Optivalue.ai

WHITE PAPER · SEPTEMBER 2026

The model that answered no longer exists

Sovereignty over time: retirement dates, silent drift, and what you will still be able to prove.

By Olivier Guillo, founder of Optivalue.ai

Classified as External

A second sovereignty, less discussed

THE ONE BEING DEBATED

It is geographic: where my data sits, under which law, in which country.

THE ONE THIS DOCUMENT ADDRESSES

It is temporal: how long will the model that produced my answer keep existing, and will it still be the same one?

Who it is for

Executives, information systems, compliance, risk and audit. No technical knowledge assumed.

What it is based on

The lifecycle pages published by three major providers, listed in the appendix and open to anyone.

What it does not do

Comment on these providers one by one. What matters is not what any one of them does, it is what your own provider's page says.

Executive summary

01 Retiring a model is not an incident: it is a documented procedure, on a schedule the provider sets and the customer does not negotiate.

02 A model does not even need to be retired for your results to change: the same service, under the same name, has already answered differently from one quarter to the next.

03 Your obligations to justify are counted in years. The service life of a hosted model is counted in months.

04 One provider commits to preserving the weights of its released models. Preserving is not serving: that commitment does not give you the model back.

05 This risk is not handled through technology: it is handled in the contract, through seven verifiable requirements.

PART 1

The schedule belongs to the provider

Lifecycle, notice periods, automatic upgrades: what the public pages say.

A lifecycle in four stages

Active

The model is generally available, open to all customers.

Deprecated

It still works but is no longer recommended, and a retirement date is assigned to it. The announcement takes effect immediately.

Notice

Sixty days to six months for a generally available model. Thirty days for a preview model; one case cited at two weeks.

Retired

Requests fail: every inference call receives the 410 Gone error.

The same pages reserve for the provider an emergency retirement on shortened notice when a model raises a compliance or safety issue.

Three findings for a risk committee

01

The retirement date may exist before you start

On one of these pages it is set at launch, at eighteen months, with no separate announcement, and drops to twelve months for models distributed on behalf of third-party publishers. You do not learn the date: it was already there.

02

The announced notice period is a minimum, and it is what you will get

Published retirement histories show intervals that stay close to the promised minimum, sometimes within a day or two. Plan on that minimum.

03

The same model can have different dates depending on the access route

Platforms operated by partners set their own retirement schedules. Your deadline therefore also depends on your reseller.

Automatic upgrades, and the choice they leave you

REGIME 1

Automatic replacement

The retired model is upgraded to its successor, region by region, even where that successor was not yet available.

REGIME 2

Refusal, and shutdown

A setting lets you never accept an upgrade. The consequence appears in the same sentence: the deployment stops working on the retirement date.

REGIME 3

Reserved capacity

Reserved-capacity deployments are not upgraded automatically, and the retirement date still applies to them.

The choice is therefore not between changing model and keeping yours, but between accepting the replacement and stopping.

And asked whether a date can be extended by exception, that page answers that retirement dates are not extendable.

PART 2

Drift, and the burden of proof

What changes without warning, what your obligations require, and what preserving is not enough to give back.

The name does not change, the answer does

Two versions of the same commercial service, three months apart, on an identical task: deciding whether a number is prime.

84%

correct answers, March 2023 version

51%

correct answers, June 2023 version, same name

Two lessons

Drift is real, and it does not run one way: on the same task, the other model in the family had improved.

Pinning a dated version limits drift without removing it: you are then exposed to the retirement of that version. The two risks do not add up, they take turns.

Source: Chen, Zaharia, Zou, arXiv 2307.09009

What drift invalidates

The prompts written to obtain an answer of a certain shape.

The thresholds tuned to decide when a human must step in.

The question set on which your experts validated the tool.

The answers already sent, which you can no longer show were the best the system could produce.

A model name does not date a performance.

If your documentation states that your system reaches a certain level of quality with a given model, without saying which version, on what date and on which test set, it proves nothing.

The clash of two calendars

THE SERVICE LIFE OF A HOSTED MODEL



12 to 18 months

YOUR OBLIGATIONS TO JUSTIFY



3 years: ISO/IEC 27001 certification cycle, with annual surveillance audits



6 years: FINRA Rule 4511 (United States); public company in the United Kingdom



7 years: audit working papers, section 802 of the Sarbanes-Oxley Act

No text we found requires replaying an AI output. The requirement comes from the burden of proof: “we produced it with a model that no longer exists” is not a breach, it is a position you cannot defend.

Preserving is not serving

WHAT ONE PROVIDER PUBLISHED IN NOVEMBER 2025

The commitment to preserve, for future deployments, the weights of all its released models, for at least the lifetime of the company.

Its deprecation page acknowledges three drawbacks of retirement: users must migrate; researchers lose access for comparative studies; and retirement raises risks related to model safety and model welfare.

A provider therefore states itself that retirement removes the ability to compare over time. Transposed to the enterprise, that is our subject.

Why that commitment does not give you the model back

A provider that archives a model inside its own systems does not give you any way to query it.

The same document limits its scope: any future availability of retired models will depend on reducing the cost and complexity of doing so.

What is committed is preservation. Returning the model to service is not: it remains the provider's decision, not the customer's right.

What gives you the model back: holding it

The smaller a model is, the more the infrastructure that runs it is within reach of a company, and the more holding it becomes practical rather than theoretical. That is why small models hold this place, not because they would answer better.

Check 1 · The licence

They are not equivalent. Some are recognised open-source licences; others set conditions, sometimes a user threshold. A licence is read, it is not summarised.

Check 2 · Provenance

Open-weights models come from different jurisdictions. In a sensitive sector, the origin of the base model can be grounds for refusal in itself. The question is asked before, not after.

Which base model does your system rely on?

Do we hold that model, or only access to it?

If you disappear, can we keep running it?

The flip side: a frozen model ages

What freezing freezes, and what it does not

Knowledge of your procedures and contracts does not live in the model: it lives in your document base, which you update without touching the model.

A frozen model querying up-to-date documents answers on up-to-date rules. What freezing fixes is the way of reasoning and writing, not the content.

The burden of keeping it current does shift hands, though: when you hold the model, nobody improves it for you.

What produces today may evolve

The version producing today's answers benefits from the best capabilities available for the work at hand.

What produced yesterday is preserved

The version that produced a signed answer is frozen and replayable for as long as that answer binds you.

One point of honesty: replay does not come from recomputation. What proves is the record: what was asked, what was answered, on which sources, and who approved. The preserved model serves to rerun the demonstration if it is challenged; it does not replace the record, and the record does not replace it.

PART 3

The clauses, and the question to ask on Monday

Seven requirements, each followed by the question that verifies it.

Seven requirements to write into the contract

REQUIREMENT	WHAT THE PROVIDER COMMITS TO DO	THE QUESTION THAT VERIFIES IT
Notice period	A stated duration between the deprecation announcement and retirement.	What is that duration, and in which clause does it appear?
Service duration	A version keeps being served for a set time after its replacement is announced.	What is that duration?
Refusing replacement	No version replacement without prior written agreement.	What contractual consequence applies to a unilateral replacement?
The right to replay	The customer may keep and run a specified version to justify a case file.	For how long, and on what technical and financial terms?
Notification	Any change liable to alter results, even without a change of name.	Within what time, and through which channel?
Exit	On exit, the customer recovers a defined set of items.	Which ones, in what format, within what time?
The schedule that governs	The contract names the retirement schedule that applies through a third-party platform.	The model developer's, or the platform's?

THE QUESTION TO ASK ON MONDAY

Which version of the model produced the answer we signed last year, and could we replay it today?

Optivalue.ai

A European software company, born out of compliance and engineering. Under the Smart Global Governance brand, our teams built and ran a governance, risk and compliance platform for large groups.

We now publish an AI specialised in the chain of evidence: the answers that commit a company before a customer, an auditor or a regulator. It is a line of work where you answer today to be questioned long afterwards.

On a real case file

If you want to know what you could produce today to justify an answer signed last year, write to us: we work through it with you, on one of your own case files.

Our site small-language-model.ai covers sovereignty in space; this document covers sovereignty in time.

www.optivalue.ai

Optivalue.ai

Legal notice

Copyright

Copyright © 2026

All rights reserved Optivalue.ai

RCS Grasse B 853 951 556

SAS with share capital of EUR 3,350,535.32

No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronic, mechanical, photocopying or otherwise, without the prior written permission of Optivalue.ai.

Disclaimer

This document is provided “AS IS”, for information purposes only.

The contractual commitments of Optivalue.ai are contained solely in the professional services agreement executed between Optivalue.ai and the Client.

Confidentiality

This document is the confidential and proprietary property of Optivalue.ai.

This document may contain approaches, techniques and other information proprietary to Optivalue.ai.

This document must not be disclosed in whole or in part to third parties without the prior written consent of Optivalue.ai.