Who needs to review code?

Human judgment in the AI era

Dr. Michaela Greiler

What you'll be able to do

01

What has changed with AI?

Understand how AI changes review and what humans and AI each add.

Explore studies, developer interviews and this room's poll responses.

02

How can code review change?

Use SCOPE to plan staged, risk-based oversight.

Explore one model for adapting how changes are reviewed.

03

When is human oversight needed?

Recognize signals that call for human oversight.

Practise making the call through scenario polls and discussion.

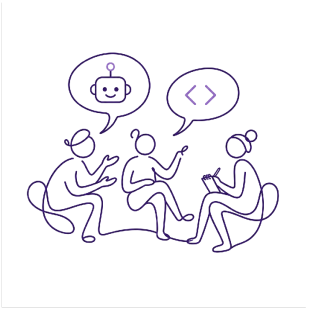
Experience, conversations & evidence

01 · DEVELOPER INTERVIEW STUDY

Over 25 developers & engineering managers

At a company investing heavily in AI since the beginning of 2026.

Their experience with AI and code review.



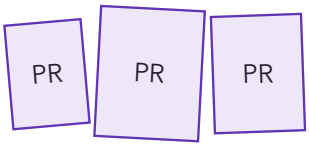
02 · PRACTICE

7 years as an independent coach
Software engineering, code reviews and developer experience.

03 · EMPIRICAL EVIDENCE

Research & industry insights
Studies, company experience and empirical data.

More code reaches the review queue



Human attention
still takes time

Read more
More PRs. Larger PRs.

Review gets framed as the bottleneck

+199.6%
average time in review
Faros · 2026

+23%
monthly merged PRs
GitHub · 2025 vs. 2024

+122%
median PR size
DX · Q3 2026

+76.3%
PRs merged without review
Faros · Q3 2026

AI can also help with code reviews

> “Yeah, it's quite easier and faster to do a code review using [AI tools]. I think it can get a lot of stuff that you couldn't think of out of the box.”

Interview participant

AI code review is gaining traction

GitHub Copilot

60M

code reviews

10× usage growth since launch
More than **1 in 5** reviews on GitHub

GitHub · 5 Mar 2026 ↗

CodeRabbit

1.8M+

open-source PRs reviewed

Since March 2026
394K in the last 30 days

Live tracker · 28 Oct 2026 ↗

Greptile

22K+

teams using Greptile

From startups to teams at
NVIDIA, Brex and Coinbase

Vendor site · 28 Oct 2026 ↗

Qodo

≈1M

PRs reviewed per quarter

Across enterprise customers
Using Qodo Merge

Case study · accessed Oct 2026 ↗

Vendor-reported adoption · Different scopes and time windows; not a quality comparison.

What does my review add?

“So what am I going to find that AI doesn't?”

Interview participant · excerpt

→ Pressure to ship faster makes this harder

The focus is shifting. But how?

People recognize that review needs to change.
There is still little guidance on how.



"today, my focus on review is on the **human understanding part.** [...] I don't care if there's an if statement or a switch statement [...] It's more about **keeping me in control** of what the **code does on a higher level.**"

Interview participant · excerpt

What do humans and AI each bring to the review?



Review needs a clear direction.

01 **Review faster**
Know what to look for.

02 **Review with purpose**
Know what value you add.

03 **Review more consistently**
Share expectations for depth and thoroughness.

Then and now: complementary review

THEN · GPT-4 TURBO

Crupi et al. · ICPC 2026
Earlier-generation model

10%

of human-reported quality issues also found by AI

~40%

of additional AI comments suggested meaningful improvements

AI suggested ≈2.4× more changes.

NOW · CLAUDE OPUS 5 AGENT

Yang et al. · CRBugBench
October 2026

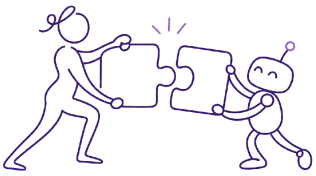
79.6%

of bugs humans caught before merge also found by the agent

64.2%

of bugs discovered after merge also found by the agent

The agent missed 20.4% of bugs humans caught.
Results count the agent's top 5 findings.



People still complement AI.

Crupi et al. · GPT-4 Turbo · Yang et al. · CRBugBench

Where AI needs help: intent & evidence

THE CHALLENGE	THE HUMAN CONTRIBUTION
Ambiguous intent	Clarify expected behaviour, forbidden outcomes and acceptable trade-offs.
Unconvincing evidence	Ask for a concrete failure scenario; check realistic cases and what the tests leave unverified.
False & low-value findings	Verify the claim. Separate bugs from preferences and decide what is worth changing.
Repeated review / fix loops	Suppress duplicates, re-review the affected code and stop cycles that add no value.

CRBugBench · Oct 2026 · GitHub · July 2026 · Claude Code team · Oct 2026

Where AI needs help: complex codebases

THE CHALLENGE	THE HUMAN CONTRIBUTION
Legacy constraints & unwritten rules	Recover business rules, historical workarounds and compatibility requirements before changing code.
Large codebases	Check that the search found the relevant code, distant callers and existing conventions.
Tightly coupled code	Trace ripple effects through dependencies. Check system behaviour beyond the edited files.
Changes across repos	Verify every affected repository, shared API and deployment dependency.

DORA · 2025 · WideSWE · Sept 2026 · CRBugBench · Oct 2026

Code Review is so much more than finding bugs

What code review builds

Correctness

Does it work as intended?

Catch behaviour and security failures.

Alignment

Are we taking the right approach?

Agree on approach, interfaces and design.

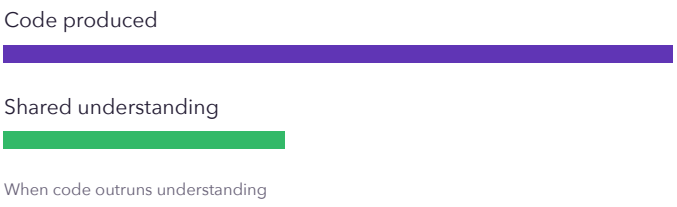
Knowledge sharing

Can the team carry it forward?

Share the system model.
Mentor, learn and grow.

Loss of Understanding

Cognitive debt · Comprehension debt · Intent debt



Can someone…

explain it?
debug it?
safely change it?

A change can pass checks while the team's understanding falls behind.

Ironies of Automation 1983

"the most successful automated systems, with rare need for manual intervention, [...] may need the greatest investment in human operator training"

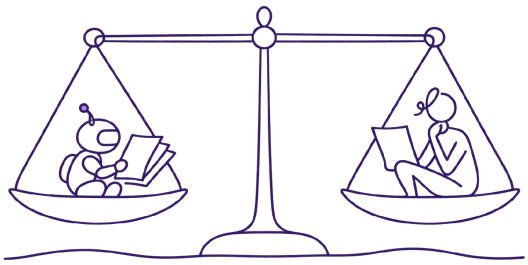
Lisanne Bainbridge



What does this mean for human review?

AI capabilities
Review pressure

Human skills
Understanding needs



What is at stake?

01 **Correctness**

failures in production

System outages · data loss · unhappy customers.

02 **Alignment**

architectural drift

Conflicting patterns · growing complexity · costly changes.

03 **Understanding**

loss of shared knowledge

Slower debugging · fragile handovers · missed learning.

Risk



Look at the specific change to understand its risk and oversight needs.

Not every change needs the same oversight.

Two small changes

A · BUTTON TEXT CHANGE

```
  function renderSaveButton(button) {
-   button.textContent = "Save changes";
+   button.textContent = "Save";
  }
```

Oversight 

B · PERMISSION CHANGE

```
  function canReadReport(user) {
-   const roles = ["admin"];
+   const roles = ["admin", "member"];
    return roles.includes(user.role);
  }
```

Oversight 

Would you choose the same oversight for both?

Who reviews the change?



Agent only

AI reviews the change.
No human checks this pass.

One person

The author checks the change with AI.

With a peer

A second person challenges assumptions and evidence.

With the team

Review together across roles and system boundaries.

How to review a change?

01 · HUMAN-DRIVEN

AI-assisted

The human directs the investigation.

Most powerful potential

02 · HUMAN-ASSISTED

AI-driven

AI chooses what to investigate. A human checks its findings.

Most standardized

03 · HANDS-OFF REVIEW

AI-autonomous

AI reviews and fixes what it deems relevant.

Most autonomous / fastest loop



SCOPE: Staged Code Oversight With proportionate Escalation

michaelagreiler.com/scope

01 Review before implementation



Agree the risk, review focus, and who needs to be involved.

02 Adaptive Oversight

Who reviews?



Agent only · One person · With a peer · With the team

How?



AI-autonomous · **AI-driven** (Human checks findings) · **AI-assisted** (Human directs investigation)

More human involvement →

03 Document Review Outcome

Record what was checked, the decision, and any remaining concerns.

AUTH-184 Bug · Ready for development

Restore report access for legacy SSO users

User problem

Users can sign in through the legacy single sign-on (SSO) provider, but see “Access denied” when opening reports they have permission to view.

Identified cause

The legacy provider sends no `role` claim. The middleware denies access, although these users have roles in our application’s membership records.

Agreed fix

For legacy SSO requests without a `role` claim, use the existing membership lookup for the verified user and current organization. Keep existing permission checks.

Acceptance criteria

- Entitled legacy users can view their reports; unentitled users remain denied.
- An existing `role` is never replaced.
- Without membership or valid role: access remains denied.
- Other providers retain existing behavior.
- Add regression tests for these cases.

LINKED PULL REQUEST

[AUTH-184] Restore report access for legacy SSO users

Author
Senior developer

Diff
14 lines · 2 files

CI
All checks passing

Tests
96% middleware line coverage

System history
Users have had sign-in and access problems in the past.

System context
An external tool used by customers.
The reports contain medical records.
Highly private and confidential information.

AUTH-184 Bug · Ready for development

Restore report access for legacy SSO users

User problem

Users can sign in through the legacy single sign-on (SSO) provider, but see “Access denied” when opening reports they have permission to view.

Identified cause

The legacy provider sends no `role` claim. The middleware denies access, although these users have roles in our application’s membership records.

Agreed fix

For legacy SSO requests without a `role` claim, use the existing membership lookup for the verified user and current organization. Keep existing permission checks.

Acceptance criteria

- Entitled legacy users can view their reports; unentitled users remain denied.
- An existing `role` is never replaced.
- Without membership or valid role: access remains denied.
- Other providers retain existing behavior.
- Add regression tests for these cases.

LINKED PULL REQUEST

[AUTH-184] Restore report access for legacy SSO users

Author
Senior developer

Diff
14 lines · 2 files

CI
All checks passing

Tests
96% middleware line coverage

System context
Internal tools. No customers.
Stable middleware.
The records are non-sensitive.

PR #482 · Open · 2,146 lines changed · All checks passing

Improve the survey editor

Description

Consolidates question creation into a shared registry and updates the callers, preserving existing defaults.

Refreshes the question toolbar: new icons, clearer labels and a more compact layout.

Stores question order in a new `position` colu surveys and switches loading and saving to use

Validation

Updated factory tests, toolbar snapshots and a migration test. CI passes.

TOO MUCH

Files changed 10

- src/surveys/questionRegistry.ts
- src/surveys/ Question.ts
- teQuestion.ts
- oolbar.tsx
- estion-toolbar.css
- src/surveys/surveyRepository.ts
- db/migrations/042_question_position.sql
- tests/questionFactory.test.ts
- tests/QuestionToolbar.test.tsx
- tests/questionPositionMigration.test.ts

What oversight would you choose for this PR?

One PR → three review questions → Three PRs

PR 1 · REFACTOR

Same behavior?

Registry + creation callers + equivalence tests

Human focus

Agree the abstraction; explain one creation path.

Oversight

PR 2 · TOOLBAR

Right experience?

Icons + labels + layout + interaction checks

Human focus

Confirm intent, keyboard use, focus and accessible names.

Oversight

PR 3 · ORDERING

Safe data transition?

Schema + backfill + read/write changes + migration tests

Human focus

Challenge ordering rules, compatibility and recovery.

Oversight

One PR → three risk levels → Three PRs

PR 3 · EXPAND

Add alongside

Add nullable `position`.
Keep old reads; maintain both representations on writes.

Before backfill

Verify mixed-version compatibility and consistent writes.

Lower risk 

PR 4 · MIGRATE & SWITCH

Backfill, then flag

Backfill in resumable batches.
Compare ordering; enable new reads gradually behind a flag.

Before wider rollout

Check results and load; test switching back after new writes.

Medium risk 

PR 5 · CONTRACT

Remove the old path

After validation and old clients retire, remove old writes, ordering data and compatibility code.

Before removal

Confirm no old consumers remain; agree recovery after deletion.

Higher risk 

Three takeaways

01 **Reduce risk in the change**
Make changes smaller and reversible.

02 **Use tools to separate concerns**
Let AI help split the work. Use stacked PR tools to coordinate dependencies.

03 **Focus attention on the highest-risk parts**
Clarify intent. Challenge evidence and recovery assumptions.

Recap



01 More PR and more help. Yet, focus matters!

Use AI's capabilities. Give review a clear direction.



02 SCOPE: risk-based, staged escalation

Decide where and when human involvement adds value.



03 Made oversight decisions

Match it to the change, system and context.
Make the work smaller and safer.

Thank you.

What questions are still open for you?

EMAIL michaela@michaelagreiler.com
X @mgreiler

UPCOMING COURSE

Code Review in the AI Era

Scan to get early access

FREE modules

awesomecodereviews.com/#course-sign-up