

The Developer Agentic Arena

Cursor vs. Windsurf vs. Claude Code vs. AGY vs. Kiro vs. Aider

Two of these six tools have been renamed, one was sold for \$60 billion, one is built for enterprises on AWS, and one has not shipped a feature release in over a year. A guide to what each has become, what it costs, and how to choose.

Research current to 10 October 2026. Prices in US dollars before tax. Bracketed numbers link to the sources listed at the end.

THE SHORT VERSION

- **The products converged.** Cursor, Devin Desktop (the former Windsurf), Claude Code, Google Antigravity and AWS Kiro now ship much the same kit: an editor or desktop agent manager, a terminal CLI, cloud-hosted agents, and extension through MCP servers, skills, hooks and subagents. They differ in who owns them, which models they favour and how usage is metered.
- **Two names in the title are out of date.** Windsurf became Devin Desktop on 2 June 2026. AGY is the command name of Google's Antigravity CLI, which replaced Gemini CLI for consumer accounts on 18 June 2026.
- **Prices lined up at about \$20, \$100 and \$200 a month.** What the money buys did not: the same \$20 gets different models, quotas and overage rules in each product.
- **Kiro is the one built around enterprise control and written specs.** It has a small public profile, AWS sign-in and billing, a GovCloud edition, and the most explicit usage pricing of the group.
- **Five of the six belong to a model lab or a cloud giant:** SpaceXAI, Cognition, Anthropic, Google and Amazon. Aider is the one independent, open-source tool, and its last feature release was in August 2025.
- **Security is the shared weak point.** Prompt injection through repository content, pull requests and tool descriptions has produced serious vulnerabilities in several of these tools this year.

First, the names

The title of this article uses the names developers still say out loud, and two of them no longer appear on a download page.

Windsurf is now **Devin Desktop**. Cognition, which makes the Devin agent, bought Windsurf in 2025 and retired the brand on 2 June 2026 [1]. The old name survives in places, including Cognition's own pricing table, which still lists a "Windsurf IDE usage quota" [2]. This article says Devin Desktop for the current product and Windsurf for its history.

AGY is **Google Antigravity**. Strictly, **agy** is the binary name of the Antigravity command-line tool, but the abbreviation has spread to the whole platform, which also includes a desktop app, an IDE and an SDK [3].

Kiro has kept its name and absorbed another. The Amazon Q Developer CLI became the Kiro CLI in November 2025 [4], and AWS will end support for the Q Developer IDE plugins on 30 April 2027 [5].

How the arena got here

In 2023 the open question was whether a language model could edit a real repository without wrecking it. Aider answered that from a terminal. Through 2024 and 2025 the contest moved into the editor, with Cursor and Windsurf both building on VS Code. In 2026 the contest is over who supervises the agents: every commercial product in this comparison now lets a developer launch, watch and review several agents at once.

The same eighteen months rearranged the ownership of the field.

When	What happened
2023	Aider's first public releases establish the terminal pair-programmer pattern [6].
May 2025	Claude Code reaches full launch after a research preview [7].
July 2025	Google hires Windsurf's founders in a licensing deal reported at \$2.4 billion; Cognition signs to acquire the product, brand and remaining team [8].
July 2025	AWS launches Kiro in preview; by Kiro's account more than 100,000 developers try it in five days [9].
17 November 2025	Kiro reaches general availability with team administration, and the Amazon Q Developer CLI becomes the Kiro CLI [10][4].
November 2025	Google launches Antigravity alongside Gemini 3 as a free, agent-first IDE [3].
February 2026	Aider 0.86.2, a patch release and still the latest on PyPI [6]. Claude Code's run-rate revenue is reported above \$2.5 billion [11].
19 March 2026	Windsurf replaces credits with quotas and raises Pro from \$15 to \$20 [12].
19 May 2026	Antigravity 2.0 at Google I/O: desktop app, Go-based CLI, SDK and a new \$100 tier [3].
2 June 2026	Windsurf becomes Devin Desktop [1].
16 June 2026	SpaceX agrees to buy Cursor's parent, Anysphere, for \$60 billion in stock [13].
18 June 2026	Gemini CLI stops serving free and individual paid accounts [3].
14 August 2026	The Cursor acquisition closes [14].
10 September 2026	Cursor launches Projects, a coordinator agent that delegates to subagents [15].
2 November 2026	Scheduled: Antigravity removes Claude 4.6 models and GPT-OSS-120b [16].
30 April 2027	Scheduled: AWS ends support for the Amazon Q Developer IDE plugins and directs their users to Kiro [5].

Dates are those given by the vendors or by the press reports cited.

The six contenders

Cursor

OWNER	FORM	ENTRY PRICE
SpaceXAI, through Anysphere	VS Code-based editor with an agents window, CLI, cloud agents and an iOS app	Free Hobby plan; Pro \$20 a month

Cursor began as an AI-first fork of VS Code and became the commercial centre of the category. Press reports put its annualised run rate at about \$2 billion in February 2026 and roughly \$4 billion by early June [17]. On 16 June SpaceX, which had absorbed xAI in February, agreed to buy Cursor's parent for \$60 billion in stock. The deal closed on 14 August [13][14].

The editor is now one surface among several. Cloud Agents run on remote machines, including self-hosted pools inside a customer's own network. Projects, in beta since 10 September, puts a coordinator agent in charge of a long-running body of work and has it plan, delegate to subagents and bring results back for checking. Bugbot reviews pull requests, and two bots added on 23 September extend that to exploitable bugs (Security Review) and post-deploy health (Rollouts). Since 6 October, agents running on a developer's own machine can be watched and answered from the iOS app [15]. Cursor also hosts repositories itself under the name Origin, and publishes a CLI and TypeScript and Python SDKs [18].

Cursor offers the broadest menu of frontier models among the commercial tools: Claude, GPT, Gemini, Meta's Muse Spark, Grok and its own Composer. How it charges for them changed with the new owner. Paid individual plans (Pro at \$20, Pro Plus at \$60, Ultra at \$200) carry two usage pools. One covers "Cursor Models", meaning Grok 4.5 to 4.7 and Composer 2.5, with what Cursor describes as significantly more included usage. The other covers third-party models, billed at each model's API price. Teams seats cost \$40 (Standard) or \$120 (Premium), and on Teams and Enterprise every third-party request carries an extra \$0.25 per million tokens that first-party models are exempt from [18].

Cursor's own guidance on real spending is unusually frank: daily agent users typically run to \$60–100 a month in total, and people running several agents or automations often pass \$200 [18].

Watch for. The pricing now tilts toward the owner's models, which matters to anyone who chose Cursor for neutrality. On security, a remote-code-execution flaw rated 9.9 out of 10 (CVE-2026-26268) was disclosed in April and fixed in version 2.5 [19].

Windsurf, now Devin Desktop

OWNER Cognition	FORM VS Code-based IDE that opens on an agent command center; plus Devin Cloud, CLI and Review	ENTRY PRICE Free plan; Pro \$20 a month
--------------------	--	--

No tool here has changed hands and names as often. The company was Codeium, then Windsurf. In July 2025 Google hired its founders, and within days Cognition signed to acquire the product, the brand and the remaining team [8]. In March 2026 the new owner replaced credits with quotas and raised Pro from \$15 to \$20, erasing the price advantage over Cursor [12]. On 2 June an over-the-air update turned Windsurf into Devin Desktop, with plans, pricing and extensions carried over [1].

The rename came with a change of emphasis. The default screen is an Agent Command Center, a Kanban-style board of local and cloud agents. Spaces group sessions, pull requests and files so that agents working on the same thing share context. Cascade, the agent that made Windsurf's reputation, gave way to Devin Local, a rewrite in Rust that Cognition says is up to 30% more token-efficient and that supports subagents. Cascade was switched off on 1 July [1][20].

The distinctive bet is openness to other vendors' agents. Devin Desktop supports the Agent Client Protocol (ACP), an open standard for running an agent inside an editor. At launch that meant Codex, Claude Agent, OpenCode and in-house agents appear on the same board and use the same Spaces as Devin [1]. Of the five commercial products, this is the one most plainly positioned as a host for agents it does not make.

Individual plans are Free, Pro at \$20 and Max at \$200; Teams is listed at \$80 a month and Enterprise is custom. A single quota covers both the cloud agent and the IDE, with pay-as-you-go beyond it, and Pro and Max allow up to ten concurrent sessions [2].

Watch for. Three names in about three years has left documentation and tutorials split across brands. An early reviewer found Spaces unfinished [20]. The pricing page does not say how large the quotas are, and the speed and efficiency figures for Devin Local and Cognition's SWE models are the vendor's own.

Claude Code

OWNER Anthropic	FORM Terminal CLI first; also VS Code and JetBrains extensions, a desktop app, web and mobile	ENTRY PRICE Included in Claude Pro at \$20 a month (\$17 billed annually); no free tier
--------------------	---	--

Claude Code launched in 2025 as a command-line agent with no editor of its own, and that is still its character. It reads a repository, edits files, runs commands and handles git, and it behaves like a Unix tool: logs can be piped into it and it can run headless in CI [21]. Its run-rate revenue was reported above \$2.5 billion by February 2026 [11], and a JetBrains survey that January, as reported, put it level with Cursor at 18% of developers, behind GitHub Copilot at 29% [22].

The same engine now runs in the terminal, in IDE extensions, in a desktop app, in the browser and on a phone, and a session can move between them. Routines run on a schedule in the cloud or fire on GitHub events. There are integrations for GitHub Actions, GitLab CI/CD and Slack, where mentioning the agent on a bug report returns a pull request [21].

Customisation runs through plain files and a small set of concepts: a CLAUDE.md instruction file per project (an existing AGENTS.md is read too), memory that the agent builds as it works, skills that package repeatable workflows, hooks that run shell commands around its actions, MCP servers for outside tools, and subagents that split a task. An Agent SDK exposes the same tools for building custom agents [21]. Much of this vocabulary has spread: the documentation for Cursor, Antigravity and Kiro now lists skills, hooks, subagents and MCP as well [18][23][24].

Claude Code runs Claude models only, reached through a Claude subscription, an API key or a third-party cloud provider [21]. On a subscription it draws from the same allowance as Claude chat, in rolling five-hour sessions with weekly limits on top. Max costs \$100 for five times Pro's usage or \$200 for twenty times. Team seats are \$20 (Standard) or \$100 (Premium) on annual billing, and Enterprise is \$20 a seat plus usage at API rates. List API prices are \$2 in and \$10 out per million tokens for Sonnet 5.5, and \$4 and \$20 for Opus 5.5 [25].

Watch for. One vendor's models, and no free tier to try it on. Because limits are shared, heavy chat use eats into coding capacity. Among agentic projects tracked by OWASP, the Claude Code repository has the second-highest count of published security advisories, at 22 [26], a number that reflects how much scrutiny it gets as well as how many flaws were found.

AGY: Google Antigravity

OWNER	FORM	ENTRY PRICE
Google	Antigravity 2.0 desktop app, the agy terminal CLI, an IDE, extensions for other IDEs, and a Python SDK	\$0 with weekly limits; Google AI Pro \$19.99 a month

Antigravity arrived in November 2025 with Gemini 3 as a free, agent-first IDE, four months after Windsurf's founders moved to Google [3][8]. At Google I/O on 19 May 2026 it was relaunched as a platform. Antigravity 2.0 is a standalone desktop app for launching, monitoring and orchestrating agents, independent of any editor [23]. Beside it sit a new CLI written in Go, an SDK, and the original IDE. The CLI replaced Gemini CLI, which stopped serving free and individual paid accounts on 18 June [3][27].

Antigravity's signature idea is the artifact. Agents report through reviewable outputs, among them implementation plans, walkthroughs, screenshots and browser recordings, and the developer reviews those instead of reading a raw transcript. Control of Chrome is built in. The CLI and the desktop app share one agent harness: settings and permissions sync between them, and a terminal conversation that grows complicated can be exported to the desktop app [28]. Extensions bring the agent into VS Code, Visual Studio, JetBrains IDEs, Zed and Xcode [23].

The platform is Gemini-first. Every plan gets Gemini 3.1 Pro and the 3.6 to 3.8 Flash models. Claude Sonnet 5.5 and Opus 5.5 are offered on paid Pro and on Ultra. The free tier currently also includes the older Claude 4.6 models and GPT-OSS-120b, which Google says it will remove

on 2 November. Enterprise accounts get Gemini only [16].

There is no Antigravity subscription as such. Capacity comes with a Google AI plan: \$19.99 for Pro, \$99.99 for Ultra at five times Pro's allowance, \$199.99 for twenty times [27]. Organisations buy through Google Cloud, from \$30 a seat or on a pay-as-you-go basis [29]. Limits follow the amount of work the agent does, not the number of prompts; Pro and Ultra refresh every five hours up to a weekly cap. Bringing your own API key is not supported [30].

Watch for. The limits are unpublished and, in Google's words, subject to modification [30]. Google retired Gemini CLI for consumers with a month's notice, and the free tier's third-party models are about to shrink. Security researchers at Pillar Security have reported prompt-injection and sandbox-escape findings affecting Antigravity among other agents [31].

Kiro

OWNER Amazon Web Services	FORM IDE on VS Code's open-source base, organised around specs; plus a CLI, the Kiro Web cloud agent, a mobile app in preview and the Crew assistant	ENTRY PRICE Free with 50 credits a month; Pro \$20 a month
------------------------------	---	---

Kiro is AWS's agentic development environment. It launched in preview in July 2025, when by Kiro's account more than 100,000 developers tried it in five days, and reached general availability on 17 November 2025 with a CLI and team administration [9][10]. That CLI was the Amazon Q Developer CLI under a new name and a proprietary licence in place of Apache 2.0 [4]. AWS has since set 30 April 2027 as the end of support for Q Developer's IDE plugins and directs their users to Kiro [5]. For organisations that standardised on AWS's earlier assistants, Kiro is now the designated path.

What sets Kiro apart is spec-driven development. Before writing code, the agent turns a request into three documents: requirements, a design and a task list. The developer edits them, and the agent then works through the tasks. Bugfix specs do the same for root-cause analysis and regression prevention. Around that sit steering files, which hold project standards; hooks, which fire on file changes, tool use or task completion; and Powers, which bundle a tool with the knowledge of how to use it and load on demand [24]. Kiro counts more than 100 curated Powers and over 15,000 made by its community [9]. General availability added property-based tests that check generated code against the spec, and checkpoints for rolling back an agent's changes [10].

The IDE, the CLI and a browser agent share one harness and one .kiro project folder [24]. Kiro Web takes a task into an isolated cloud sandbox, plans it, writes the code and opens a pull request on GitHub or GitLab [32]. Crew is a persistent assistant that you host yourself and reach from Slack, Teams, Discord and other channels, with schedules and memory [33]. Since 30 September, Workflows run multi-step, multi-agent plans in the background across the IDE, the CLI and the web [34].

Kiro has no house model. The default is Auto, a router that mixes models and sets the baseline price. Named models cost a published multiple of it: 1.3 times for Claude Sonnet 5.5, 2.0 for

Claude Opus 5.5, 4.4 for GPT-5.6 Sol, and as little as 0.05 for the open-weight Qwen3 Coder Next [35]. Plans are sold in credits: 50 a month free, 1,000 for \$20 (Pro), 2,000 for \$40 (Pro+), 5,000 for \$100 (Pro Max) and 10,000 for \$200 (Power), with extra credits at \$0.04 each and no rollover [36].

The enterprise footing is the real difference. Sign-in runs through AWS IAM Identity Center, Okta or Entra ID. Administrators decide from central registries which models, MCP servers and extensions developers may use. Paid plans exclude customer content from model training and include indemnity for generated output. Kiro is a HIPAA-eligible service and is offered in AWS GovCloud (US), at about 20% above commercial pricing. Named customers include Siemens, FINRA, Rackspace, NTT Data and Mondelez [37][36]. That reach is hard to see from outside: one tracker ranks Kiro 20th of 44 coding agents by public footprint and notes that it has no public repository, no package downloads and no default commit attribution to count [38].

Watch for. Kiro Web runs in one AWS region, US East, and an administrator must enable cloud sessions before an organisation can use it [32]. The model list has no Gemini or Grok, and availability varies by country [36]. On security, CVE-2026-0830 allowed a crafted workspace to run commands and was fixed in version 0.6.18 [39], and researchers at Mindgard showed that content in a malicious repository could make the agent send local data to an outside server, which Amazon says it fixed in version 0.8.140 [40].

Aider

OWNER	FORM	ENTRY PRICE
Independent (Paul Gauthier); Apache 2.0 licence	Terminal pair programmer that works inside your git repository	Free; you pay your model provider directly

Aider is the oldest of the six and the template for the terminal agents that followed. It builds a map of the repository's classes, functions and their relationships to give the model context, applies edits in a diff format chosen for each model, and commits every change to git with a generated message, so that /undo and ordinary git tools are the safety net [41]. An architect mode pairs a strong reasoning model with a cheaper one that writes the edits. It can lint and test after each change and try to fix what fails [42]. It connects to almost any model, hosted or local, using keys you supply [6].

The caveat is the state of the project. The last feature release, 0.86.0, shipped in August 2025. The only release since is 0.86.2, on 12 February 2026 [6]. The main branch has kept receiving changes, including definitions for Claude 4.5 and 4.6, Gemini 3 previews and GPT-5.3 and 5.4, but they are unreleased [42], and one tracker reports that the project's model leaderboards were last updated in November 2025 [43]. The release notes record no MCP client, no subagents, no cloud runner and no agent manager. Newer models can generally be addressed by name, but without the tuned per-model settings Aider is known for.

It still earns its place for three reasons. Cost is transparent: there is no subscription, and every token appears on your own provider bill at list price. Behaviour is auditable: nothing changes

that is not a git commit. And it goes where the others cannot easily follow, over SSH, into a container, or against a local model on a machine with no outside network. Its Polyglot benchmark, 225 Exercism problems across six languages, was for two years a standard reference for how well models edit code [44].

Watch for. Maintenance risk. Check for a new release or an active fork before standardising a team on it.

Head to head

The tables below compress the profiles. Product details come from each vendor's documentation and pricing pages as read on 10 October 2026.

	Cursor	Devin Desktop (Windsurf)	Claude Code	Antigravity (AGY)	Kiro	Aider
Owner	SpaceXAI, via Anysphere	Cognition	Anthropic	Google	Amazon Web Services	Independent, open source
Home surface	Editor built on VS Code, with an agents window	Agent command center inside a VS Code-based IDE	Terminal	Desktop agent manager; agy in the terminal	IDE on VS Code's open-source base, organised around specs	Terminal
Other surfaces	CLI, cloud agents, iOS app, Slack, JetBrains and Xcode integrations	Devin Cloud, CLI, Review, JetBrains plugin	VS Code, JetBrains, desktop app, web, mobile, Slack, CI	IDE; extensions for VS Code, Visual Studio, JetBrains, Zed, Xcode; SDK	CLI, Kiro Web, mobile app (preview), Crew assistant	Watch mode for any editor; experimental browser UI
Models	Grok and Composer (first-party pool); Claude, GPT, Gemini, Muse Spark	Cognition's SWE models, plus OpenAI, Claude and Gemini	Claude only	Gemini on every plan; Claude on paid consumer tiers	Auto router; Claude, GPT-5.6 and open-weight models	Almost any, including local models
Cloud agents	Yes: Cloud Agents and Projects, with a self-hosted option	Yes: Devin Cloud	Yes: web sessions and scheduled routines	Managed Agents through the Gemini API; scheduled tasks	Yes: Kiro Web cloud sessions, US East only	No
Extension points	MCP, rules, skills, hooks, subagents, plugins	MCP, skills, subagents, third-party agents over ACP	MCP, skills, hooks, subagents, plugins, Agent SDK	MCP, skills, rules, hooks, subagents, plugins, SDK	MCP, specs, steering, hooks, Powers, skills, subagents	Conventions file, YAML config, scripting
Enterprise identity	SAML/OIDC SSO on Teams; SCIM on Enterprise	SAML/OIDC SSO on Enterprise	SSO on Team; SCIM and audit logs on Enterprise	Admin controls through Google Cloud plans	IAM Identity Center, Okta, Entra ID; GovCloud edition	None; runs on your own keys
Licence	Proprietary	Proprietary	Proprietary	Proprietary	Proprietary	Apache 2.0

Table 1. What each tool is. Sources: vendor documentation [18][45][2][21][25][23][29][24][37][42].

	Cursor	Devin Desktop (Windsurf)	Claude Code	Antigravity (AGV)	Kiro	Aider
Free tier	Hobby: limited agent requests	Free: limited usage	None for Claude Code	\$0 with weekly limits	50 credits a month	The tool itself is free
Entry paid	Pro \$20	Pro \$20	Pro \$20 (\$17 billed annually)	Google AI Pro \$19.99	Pro \$20 (1,000 credits)	–
Middle	Pro Plus \$60	–	Max 5x \$100	Ultra 5x \$99.99	Pro+ \$40; Pro Max \$100	–
Top	Ultra \$200	Max \$200	Max 20x \$200	Ultra 20x \$199.99	Power \$200 (10,000 credits)	–
Team	\$40 or \$120 per seat	Teams \$80 a month	\$20 or \$100 per seat, annual billing	From \$30 per seat through Google Cloud	Same per-user plans up to 500 users; larger by quote	–
How usage is metered	Two monthly pools: first-party models, and third-party at API price	One quota shared by the cloud agent and the IDE	Five-hour sessions plus weekly limits, shared with Claude chat	Agent work done; five-hour refresh up to a weekly cap on paid plans	Monthly credits; each model has a published multiplier	Per token, on your own API key
Past the limit	On-demand at API rates	Pay-as-you-go	Usage credits at API rates	Purchased AI credits; no own-key option	\$0.04 per extra credit	Not applicable

Table 2. What a month costs. Monthly prices per individual unless stated. Sources: vendor pricing pages [45][18][2][25][29][27][36].

Two things in Table 2 deserve a second look. First, list prices are close to identical, so the real comparison is how fast each plan's allowance runs out on your work. Only Kiro states its allowance as a number, and that number is in credits, a unit of its own. Every Kiro plan prices included credits at two cents each and extra credits at four, so there is no volume discount and running over costs double. Second, the free row is misleading for Aider: an afternoon of agentic work on a frontier model can cost more in tokens than a month of a \$20 plan, and less on a small or local model. With Aider the bill is whatever you choose to make it.

Three fault lines

Who picks the model

Four of the five commercial tools have a house model and a financial reason to prefer it. Cursor gives Grok and Composer a larger included allowance and exempts them from its per-token fee [18]. Antigravity runs Gemini on every plan and restricts other models by tier, with none at all for enterprise accounts [16]. Claude Code runs Claude and nothing else. Devin Desktop promotes Cognition's SWE models for its own agent, while using ACP to let rivals' agents run inside its interface.

Kiro is the exception. It resells Claude, GPT and open-weight models, each at a published credit multiplier, and steers through price: its Auto router sits at 1.0, below the 1.3 of Claude Sonnet 5.5 and the 2.0 of Opus 5.5 [35]. Its menu has gaps of its own, with no Gemini or Grok.

For a buyer the practical question is how much it would cost to change your mind. If the best coding model next spring comes from a lab your tool does not favour, Cursor will probably offer it at API price, Devin Desktop may host that lab's agent, Kiro's answer will depend on whether it already carries that lab, Antigravity and Claude Code will most likely not offer it, and Aider will accept it as soon as you paste in a key.

Where the agent runs, and who is watching

The 2026 feature race was about supervision. Cursor's Projects, Devin Desktop's command center, Antigravity 2.0, Kiro's Workflows and Claude Code's background agents all assume a developer overseeing several agents in parallel, some on the laptop and some in the cloud, and checking in from a phone. Cursor says a Project coordinator can delegate to thousands of subagents [15].

Nobody has shown that review capacity scales the same way. A person can launch ten agents far faster than they can read ten diffs. Most vendors answer with another agent that reviews: Bugbot and Security Review at Cursor, Devin Review at Cognition, code review on pull requests in Claude Code. Kiro's answer is to move the human review earlier, onto the requirements and design documents of a spec, before any code exists [24]. Teams adopting parallel agents should decide in advance what a human must read before merge, and measure cost per merged change instead of cost per seat.

The year's clearest cautionary tale involves the maker of one of these tools. The Financial Times reported in February 2026 that AWS engineers had let Kiro make changes to a live system the previous December, and that the agent chose to delete and recreate the environment, causing a 13-hour disruption. Amazon disputed the framing. It said the disruption was confined to Cost Explorer in one region, traced it to a misconfigured role and therefore to user error, and said it had since required peer review for production access [46]. Either account leads to the same lesson: an agent acts with whatever permissions it is handed.

What can go wrong

An agent that reads files, runs shell commands and browses the web will sometimes read text written by an attacker, and it cannot reliably tell that text apart from its instructions. OWASP's 2026 report on agentic security found that 28 of the 53 agentic projects it tracks are coding agents, and it ties prompt injection to six of the ten categories in its risk list for agentic applications [26].

The year's incidents show the pattern. Cursor's CVE-2026-26268 let a project containing an embedded bare git repository run code through git hooks. Kiro's CVE-2026-0830 ran commands hidden in the folder names of a crafted workspace [39]. A technique named Comment and Control hijacked Claude Code, Gemini CLI and GitHub Copilot running in GitHub Actions with nothing more than a pull-request title or body [19]. Reported findings against Cursor and Antigravity described ways out of the sandbox that never attacked the sandbox directly [26][31]. Aider's exposure is smaller mainly because it does less by itself: it asks before running commands and has no cloud runner.

A WORKING BASELINE FOR ANY OF THE SIX

- Treat repository content, issues, pull-request text and web pages as untrusted input to the agent, not only to your application.
- Keep the sandbox and command-approval settings on for local agents, and give cloud agents the narrowest repository and secret scopes that let them finish the task.
- Do not hand an agent a role that can change production without a second person's approval.
- Do not let an agent in CI act on pull requests from forks with write permissions or secrets in scope.
- Install MCP servers, skills and plugins from sources you would accept a dependency from, and pin their versions.
- Update promptly. The serious flaws of 2026 were fixed quickly once disclosed.

How to choose

No one tool leads on every count, and no independent benchmark compares all six like for like, so the sensible starting point is how you work.

- **You live in an editor and want one product for completion, chat and agents.** Cursor. It has the most mature editor and the broadest frontier-model menu, at a cost that Cursor itself puts at \$60–100 a month for daily agent use.
- **You expect to supervise many agents, from more than one vendor, in one place.** Devin Desktop. Its board, Spaces and ACP support are built for that, and the cloud agent is included from the \$20 plan.
- **You work in the terminal, script your tooling, or want agents in CI and chat.** Claude Code. It is built to be composed with other tools and has the widest spread of surfaces, in exchange for a single model family.
- **You want to start at \$0, already pay for a Google AI plan, or build on Google Cloud.** Antigravity. The free tier is real, the artifact-based review is worth trying, and the limits and model list are the things to keep an eye on.
- **Your organisation runs on AWS, needs central control of models and tools, or works under compliance rules.** Kiro. Sign-in, billing and governance come through AWS, there is a GovCloud edition, and the spec workflow leaves written requirements and a design beside the code.
- **You need auditability, local models or strict cost accounting.** Aider, with the maintenance question answered first.

These choices combine more easily than the rivalry suggests. Claude Code installs as an extension inside Cursor [21]. Devin Desktop runs Claude Agent and Codex beside its own agent [1]. Aider's watch mode works alongside any editor. A common pattern is an editor for reading and small edits and a terminal or cloud agent for longer tasks.

Before committing a team, run a two-week trial on your own repository. Give each candidate the same three tasks: a bug with a failing test, a change that touches several files, and a dependency upgrade. Record what each cost, how much of the output was merged without rework, and how long review took. Those three numbers will tell you more than any feature table, including the ones above.

What to watch next

- **Cursor under SpaceXAI.** Whether third-party models stay available at parity, or the first-party pool keeps growing at their expense.
- **Antigravity's free tier after 2 November.** The removal of Claude 4.6 and GPT-OSS-120b will show how far Google intends to narrow the free plan to Gemini.
- **ACP adoption.** If more editors and agents support it, the choice of interface and the choice of agent come apart, which favours Devin Desktop's strategy.
- **The Amazon Q Developer wind-down.** Support for the Q Developer IDE plugins ends on 30 April 2027, which pushes their remaining users toward Kiro. Whether Kiro Web expands beyond the US East region will matter to teams outside it.
- **Aider's next release.** A tagged release from the main branch, or a fork that takes over, would settle the main doubt about it.
- **Quota transparency.** Kiro publishes credit counts, per-model multipliers and a flat overage price. The other four commercial vendors describe their limits in relative terms. Limits stated in comparable units would make buying easier for everyone.

Method, limits and disclosure

Prices, plan contents, model lists and feature descriptions were taken from the vendors' own pricing and documentation pages on 10 October 2026 wherever those were available. Corporate history, revenue and survey figures come from press and analyst reports and are marked as reported; several rest on a single secondary source. Where sources disagreed, for example on the size and timing of the Windsurf acquisition, the text gives only what the cited source states.

No hands-on benchmarking was done for this article, and vendor performance claims are attributed, not endorsed. This market changes monthly, so prices and limits should be confirmed on the vendor's page before buying.

This article was researched and drafted with Claude, an AI model made by Anthropic. Claude Code is an Anthropic product, Kiro relies heavily on Anthropic's Claude models, and Amazon is an investor in Anthropic. Readers should weigh those sections accordingly.

Sources

1. Runtime Wire, Cognition turns Windsurf into Devin Desktop. <https://runtimewire.com/article/cognition-turns-windsurf-into-devin-desktop-a-command-center-for-coding-agents>
2. Devin (Cognition), Pricing. <https://devin.ai/pricing>
3. The Next Web, Antigravity turns into a full agentic development platform. <https://thenextweb.com/news/google-antigravity-2-desktop-cli-sdk-io-2026>
4. Kiro Docs, Migrating from Amazon Q Developer CLI. <https://kiro.dev/docs/upgrade-guides/migrating-from-q/>
5. AWS Docs, Amazon Q Developer IDE plugins end of support. <https://docs.aws.amazon.com/amazonq/latest/qdeveloper-ug/q-developer-ide-end-of-support.html>
6. PyPI, aider-chat. <https://pypi.org/project/aider-chat/>
7. Panto, Claude AI statistics 2026. <https://www.getpanto.ai/blog/?p=4129>
8. Claude Market Blog, Windsurf in 2026: the Devin Desktop rebrand, explained. <https://claudemarket.ai/blog/windsurf>
9. Kiro Blog, One year of Kiro: a look back, and a look ahead. <https://kiro.dev/blog/one-year/>
10. SiliconANGLE, AWS launches Kiro into general availability. <https://siliconangle.com/2025/11/17/aws-launches-kiro-general-availability-team-features-cli-support/>
11. Amplifying, Claude Code agent profile. <https://amplifying.ai/coding-agents/claude-code>
12. CloudZero, Windsurf pricing. <https://www.cloudzero.com/blog/windsurf-pricing>
13. The Next Web, SpaceX makes its \$60bn Cursor takeover official. <https://thenextweb.com/news/spacex-cursor-acquisition-official-60-billion>
14. tbreak, Grok 4.6 lands in Cursor as SpaceXAI closes the acquisition. <https://tbreak.com/grok-4-6-cursor-spacexai-acquisition/>
15. Cursor, Changelog. <https://cursor.com/changelog>
16. Google Antigravity Docs, Models. <https://antigravity.google/docs/models>
17. Technology Magazine, SpaceX targets AI dominance with US\$60bn Cursor acquisition. <https://technologymagazine.com/news/spacex-targets-ai-dominance-with-us-60bn-cursor-acquisition>
18. Cursor Docs, Models & Pricing. <https://cursor.com/docs/models-and-pricing>
19. Daniel Vaughan, Spring 2026 AI coding agent vulnerabilities. <https://codex.danielvaughan.com/2026/05/02/spring-2026-ai-coding-agent-vulnerabilities-cursor-rce-comment-control-codex-cli-defence/>
20. Danilchenko, Devin Desktop review. <https://www.danilchenko.dev/posts/devin-desktop-review/>
21. Claude Code Docs, Overview. <https://code.claude.com/docs/en/overview>
22. Pasquale Pillitteri, GitHub Copilot under pressure (survey round-up). <https://pasqualepillitteri.it/en/news/3392/github-copilot-cursor-claude-code-ai-coding-showdown-2026>
23. Google Antigravity Docs, Antigravity 2.0 overview. <https://antigravity.google/docs/overview>
24. Kiro Docs, Overview. <https://kiro.dev/docs/>
25. Claude (Anthropic), Plans & pricing. <https://claude.com/pricing>
26. Help Net Security, Prompt injection still drives most agentic AI security failures. <https://helpnetsecurity.com/2026/06/11/owasp-prompt-injection-ai-security-failures>
27. CloudZero, Google Antigravity pricing. <https://www.cloudzero.com/blog/google-antigravity-pricing/>
28. Google Antigravity Docs, CLI overview. <https://antigravity.google/docs/cli/overview>
29. Google Antigravity, Pricing. <https://antigravity.google/pricing>
30. Google Antigravity Docs, Plans. <https://antigravity.google/docs/plans>
31. ThreatCluster, CVE-2026-22708 intelligence summary. <https://threatcluster.io/entities/cve/CVE-2026-22708>
32. Kiro Docs, Web. <https://kiro.dev/docs/web/>

33. Kiro Docs, Crew. <https://kiro.dev/docs/crew/>
34. Kiro, Changelog. <https://kiro.dev/changelog/>
35. Kiro Docs, Models. <https://kiro.dev/docs/models/>
36. Kiro (AWS), Pricing. <https://kiro.dev/pricing/>
37. Kiro, Enterprise. <https://kiro.dev/enterprise/>
38. Amplifying, Kiro agent profile. <https://s.amplifying.ai/coding-agents/kiro>
39. AWS Security Bulletin 2026-001-AWS (CVE-2026-0830). <https://aws.amazon.com/security/security-bulletins/2026-001-AWS/>
40. The Hacker News, Amazon Kiro prompt injection finding. <https://thehackernews.com/2026/08/amazon-kiro-prompt-injection-can.html>
41. Real Python, Aider reference. <https://realpython.com/ref/ai-coding-tools/aider/>
42. Aider, Release history. <https://aider.chat/HISTORY.html>
43. Ry Walker Research, Aider. <https://rywalker.com/research/aider>
44. YesPress, Aider product profile. <https://yespress.io/products/aider.md>
45. Cursor, Pricing. <https://cursor.com/pricing>
46. GeekWire, Amazon pushes back on Financial Times report blaming AI coding tools for AWS outages.
<https://www.geekwire.com/2026/amazon-pushes-back-on-financial-times-report-blaming-ai-coding-tools-for-aws-outages/>