SAT 10 OCT

14:00

14:00 16:00 18:00 20:00 22:00 00:00 01:01

SATURDAY, 10 OCTOBER 2026

One day of AI-driven software delivery.

It starts with a voice note.

VOICE NOTE · WISPR FLOW DICTATION



“The AI assistant on my site should learn from my articles.”

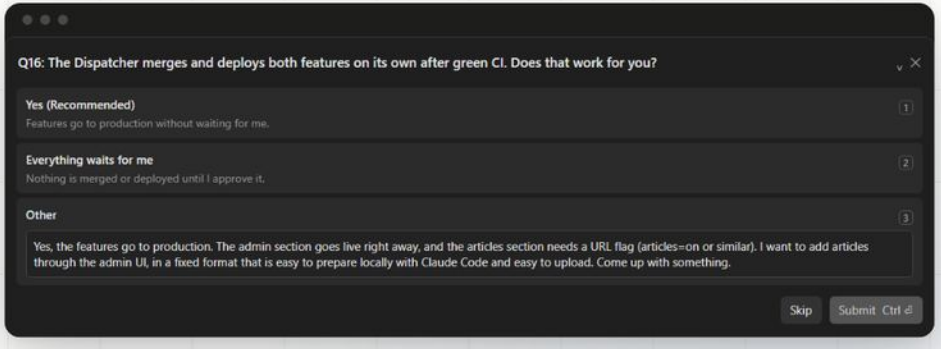
SAT 10 OCT

14:05

14:00 16:00 18:00 20:00 22:00 00:00 01:01

14:05 · GRILL ME

An agent interviews me, one decision at a time.



Q The Dispatcher merges and deploys both features on its own after green CI. OK?

A Yes, and the admin panel goes live now; the articles section stays hidden behind a URL flag.

ONE FEATURE BECOMES THREE

01
An articles section

02
An admin panel to publish

03
The assistant learns from articles

SAT 10 OCT

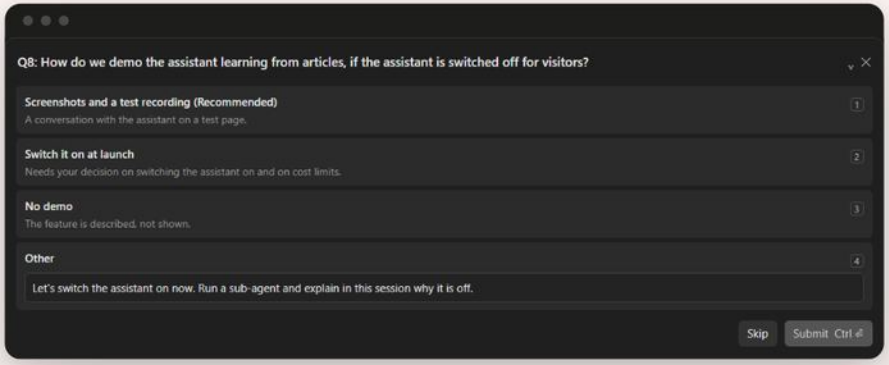
14:05

14:00 16:00 18:00 20:00 22:00 00:00 01:01

FAILURE 1 OF 3

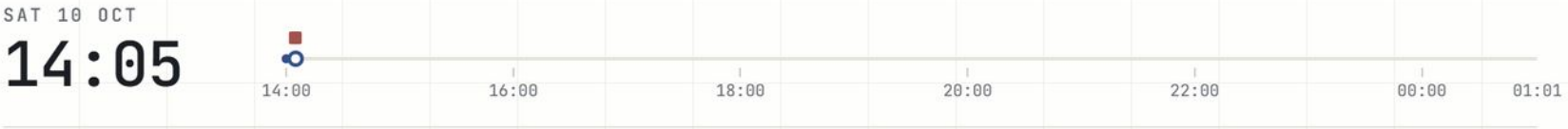
The agent believed a stale document.

It said the assistant was off. It had been live for 11 days.



THE QUESTION ASSUMED: ASSISTANT OFF

ACTUALLY LIVE FOR 11 DAYS



14:05 TO 14:28 · EVERY ANSWER ON RECORD

Every question and answer is recorded.

One answer added a requirement nobody asked for.



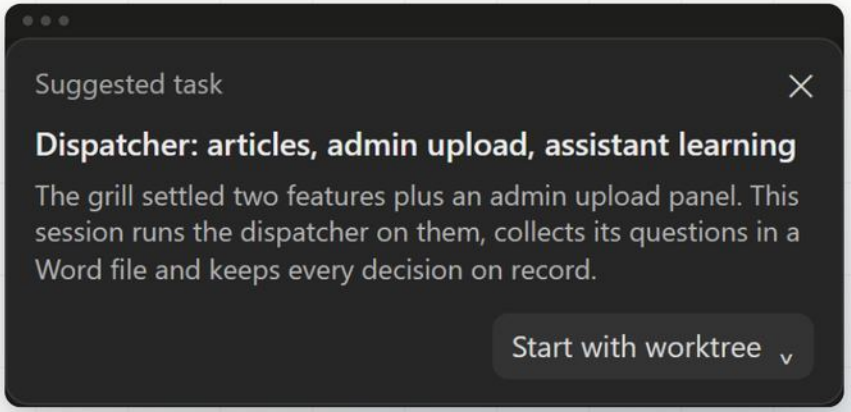
SAT 10 OCT

14:28

14:00 16:00 18:00 20:00 22:00 00:00 01:01

14:28 · THE DISPATCHER STARTS

My own AI delivery-lead agent takes over.



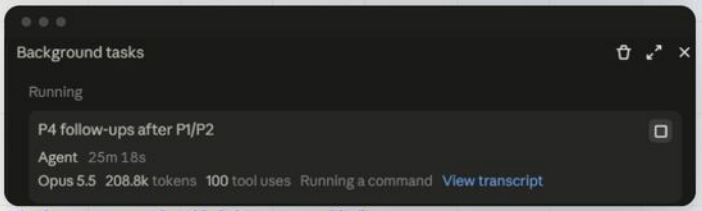
6 MINUTES LATER, ONE DOCUMENT

5

QUESTIONS FOR ME

7

DECISIONS ALREADY TAKEN



Sub-agents build in parallel.

SAT 10 OCT

15:12



15:04 TO 15:12 · FIRST PULL REQUEST

15:04 the first pull request opens. 15:12 it is merged, automatically, after every check passed.



No human in between.

SAT 10 OCT

15:12



15:12 · THE GATES

The agents test their own work. Every time.

Every pull request: first locally by each sub-agent, then again in CI.

ONE RED CI RUN

The Polish header at 1024 px ran 5 px past the margin. Linux renders labels wider than Windows.

FIXED, GREEN ON THE NEXT RUN

- **Lint**
- **Types**
- **Unit tests**
- **End-to-end** *in headless browsers*
- **Build** *a production build*
- **Content** *no em dashes, no number without a source*
- **Deploy** *only a green main reaches the server*

294 / 294

LIVE CHECKS AFTER EACH RELEASE. A FAILED CHECK ROLLS IT BACK.

OVER THE DAY: 29 CI RUNS · 10 RELEASES · 294 OF 294 LIVE CHECKS EVERY TIME · 0 ROLLBACKS

SAT 10 OCT

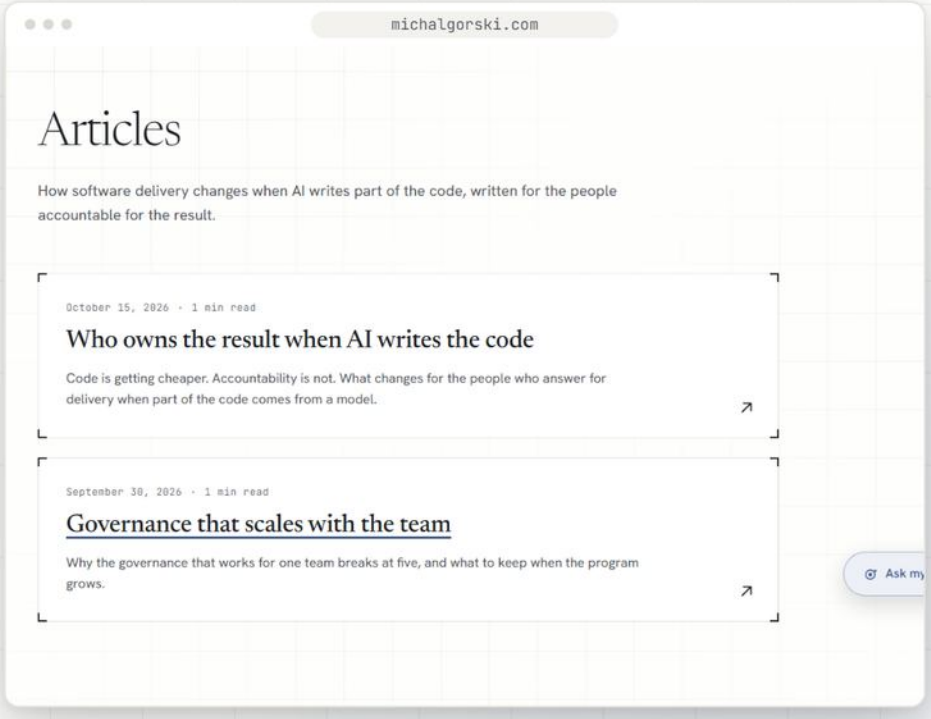
16:49

14:00 16:00 18:00 20:00 22:00 00:00 01:01

16:49 · IN PRODUCTION

Feature 1: the articles section.

All three features live, behind a hidden-section flag.



SAT 10 OCT

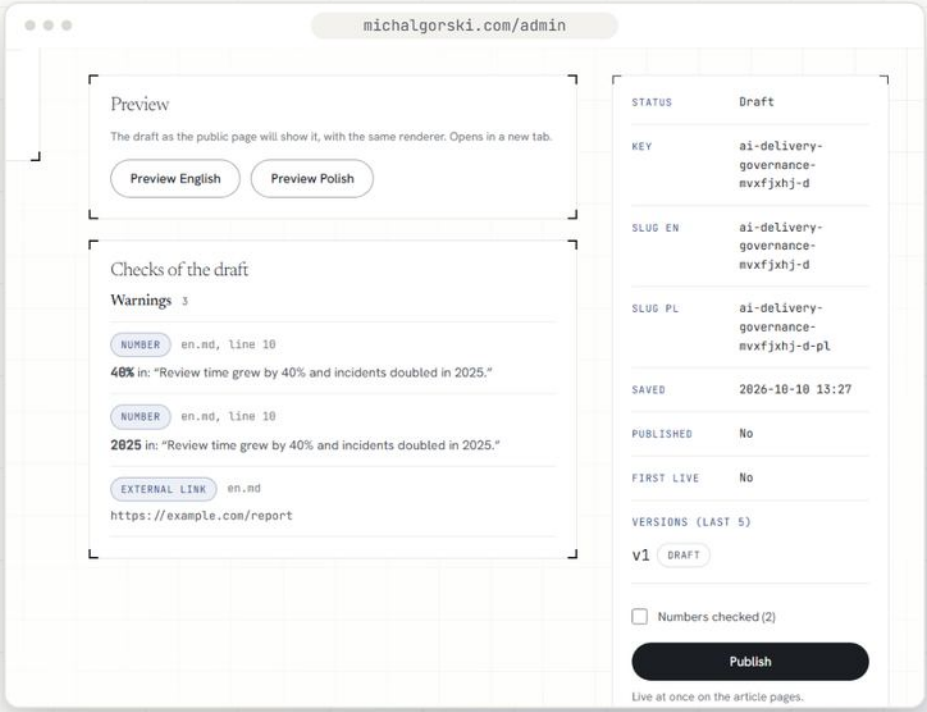
16:49

14:00 16:00 18:00 20:00 22:00 00:00 01:01

FEATURE 2 · THE ADMIN PANEL

It checks an article before publishing.

Every number and external link is listed.
Publish waits until I confirm the numbers.



SAT 10 OCT

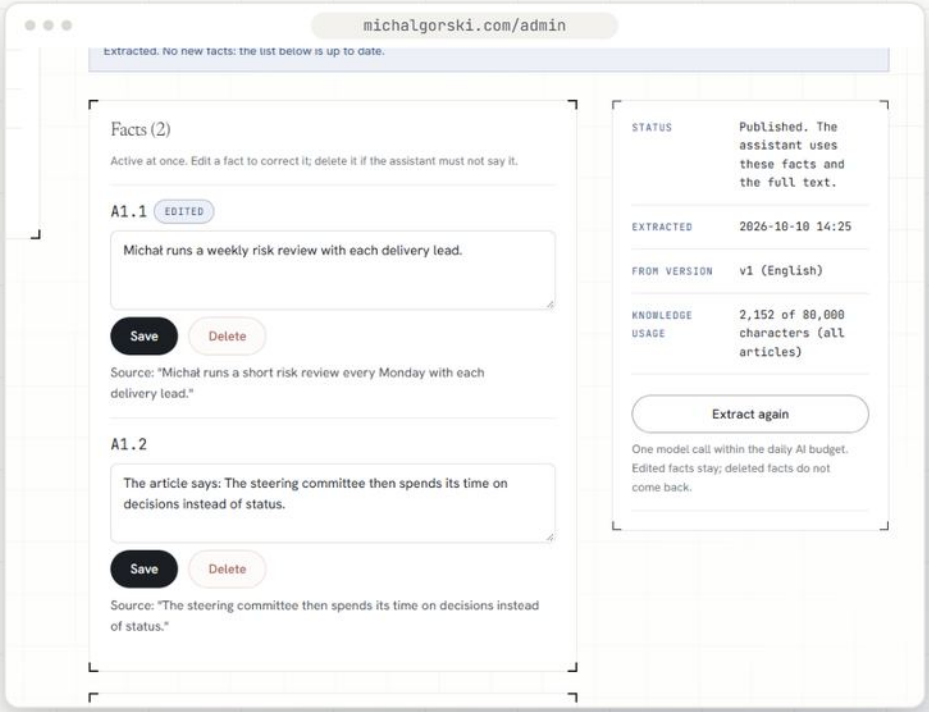
16:49

14:00 16:00 18:00 20:00 22:00 00:00 01:01

FEATURE 3 · THE ASSISTANT

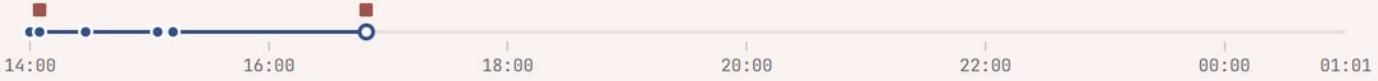
The facts the assistant learned.

Each fact with its source sentence. Edit or delete. *A deleted fact never comes back.*



SAT 10 OCT

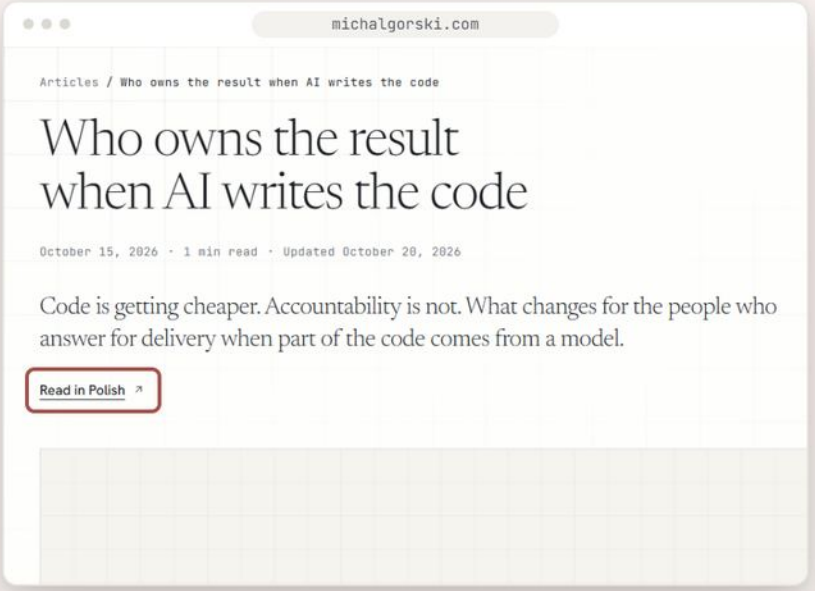
16:49



FAILURE 2 OF 3

A “Read in Polish” link nobody wanted.

Caught in a screenshot review, already deployed behind the flag. *Shipping first was by design.*



REMOVED IN THE NEXT FIX

SAT 10 OCT

18:33

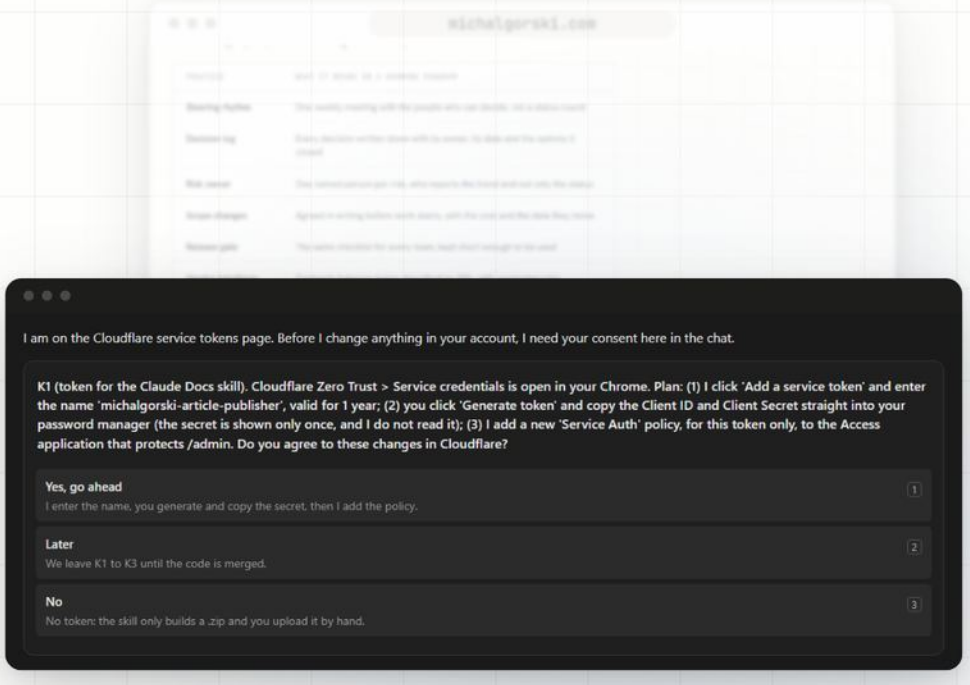


18:33 · RUN 2

Tables, downloads, publishing from my writing tool via an API.

The agent stops and asks before touching an external account.

The secret stays with the human.



SAT 10 OCT

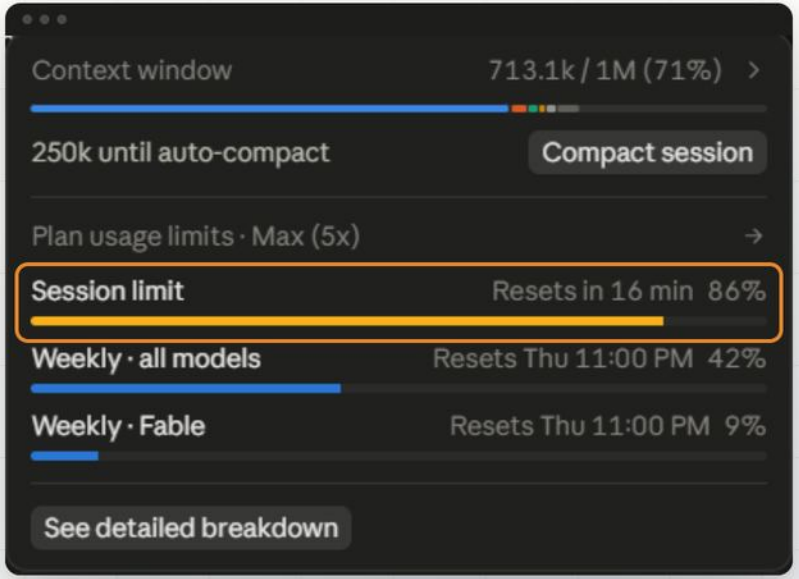
18:43



18:43 · USAGE

86% of the 5-hour limit.

My current plan costs 100 USD a month. The next one costs 200 USD.



SAT 10 OCT

~19:50



~18:55 · AN HOUR OFFLINE

Pancakes with my daughter.

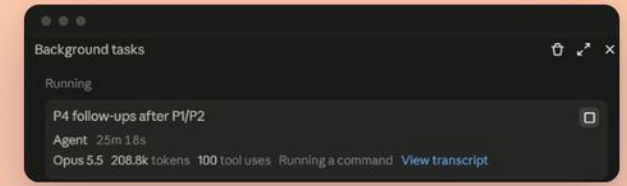
The best hour of the day.

One hour. The agents keep working.





● AGENTS STILL WORKING



SAT 10 OCT

20:18



20:18 · RUN 3, A SMALL FIX

A contact request from an article now says which article.

CONTACT REQUEST

SENT FROM THE ARTICLE

Who owns the result when AI writes the code

SAT 10 OCT

20:25



20:25 · ACCEPTANCE TEST

The agent writes me a test plan.

32

STEPS FOR ME TO RUN

SAT 10 OCT

21:07



21:07 · MY RESULTS

My results against the agent's plan.

22	2	5	1
STEPS PASSED	DEFECTS	NEW REQUESTS	MISTAKE IN THE AGENT'S OWN TEST PLAN

SAT 10 OCT

21:07



21:07 · A RISK SHOWS UP

Every demo grows the wish list.

The moment people see a working feature, they see what else they want.

My own test turned into five new requests.

Each looks like two hours of agent work.

Two hours here, two hours there, and the next ten hours are gone.



SAT 10 OCT

21:07



FAILURE 3 OF 3

A credential on the server picked up invisible characters.

The next release's config check stopped it. *Production untouched.*

CREDENTIAL AS WRITTEN

••••••••••

CONFIG CHECK: RELEASE STOPPED

SAT 10 OCT

22:05



22:05 · RUN 5

Long material gets published as a series: a map and five parts.



A series needs navigation between parts: *one more feature.*

SAT 10 OCT

22:20



22:20 · I GO TO BED. RUN 6 STARTS.

More agents chase unstable tests. They find a real server bug.

Before the first real series, the whole publishing path runs on a test series in production.

UPLOAD · PREVIEW · PUBLISH ALL PARTS · EN/PL NAVIGATION · WITHDRAWAL · EVERY ADDRESS BACK TO 404

SAT 10 OCT

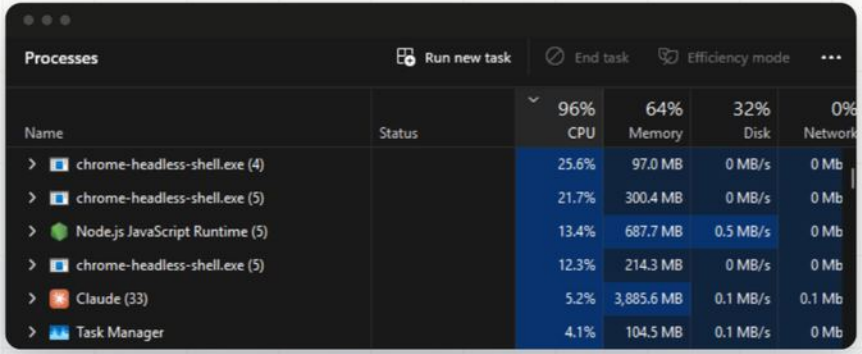
~23:20

14:00 16:00 18:00 20:00 22:00 00:00 01:01

~23:20 · THE HIDDEN COST

My laptop at 96% CPU.

Headless browsers, builds and the Claude app on one machine. About 3.5 hours of heavy local test and build runs over the day.



96% *CPU*

SUN 11 OCT

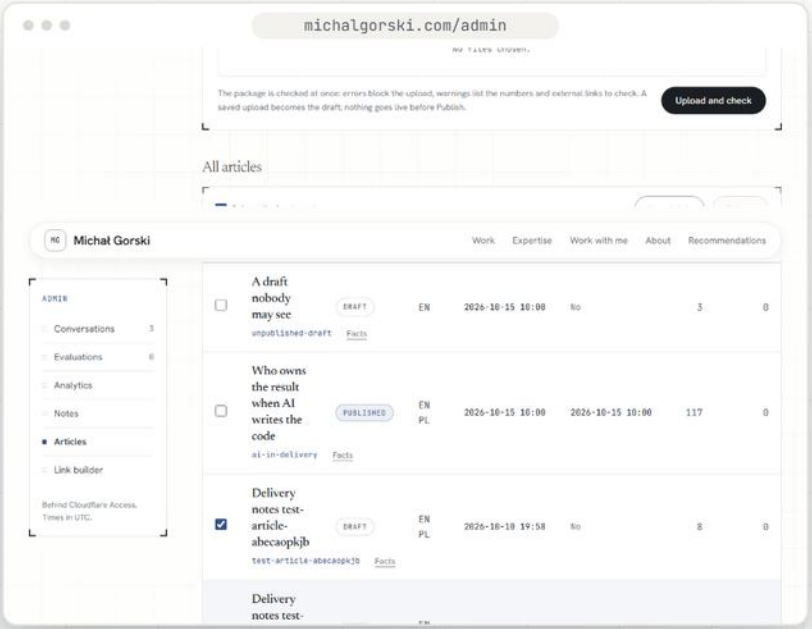
00:15



00:15 · UPLOADED

The series goes up as a draft.

Publishing stays my decision.



SUN 11 OCT

01:01



01:01 · THE DISPATCHER

Done.

A written report, with what it did not check.

DONE

All three features are live in production. The articles section is hidden behind `?articles=on`, and there is no article on the server yet.

What works:

- **A, article pages (PR #126):** `/articles` and `/pl/artykuly`. Without the flag the pages return 404. With the flag they work, but search engines do not index them and they are not in the navigation, sitemap or llms.txt. Checked in production.
- **C, admin panel (PR #125):** in `/admin/articles` you drag in .md files and images or a .zip, up to 5 MB. Only this route has that limit; the rest of the server still accepts 64 KB. The panel checks the file on upload, then there is a preview, Publish (live at once) and Unpublish.
- **Local format (PR #124):** you check a file on your machine with `npm run article:check`. There are also templates and an `article-package` skill.
- **B, assistant (PR #127):** reads the full text of published articles and extracts facts from them. You get the list of facts by email, with a link to edit and delete them. The assistant's daily cap is now 2 USD. Test on the real model with a sample article: 5 of 5 answers correct, cost 0.03 USD.
- **Fixes (PR #128):** I removed the "Read in Polish" link. The EN | PL switch in the header already led to the other version of the article.

Each of the four deployments passed the live check, 294 of 294 tests, with no rollback.

What I did not check: the assistant's answers in production. Asking it needs a Turnstile token, and the repo has no script for that test. Two tests outside this scope sometimes fail under heavy machine load; noted for the next run.

For you, in the report `D-07-dispatcher-report.docx` (open in Word):

- Q6: do you accept the sentence under the index heading? My proposal stands for now, visible only with the flag.
- D8 and D9: two small decisions you can change. D8 is the contact link from the article page, D9 the language switch on an article without a translation.
- You upload and publish the article yourself in the panel, after your yes. I do not publish it.
- Switching the section on for visitors waits as NEW-57 in open items. It is one configuration change and one deployment.

Measurements are in `D-99-measurements.md`: 3 h 41 min from start to the last deployment. 13 CI runs, about 142 GitHub Actions minutes in total. Sub-agents used about 1.9 million tokens. Failures with their causes are in `D-00-status.md` and in the report.

I removed the package worktrees. You can archive the Dispatcher session whenever you like.

SUN 11 OCT

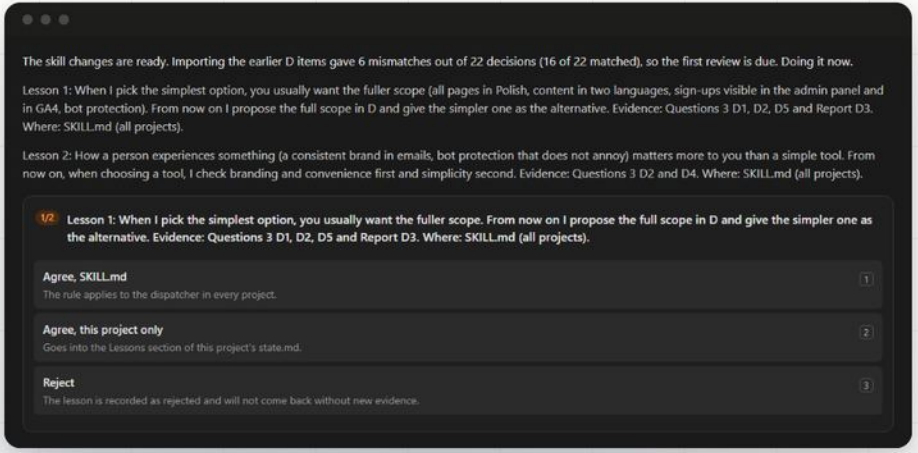
01:01



THE AGENT LEARNS HOW I DECIDE

It compared 22 earlier decisions with mine.

16	6	2
MATCHED	MISSED	LESSONS DRAWN



Lesson: when it picks the simplest option, I usually want the fuller scope.

It asks me one lesson at a time before adopting it.

SUN 11 OCT

01:01



THE PARKING LOT

Seven ideas parked for later.

01 Switch the articles section on for visitors

02 RAG for the assistant

03 Comments under articles

04 An article checker that explains errors

05 Short links with campaign tags

06 A pre-publication version for friends

07 The assistant suggests what to read next

SUN 11 OCT

01:01



THE RISKS

Two risks. Both old. Both much faster now.

RISK 1

Deciding what to build faster than anyone can check it

WHERE IT SHOWED UP:

Unsigned requirements
The wrong weight of documentation
Agents trust stale documents
Scope decided in a click
Every demo grows the wish list
Scope that never stops growing
Decisions nobody made

RISK 2

Building and running it without seeing the full cost

WHERE IT SHOWED UP:

Review after release, by design
No definition of done
Green checks, wrong feature
Agents colliding in the same code
The manual step nobody verifies
Token spend that keeps running
The machine is part of the budget

SUN 11 OCT

01:01



THE QUESTION

Was it cheaper?

A TEAM, BY MY ESTIMATE

0.3 business analyst

+ 2 developers

+ 0.5 QA

about 3 weeks

My estimate from 20 years in software delivery.

THIS RUN

11 hours from voice note to production

~2.5 to 3 h of my time

~38 h of AI agent time

~270 to 330 USD of tokens at API prices

SUN 11 OCT

01:01



THE DAY IN NUMBERS · TIME

~11

HOURS, 14:00 TO 01:01

6

RUNS

11

PULL REQUESTS

10

RELEASES

0

ROLLBACKS

SUN 11 OCT

01:01



THE DAY IN NUMBERS · AGENTS AND ME

~38 h

OF AI AGENT RUNNING TIME IN ONE 11-HOUR DAY

Main session ~11 h · Dispatcher 10 h 33 min · 15 sub-agents ~15.3 h · two side sessions ~1.6 h

~15 h

OF SUB-AGENT WORK

15

CODING SUB-AGENTS

~2.5 to 3 h

MY TIME ON THE FEATURE

1 h

OF PANCAKES WITH MY DAUGHTER

The best hour of the day.

SUN 11 OCT

01:01



THE DAY IN NUMBERS · COST

~720M tokens

processed, 97% from cache

~270 to 330 USD

at API prices, on a flat-rate subscription

~30% to 52%

of the weekly limit

1.91 USD

of CI usage, 0 billed

~3.5 h

of heavy local runs on my laptop (96% CPU at peak)

Code got cheaper.
Accountability didn't.

AI SDLC: cheaper code & new risks to manage

michalgo.ski.com