

Source dossier

It removes the reader — code review agents reach 71.9% of bugs in their top five comments, and the German act that makes that record discoverable lands in 56 days

Edition 2026-10-14 · narrative: Evidence you generate is evidence you hold · compiled 2026-10-14

Claims index

Claim	Statement	Status	Sources
CLAIM-2026-10-14-01	CRBugBench mines human-verified bugs from repository history and pairs each with the pull request that introduced it, comprising 1,000 review tasks across 98 open-source GitHub repositories plus 800 tasks from proprietary repositories at one large company.	verified	arxiv2026when
CLAIM-2026-10-14-02	On CRRugBench the best agent, claude-opus-5 with copilot-cli, catches 71.9% of bugs within its top five comments, up from 47.7% for opus-4.6 over five months.	verified	arxiv2026when
CLAIM-2026-10-14-03	The CRRugBench headline figure is a top-five measure over the agent's comments rather than a single verdict, and the paper reports agents weaker on complex changes, on bugs requiring reasoning beyond the local code, and noticeably worse on proprietary repositories where project context matters more.	verified	arxiv2026when
CLAIM-2026-10-14-04	The paper reports that scaling compute helps most through prompt design, with exploration-nudging prompts improving performance by up to 5.8% at 10% higher cost.	verified	arxiv2026when
CLAIM-2026-10-14-05	The paper's own adoption guidance is to scale review effort to pull-request complexity, steer agents toward project priorities, provide repository knowledge, choose models by pull-request characteristics and monitor quality as models change.	verified	arxiv2026when
CLAIM-2026-10-14-06	The Bundestag passed the Gesetz zur Modernisierung des Produkthaftungsrechts on Thursday 8 October 2026, on the Beschlussempfehlung of the Ausschuss fuer Recht und Verbraucherschutz dated 7 October 2026 (Drucksache 21/8429), and under Art. 17(1) the act takes effect on 9 December 2026.	verified	deutscherbu2026erste, deutscherbu2026beschluss

Claim	Statement	Status	Sources
CLAIM-2026-10-14-07	Paragraph 2(1) Nr. 3 of the new German product liability act includes software in the definition of product, excluding free and open-source software developed or supplied outside a business activity, and Paragraph 7(3) requires considering a product's ability to learn or acquire new functions after it is placed on the market.	verified	deutscherbu2026be schluss
CLAIM-2026-10-14-08	Paragraph 19 lets a court order disclosure of evidence on a plausible claim, and the committee amendment extends this to evidence held by third parties, with the ZPO right to refuse testimony applying and separate safeguards for trade secrets.	verified	deutscherbu2026be schluss
CLAIM-2026-10-14-09	Paragraph 20(1) presumes a defect where the defendant fails to disclose evidence despite a court order, where the product does not meet mandatory safety requirements, or where an obvious malfunction caused the harm; Paragraph 20 was unchanged by the committee.	verified	deutscherbu2026be schluss
CLAIM-2026-10-14-10	Under Paragraph 23, products placed on the market by 8 December 2026 remain subject to the old Produkthaftungsgesetz.	verified	deutscherbu2026be schluss
CLAIM-2026-10-14-11	AI Act Article 14's human-oversight duty applies to high-risk systems only, from 2 December 2027 for Annex III and 2 August 2028 for Annex I after Regulation (EU) 2026/1744.	verified	euaiactexpl2026hu man
CLAIM-2026-10-14-12	Art. 22(1) of Directive (EU) 2024/2853 sets the transposition deadline at 9 December 2026, which the German act's entry into force matches.	verified	deutscherbu2026be schluss, eurlex2024directive

Only claims marked **verified** may carry a post. **single-source** may appear as hedged context; **disputed** must not be published at all.

(When) Should a Software Factory be Lights-off?

Benchmarking Code Review Agents on Bug Detection

PRIMARY

Publisher	arXiv
Published	2026-10-09
Retrieved	2026-10-14
Cite key	arxiv2026when
URL	https://arxiv.org/abs/2610.13668

arXiv 2610.13668 [cs.SE], v1 submitted 9 October 2026, 14:37:15 UTC, announced in the 12 October listing. CC BY 4.0.

Authors as printed on the abstract page: Chenyang Yang, Hao He, Ti-Chung Cheng, Jenna Butler, Emerson Murphy-Hill, Carmen Badea, Christian Bird. No affiliations are shown on the abstract page and the HTML full text returned HTTP 429 to the fetch proxy, so no institution is attributed anywhere in this edition.

From the abstract record: developers spend more time reviewing than writing, and human review lags AI code generation, which is what makes review agents and 'lights-off' software factories attractive. CRBugBench mines bugs from repository history and pairs each with the pull request that introduced it. 1,000 review tasks across 98 open-source GitHub repositories, plus 800 tasks from proprietary repositories at one large company. Best agent: claude-opus-5 with copilot-cli catches 71.9% of bugs within its top five comments, up from 47.7% for opus-4.6 over five months. Agents are weaker on complex changes and on bugs requiring reasoning beyond the local code, and noticeably worse on proprietary repositories where project context matters more. Scaling compute helps most through prompt design: exploration-nudging prompts improve performance by up to 5.8% at 10% higher cost. The authors close with adoption guidance: scale review effort to PR complexity, steer agents toward project priorities, provide repository knowledge, choose models by PR characteristics, monitor quality as models change.

Note on the metric: top-five is a recall-at-k over the agent's comments, not a single verdict. No per-repository or per-complexity table is claimed in this edition.

Beschlussempfehlung und Bericht des Ausschusses fuer Recht und Verbraucherschutz

PRIMARY

Publisher	Deutscher Bundestag
Published	2026-10-07
Retrieved	2026-10-14
Cite key	deutscherbu2026beschluss
URL	https://dserver.bundestag.de/btd/21/084/2108429.pdf

Drucksache 21/8429, 21. Wahlperiode, Vorabfassung, 7 October 2026. Beschlussempfehlung und Bericht of the Ausschuss fuer Recht und Verbraucherschutz (6. Ausschuss) on Drucksache 21/4297, 'Entwurf eines Gesetzes zur Modernisierung des Produkthaftungsrechts', implementing Directive (EU) 2024/2853 and replacing the 1989 Produkthaftungsgesetz.

Committee result: adoption as amended, with CDU/CSU, SPD and Buendnis 90/Die Gruenen in favour; AfD and Die Linke abstaining. A resolution asks the government to examine liability insurance for authorised representatives of non-EU manufacturers.

Provisions relied on here:

- Paragraph 2(1) Nr. 3 includes software in the definition of product. Free and open-source software developed or supplied outside a business activity is excluded.
- Paragraph 7(3) requires considering a product's ability to learn or acquire new functions after it is placed on the market.
- Paragraph 19 lets a court order disclosure of evidence on a plausible claim. The committee amendment extends this to evidence held by third parties, with the ZPO right to refuse testimony applying. Trade secrets get a separate evidentiary procedure and natural persons of the parties can be excluded from access.
- Paragraph 20(1) presumes a defect where the defendant fails to disclose despite a court order, where the product does not meet mandatory safety requirements, or where an obvious malfunction caused the harm. Paragraph 20 was unchanged by the committee.
- Paragraph 23: products placed on the market by 8 December 2026 remain under the old law.
- Art. 17(1): the law takes effect 9 December 2026, matching Art. 22(1) of the Directive. Art. 2 and Arts. 5-16 take effect the day after promulgation.

The explanatory text says the law reaches manufacturers of AI systems through the software inclusion; the operative text does not name AI.

Erste Lesung zur Modernisierung des Produkthaftungsrechts

PRIMARY

Publisher	Deutscher Bundestag Textarchiv
Published	2026-03-04
Retrieved	2026-10-14
Cite key	deutscherbu2026erste
URL	https://www.bundestag.de/dokumente/textarchiv/2026/kw10-de-produkthaftungsrecht-1150448

Bundestag textarchiv entry for calendar week 10 of 2026. First reading of the government bill (Drucksache 21/4297) on Wednesday 4 March 2026, after which it was referred to the Ausschuss fuer Recht und Verbraucherschutz. The page records that the Bundestag passed the bill on Thursday 8 October 2026 (calendar week 41), and carries a 2./3. Lesung tab.

Used deliberately in place of law-firm coverage: the passage date is taken from the acting institution's own record, not from a commentary publication date.

Directive (EU) 2024/2853 on liability for defective products

PRIMARY

Publisher	EUR-Lex
Published	2024-10-23
Retrieved	2026-10-14
Cite key	eurlex2024directive
URL	https://eur-lex.europa.eu/eli/dir/2024/2853/oj/eng

Document date 23 October 2024; first entry into force 8 December 2024, per the EUR-Lex record. The transposition deadline of 9 December 2026 is relied on from Drucksache 21/8429, which cites Art. 22(1), because this retrieval returned only site navigation and metadata to the fetch proxy and the article text could not be read. Article wording is therefore not quoted from this source in this edition. Cite key reused from the vault, same document.

Human oversight - Article 14

SECONDARY

Publisher	EU AI Act Explorer
Published	2026-08-02
Retrieved	2026-10-14
Cite key	euaiactexpl2026human
URL	https://artificialintelligenceact.eu/article/14/

Article 14 human oversight, the duty the practitioner reflex reaches for. High-risk only. Application dates after Regulation (EU) 2026/1744: 2 December 2027 for Annex III, 2 August 2028 for Annex I. Cite key reused from the 13 October edition, same page.