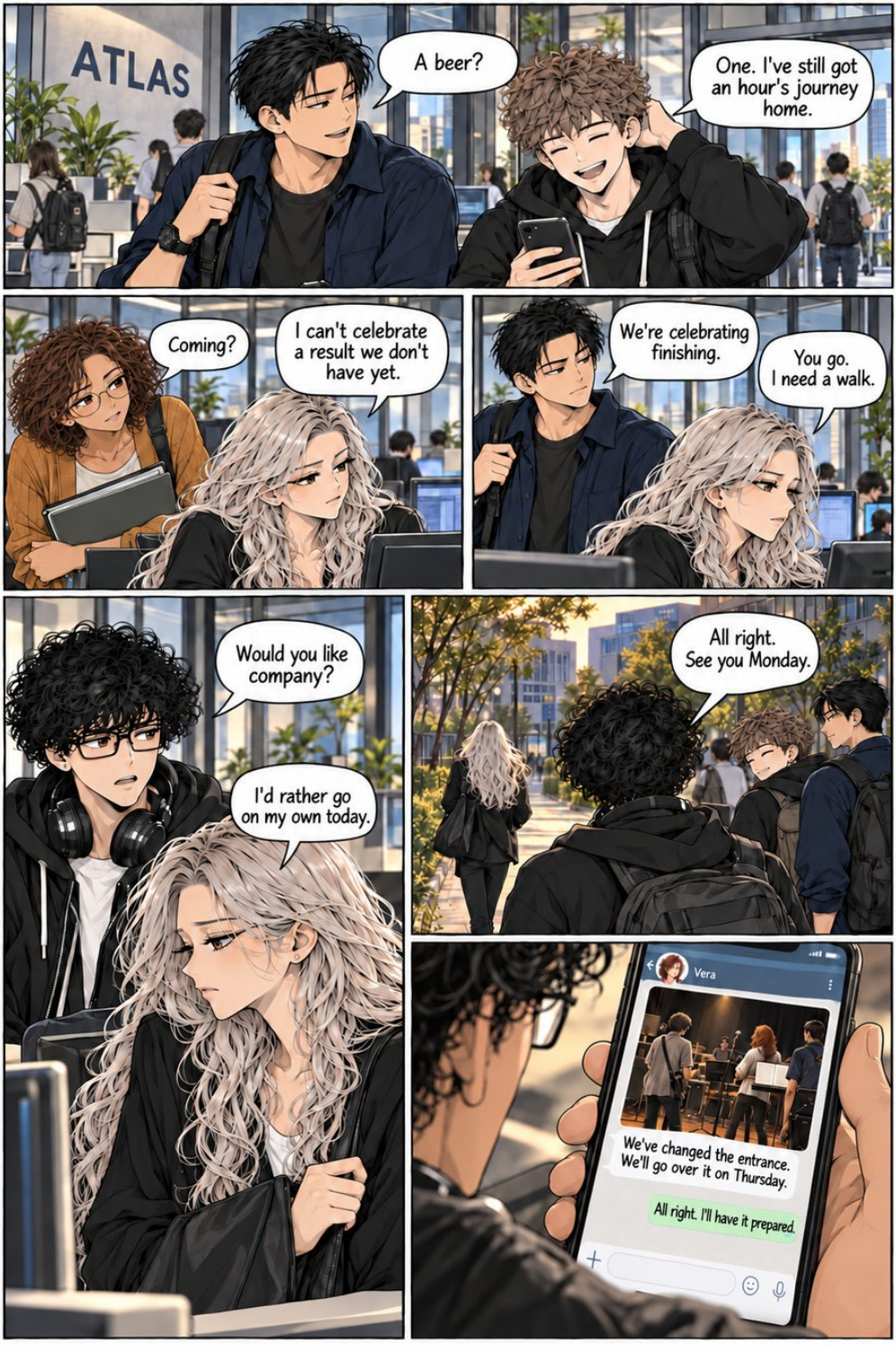

Saturday, 31 October · Competition deadline
I'm on the bus. There's one every half hour today.
ATLAS
I'll wait for you upstairs. I've brought coffee.
Have you slept?
Some. Don't ask how often I checked my phone.
Matías
Good luck today. Call me when you're done.
Has it finished?
It's finishing the last part. We've got outputs to review already.
It hasn't put any matches here.
Let me see.


Does that mean it hasn't processed it?
query
matches
No results
Processed
It has. It hasn't found any that pass the minimum.
There might not be any from the same patient. The challenge allows zero, one or several.
query
matches
No results
Quite a few queries are empty. If we lower the minimum a bit...
Because we've learnt something new, or because seeing it empty scares you?
We don't know which ones should match.
Then we can't tell whether it's right just by looking at its own predictions.
We keep the minimum we chose in validation.
And I'll check that all the queries are there. We can know that.


It's finished. The output's complete.
I'm here for the review we agreed.
What if the others have improved more than us?
It could happen.
Thanks. Just what I needed.
Sorry. I don't know how to reassure you without making it up.
We don't know where we'll place. We know what we've chosen and what we've left untested.
It annoys me that that's all we can do now.
Bring me the final file. I'll send it with Inés.
Here it is.

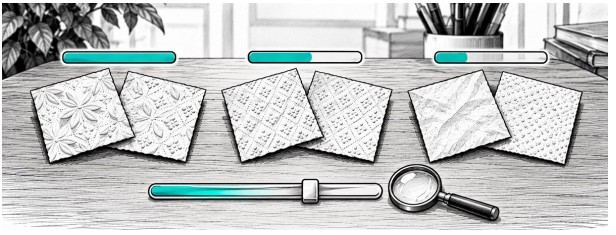

Complete output. One row per query, with zero, one or several matches.
Final submission · Matches CSV
One model: the adapted one. We adjusted the last block with ordinary and difficult pairs, choosing patients first.
Adapted model
Sampling by patient
Threshold chosen in validation: fixed
Are we changing anything?
No. No discarded ensemble, and no new threshold because we're scared. This is the version we chose.
Sending.
Submission received
31 October 2026
Before the deadline.
Bloody hell.
That sums it up pretty well.


Thanks for leaving it in Nico's hands.
He did the last preparation. It's not just mine any more.
This wasn't in the handover.
That I am taking with me.
If you feel like coming to...
I'm going home today. Another day, we'll see.
Of course.
But tell me when the rankings come out. I do want to see that.
See you.
ATLAS


ATLAS
A beer?
One. I've still got an hour's journey home.
Coming?
I can't celebrate a result we don't have yet.
We're celebrating finishing.
You go. I need a walk.
Would you like company?
I'd rather go on my own today.
All right. See you Monday.
Vera
We've changed the entrance. We'll go over it on Thursday.
All right. I'll have it prepared.


Don't talk to me about the model now.
Show him the mix. The short version, though.
I think I'm out of time.
A bit. Like last week's drums.
We've submitted. I don't know if it'll be enough. I nearly changed the minimum because I was scared of submitting queries with no matches.
Mara has left. Alicia wanted to be alone. It was hard not to try to fix both things by talking.
That's enough for today.

What a Query with No Matches Means



Illustrated textures: similarity scores and an acceptance selector.

The run ends and some queries are empty. Nico asks whether they were processed. Before touching the model, we should settle that doubt: an answer may be missing, or there may be an answer saying it found no pair.

Checking it went through the program

The gallery may contain zero, one or several images from the same patient. An empty result is correct if there were none and an error if the program missed a genuine pair. It is another matter if the query does not appear in the file at all. The team can check coverage and completion: every expected query must have an output, even if that output is empty.

What looking at predictions tells us

ATLAS does not know the final exam's correct pairs. Leo can lower the minimum similarity and get a file full of proposals, but he would still lack the reference answer to measure whether it improved. That change would respond to the unease the empty results cause him. They keep the value chosen in validation, where they could compare correct answers and errors before reaching this point.

The method and file they submit

The team keeps a pretrained network and adjusts its last block with ordinary and difficult pairs. Patient selection prevents those with more images from dominating. It compares representations and applies the threshold chosen in validation. The model ensemble remained unfinished and is excluded from the submission. Alicia and Inés send a CSV containing each query's matches; they also keep the model, preparation and configuration to repeat the process. Nexo receives its report separately.

Sources: · WSI challenge · Thresholds

Keeping the Decision Until Submission



Preparing an activity and reserving an hour for it in the week.

The run has finished and Leo has no further test to take away his fear. He has a chosen version, answers he does not know and a submission deadline. Submitting makes him finish before feeling sure of his ranking.

Leaving a handover before going home

Leo prepares what is needed and still wants to stay and watch. At home, he nearly asks Jun about something he can check himself. He stops the message and sleeps a little, though he still watches his phone. That rest rests on earlier work: someone else can continue without him being there and without every doubt of Leo's becoming an interruption for his colleague.

Looking at empty results without filling them out of fear

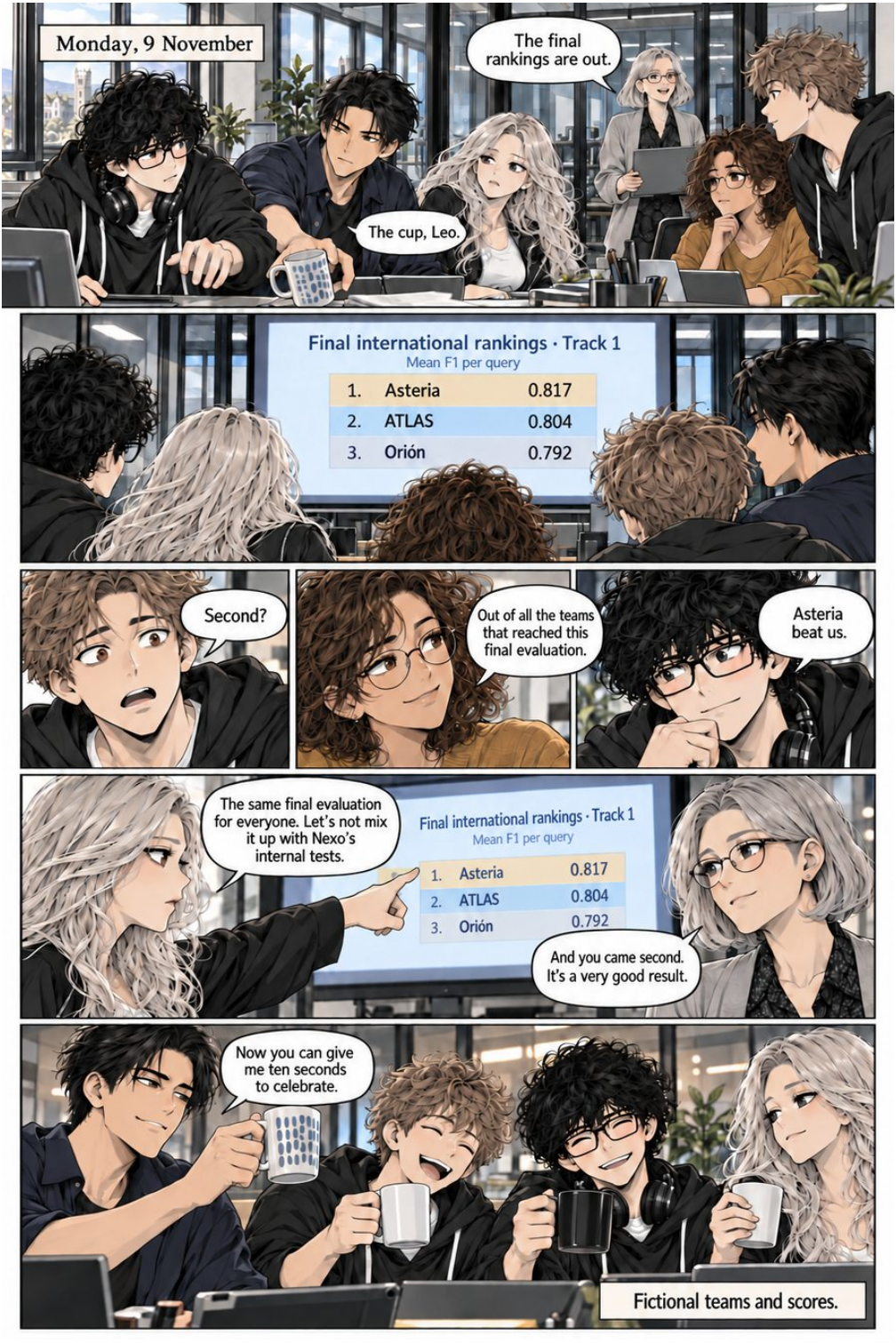
He is about to lower the threshold to obtain more matches. Then he acknowledges he does not know which ones should appear. With Alicia, he cannot find a sentence guaranteeing all will be well either. He can share the unease they both feel without inventing a result. Keeping their decision requires accepting a discomfort another program change would only have concealed.

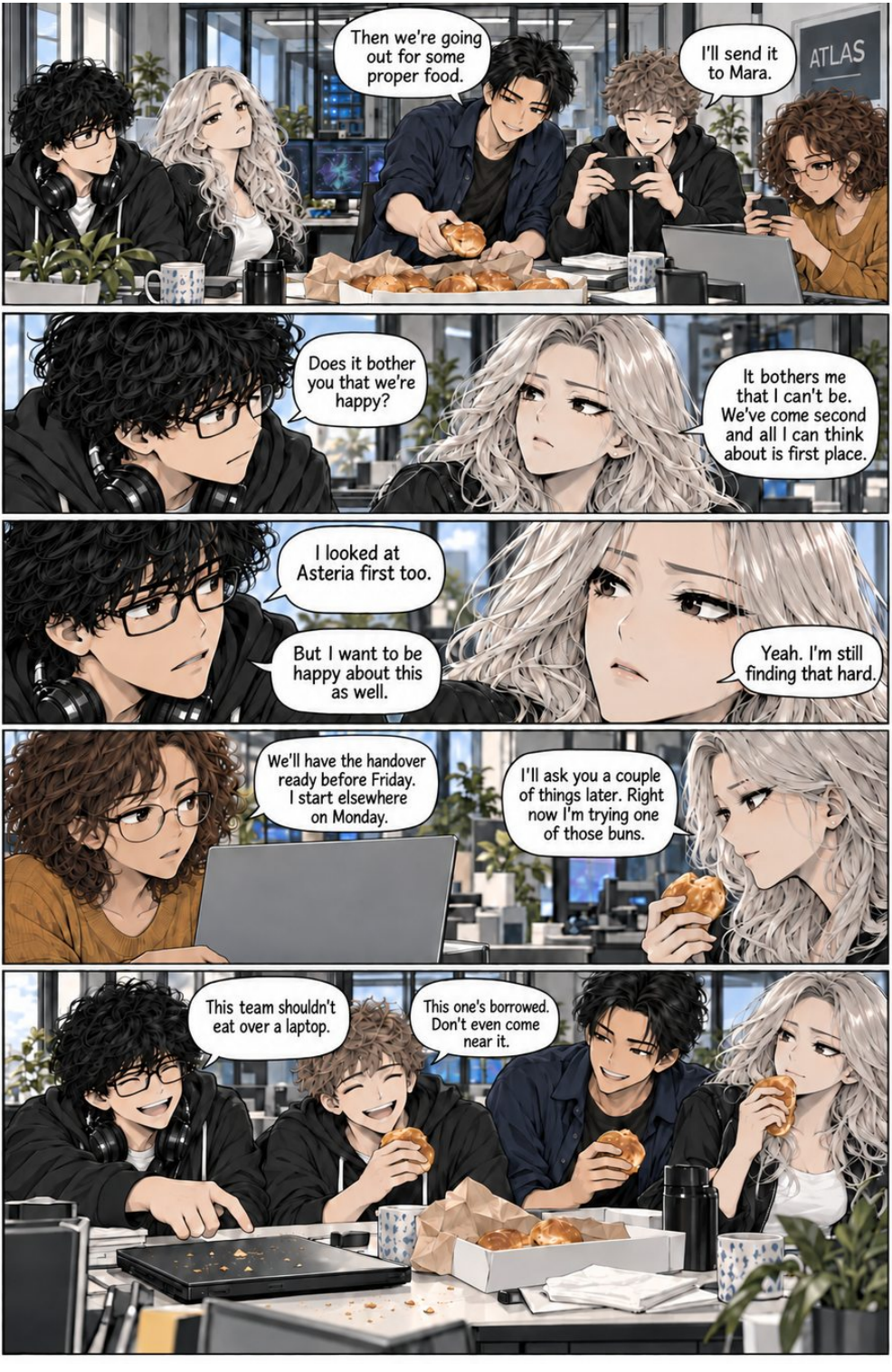
What he does with the remaining afternoon

Mara finishes her last review and leaves. Alicia prefers to walk alone. Leo offers company and accepts that answer, though he would like to keep talking to fix what is unresolved. Then he celebrates with Jun and Nico, each still at their own pace. He has managed to spend some time with them that afternoon. He does not need to wait until everyone is happy or the rankings are known to stay with them for a while.

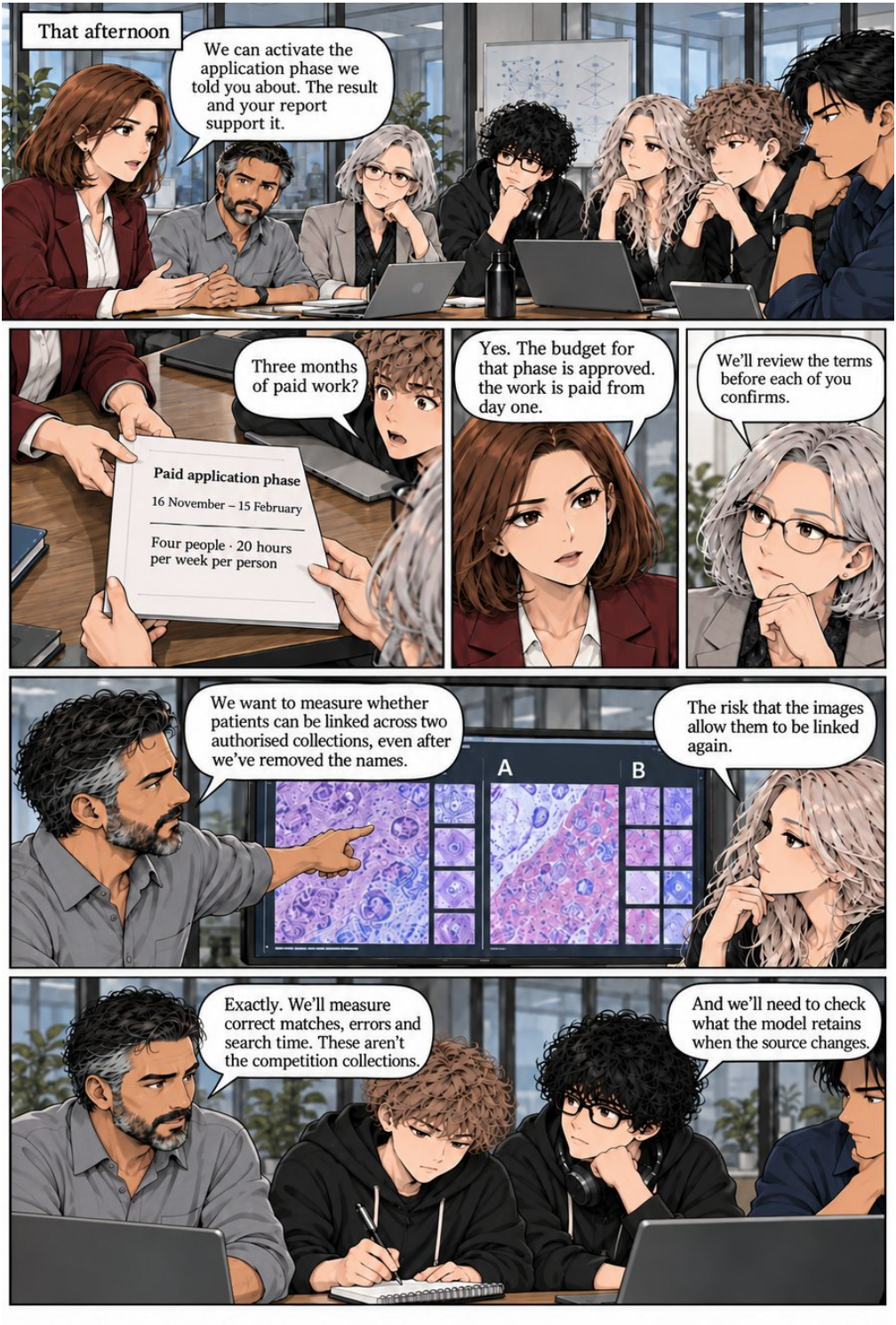

Thursday, 5 November · 19:00
I've prepared the change. From here?
No. You enter from the back now.
Sorry. I missed the mark.
Go back to the door.
There. That time, yes.
How's the show coming along?
We still need more rehearsals. But it's starting to look like what I wanted.
We can go through it again if you like.
Not today. Will you stay for a drink?
Don't you need to keep rehearsing?
I want to spend some time with you, Leo.
All right. I'd like that.

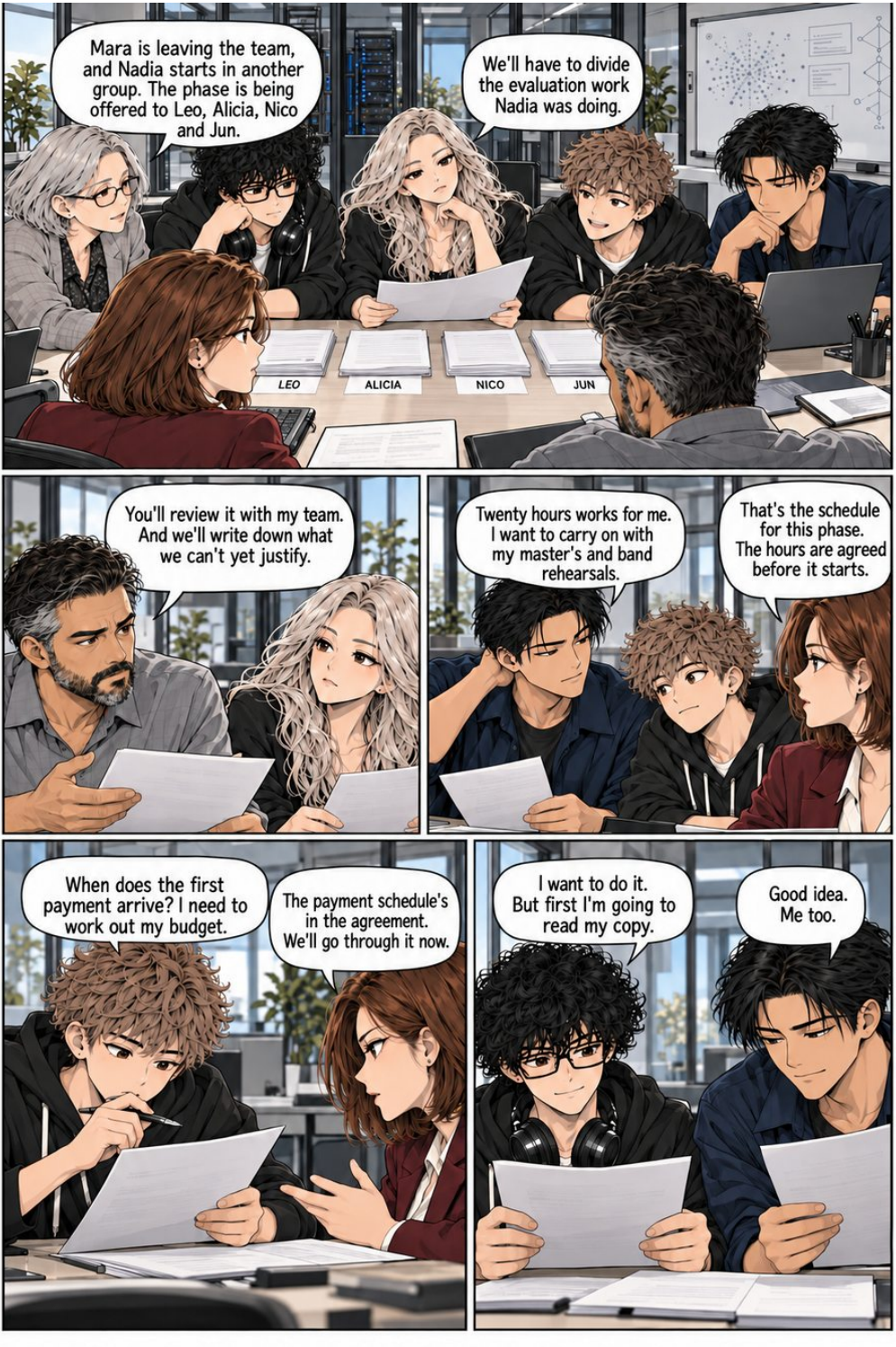

Saturday, 7 November · Group training
50
How are you doing?
I can keep going longer now. At first I had to stop all the time.
That's as far as I go.
Come on, Leo!
That's it!
Another round?
Give me a minute.


Monday, 9 November
The final rankings are out.
The cup, Leo.
Final international rankings · Track 1
Mean F1 per query
1. Asteria 0.817
2. ATLAS 0.804
3. Orión 0.792
Second?
Out of all the teams that reached this final evaluation.
Asteria beat us.
The same final evaluation for everyone. Let's not mix it up with Nexo's internal tests.
Final international rankings · Track 1
Mean F1 per query
1. Asteria 0.817
2. ATLAS 0.804
3. Orión 0.792
And you came second. It's a very good result.
Now you can give me ten seconds to celebrate.
Fictional teams and scores.


Then we're going out for some proper food.
I'll send it to Mara.
ATLAS
Does it bother you that we're happy?
It bothers me that I can't be. We've come second and all I can think about is first place.
I looked at Asteria first too.
But I want to be happy about this as well.
Yeah. I'm still finding that hard.
We'll have the handover ready before Friday. I start elsewhere on Monday.
I'll ask you a couple of things later. Right now I'm trying one of those buns.
This team shouldn't eat over a laptop.
This one's borrowed. Don't even come near it.

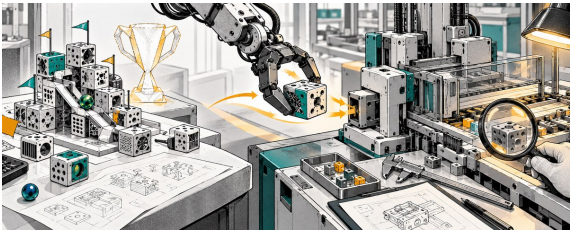

That afternoon · La Segunda Taza
Nico
Second. Your preparation made it all the way to the end.
We've come second.
We?
Until a week ago, I was there.
Then I'll let you show off today. But only until my coffee arrives.
Do you want to go and celebrate with them?
I've got class today. We'll arrange something.
Nico
I'm so happy. Let me know when you're getting together for a proper catch-up.


That afternoon
We can activate the application phase we told you about. The result and your report support it.
Three months of paid work?
Paid application phase
16 November – 15 February
Four people · 20 hours per week per person
Yes. The budget for that phase is approved. the work is paid from day one.
We'll review the terms before each of you confirms.
We want to measure whether patients can be linked across two authorised collections, even after we've removed the names.
A
B
The risk that the images allow them to be linked again.
Exactly. We'll measure correct matches, errors and search time. These aren't the competition collections.
And we'll need to check what the model retains when the source changes.


Mara is leaving the team, and Nadia starts in another group. The phase is being offered to Leo, Alicia, Nico and Jun.
We'll have to divide the evaluation work Nadia was doing.
LEO
ALICIA
NICO
JUN
You'll review it with my team. And we'll write down what we can't yet justify.
Twenty hours works for me. I want to carry on with my master's and band rehearsals.
That's the schedule for this phase. The hours are agreed before it starts.
When does the first payment arrive? I need to work out my budget.
The payment schedule's in the agreement. We'll go through it now.
I want to do it. But first I'm going to read my copy.
Good idea. Me too.


That night
Second. I've spent all day looking at it again.
NEXO
GUIDANCE
GUIDANCE
I was happy and envied first place almost at the same time.
Nico saw something I wasn't looking at. It was hard to let him have the keyboard.
Mara
Would you like to meet on Saturday? I'd like to see you without it being to ask for help with work.
I don't know what she'll say.
GUIDANCE
The lines stick in my head. I'll have to keep trying with the rest.
I've Finished My Studies — What Now?
Roberto Canales Mora
The Climber's Technique
Accept Diversity
Think for Yourself
Take Care of Yourself
Love Yourself

What Second Place Tells Us



A technique moves to another task and needs new tests.

In the rankings invented for this story, Asteria scores 0.817, ATLAS 0.804 and Orión 0.792. These are mean F1 scores per query, calculated on the common final evaluation. They are not published results from the real competition.

What 0.804 means

F1 combines correct, false and missed matches. A score of 0.804 does not mean they identified 80.4% of patients. Nor can we measure improvement by subtracting a result from an internal test with different queries. It does let us rank these teams when all are evaluated under the same conditions.

What they have managed to demonstrate

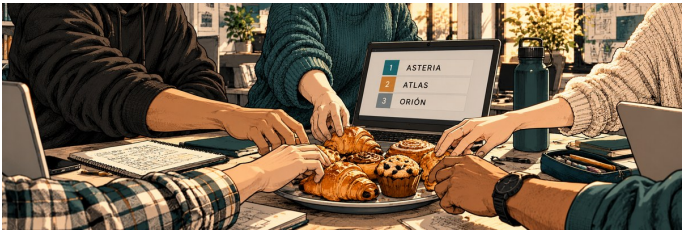
ATLAS finishes with a solution combining image preparation, partial model adjustment, example selection and a threshold fixed before seeing the final answers. The score lets us assess the whole. Finding how much each part contributes would require controlled comparisons. The rankings do not explain which method Asteria used or which decision produced the difference.

Why Nexo keeps paying

The fictional company wants to apply what has been learnt to its own problem and get to know the people who can solve it. It offers three paid months, four people and twenty hours a week. They will need to check permissions, data and usefulness for that new assignment. Second place gives the team credibility; the budget comes from Nexo, not a cash prize from the real challenge.

Sources: · Attention · WSI challenge

Second, and Wanting to Continue



A shared result, with different joys and concerns. Fictional rankings.

ATLAS comes second in an international competition. They built something that works and submitted it on time. That deserves happiness, even if Leo immediately looks at the gap to first place.

Each receives different news

Alicia finds it hard to enjoy: she came second and sees only first. Mara receives the result at the café and is happy about work that is hers too. She has class that day and does not rush off to celebrate with the team. The rankings brought them shared good news, but each arrives at it with a different life.

The opportunity now has terms

Nexo offers a paid three-month application phase, with four people and twenty hours a week. The strong ranking has given the team credibility. Now they can ask what they will have to do and decide whether they want it. We do not yet know what jobs will follow or whether the method will work the same way on the new assignment.

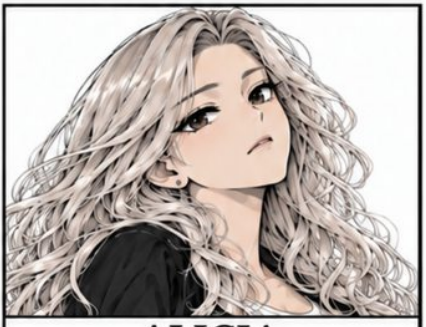
What Leo takes with him

He has learnt to adjust a model and has had to leave the keyboard in Nico's hands. Outside ATLAS, he still exhausts himself in the gym, although he can now get a little way up the rope and sit down to rest with the group. In theatre he learns the lines easily and still struggles to understand what Vera expects. When he offers to rehearse again, she suggests a drink: she wants to spend some time with him. Leo is finding it hard to understand that she might like his company too. He does not leave here with his whole life solved. He writes to Mara without knowing what she will reply and returns to the book. ‘Love Yourself’ and ‘Take Care of Yourself’ remain unfinished business. I like that he ends wanting to carry on, even with what still goes wrong.

CHARACTERS



LEO



ALICIA



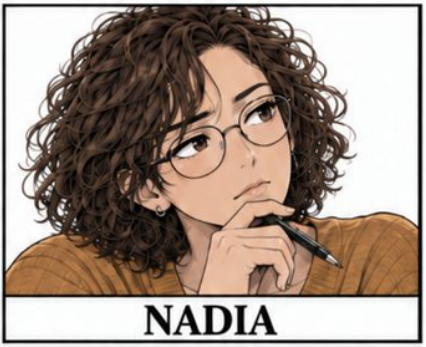
MARA



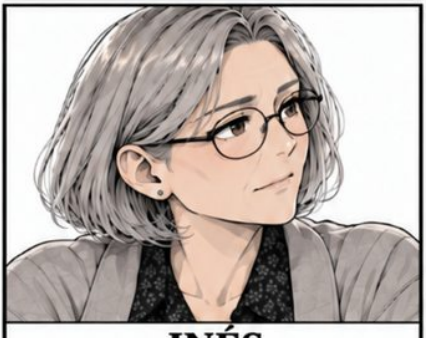
NICO



JUN



NADIA



INÉS



ÁLVARO

Sources and Scope

The article behind the story

This work grew from Roberto Canales Mora's article 'Giftedness, hyperactivity and attention deficit: like a vampire in a world of humans'. It draws questions about impulsivity, attention, belonging, listening and a functional life with imperfections from it. The adaptation does not turn a personal experience into a universal recipe.

The book Leo finds

I've Finished My Studies — What Now?, by Roberto Canales Mora, appears as a prompt to think and choose. The titles seen in the contents come from the book. Leo's thoughts, decisions and notes belong to this fiction; they are not quotations from its pages.

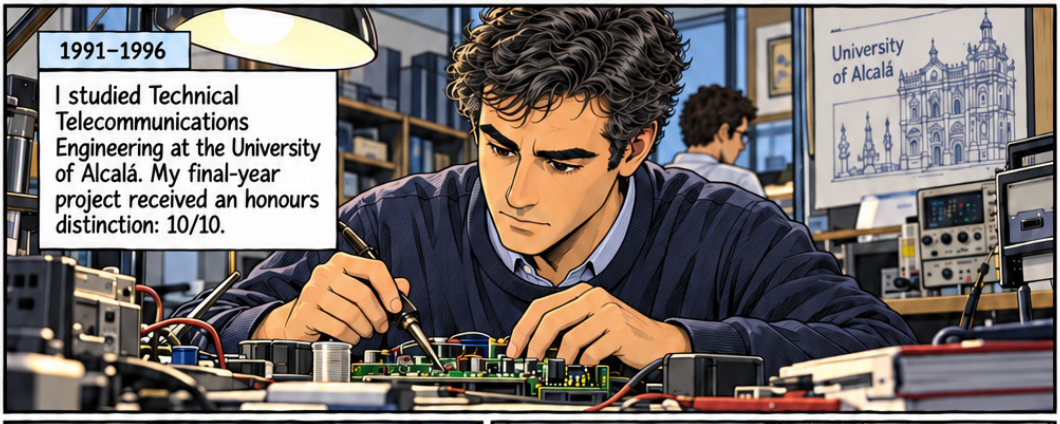
The final mountain diagram is a visual adaptation for this story.

The technical explanations

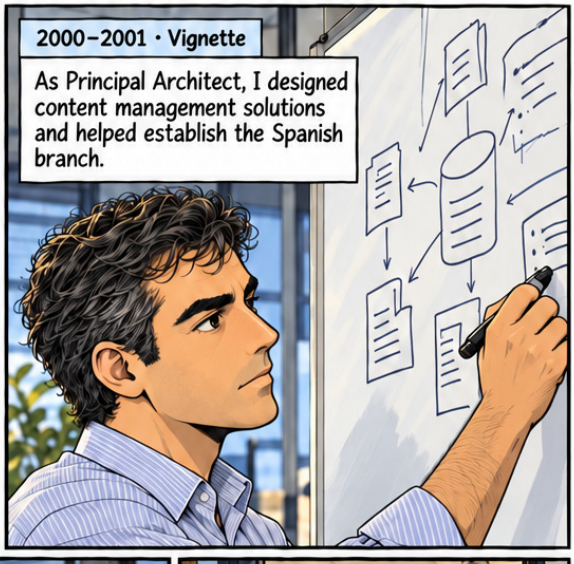
The commentaries draw on the scenes and primary technical documentation. Each explanation's references are links for further reading. They do not aim to reconstruct real competitors' solutions, publish medical data or provide official participation instructions.

Authorship and review

Roberto Canales Mora directs this work. Codex helped structure, write and review it; the illustrations were generated or adapted with AI tools. Book 1, Version 1.0, is published on HardMent. More about the work: hardment.com/en/comic.html.


1991–1996
I studied Technical Telecommunications Engineering at the University of Alcalá. My final-year project received an honours distinction: 10/10.
University of Alcalá

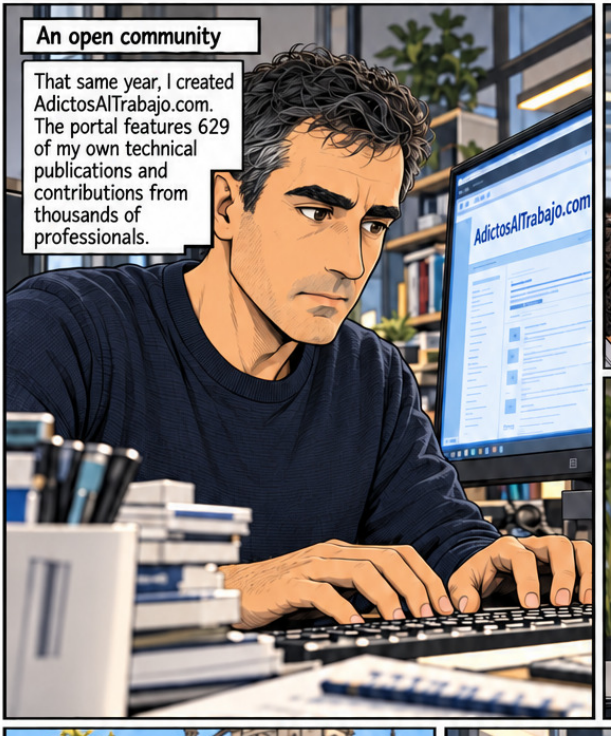

1995–2000/01 · Software AG
At Software AG, I worked in consultancy, development, architecture, and technical pre-sales. I also delivered training.

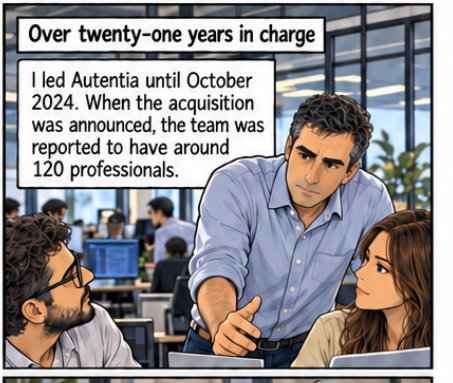

2000–2001 · Vignette
As Principal Architect, I designed content management solutions and helped establish the Spanish branch.


2001–2003 · Credit Suisse
I subsequently became Head of Development in Spain for Credit Suisse, overseeing development work, suppliers, and budgets.


From technology to business
In May 2003, I founded Autentia. Alongside developing software, I now had a company to run.


2003 · Autentia
I founded Autentia Real Business Solutions, a consultancy specialising in quality software, architecture, and agile transformation.

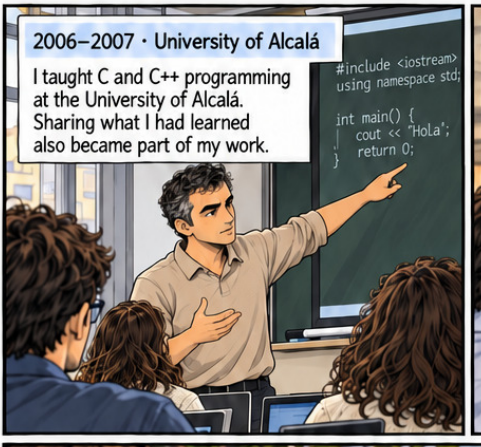

An open community
That same year, I created AdictosAlTrabajo.com. The portal features 629 of my own technical publications and contributions from thousands of professionals.
AdictosAlTrabajo.com

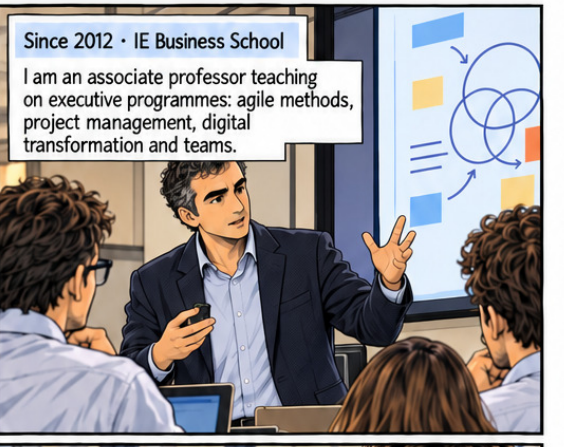

Over twenty-one years in charge
I led Autentia until October 2024. When the acquisition was announced, the team was reported to have around 120 professionals.


2016–2018 · TEAMS
I was also CEO of TEAMS. The startup did not succeed; it is part of my experience as an entrepreneur.

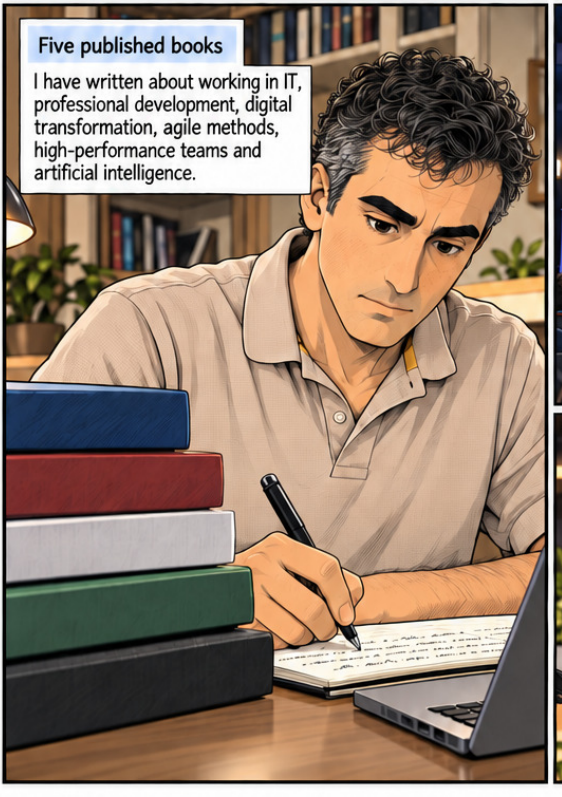

2023 · Joining Izertis
Autentia became part of Izertis in 2023. That year, I received the Senior Entrepreneur Award from the University of Alcalá.


2024–2026 · A new role
Between October 2024 and January 2026, I served as Director of Business Development at Izertis.


2006–2007 · University of Alcalá
I taught C and C++ programming at the University of Alcalá. Sharing what I had learned also became part of my work.
#include <iostream>
using namespace std;
int main() {
cout << "Hola";
return 0;
}

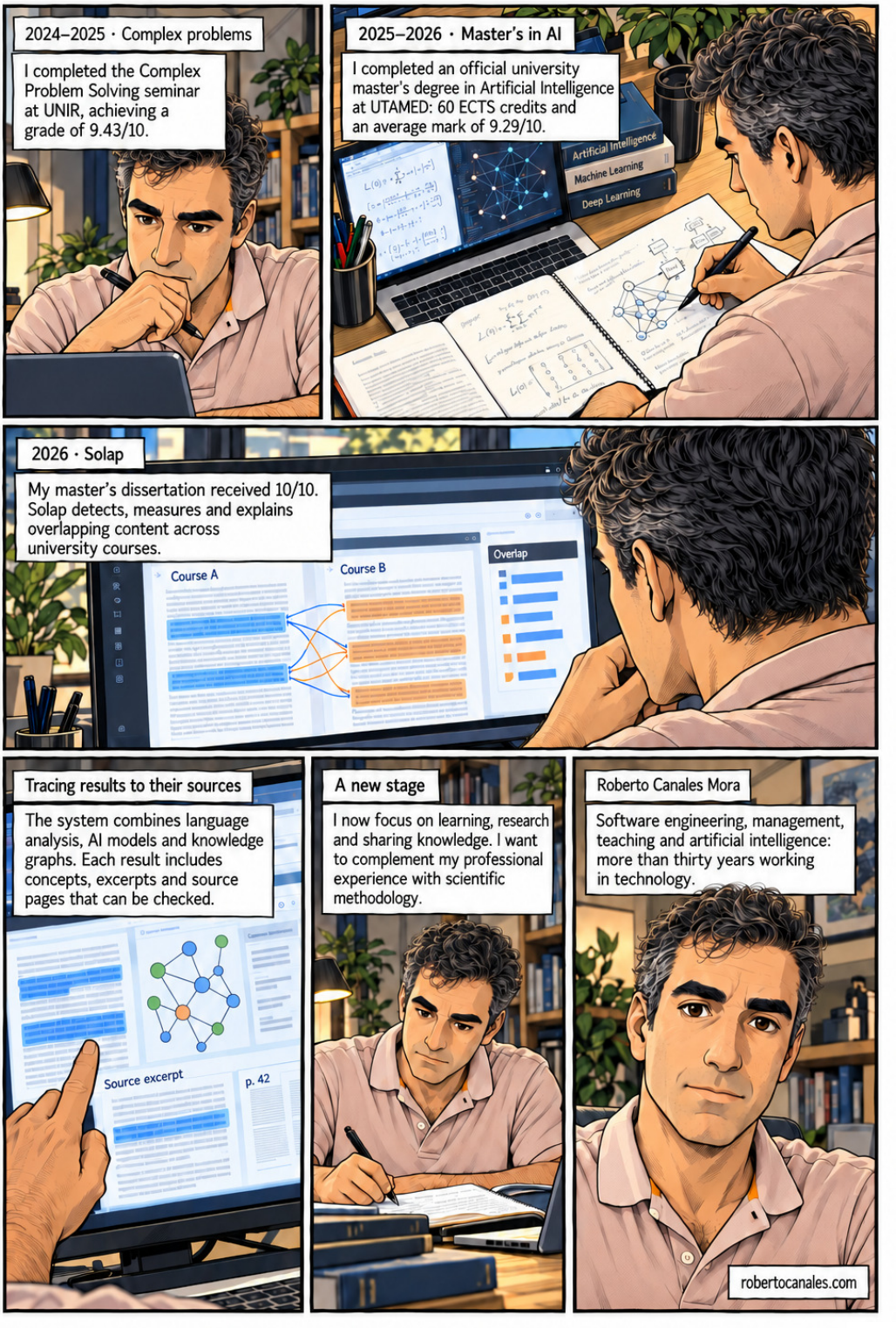

Since 2012 · IE Business School
I am an associate professor teaching on executive programmes: agile methods, project management, digital transformation and teams.

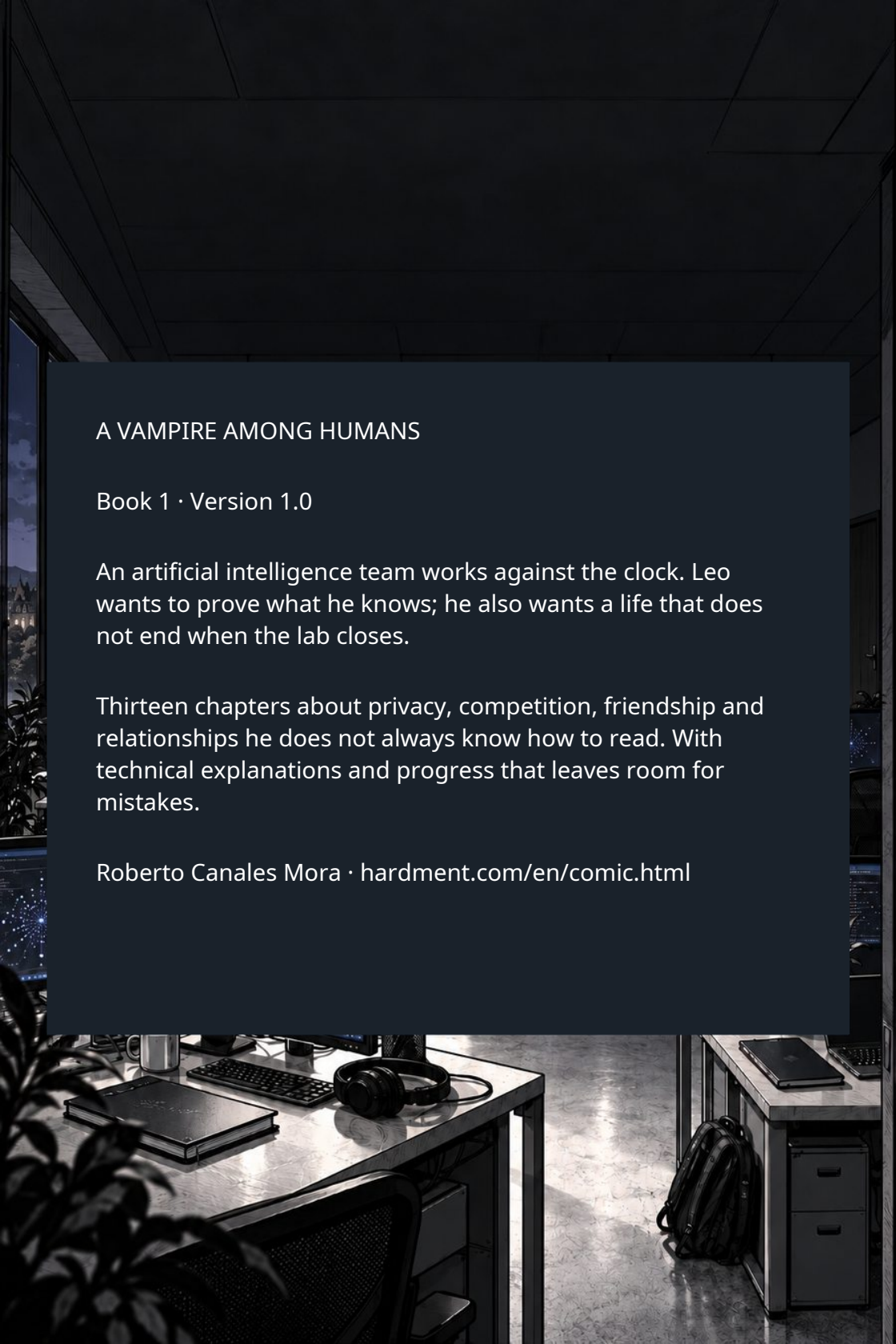

Universities and business
I have also taught at ENAE Business School, and since 2024, I have taught entrepreneurship at CEU San Pablo.


Five published books
I have written about working in IT, professional development, digital transformation, agile methods, high-performance teams and artificial intelligence.


Conferences and events
I have shared these experiences at CAS, Codemotion, Commit Conf and project management conferences.


Since 2024 · JuegoConSentido
I co-created JuegoConSentido with Iván Carrillo: conversations, interviews and sharing knowledge as we keep learning.


2024–2025 · Complex problems
I completed the Complex Problem Solving seminar at UNIR, achieving a grade of 9.43/10.
2025–2026 · Master's in AI
I completed an official university master's degree in Artificial Intelligence at UTAMED: 60 ECTS credits and an average mark of 9.29/10.
Artificial Intelligence
Machine Learning
Deep Learning
2026 · Solap
My master's dissertation received 10/10. Solap detects, measures and explains overlapping content across university courses.
Course A
Course B
Overlap
Tracing results to their sources
The system combines language analysis, AI models and knowledge graphs. Each result includes concepts, excerpts and source pages that can be checked.
Source excerpt
p. 42
A new stage
I now focus on learning, research and sharing knowledge. I want to complement my professional experience with scientific methodology.
Roberto Canales Mora
Software engineering, management, teaching and artificial intelligence: more than thirty years working in technology.
robertocanales.com



A VAMPIRE AMONG HUMANS

Book 1 · Version 1.0

An artificial intelligence team works against the clock. Leo wants to prove what he knows; he also wants a life that does not end when the lab closes.

Thirteen chapters about privacy, competition, friendship and relationships he does not always know how to read. With technical explanations and progress that leaves room for mistakes.

Roberto Canales Mora · hardment.com/en/comic.html