



Where the Work Runs

A plain architecture paper for businesses deciding where AI should run

Justice Apex Inc., Collierville, Tennessee

Version 1.0

October 18, 2026

Contents

SECTION ONE. WHAT THIS PAPER COVERS AND WHAT IT LEAVES OUT	3
SECTION TWO. THE SHAPE OF AN ENGAGEMENT	3
SECTION THREE. TESTING BEFORE WE PROMISE	4
SECTION FOUR. MODELS	5
SECTION FIVE. DATA AND ACCESS	5
SECTION SIX. THE LOG AND THE OFF SWITCH	6
SECTION SEVEN. HOW AN ENGAGEMENT STARTS	7
SECTION EIGHT. WHAT WE DON'T CLAIM	7
APPENDIX. A CHECKLIST FOR YOUR IT PERSON	7
NEXT STEP	9

SECTION ONE.

WHAT THIS PAPER COVERS AND WHAT IT LEAVES OUT

In this section you'll learn what the paper covers and what it leaves out. This paper explains where each part of our work for a business runs and what crosses between the parts. It also explains how we test a private model on your own hardware before we promise anything about it.

It's a plain description, not a certification, an audit report or a guarantee.

It describes what we can state today, and we have left out what we can't yet back with a test or a record. It doesn't cover prices, the details of any one engagement, or the network of any customer. It doesn't describe how your records are kept apart from other businesses, how keys and passwords are kept, what happens to your data if you leave, or whether any model is trained on your records.

You may have been sold something before that covered less than the brochure said. That's why the limits come first. The appendix has a question on each of the four things we don't describe, so you can put them to any vendor, us included.

SECTION TWO.

THE SHAPE OF AN ENGAGEMENT

This section shows where each part of the work runs and what crosses between the parts. Knowing which part sits where is the first thing to settle with any vendor.

We build and run software for a business. Today that is FORGE, our ERP and CRM system, and the receptionist agent.

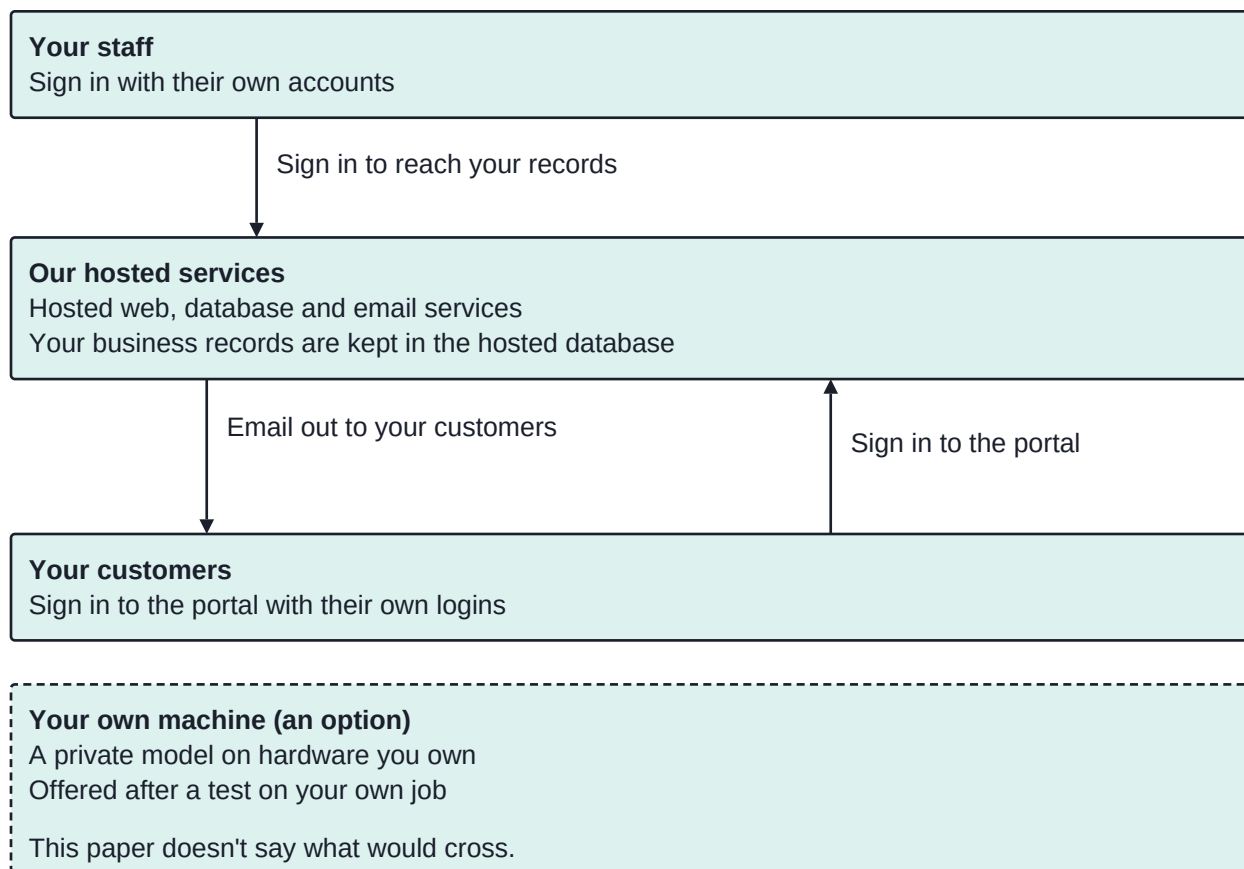
Start with your staff. Your staff sign in with their own accounts.

Next, your customers. Your customers sign in to the portal with their own logins, kept separate from your staff accounts. If the portal is on for your business, each of your customers sees their own price list when they sign in.

Then our hosted services. Hosted means the software runs on computers that a provider looks after, and you reach it over the internet. The hosted version of FORGE runs on hosted web, database and email services. It doesn't run on a computer in your office. In FORGE, your business records are kept in a hosted database, and you reach them by signing in. Email crosses a boundary, out to your customers.

Last, the option of a private model. A private model is an AI model that runs on hardware you own. A private model on your own hardware is an option you can ask for, and this paper doesn't describe it as an installed product. This paper doesn't say what would cross between a private model on your own hardware and our hosted services. Section Four says how we test one.

Figure 1. The shape of an engagement



Solid outline: described in this paper. Dashed outline: an option, promised only after a test.

Your staff and your customers sign in to our hosted services, and email goes out from them to your customers.

SECTION THREE.

TESTING BEFORE WE PROMISE

This short section says when we'll make a promise about how a machine will perform. A list of hardware parts doesn't tell you whether a machine can do your work. We test on your own job before we promise anything about how a machine will perform. The appendix has a question about this too.

SECTION FOUR.

MODELS

Here's what didn't work when we tried a small model, and the design we're building toward.

An open-weight model is one whose files are published, so you can run it on your own machine. On September 15, 2026, we ran a small open-weight model as a complete assistant through a large AI agent program, and that was not a working shape that day. An AI agent program is software that lets a model take steps on its own, such as reading a file or sending a message.

A context window is how much text a model can read at one time. At the model server's default context window, the program's opening instructions were longer than the window and were cut off before the model saw them. The lesson is to set the window on purpose and not to leave it at the default. After we raised the window, the opening instructions only just fit, and larger ones would not.

The shape we're building toward is a design, not a result.

In that design, the model is given a short set of instructions, the hard rules are written as ordinary code, and the data is held in a lookup. A lookup is a table the program checks for the answer, much like a price list. In that design, the model decides and writes, and ordinary code does the counting and pricing.

A local model on one machine is a different thing from a hosted model, and the two must be tested on the same job before either is chosen. If you ask for a private model on your own hardware, we'll run it through a measured test before we promise anything, and we'll only promise what it passes.

SECTION FIVE.

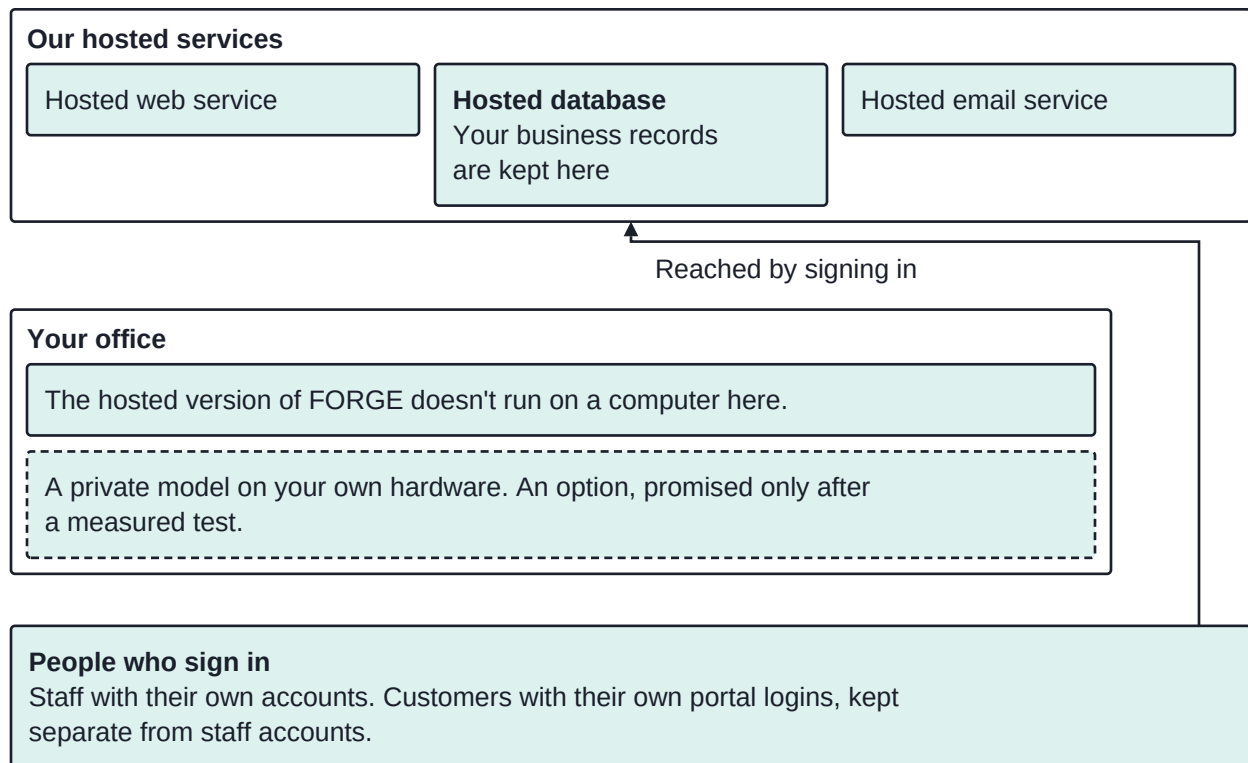
DATA AND ACCESS

This section covers where your records are kept and how people get to them. Who can sign in is one of the first things an IT lead asks about any system.

In FORGE, your business records are kept in a hosted database, and you reach them by signing in. Your staff sign in with their own accounts. Your customers sign in to the portal with their own logins, kept separate from your staff accounts. If the portal is on for your business, each of your customers sees their own price list when they sign in.

Businesses don't all need the same features. We set which FORGE features are switched on for your business, such as the customer portal or invoicing.

Figure 2. Where your records are kept



In the hosted version of FORGE, your business records are kept in a hosted database and you reach them by signing in.

SECTION SIX.

THE LOG AND THE OFF SWITCH

This section says what the paper does and doesn't say about a log and an off switch. A log is a written record of what a system did. An off switch is a way to stop it. Both matter to anyone who lets software act for their business.

This paper doesn't describe a log you can read or an off switch you hold as part of any product.

If a vendor tells you their product has them, ask to see them before you count on them.

SECTION SEVEN.

HOW AN ENGAGEMENT STARTS

This section tells you what the first step looks like. A first engagement starts with a short call. We agree the job and price it before we start. If you ask for a private model on your own hardware, a test on your own job comes before any promise about how it will perform. Bring your IT person to the call if you can. A list of questions to hand to your IT person is at the end of this paper.

SECTION EIGHT.

WHAT WE DON'T CLAIM

This section lists what we don't claim. A vendor that names what isn't covered is easier to trust.

This paper makes no claim that any product of ours meets a compliance standard.

This paper makes no promise about how fast any of it runs, how often it's right or how long it stays up. This paper names no outside standard, audit or test. The sections in this paper are the whole of what we describe today, and we will add to them as each part is proven.

APPENDIX.

A CHECKLIST FOR YOUR IT PERSON

Here's a two-page list for your IT person. These questions suit any AI vendor, including us. Ask every vendor the same questions in the same words, and write down the answers.

1. Where are my records kept, and who runs the machines they sit on?

2. Which outside services handle my data, such as hosting, email and model providers?

3. Does any of my data leave my network to reach a model, and which data?

4. Can the model server be reached from outside my network, and who checked that, and when?

5. How are my records kept apart from your other customers, and what test shows it?

6. What does the system record about what the AI did, and can I read that record myself?

7. How do I switch it off, who else can, how long does it take, and what do callers or visitors get while it's off?

8. Who can sign in, how are roles set, and is there a second factor?

9. Where are keys and passwords kept, and does anything you hand me carry them?

10. Is any of my data used to train a model, yours or anyone's?

11. If I leave, can I take my data, in what form, and what do you delete, and when?

12. What certifications do you hold, and if you name one, who issued it and may I read the report?

13. What did you test on my own job, what was measured, and what was not tested?

14. What uptime do you promise, how is it measured, and what do I get if it's missed?

NEXT STEP

To start with a call, book a time at <https://justiceapex.ai/go/book>, or write to chad@justiceapex.ai.

