Google DeepMind AlphaEvolve

IMPLEMENTED 2024–2026 · TECHNOLOGY AND SOFTWARE · GLOBAL · BAND C · PARTLY PROVED
VERSION 1.0 · PUBLISHED 15 NOVEMBER 2026 · UPDATED 15 NOVEMBER 2026 · FACTS CHECKED 15 NOVEMBER 2026

A coding agent that pairs Gemini models with automated checks to evolve algorithms. Google reports that one of its results, a scheduling rule for its Borg data centre system, has been in production fleet-wide, and a matrix multiplication result has been checked and extended by outside mathematicians.

ABOUT THIS SYNOPSIS

This is MarLuc's summary and grading of publicly available sources, for information only. It is not investment, legal, financial or professional advice, and should not be relied on for any decision. Readers must validate it against the original sources and, where appropriate, challenge its conclusions.

Facts and figures are as reported by the sources listed below, which may be incomplete, self-reported or out of date. Check the original source before relying on any figure.

The impact band, proof marker and SEAM view are MarLuc's opinion. The organisation is not a MarLuc client and has not reviewed or endorsed this synopsis. Names and trademarks belong to their owners.

NOTICE WORDING LAST UPDATED 16 NOVEMBER 2026

AT A GLANCE

Frontier case: AI used for a scientific, medical or engineering breakthrough. The transformation is potential and unproven; MarLuc grades what has changed in practice so far.

AI used: Gemini-powered evolutionary coding agent that proposes and tests changes to code against automated evaluators.

Implemented: 2024–2026. Grade: band C, Process change, Partly proved.

Headline result, as reported: By Google's account, an AlphaEvolve scheduling heuristic has run across Google's fleet and recovers on average 0.7% of worldwide compute; its 48-multiplication matrix algorithm has been independently built on

1 · The case

The facts, as reported by the 5 sources listed at the end. Where a source is silent, the field says "Not reported".

Organisation

Google DeepMind and Google. Google DeepMind is Google's AI research laboratory. The white paper's authors are DeepMind researchers, and the operational results come from Google's own infrastructure and Gemini training. Canvas block: 4. Operations and process. Also 6. Data and technology.

Business challenge

The challenge. Per the white paper, discovering new algorithms normally needs a long cycle of ideas, testing and backtracking. For data centre scheduling, inefficient assignments leave machines with stranded resources: a machine can run out of memory while CPU is still free. Kernel tuning for Gemini training relied on search-based autotuning or manually written heuristics, which the paper calls a major engineering bottleneck.

How it was addressed. AlphaEvolve pairs Gemini models with automated evaluators and an evolutionary framework. A person supplies a seed program and a scoring function; Gemini models propose changes to the code, the evaluator scores them, and the best versions seed the next round. At Google it was applied to Borg scheduling, a Gemini kernel tiling heuristic, a TPU arithmetic circuit and matrix multiplication algorithms. It was also tested on over 50 open mathematics problems. [1] [2] [3] [4] [5]

Who implemented it

In-house (In-house): Google DeepMind: Built AlphaEvolve and applied it to Google's data centre scheduling, Gemini kernel tuning and TPU circuit work..
Google (Gemini models) (Model or cloud provider): Gemini Flash and Pro models generate the candidate code changes..
Operational results are from inside Google; the 2026 general availability announcement lists outside customers but this case grades Google's own use.

Technology (SEAM pillar E)

SEAM types: T7 AI agents, T3 API integrations. Deployment: Not reported. Vendors and products: Google DeepMind, AlphaEvolve, Gemini 2.0 Flash and Gemini 2.0 Pro.

Where it was built in: Run as an evolutionary loop around Gemini models with automated evaluators; outputs are code (a Borg scheduling heuristic, a Pallas kernel tiling heuristic, Verilog for a TPU circuit) deployed into Google's own systems.

What changed

Before: Borg ranked machines with an existing production heuristic; kernel tiling used expert-designed heuristics. After, by Google's account: AlphaEvolve's scheduling heuristic was rolled out fleet-wide and the kernel heuristic was deployed in production. The paper says kernel optimisation time fell from several months of engineering effort to days of automated experimentation. Google released the tool to Google Cloud customers in private preview in 2025 and general availability in July 2026. [1] [2] [3] [4] [5]

What it cost

Not reported

What it saved

Scheduling: recovers on average 0.7% of Google's fleet-wide compute, which would otherwise be stranded, per post-deployment measurements reported by Google. Gemini training: a 23% average kernel speedup and a 1% reduction in overall training time. Money saved: Not reported. [1] [2] [3] [4] [5]

Governance and control

The scheduling heuristic only changes the ranking of machines that Borg has already judged able to run a job, so the paper says decisions are correct by construction, and it was tested on unseen workloads before rollout. A TPU circuit change was validated by TPU designers. Accountable owner and monitoring after rollout beyond the fleet measurements: Not reported.

S2 · Accountable owner · Not reported
Accountable owner: Not reported for the Google deployments.

A1 · Safe use · Claimed
The scheduling heuristic only reranks machines Borg has already found suitable, so the paper says decisions are correct by construction. [1]

A2 · Human in the loop · Claimed
TPU designers validated a circuit change for correctness, per the white paper; the Cloud GA post quotes a customer saying engineers still own review and release. [1] [3]

A3 · Monitoring and assurance · Claimed
The scheduling heuristic was tested on unseen workloads and post-deployment measurements confirmed simulator results, by Google's account. [1]

A4 · Vendor AI · Not reported
Not reported.

A5 · Agent controls · Claimed
The agent proposes code only; the white paper says automated evaluation lets it avoid incorrect suggestions from the base model. [1]

A6 · Standards and regulation · Not reported
Not reported.

What worked, what did not, what is unproven

WORKED

– A scheduling heuristic found by AlphaEvolve beat the one in production on an unseen test set and was rolled out to the whole fleet, where post-deployment measurements confirmed the simulator results (white paper).

– A kernel tiling heuristic gave an average 23% kernel speedup and a 1% shorter Gemini training time, and has been deployed in production (white paper).

– Dumas, Pernet and Sedoglavic, who are outside Google, published a variant of the 48-multiplication algorithm that uses only rational coefficients (arXiv).

DID NOT

– The white paper says a TPU circuit improvement was also independently caught by downstream synthesis tools.

– Of 54 matrix multiplication targets, AlphaEvolve matched the state of the art in 38 cases, surpassed it in 14 and fell behind in 2 (white paper).

– Tao and co-authors report AlphaEvolve was not able to match or exceed previous results in all cases, and that some improvements could likely have been matched by more traditional methods.

UNPROVEN

– The compute figures are Google's own measurements on its own fleet; no independent audit is reported.

– How far results carry beyond problems that have a fast automated evaluator; the white paper

– Whether outside customers' reported gains hold in routine use; the named customer results in Google's July 2026 blog are customers' own statements published by Google.

Timeline

May 2025 Google researchers say the AlphaEvolve scheduling heuristic for Borg has been in production for over a year (per The Register); the white paper says it was rolled out to the entire fleet [2]

May 2025 Google DeepMind announces AlphaEvolve, including the 48-multiplication result for 4 × 4 complex-valued matrices [1]

Jun 2025 White paper posted to arXiv; the same month Dumas, Pernet and Sedoglavic post a version of the 48-multiplication algorithm using only rational coefficients [4]

Nov 2025 Georgiev, Gomez-Serrano, Tao and Wagner publish results on 67 mathematical problems, including cases where AlphaEvolve did not match earlier results [5]

2025 AlphaEvolve introduced on Google Cloud in private preview (per Google's July 2026 post, 'last year') [3]

Jul 2026 AlphaEvolve becomes generally available on Google Cloud, with named customers reporting results [3]

2 · Grade and SEAM assessment

MarLuc's opinion, made on the evidence above. How cases are graded: marluc.co.uk/case-studies/how-we-rate

Level of change and impact

Claimed: A. Grade: C, Process change. Proof: Partly proved. Route: Model-led.

Process change (Level 2, band C), claimed as transformation (band A). Two Google processes now use code AlphaEvolve found, with measured results reported by Google. Any wider transformation of engineering or discovery work is potential and unproven.

Graded C: two specific processes, data centre job scheduling and a Gemini kernel tuning step, now use code that AlphaEvolve found, and Google reports measured results after rollout. Claimed A: Google and its customers describe AlphaEvolve as able to transform engineering and discovery work. That transformation is potential and unproven. The operational results are small percentages of Google's own compute and training time, and the main limit is that a problem needs an automated evaluator.

The SEAM view and lessons learnt

The method has a clear pattern: a person defines the problem and the score, the agent searches, and the result is ordinary code that engineers can read and test. That is why Google chose it over a deep reinforcement learning approach for scheduling, citing interpretability and ease of deployment. The same pattern fits any process with a trustworthy automated score. The 0.7% figure is large at Google's scale but small as a share, and it has not been checked outside Google. The matrix result is the best checked part, because it is a mathematical object that others have verified and built on.

2. Test candidate changes on unseen data and keep humans validating production changes.
3. Check headline gains against independent results and the cases where the method fell short.

Evidence

Score 2 of 4. Mixed: the operational figures are Google's own (white paper and Google Cloud blog). The matrix multiplication result is a checkable mathematical result that outside researchers have built on. No independent audit of the compute savings.

Evaluation method. Company white paper (arXiv preprint, not peer reviewed in the source opened) and a company blog, with an independent follow-up preprint on the matrix algorithm and an outside mathematicians' paper on the mathematics tests.

Sources disagree. The Dumas paper says the 48-multiplication result is over the complex numbers and offers a rational-coefficient variant, so the 'first improvement after 56 years' claim applies to the complex-valued case.

Status. Ongoing: generally available on Google Cloud from July 2026. Google researchers said in May 2025 that the Borg heuristic had been in production for over a year.

3 · Original sources

Read these in full. To challenge a fact or a grade, use the contact form at marluc.co.uk or the link on https://marluc.co.uk/case-studies/google-deepmind-alphaevolve

[1] Google DeepMind: AlphaEvolve: A coding agent for scientific and algorithmic discovery (white paper, arXiv)
https://arxiv.org/pdf/2506.13131

[2] The Register: Google DeepMind promises to help you evolve your algos (15 May 2025)
https://theregister.com/2025/05/15/google_deepmind_debuts_algorithm_evolving

[3] Google Cloud Blog: Solve harder problems with AlphaEvolve, now available to everyone on Google Cloud (10 July 2026)
https://cloud.google.com/blog/products/ai-machine-learning/alphaevolve-is-available-for-everyone

[4] arXiv: Dumas, Pernet and Sedoglavic, A non-commutative algorithm for multiplying 4x4 matrices using 48 non-complex multiplications
https://arxiv.org/abs/2506.13242

[5] arXiv: Georgiev, Gomez-Serrano, Tao and Wagner, Mathematical exploration and discovery at scale
https://arxiv.org/pdf/2511.02864

Google DeepMind and Google website: https://deepmind.google

Version 1.0 · Published 15 November 2026 · Updated 15 November 2026 · Facts checked against sources 15 November 2026 · Downloaded 26 November 2026. This PDF may be out of date: the latest version, and a way to report an error or a better source, is at https://marluc.co.uk/case-studies/google-deepmind-alphaevolve

VERSION HISTORY

v1.0 15 November 2026 First published