MARLUC

CASE STUDY · MARLUC SYNOPSIS

GOV.UK Chat

IMPLEMENTED 2023–2026 · GOVERNMENT AND PUBLIC SERVICES · UK · BAND D · CLAIMED ONLY

VERSION 1.0 · PUBLISHED 15 NOVEMBER 2026 · UPDATED 15 NOVEMBER 2026 · FACTS CHECKED 15 NOVEMBER 2026

GOV.UK Chat is an AI assistant that answers questions using guidance published on GOV.UK. After two public pilots, GDS reports 90% answer accuracy on its own assessment, and launched it in the GOV.UK app in May 2026.

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Facts and figures are as reported by the sources listed below, which may be incomplete, self-reported or out of date.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

AI used: Generative AI assistant answering questions from published GOV.UK guidance, using Anthropic's Claude models on Amazon Bedrock.

Implemented: 2023–2026. Grade: band D, Task support, Claimed only.

Headline result, as reported: GDS reports answer accuracy up from 76% to 90% across two public pilots, by its own assessment; launched in the GOV.UK app on 14 May 2026

1 · The case

The facts, as reported by the 4 sources listed at the end. Where a source is silent, the field says "Not reported".

Organisation

Government Digital Service (GDS), Department for Science, Innovation and Technology, UK Government. UK central government. GDS is the digital centre of government and runs GOV.UK. The GOV.UK app had 563,000 registered users in May 2026 (Civil Service World). Canvas block: 3. Customers and proposition. Also 6. Data and technology, 8. Governance and risk.

Business challenge

The challenge. The government says its guidance runs to about 80,000 pages across the GOV.UK site. DSIT told Civil Service World that some government call centres take around 100,000 calls a day and that up to half relate to queries an automated tool could handle. GDS began developing a chatbot in 2023.

Who implemented it

In-house (In-house): Built and runs GOV.UK Chat, with its data scientists and researchers.
Anthropic (Model or cloud provider): Supplies the Claude models used by the latest version.
Amazon Web Services (Model or cloud provider): Hosts the models through the Bedrock platform.
Civil Service World reports the first version used the OpenAI technology behind ChatGPT; the vendor is not named in the GDS posts read.

Technology (SEAM pillar E)

SEAM types: T3 API integrations. Deployment: Cloud APIs. Vendors and products: Anthropic, Amazon Web Services, Claude models (Anthropic), Amazon Bedrock.

Where it was built in: Inside the GOV.UK app; answers draw exclusively on guidance published on GOV.UK and link back to it. Earlier versions used the OpenAI technology behind ChatGPT (per Civil Service World).

What changed

Before: people searched or navigated GOV.UK pages. After: app users can ask a question in plain language and get a conversational answer with links to the source guidance. The tool does not attempt to provide advice and cannot use a person's own records. It is available only in the app; GDS says it plans to explore the web in future. [1] [2] [3] [4]

What it cost

Cost of development and running: Not reported.

What it saved

Time and call-centre savings: Not reported as measured. In pilot research participants said the answers helped them decide what to do next and, for some, saved a phone call to a department. DSIT's estimate that up to half of some call-centre queries could be handled is a hope, not a result. [1] [2] [3] [4]

Governance and control

GDS says answers are assessed by subject matter experts and automated evaluation tools, the tool answers only from published guidance, users are warned that it can make mistakes, and guardrails blocked all 508 jailbreak attempts in the pilots. GDS says it worked with security experts including the AI Security Institute before release. GDS says personal information is filtered out.

S2 · Accountable owner · Claimed
GDS leads the work; the launch post is written by its Deputy Director for GOV.UK AI. No named accountable owner for outcomes is reported. [2]

A1 · Safe use · Claimed
By GDS's account the tool signposts when to check the original guidance, does not attempt to provide advice, warns users that it can make mistakes and filters personal information. [1] [2]

A2 · Human in the loop · Not reported
No human review of individual answers before they reach users is described; humans assess samples, and handover to departmental support is a future plan. [1]

A3 · Monitoring and assurance · Claimed
GDS says answers are assessed by subject matter experts and automated tools, and that the tool is monitored and evaluated for accuracy and safety. [1] [2]

A4 · Vendor AI · Claimed
GDS says it tested a range of models including open source ones and chose on accuracy and speed; it built the system so it can change models, and says model choice is kept under review. [1]

A5 · Agent controls · Not reported
The launched tool answers questions only. GDS says it is experimenting separately with agentic AI. [1]

A6 · Standards and regulation · Claimed
GDS says it used the Scan, Pilot, Scale approach in the AI Opportunities Plan and worked with security experts, including the AI Security Institute, on safety checks before release. [2]

What worked, what did not, what is unproven

WORKED

– GDS reports accuracy rising from 76% to 90% across all topics, using a strict test: an answer counts as accurate only if it meets all the standards of the published content.

– 73% of surveyed GOV.UK app users found it useful and 64% were satisfied.

– GDS reports the system coped comfortably with demand and that guardrails prevented all 508 jailbreak attempts.

– After the soft launch, GDS says performance matched the pilot.

DID NOT

– The 2023 private pilot did not meet GDS's required accuracy and in a few cases produced outright mistakes (per The Register).

– At 90% accuracy, about one answer in ten does not meet GDS's standard, often because only part of the question is answered, by GDS's account.

– Answers take 10.7 seconds on average, slower than GDS wants; satisfaction rose in tests when faster speeds were simulated.

– Some users wanted an adviser with access to their personal information, which the tool does not support.

– Web launch: GDS says it plans to explore GOV.UK Chat on the web in future; in March 2026 The Register reported a plan to implement it across the website later in 2026.

UNPROVEN

– Accuracy: the figures are GDS's own assessment against its own standard; an independent measurement: Not reported.

– Savings to call centres or departments: Not reported.

– How accuracy holds across the full app audience and over time: GDS says performance after the soft launch matched the pilot; a published figure: Not reported.

– Results of the AI Security Institute checks: Not reported.

2023 A private pilot did not meet GDS's required accuracy levels and in a few cases produced mistakes [3]

2024 First public pilot on a web version of GOV.UK Chat (late 2024, per The Register): 10,136 users asked 23,838 questions [3]

2025 Second pilot in the GOV.UK app (autumn 2025, per The Register): 641 people asked 2,670 questions in four weeks [3]

Mar 2026 GDS publishes pilot findings: 90% accuracy, 73% useful, 64% satisfied; soft launch in the GOV.UK app on 26 March [2]

May 2026 GOV.UK Chat officially launched in the GOV.UK app on 14 May [2]

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: Not applicable. Grade: D, Task support. Proof: Claimed only. Level of change: Task support.

Task support (Level 1, band D). The tool adds an optional route to existing guidance. No change to departmental processes or staffing is reported.

GOV.UK Chat adds an optional conversational route to government guidance inside the GOV.UK app. GDS reports good accuracy and user satisfaction. The sources do not report any change to departmental processes, staffing or call-centre volumes, so this is graded as task support at the citizen level, with the results claimed by the developer rather than checked by anyone else.

The SEAM view and lessons learnt

GDS limited the assistant to published guidance, tested over 18 months with two public pilots and held back the launch to the app first. Its accuracy test is strict and it reports where the tool falls short, including speed. The gap is evidence outside GDS: accuracy, savings and effects on call centres are reported by the builder or by DSIT as hopes. For a larger organisation, the point is to design the accuracy test first, publish it and have it repeated by someone else.

1. Restrict an assistant to approved content and link to the source.
2. Run public pilots before launch and publish the accuracy test.
3. Have someone outside the team repeat the accuracy test.

Evidence

Score 1 of 4. Own account: the delivery team's published findings and launch post. Independent press reports GDS's figures; no independent test of accuracy found.

Evaluation method. GDS pilot research: web pilot with 10,136 users (99 usability tests, a diary study, 576 surveys, manual scoring of 419 conversations and 1,084 question-answer pairs) and an app pilot with 641 users (1,338 in-scope answers reviewed); accuracy rated by subject matter

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/gov-uk-chat-government-digital-service

[1] Inside GOV.UK (GDS): 5 things we learned testing GOV.UK Chat: an AI assistant for government
https://insidegovuk.blog.gov.uk/2026/03/16/5-things-we-learned-testing-gov-uk-chat-an-ai-assistant-for-government

[2] GDS blog: Answers in seconds, 24/7: GOV.UK Chat launches in the GOV.UK app
https://gds.blog.gov.uk/2026/05/14/gov-uk-chat-launches/

[3] The Register: GOV.UK chatbot gets smarter but slower as LLMs improve
https://www.theregister.com/on-prem/2026/03/19/govuk-chatbot-gets-smarter-but-slower-as-llms-improve/5229770

[4] Civil Service World: AI chatbot fully integrated into GOV.UK app
https://www.civilserviceworld.com/professions/article/govuk-ai-chatbot-launched

Government Digital Service (GDS), Department for Science, Innovation and Technology, UK Government website: https://www.gov.uk

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/gov-uk-chat-government-digital-service

VERSION HISTORY

v1.0 15 November 2026 First published