Online Resource 1

How Developers Discuss Generative AI Across the AI-Integration Gradient: A Longitudinal, Multi-Repository Study of GitHub Issues

Panida Rumriankit, Akito Monden, Peerapas Chinnapatpakdeekul

International Journal of Networked and Distributed Computing

Corresponding author: Panida Rumriankit, Faculty of Engineering, Kasetsart University, Bangkok, Thailand (panida.rum@ku.th)

This file reports the interrupted time-series coefficients and sensitivity analyses (Tables S1–S3), the AI-event timing checks for the closure and emergence analyses (Tables S4–S5), and the sensitivity of the RQ1 topics to repeated boilerplate text (Tables S6–S7). All time-series models use the 67 monthly VS Code observations (January 2021–July 2026) and HAC standard errors with three lags. Copilot Chat issues have been tracked in the VS Code repository since 30 June 2025; the "Copilot issues moved to tracker" terms model this change.

Table S1: Primary interrupted time-series specification: agent-mode preview (February 2025) plus the July 2025 tracker consolidation. Left: $\log(n{+}1)$ monthly AI-keyword issues; right: logit of the AI-keyword share of all monthly issues. Bottom: treatment-by-period terms of the difference-in-differences model with non-AI-keyword issues as the control series, fit to the paired monthly contrast $\log(n_{\mathrm{AI}}{+}1) - \log(n_{\mathrm{other}}{+}1)$ so that the HAC errors respect the month pairing; up to the +1 offset it equals the share model

	Count model			Share model		
Term	Coef.	z	p	Coef.	z	p
Intercept	4.423	58.99	<0.001	−3.105	−28.69	<0.001
Baseline slope	0.029	4.74	<0.001	0.035	3.91	<0.001
Level: ChatGPT (2022-11)	−0.045	−0.35	0.725	−0.022	−0.16	0.873
Slope: ChatGPT	0.032	1.30	0.195	0.027	1.28	0.201
Level: Copilot Chat beta (2023-09)	−0.253	−1.55	0.121	−0.191	−1.27	0.203
Slope: Copilot Chat beta	−0.040	−1.80	0.071	−0.022	−1.12	0.264
Level: agent-mode preview (2025-02)	**0.587**	**5.31**	**<0.001**	**0.570**	**4.60**	**<0.001**
Slope: agent-mode preview	0.056	0.99	0.324	0.106	2.17	0.030
Level: Copilot issues moved to tracker (2025-07)	**1.385**	**5.87**	**<0.001**	**0.884**	**5.10**	**<0.001**
Slope: Copilot issues moved to tracker	−0.126	−2.25	0.025	−0.195	−4.63	<0.001
	DiD level			DiD slope		
ChatGPT	−0.021	−0.15	0.878	0.027	1.29	0.199
Copilot Chat beta	−0.190	−1.27	0.204	−0.022	−1.11	0.267
Agent-mode preview	0.568	4.60	<0.001	0.106	2.17	0.030
Copilot issues moved to tracker	0.883	5.10	<0.001	−0.195	−4.63	<0.001

Table S2: Agent-mode level shift (p) by breakpoint date, with and without the July 2025 consolidation term. ChatGPT (2022-11) and Copilot Chat beta (2023-09) breakpoints held fixed. *Retained*: classifier-retained monthly counts. AIC refers to the count model

Agent-mode breakpoint	Consol.	Count	Share	DiD	Retained	Consol. level (count)	AIC
Feb 2025 (Insiders preview)	no	1.08 (0.007)	1.15 (0.004)	1.15 (0.004)	1.58 (0.005)	—	50.2
Feb 2025 (Insiders preview)	yes	0.59 (<0.001)	0.57 (<0.001)	0.57 (<0.001)	0.67 (<0.001)	1.38 (<0.001)	-17.8
Mar 2025 (Stable 1.98)	no	1.04 (0.044)	1.09 (0.024)	1.09 (0.024)	1.68 (0.008)	—	53.8
Mar 2025 (Stable 1.98)	yes	0.28 (0.026)	0.34 (0.046)	0.34 (0.046)	0.64 (<0.001)	1.20 (<0.001)	-7.0
Apr 2025 (general availability)	no	1.14 (0.033)	1.15 (0.016)	1.15 (0.016)	1.78 (0.009)	—	50.1
Apr 2025 (general availability)	yes	0.16 (0.305)	0.28 (0.164)	0.28 (0.164)	0.51 (0.011)	0.98 (<0.001)	-4.9

Table S3: Month-by-month scan of the third breakpoint (ChatGPT and Copilot Chat beta fixed; no consolidation term). The best-fitting month (lowest AIC) is in bold

	Count model			Share model		
Breakpoint	Level	p	AIC	Level	p	AIC
2024-07	−0.06	0.788	64.9	0.09	0.722	57.6
2024-08	0.13	0.594	64.9	0.33	0.197	56.6
2024-09	0.23	0.395	64.9	0.51	0.066	54.9
2024-10	0.41	0.167	63.9	0.74	0.011	50.5
2024-11	0.46	0.185	64.1	0.77	0.029	50.3
2024-12	0.68	0.053	61.1	0.92	0.011	45.7
2025-01	0.94	0.008	54.6	1.10	0.002	37.0
2025-02	1.08	0.008	50.2	1.15	0.004	33.9
2025-03	1.04	0.044	53.8	1.09	0.024	36.7
2025-04	1.14	0.033	50.1	1.15	0.016	32.0
2025-05	1.33	0.004	39.4	1.29	0.001	18.9
2025-06	1.60	<0.001	15.5	1.37	<0.001	6.3
2025-07	**1.70**	**<0.001**	**0.6**	**1.36**	**<0.001**	**2.1**
2025-08	1.32	<0.001	38.0	1.07	<0.001	23.7
2025-09	1.06	0.019	50.3	0.78	0.024	35.4
2025-10	0.88	0.055	55.0	0.57	0.087	39.2
2025-11	0.73	0.095	57.2	0.30	0.388	42.7
2025-12	0.57	0.162	58.5	0.16	0.595	42.6

Table S4: AI-event timing. The AI event is the first keyword occurrence: issue creation if a keyword is in the title or body, otherwise the first comment containing one (core keywords first, then "cursor"/"gemini"). Top: lag from creation to the AI event (days). Bottom: pooled Cox hazard ratios (repository fixed effects + creation year; > 1 means AI-relevant issues close faster than incidental keyword matches), as reported in the article, keeping only AI-relevant issues whose AI evidence is in the title or body, and with AI relevance as a time-varying covariate that starts at the AI event (for issues whose first AI evidence appears after closure, exposure never starts while the issue is open). *After opening*: AI-relevant issues whose AI evidence appears only in the comments; in parentheses, those whose first AI evidence appears after closure

Label set	Group	n	Median lag	>30 days	>365 days	Later calendar year
A1	Contributor class, six comparison repos	130	0.3	9	3	6
A1	AI-relevant, AI-feature tiers	409	0.0	3	0	0
A1	AI-relevant, AI-bystander tier	185	0.1	10	4	7
A2	Contributor class, six comparison repos	10	0.5	1	1	1
A2	AI-relevant, AI-feature tiers	377	0.0	4	1	1
A2	AI-relevant, AI-bystander tier	34	0.0	2	1	2
Intersection	Contributor class, six comparison repos	8	0.5	0	0	0
Intersection	AI-relevant, AI-feature tiers	356	0.0	3	0	0
Intersection	AI-relevant, AI-bystander tier	20	0.0	1	1	1

Label set	Tier	AI n	After opening (closure)	As reported	Title/body only	Time-varying (p)
A1	AI-feature	408	35 (4)	1.44 [1.22, 1.70]	1.49 [1.26, 1.77]	1.51 [1.28, 1.79] (<0.001)
A1	AI-bystander	185	121 (10)	0.99 [0.80, 1.22]	0.92 [0.65, 1.30]	1.20 [0.96, 1.50] (0.108)
A2	AI-feature	376	24 (2)	1.41 [1.20, 1.66]	1.48 [1.25, 1.75]	1.50 [1.27, 1.76] (<0.001)
A2	AI-bystander	34	17 (2)	1.73 [1.18, 2.54]	2.37 [1.41, 3.97]	2.04 [1.38, 3.02] (<0.001)
Intersection	AI-feature	355	21 (2)	1.37 [1.16, 1.62]	1.42 [1.20, 1.67]	1.42 [1.20, 1.67] (<0.001)
Intersection	AI-bystander	20	9 (1)	2.19 [1.34, 3.56]	2.77 [1.47, 5.22]	2.66 [1.61, 4.41] (<0.001)

Table S5: Design-weighted estimates of contributor-class issues in the six comparison repositories, dated by issue creation (as in the article) / by AI event. 95% stratified-bootstrap intervals for the AI-event dating in 2026: A1 [89.0, 146.4], A2 [4.0, 34.7], intersection [3.0, 34.6]. 2026 covers January–August. For census repositories the intervals treat the 2021–2026 issues as one realization of an ongoing process; no interval includes annotation error

Label set	2021	2022	2023	2024	2025	2026
A1	1.0 / 0.0	2.0 / 1.0	6.0 / 6.0	5.0 / 3.0	33.0 / 33.0	112.7 / 116.7
A2	0.0 / 0.0	0.0 / 0.0	1.0 / 0.0	0.0 / 0.0	1.0 / 2.0	17.4 / 17.4
Intersection	0.0 / 0.0	0.0 / 0.0	0.0 / 0.0	0.0 / 0.0	1.0 / 1.0	16.4 / 16.4

Table S6: RQ1 sensitivity to boilerplate. Boilerplate is every 8-token sequence of the preprocessed text that occurs in at least 1% (321) or 0.5% (161) of the 32,096 classifier-retained VS Code documents; tokens covered by such a sequence are removed (at 1%: 257 sequences, 62% of documents affected, 857 documents left with fewer than five tokens and dropped). The most frequent are the triage bot's request for missing information and the Copilot reporting guidelines (about 8,300 documents), closing, stale-issue, old-version and duplicate notices (1,900–5,600 each), and service-outage and screencast instructions (about 340). Each condition refits the article's BERTopic configuration with the same all-MiniLM-L6-v2 embeddings and UMAP seeds 1–5 (the published model was unseeded); values are medians (ranges) over seeds. *MK*: volume-restricted Mann–Kendall trends as in the article. *Kept*: published top-10 topics with at least half of their documents in one rerun topic. *AI*: topics among the ten largest that concern operating AI assistants; the others are Windows Subsystem for Linux, settings sync, non-English reports, and clusters organized around bot or template text (rules on each topic's top terms, `composition.py`; every call is listed in `top10_topic_categories.csv` in the archive). In all ten boilerplate-removed runs the largest topic is rate limits and quotas (4.8–5.3% of documents)

Condition	Docs	Tokens rem. (%)	Topics	Outl. (%)	MK up / down of tested	Kept	AI
Published (unseeded)	32,096	—	107	41.8	79 / 7 of 105	—	6
Original text, seeded	32,096	—	102 (99–104)	39.7	73 (71–77) / 8 (7–9) of 99 (98–102)	10 (9–10)	6
Removed ≥1%	31,239	16.3	93 (91–94)	39.0	61 (60–65) / 8 (7–8) of 88 (86–90)	7	7 (7–8)
Removed ≥0.5%	30,579	18.6	83 (83–85)	39.7	59 (57–61) / 7 (6–8) of 81 (81–83)	7	7

Table S7: Share of each published top-10 topic's documents that fall in its best-matching topic of a rerun, median (range) over five seeds. Labels as in Table 2 of the article

ID	Published topic	Original text	Removed ≥1%	Removed ≥0.5%
0	Agent file tools and skills	0.81 (0.74–0.83)	0.65 (0.59–0.71)	0.66 (0.54–0.73)
1	Agent terminal commands	0.96 (0.94–0.98)	0.93 (0.93–0.95)	0.93 (0.92–0.96)
2	Windows Subsystem for Linux	0.99 (0.99–0.99)	0.86 (0.85–0.89)	0.54 (0.53–0.56)
3	MCP server configuration	0.99 (0.98–0.99)	0.97 (0.96–0.98)	0.97 (0.96–0.98)
4	Inline completion feedback	0.97 (0.95–0.98)	0.42 (0.41–0.43)	0.41 (0.40–0.43)
5	Non-English reports	0.91 (0.88–0.94)	0.78 (0.75–0.79)	0.76 (0.74–0.80)
6	Chat response errors (boilerplate)	0.95 (0.93–0.99)	0.11 (0.09–0.16)	0.10 (0.08–0.33)
7	Rate limiting	0.98 (0.98–0.99)	0.90 (0.87–0.91)	0.89 (0.88–0.92)
8	Custom model endpoints	0.90 (0.47–0.90)	0.85 (0.81–0.88)	0.85 (0.80–0.90)
9	Chat request failures (boilerplate)	0.90 (0.87–0.91)	0.15 (0.13–0.16)	0.15 (0.09–0.17)