Ferrous Labs

THE AI PRODUCT OPPORTUNITY GUIDE

Could your expertise become an AI product?

Five signs your business may already have the foundations — and what to do if it has.

FERROUSLABS.CO.UK

LONDON AI ENGINEERING STUDIO

What's inside

The checklist

You may be closer to an AI product than you think

Some of the strongest AI product opportunities don't start with a new idea. They start with something a business already has: specialist expertise, valuable data, a proven process, recurring client demand, or a service that has become difficult to scale.

This checklist highlights five signals that could indicate there's an opportunity to turn some of that existing value into a product.

It is not a product-readiness test. Think of it as a way to identify where it is worth looking more closely. As you go through the five signs, ask: **does this sound like our business?**

Tick every box that is true. There are fifteen. Part 1 asks you to count them.

Sign 1: Your expertise is valuable, but difficult to scale

- [ ] Customers repeatedly pay for specialist knowledge or analysis.
- [ ] Demand often exceeds your capacity to deliver.
- [ ] Growth currently means more people.

Ask yourself. If demand doubled tomorrow, would this expertise require twice as many people?

Sign 2: You turn away work that's too small to service profitably

- [ ] You have a minimum engagement size.
- [ ] Smaller prospects want what you offer, but can't justify the full service.
- [ ] Serving the long tail manually would cost more than it's worth.

Ask yourself. How much demand are you currently saying no to because your service model can't profitably support it?

Sign 3: Clients ask for your framework, not just the outcome

- [ ] Clients ask for your methodology, templates or tools.
- [ ] Your frameworks have value independently of the final service.

- [ ] People want to apply your thinking themselves.

Ask yourself. What do clients repeatedly ask you to give them so they can do more without you?

Sign 4: Your market is beginning to productise around you

- [ ] New software products are appearing in your category.
- [ ] Competitors are introducing self-serve or AI-enabled services.
- [ ] Customers are beginning to expect faster, technology-enabled delivery.

Ask yourself. If someone built the product version of your service tomorrow, what would your domain expertise allow you to do better?

Sign 5: You have the data nobody else can create

- [ ] You hold years of specialist decisions or outcomes.
- [ ] Experts within your business have classified, diagnosed, approved or rejected thousands of cases.
- [ ] Your historical data contains judgement that couldn't simply be scraped from the internet.

Ask yourself. What does your organisation know, through its data and past decisions, that a competitor couldn't easily recreate?

So, what did you recognise?

0 to 1 signs — keep watching. There may not be a strong productisation case yet, and that's completely fine. Look for recurring demand, bottlenecks and proprietary knowledge before forcing an AI use case.

2 to 3 signs — worth exploring. There may be something here. The next step isn't building software; it's identifying the strongest problem and testing whether a product could solve it commercially.

4 to 5 signs — strong signals. You potentially have several of the ingredients required for productisation. The priority now is validating the proposition before committing significant time or investment.

The five signs are only the start

Other shifts can make an AI product opportunity significantly more compelling, or more urgent. Here are three more signals worth watching.

Signal 1: The cost of the technology just collapsed

The signal. Tasks involving vision, OCR, transcription or document analysis that were previously prohibitively expensive may now be commercially viable.

The opportunity. Combine cheaper underlying technology with your domain expertise.

The catch. Everyone has access to the same models. Your moat has to come from what you add.

Ask yourself. What has recently become technically or commercially possible in your industry that wasn't viable five years ago?

Signal 2: Regulation is creating a repeatable process

The signal. A new standard, reporting requirement or compliance regime can suddenly force an entire industry to perform the same structured task.

The opportunity. Instead of repeatedly delivering the compliance process manually, could you build the infrastructure that helps others complete it?

Ask yourself. When an industry has to follow the same new process, could part of that process become a product?

Signal 3: A funded startup just entered your niche

The signal. A VC-backed entrant is productising something your industry has traditionally delivered manually.

The opportunity. Their investment is uncomfortable validation that a market may exist. Your advantage may be the domain expertise, data and client relationships they've yet to build.

Ask yourself. If a software-first competitor entered your market tomorrow, what could they replicate and what couldn't they?

Before you build anything, ask

Problem. Is there a clear, recurring problem worth solving?

Demand. Are customers already demonstrating demand for a better way?

Advantage. Do you have expertise, data or IP that's difficult to replicate?

Economics. Could a product deliver this more efficiently than your current service model?

Evidence. What's the smallest experiment that could prove or disprove the opportunity?

Don't start by building software

Spotting an opportunity doesn't mean committing to months of product development. The objective is to reduce uncertainty before increasing investment.

Hypothesis → Experiment → Build → Improve & maintain

That sequence is the whole of this guide. Everything that follows tells you how to run it.

A real-world example: Growth Gorilla

The opportunity. Iconic Digital already had the expertise. The opportunity was to turn more of that expertise into a scalable AI-powered product.

What it does. Ferrous Labs helped build Growth Gorilla: an AI-powered platform designed around marketing expertise and delivery, bringing together capabilities across content, social, SEO and GEO, and strategy.

The interesting part. It isn't just the AI. It's taking expertise that already exists inside a business and finding a way to productise it.

How to use this guide

The checklist you have just read asks whether your expertise could become a product. This guide is about what to do with a "yes".

It is written to be worked through, not read. Each part ends with something to fill in, decide or check inside your own business. By the end you will have a scored opportunity, a list of things to verify before spending money, a four-week validation plan, a realistic budget, and a set of conditions under which you would stop. None of it requires a technical brief. All of it starts with the business opportunity.

Part 1: Score yourself

The five signs contain fifteen tick boxes. Count yours.

SIGN	TICKS
Sign 1: Valuable but hard to scale	___ / 3
Sign 2: Turning away small work	___ / 3
Sign 3: Asked for the framework	___ / 3
Sign 4: Market productising around you	___ / 3
Sign 5: Data nobody else has (count double)	___ / 6
Total	___ / 18

0 to 6: Not yet. Your business is a service business and that is a perfectly good thing to be. Revisit in twelve months. Sign 4 is the one to keep watching.

7 to 11: One or two signs are strong. Do not build anything yet. Work through Parts 4 to 6 for the single strongest sign. That takes about a month and costs almost nothing.

12 to 18: Strong foundations. The question is no longer "could we" but "which one first, and what would it cost to find out". Work through the whole guide. Give someone in the business two days a week for the next six weeks to own it.

WHY SIGN 5 COUNTS DOUBLE

Proprietary data that contains expert judgement is the hardest asset to acquire and the one investors in vertical AI say compounds most [17]. It is also the one SME owners most often overestimate (Part 5), so it counts double only if the Sign 5 boxes are honestly ticked.

WHICH SIGN IS STRONGEST?

Circle the sign with the most ticks. If there is a tie, choose the one where you can name a specific client, a specific piece of work and a specific price paid. Specificity is what you will need in Part 6.

Part 2: Why "could we?" is the wrong first question

The checklist deliberately asks "could your expertise become a product?". Before going further, it is worth knowing why the honest answer is usually "probably, and that is not the point".

Plenty of UK businesses use AI. Few are getting much from it.

Self-reported AI use among UK businesses with 10 or more employees rose from around 12% to around 35% between late 2023 and mid 2026 [1]. Among businesses with 0 to 9 employees it is 28% [1]. The British Chambers of Commerce found B2B service businesses (finance, law, marketing) at 46% adoption against 26% for B2C and manufacturing [4].

Depth is the problem, not breadth. Adopters use on average 1.6 AI technologies and only 10% report extensive use [1]. In DSIT's survey of 3,500 businesses, 77% of adopters reported no change in revenue and only 12% reported an increase [3]. The barriers reported most often are the same few everywhere: not having identified a use case (71% of DSIT respondents), lack of skills and expertise, and cost [136].

The takeaway: the businesses around you have tools. Very few have products. The gap between the two is the opportunity, and it is a business gap, not a technology gap.

AI projects fail for business reasons, not model reasons

RAND interviewed 65 practitioners and found the root causes of AI project failure were, in order: leadership not defining the problem and the metric; insufficient good-quality data; chasing the newest technology rather than the right one; underinvesting in infrastructure; and applying AI to problems beyond what it can do [9]. None of those is about the model.

BCG's survey of 1,000 executives put it as a ratio: around 70% of the difficulty is people and process, 20% is technology and data, and 10% is the algorithm [11]. The businesses that got value pursued fewer opportunities and scaled more of them.

A widely quoted MIT figure says 95% of generative AI pilots deliver no measurable profit-and-loss impact [8]. That study's method has been criticised (small sample, undisclosed demographics, not peer reviewed) [8a], so treat it as a warning rather than a statistic. The more defensible finding from the same report is that tools built with an external specialist reached deployment roughly twice as often as internal builds [8].

Service businesses that build products fail in a specific way

The only large-sample study of this exact question is old but blunt: of product ventures started by 134 software-service companies in the late 1990s, roughly 87% were unsuccessful. The author's explanation was that the businesses failed to recognise how different the two sectors' business

models are, and carried service-sector practices and culture into the product venture [12]. The technology has changed since 2001. The organisational trap has not.

A consultancy co-founder who spent around $30m on two failed product pivots summarised it this way: "In consulting companies, people are the heroes. In product companies, the product is the hero." Services businesses optimise for utilisation; product businesses obsess over recurring revenue and churn. "Half measures don't work." [13]

A 2026 investor analysis adds the commercial version: integrated service businesses rarely exceed low single-digit market penetration, while vertical software reaches 10% to 80% share; the real moats are "proprietary data loops, workflow integration, distribution, network effects", not labour hours [14].

So what is the right first question?

Not "could we build this?" but "would anyone pay for it, would it cannibalise what pays the bills, and who would own it?". Parts 4 to 6 answer those three, in that order, before any money is spent on code.

Part 3: The three shapes a product can take

”AI product” makes people picture a venture-backed software start-up. That is the least likely and most expensive shape. There are three, and they form a ladder.

SHAPE 1: INTERNAL TOOL

Your team uses it. Clients see faster, cheaper, more consistent delivery. It changes margin, capacity and how much senior time each job needs, not revenue directly. Risk is lowest: no new sales motion, no support desk, no churn. Typically an exploration sprint and one or two engineering sprints (Part 7.1). Best when Sign 1 is strongest.

SHAPE 2: CLIENT ADD-ON OR PORTAL

Existing clients use it, self-serve, for part of what you do. It adds recurring revenue to relationships you already have and keeps you in the relationship between engagements. Risk is moderate: you now have users to support, a price to defend and a product someone has to keep working. Typically a full four-stage programme, so a year-one build envelope in the £43k to £85k range (Part 7.1). Best when Sign 2 or Sign 3 is strongest.

SHAPE 3: STANDALONE PRODUCT

People who were never your clients use it. It changes everything: you are running a second business with marketing, sales, onboarding, support and a roadmap. Go-to-market cost usually exceeds build cost within the first year (Part 7.2), and that is the number people forget. Best when Sign 4 and Sign 5 are strong and Part 4 has found a market gap.

THE DECISION RULE

Start one rung lower than you think. Shape 1 proves the workflow works and produces the usage data and error rates you will need to sell Shape 2. Shape 2 proves people will pay and tells you what they actually use. Shape 3 is what you do once Shape 2 is boring.

Investors say the same thing from their side. Bessemer’s playbook for vertical AI advises starting with ”tertiary workflows” (fastest sales, lowest resistance), moving to adjacent workflows, and only attacking the core workflow once you can demonstrate something close to ”an undeniable miracle” [17].

TWO KINDS OF SERVICE BUSINESS

One more distinction decides which shape is realistic. Is the human in your service a wedge (a way in, which software can later replace) or a delivery engine (a permanent layer, because a qualified, insured, accountable person has to sign)? [14]

Inspection, audit, legal opinion, structural sign-off and regulated advice are delivery-engine services. The product there does not replace the expert; it gets the expert to the signature faster and lets them handle three times the volume. A UK inspection-AI company put it well: the industry "doesn't need AI that replaces people; it needs to empower experts to work faster and smarter" [42]. That is Shape 1 grown into Shape 2, with the credential staying the moat (Part 4, criterion 5).

Part 4: Pick one. The opportunity scorecard

Take the strongest sign from Part 1 and write down up to three specific product ideas it suggests, one sentence each:

"[Who] pays [roughly what] to get [outcome] without [the thing they currently hate]."

Score each idea 0, 1 or 2 against the eight criteria below. They are drawn from what investors in vertical AI look for and, more usefully, from what the counter-arguments say survives when the models become commodities [15][16][17][18][19][20].

1. Toil. Is the work repetitive, necessary and unloved? Reviewing, triaging, extracting, classifying, drafting the standard version. (0 = every job is different; 2 = the same structured task, daily.)

2. Compulsory spend. Does the customer have to do this? Tax, compliance, insurance, regulatory reporting, safety. (0 = nice to have; 2 = legally or contractually required.)

3. Repeatable and definable. Can you write the process down and say what "done" looks like? "I do marketing strategy" is not an offer, it is a category [20]. (0 = depends on the senior's judgement each time; 2 = there is a checklist and a template already.)

4. Fragmented customer base with no system of record. Are there hundreds or thousands of potential customers, none of whom already run their lives in one dominant piece of software? Fragmented markets favour new entrants; markets with a dominant platform force you into being an add-on to it [18]. (0 = ten big buyers, or everyone lives in one incumbent tool; 2 = thousands of small buyers on spreadsheets and email.)

5. An accountability layer you own. Do clients pay you partly for your credential, insurance or willingness to own the outcome? That layer is not commoditised by a better model, and a product that sits underneath it inherits its defensibility [19]. (0 = no; 2 = yes, and the product would sit under that signature.)

6. Data that contains decisions. Do you hold historical records that include the expert's decision, not just the inputs? Raw data is rarely a moat on its own [15][16]; curated, labelled, high-quality domain data compounds [17]. (0 = inputs only, or nothing structured; 2 = years of records with the decision and outcome attached.)

7. A small market made viable by AI. Was this market "too small" for software because willingness to pay was tied to software budgets, and does it become viable when the product replaces labour? [16] (0 = the customer already buys software for this; 2 = the customer currently pays people for this, not software.)

8. The 10x model test. If a frontier lab releases a model ten times better tomorrow, does this product still have a reason to exist? [20] It does if criteria 4, 5 or 6 scored 2. (0 = the product is the model; 2 = the product is the workflow, the data, the credential and the distribution.)

CRITERION	IDEA A	IDEA B	IDEA C
1. Toil			
2. Compulsory spend			
3. Repeatable and definable			
4. Fragmented market, no system of record			
5. Accountability layer you own			
6. Data that contains decisions			
7. Small market made viable by AI			
8. The 10x model test			
Total (out of 16)			

Anything under 8, park. Take the top scorer into Part 5. If two are close, take the one nearer Shape 1.

A NOTE ON MARKET SIZE

You will be tempted to size the market in labour spend ("accountants spend £X billion a year on this"). Investors do it; it is also the most common way to fool yourself. Only a fraction of any service budget is addressable, because much of it buys liability transfer and political cover, not output [19]. Size the market as the number of businesses with this exact problem, multiplied by what a stripped-back version is worth to them per year. Part 6 gets you the second number from real conversations.

Part 5: The foundations audit

Before talking to a single customer, spend a week checking four things inside the business. Each has killed products after the money was spent. Each takes hours to check now.

5.1 Process: can it be written down?

- [ ] There is a documented process, checklist or template for the work (even if nobody follows it exactly).
- [ ] Two different people in the business would produce broadly the same output from the same inputs.
- [ ] You can list the inputs the work needs (documents, data, client answers) and where each one comes from.
- [ ] You can describe what "good" looks like in a way a stranger could check.
- [ ] You can name the 20% of cases that need a senior and say what makes them different.

If fewer than three are ticked, the product is not ready to be specified. Spend two weeks having the people who do the work write it down. That document is the product specification, whoever builds it.

5.2 Data: is it really there?

This is where Sign 5 gets tested. Fill in the table for the main dataset the product would rely on.

QUESTION	YOUR ANSWER
Where does it live? (a system with an export / spreadsheets / PDFs and Word files / email / several)	
Does each record contain the inputs?	Y / N
Does each record contain the expert's decision?	Y / N
Does it contain the reasoning?	Y / N
Does it contain what happened afterwards?	Y / N
How many records, over how many years?	
Consistent format throughout?	Y / N
Who owns it? (see 5.3)	
Does it contain personal data?	Y / N

Reading the answers:

Records without the decision attached are inputs, not training data. They can still support extraction and search products, but not anything that claims to replicate judgement.

Unstructured data (PDFs, scans, emails) is no longer a blocker. Vision models now read a clean page for a fraction of a penny, well under £1 per thousand pages at current small-model prices [25][26]. The cost of getting the data into shape is people and checking, not compute.

Quality beats quantity. A few hundred carefully checked examples of the decision your experts make, with the reasoning, is worth more than years of unchecked records [17]. Part 9 explains why: you need them for testing before you need them for training.

5.3 Contracts and rights: is it yours to use?

This is the check SME owners most often skip. Pull three recent client engagement letters or terms and look for the following.

- [ ] **IP clause.** Under UK law an independent consultancy is the first owner of copyright in what it creates unless it has assigned it in writing (Copyright, Designs and Patents Act 1988, ss.11 and 90) [35]. But client-drafted terms usually contain a "hereby assigns" clause handing "Deliverables" to the client, with "background IP" carved out for you [36]. Which do yours say? Your methodology is normally background IP. The reports you produced may not be yours.

- [ ] **Database right.** A database built with substantial investment attracts its own 15-year right, and for a commissioned database the first owner is typically the commissioner, not the person who built it [37]. If a client paid you to compile it, check.
- [ ] **Confidentiality.** Scope and duration. Does it cover "derived" or "aggregated" data? Newer terms increasingly do [36].
- [ ] **Data processing terms.** If you were a processor for the client, the data processing agreement restricts purposes. Product development is a new purpose.
- [ ] **Regulated-sector terms.** Solicitors, surveyors, accountants and engineers have professional-body rules that reach into this (Part 11).

The practical rule: anything you plan to use that came from a client engagement gets checked against that engagement's terms, and where the terms are silent, ask. A short letter to a good client requesting a licence to use anonymised historical data for product development costs nothing to send and is often granted. A product built on data you turn out not to own is worth nothing.

On AI-generated output: the government's March 2026 report on copyright and AI proposed removing s.9(3) of the CDPA (which gives authorship of computer-generated works to the person who made the arrangements) while keeping protection for AI-assisted human work [38]. No legislation is timetabled [39]. Do not rely on owning copyright in what the model produces. Rely on owning the workflow, the prompts, the evaluation data and the brand.

5.4 Personal data: the UK GDPR route

If the data includes anything about identifiable people (clients' staff, customers, tenants, patients, employees), the following applies. The Data (Use and Access) Act 2025 came into force in stages from February 2026 and did not create a shortcut for reusing client data to build a product [31][32][33].

What the ICO expects [30][31][33][34]:

1. Development and deployment are separate purposes and each needs its own lawful basis. "Building an AI product" is not a specific enough purpose; "a model that classifies X for Y" is [30][31].
2. Legitimate interests is the usual basis for development, and it needs a documented three-part assessment (purpose, necessity, balancing). "The data might be useful" fails the necessity test [30]. The Act's new "recognised legitimate interests", which skip the balancing test, cover things like crime prevention and safeguarding, not commercial product development [32][33].
3. A Data Protection Impact Assessment is expected. The ICO's own list of triggers includes "innovative technology" such as AI and machine learning [34].
4. Anonymise before you train where you can. Pseudonymised data is still personal data. Anonymising is itself processing and needs a lawful basis. The test is "reasonably likely" re-identification by a motivated intruder, not zero risk [33].

5. If the product will make solely automated decisions with significant effects on people, the Act now allows this on a legitimate-interests basis (unless special-category data is involved) with safeguards: tell the person, let them make representations, allow human intervention, allow contest [32].

A statutory ICO code of practice on AI and automated decision-making is due in 2027 [32a]. Until then, the 2024 ICO AI guidance is the reference.

Time and cost to do this properly for an SME: a few days of a data protection adviser's time, in the low thousands. Do it before Part 6, not after the build.

5.5 People: who owns it?

- [ ] Name one person who will own the product. Not a committee.
- [ ] Can that person give it two days a week for six months without their billable work collapsing?
- [ ] Is there budget ring-fenced so that a busy quarter cannot eat it?

In Vecteris's 2025 survey of professional services businesses, those with at least one person dedicated to productisation were 36% more likely to report strong outcomes, and only 15% ran a separate sales structure [46]. Their operating-model advice: one senior owner, funding released in stage-gated tranches, and a small squad of builders plus subject-matter experts from delivery [47]. In a business of ten to fifty people the "squad" is usually the owner plus an external product partner who brings the product discipline the business does not have in-house (Part 8).

If nobody inside the business can be named as owner, stop here. That is not a failure; it is the most common and cheapest place to find out.

Part 6: Validate before you build. A 30-day plan

The checklist's closing line is right: the next step is to determine whether it is worth building. Here is how, in four weeks, for the cost of one person's part-time attention. This is the work a good product partner will do with you; it is also work you can start tomorrow on your own.

Week 1: Write it down and find twenty people

Days 1 to 2. Take the winning idea from Part 4 and write:

- The one-sentence offer (Part 4 format).
- The three things it will do in version one. No more.
- The one thing it will not do, and who handles that (usually you).
- Who the user is, by job title.
- What they do today instead, and what that costs them.

Days 3 to 5. Build a list of twenty people to talk to:

- Ten existing clients who have bought the service this replaces.
- Five prospects you turned away or lost on price (Sign 2).
- Five people who match the profile but have never heard of you.

The last five matter most. Clients like you. Strangers tell the truth.

Weeks 2 and 3: Twenty conversations

How many is enough? Five to eight interviews within one homogeneous segment surface the main problems; fifteen to twenty give confidence before a serious investment [28]. Steve Blank's Lean LaunchPad sets ten to fifteen a week [27a]; the US NSF I-Corps programme requires 100 over seven weeks [27]. Twenty in two weeks is a realistic SME target.

Use The Mom Test rules [27b]: talk about their life, not your idea; ask about specific things that happened in the past, not what they would hypothetically do; listen more than you talk.

A thirty-minute script that works:

1. "Walk me through the last time you had to [the task]. What happened, start to finish?"
2. "What did that cost you, in time or money or grief?"
3. "What did you try? What did you use? Why did you stop?"
4. "Who else was involved? Who signed it off?"
5. "If it went wrong, what would happen?"
6. Only now: "We're thinking about a tool that does [three things]. What's the first thing that worries you about that?"

7. "What would you need to see to trust it?"
8. "Would you be willing to try an early version? What would you expect to pay for it, roughly, per month?"
9. The ask (see the commitment ladder).

Three kinds of answer to discount: compliments ("that sounds great"), hypotheticals ("I would definitely use that") and feature wishlists. The signal is specific past pain plus a commitment [27b].

THE COMMITMENT LADDER

A good conversation ends with the other person giving something up: time, reputation or money [27b]. In rising order of value:

1. They agree to a second conversation.
2. They introduce you to someone else with the problem.
3. They agree to share example data or documents.
4. They sign a non-binding letter of intent (common in B2B, a real filter even though investors discount it) [29].
5. They agree to a paid pilot.

Log which rung each of the twenty reaches. Fewer than five reaching "share data" or above is a stop signal.

PRICING WITHOUT ASKING "WHAT WOULD YOU PAY"

The Van Westendorp method uses four questions [29a]: at what monthly price would this be so expensive you would not consider it; so cheap you would doubt it; getting expensive but you would still think about it; a bargain. Ask all four. Twenty answers give you a defensible range and a point where "too cheap" and "too expensive" cross. That is your starting price.

INTERVIEW LOG

NAME	SEGMENT	LAST TIME THEY DID THE TASK	COST TO THEM	WORKAROUND	MAIN WORRY	TRUST CONDITION	VAN W. (4 PRICES)	TOP RUNG

Week 4: Decide, and pre-sell if you can

Count from the log: people who described the same two or three problems in similar words, unprompted (more than half means you have found something real [28]); people who reached "share data" or above; the Van Westendorp range; and the trust conditions, which become the acceptance criteria in Part 9.

Then choose one of four outcomes and write the date next to it.

Build a concierge version. Deliver the product by hand to three to five paying customers before any software exists. The customer knows it is manual. Meal-planning service Food on the Table ran this way at first; Zappos' founder photographed shoes in shops and bought them at retail when someone ordered [29b]. In an expertise business this is easy: it is the service, sold in the product's shape, at the product's price, with the product's turnaround. If you cannot sell that, software will not fix it.

Run a paid pilot. If the offer needs some software to be credible, charge for the pilot. Paid pilots make the customer invested and stop the "signed up and never logged in" pattern; keep them to around a month for a simple tool, set a baseline before you start, and agree in writing what success looks like [29c].

Park it, with a date. The idea is real but the timing, data or owner is not. Write the date and the condition that would reopen it.

Drop it. Write down why. That page is worth money the next time someone in the business has the same idea.

One warning on what counts as a sale. A US AI sales-tool company reported nearly $10m of annual recurring revenue while retaining around $3m, partly by counting three-month trials as annual contracts [45]. Decide now: a customer counts when they have paid for a second period.

Part 7: The money

7.1 What it costs to find out, and what it costs to build

Most businesses ask "what would the product cost?" before they have spent anything on knowing whether to build it. That is the wrong order, and it is why the figure feels frightening: you are being asked to price a thing nobody has tested.

Work is bought in sprints, and a sprint is a fixed block of days with something at the end of it. Below is the Ferrous Labs rate card. It is published here so you can budget honestly rather than guess. All figures exclude VAT.

SERVICE	UNIT	THE FRACTIONAL TEAM IT PUTS ON THE WORK	FEE
Data Exploration Sprint	5 days	Senior Data Scientist	£8,000
Research and Design Sprint	5 days	Senior UX Designer, Senior AI Engineer	£8,000
Model Training Round	5 days	Senior Data Scientist	£12,000
Engineering Sprint	10 days	Senior UX Designer, Senior AI Engineer, Senior Software Engineer, QA Engineer	£15,000
Application Maintenance	per day	Senior UX Designer, Senior AI Engineer, Senior Software Engineer, as needed	£1,500
Model Maintenance	per day	Senior AI Engineer, Senior Data Scientist, as needed	£2,000

Cloud infrastructure and any third-party data licences are passed through at cost, typically £100 to £400 a month for a product at SME scale.

The third column is the part that is easy to miss. A product like this needs a data scientist, an AI engineer, a software engineer, a designer and a tester, and it needs each of them for a different and quite short stretch of the work. Hiring that team is out of reach for a business of ten to fifty people, and hiring one person who claims all five skills is how products end up with no evaluation discipline and no design. Buying the sprint buys the specialists for the days the work actually needs them, and stops paying for them when it does not.

WHAT A FULL PROGRAMME LOOKS LIKE

Those units combine into the four stages in Part 8. A representative programme for an expertise product with a real AI component:

STAGE	WHAT IT TAKES	COST
Product Hypothesis	1 × Data Exploration Sprint	£8,000
Product Experiment	1 × Research and Design Sprint plus 1 to 2 × Model Training Round	£20,000 to £32,000
Product Formulation	1 to 3 × Engineering Sprint	£15,000 to £45,000
Indicative full programme to a working product		£43,000 to £85,000
Product Execution	2 maintenance days a month (one model, one application)	£3,500 a month

Two things matter more than the totals.

You only ever commit to the first sprint. £8,000 buys the Hypothesis stage. Everything after it is re-forecast once that sprint has told you what is actually there, and you approve each stage separately. Stopping after any stage is a normal outcome, not a failure.

Treat every figure beyond the first sprint as plus or minus 50% until that sprint is done. They are useful for budgeting, not for signing. Anyone who quotes a precise price for the whole programme before looking at your data is guessing, and the guess will be wrong in the expensive direction.

The honest version of the arithmetic: you are spending roughly a tenth of the programme to find out whether the other nine tenths are worth spending. If the answer in week one is that the signal is not there, you have spent £8,000 rather than £70,000, and that is the sprint doing its job.

HOW THIS COMPARES WITH THE MARKET

For context, here is what the wider UK market charges for equivalent work [21][22]:

STAGE	TYPICAL UK MARKET PRICE	TYPICAL TIME
Validation prototype	£5,000 to £15,000	3 to 4 weeks
Simple single-platform MVP	£8,000 to £30,000	8 to 12 weeks
Production-ready MVP	£15,000 to £35,000	8 to 12 weeks
MVP with payments, integrations or AI features	£35,000 to £80,000	3 to 6 months
Complex multi-system product	£70,000 to £150,000	5 to 8 months

A senior UK full-stack developer's median day rate is around £625; London agencies charge £600 to £900 a day and regional agencies £350 to £550 [21]. Ignore quotes built on public-sector rate cards, where £125 to £170 an hour and six-figure twelve-week builds are normal [23]; those reflect government procurement, not the SME market.

The market ranges and the rate card broadly agree on what a build costs. Where they differ is at the front. Most of the market has no priced way to answer "should we build this at all", so the first cheque is the build cheque and the question never gets asked.

THE REST OF THE YEAR-ONE BUDGET

The sprints are not the only cost. Whoever builds it, budget for these too:

ITEM	COST
Data protection advice and a DPIA where personal data is involved (Part 5.4)	£2,000 to £5,000
Legal: terms of service, privacy policy, data processing terms	£2,000 to £6,000
Design and onboarding beyond what the sprints cover	£3,000 to £10,000
Contingency	20% of the programme

7.2 What it costs to run

This is the part that surprises people, in a good way and a bad way.

Hosting is nearly free at the start. An MVP with under 100 users typically costs £0 to £15 a month; 100 to 1,000 users, £20 to £150 a month; managed databases from around £20 a month [24]. At Ferrous Labs cloud infrastructure is passed through at cost, typically £100 to £400 a month for a product at this scale.

Model costs are small at SME volume. Extracting structured data from 1,000 documents a month costs in the region of £1 to £6 a month on a small model, £5 to £10 on a mid-tier model and £30 to £40 on a frontier model at September 2026 list prices [25]. Inference cost for a fixed level of capability has been falling roughly ten times a year [25a]; the cost to reach GPT-3.5-level performance fell more than 280-fold in about eighteen months [25b]. Model cost is not the cost driver.

Maintenance is the cost driver. The industry rule of thumb is 15% to 25% of the initial build cost per year to keep a product working [24a] [24b], and across a product's lifetime maintenance has been estimated at around 60% of total cost, with a range of 40% to 80% [24c]. In practice it is bought as days. A product with one application and one trained model takes around two days a month (£3,500) covering drift, data refresh, service upkeep and seasonal retraining. A product with no trained model needs application maintenance only, so one to two days a month, £1,500 to £3,000. Unused days roll over.

AI products carry that line because roughly once a year a model you depend on will be retired and the product has to be re-tested on its replacement [69] [70]. That is what the maintenance days are for; it is not an optional extra.

Go-to-market is the big one for Shape 3. Sales and marketing efficiency for B2B software start-ups roughly halved in 2025: each £1 of sales and marketing generated 3.2 of revenue against 6.1 the year before [24d]. If you are selling to strangers rather than to your own clients, expect the first year's marketing and sales spend to exceed the build.

Year-one running budget for a Shape 2 product:

ITEM	COST
Cloud infrastructure and model calls	£1,200 to £5,000
Maintenance (1 to 2 days a month, rising to 2 with a trained model)	£18,000 to £42,000
Product owner time inside your business (2 days a week, opportunity cost)	£25,000 to £40,000
Support, in-house or bought in while you build the capability	£5,000 to £12,000
Marketing to existing clients	£2,000 to £5,000
Year-one running total	£51,000 to £104,000

Build plus run in year one is therefore roughly £110,000 to £190,000 for a product you intend to sell, of which the code is under half. That is the number to test the Van Westendorp price range against, and it is precisely why the £8,000 first sprint exists.

7.3 Pricing what you build

Three findings to have before setting a price.

Cheap AI products churn. Across 3,500 companies, AI products priced above $250 a month retained 70% of gross revenue year on year; those at $50 to $249 retained 45%; those under $50 retained 23% [44]. A low price does not buy loyalty, it buys trial-and-abandon. For a product carrying real expertise, start above £200 a month per customer, not per seat, and let the Van Westendorp range tell you how far above.

Subscription is still the default; hybrid is growing. In High Alpha's 2025 benchmarks, 53% of software companies were pure subscription, 31% subscription plus usage, 11% usage only and 5% outcome-based [43]. Among businesses under $5m revenue a flat fee is the commonest model [43a]. Enterprise buyers increasingly say they prefer consumption or outcome pricing over per-user [43b], but even Salesforce has moved back towards seats with usage caps because customers wanted predictable bills [43c]. For an SME product, a flat monthly fee per customer with a fair-use volume cap is easy to explain and easy to buy.

Professional services products land in the five figures. Subscription products from professional services businesses typically sell for low to high five figures a year depending on client size [48]. UK anchors: a construction project-management tool built for SME contractors sells at £100 per seat per month and won nearly twenty customers in its first two months [41]; enterprise legal AI tools sell for hundreds of thousands a year [40]. SMEs typically pay £10 to £50 per user per month for generic software and waste nearly 40% of it [44a], which is exactly why an expertise product should be priced on the outcome it produces, not on seats.

Pricing worksheet:

LINE	VALUE
Van Westendorp range from Part 6 (per month)	£____ to £____
What the customer pays today (people, outsourcing, your service), per month	£____
Your cost to serve one customer (infrastructure, model, support, share of maintenance), per month	£____
Gross margin at proposed price	____ %
Customers needed to cover year-one running cost	____
Customers reachable from existing relationships in 12 months	____

If the last number is smaller than the one above it, you are building Shape 3 and should price and budget as such.

7.4 Cannibalisation: the question nobody wants to ask

If the product replaces the routine, lower-margin part of your service, it frees senior time and you win twice. If it replaces the high-margin senior work, you are trading £1 of fees for pennies of subscription. Answer before building:

- Which service line does this product replace, wholly or partly?
- What margin does that line earn today?
- What does the senior time freed by the product get spent on instead? If the answer is "nothing yet", that is the problem to solve first.

The productisation literature's answer is to aim products at the price-sensitive segment the service cannot reach profitably (Sign 2) and at enhancing the service rather than substituting for it [48a]. Law practices provide the cautionary version: the billable hour survives because compensation and performance systems depend on it, and productivity gains "threaten revenues and profits" under that model [48b]. If your business bills by the hour, decide how the product's users and revenue will be credited before the product exists.

7.5 Funding and support (current September 2026)

The grant landscape for general SME software innovation has thinned in 2026. What is live :

TAX RELIEF: THE MOST RELIABLE LEVER

- Merged R\&D expenditure credit: 20% credit on qualifying R\&D spend, worth around 15% net after corporation tax, for accounting periods from April 2024. New claimants must notify HMRC in advance and file an Additional Information Form [49].
- Enhanced R\&D Intensive Support (ERIS): for loss-making SMEs spending at least 30% of total expenditure on R\&D, worth roughly 27p per £1 of qualifying spend [49].
- Data Exploration Sprints and Model Training Rounds, where the technical outcome is genuinely uncertain, are usually the most clearly qualifying spend in the whole programme. Speak to an adviser before the work starts rather than at year end, because the claim depends on how the work is recorded while it happens.

DEBT

- Growth Guarantee Scheme: loans up to £2m with a 70% government guarantee, extended in July 2026 with the turnover cap raised to £54m and terms up to ten years [51].
- Start Up Loans: up to £25k per founder at 7.5% fixed, available to businesses trading under 60 months, with twelve months of mentoring [52].

EQUITY INCENTIVES (IF THE PRODUCT BECOMES ITS OWN COMPANY)

- SEIS: up to £250k lifetime for companies under three years old with under £350k gross assets; investors get 50% relief [53].
- EIS: annual limit doubled to £10m and lifetime to £24m from April 2026; investors get 30% relief [53].

GRANTS AND ADVICE

- Innovate UK Smart Grants: paused in 2025 and not replaced; Innovate UK's 2026 strategy concentrates on fewer, larger bets in the Industrial Strategy sectors [50].
- Innovate UK Business Growth (adviser service): new requests paused [50a].
- BridgeAI: Innovate UK's AI adoption programme, expanded to eight sectors in late 2025 [50b].
- Made Smarter: manufacturers only; national rollout across England 2025 to 2026 [50c].
- Local Growth Hubs: national funding ended March 2026; some regions have refunded locally, so check yours [50d].
- Help to Grow: Management: still running twelve-week cohorts [50e].

Realistic advice: fund the first sprint from cash flow, claim the R\&D relief on the qualifying parts, and decide how to fund the rest once you have evidence rather than a hope. Treat grants as a bonus, not a plan. Where the opportunity is real but the cash is not, the commercial models in Part 8.4 exist for exactly that situation.

Part 8: How a build actually runs

Most product failures described in Part 2 were not failures of code. They were failures of sequence: the business committed to a build before it knew whether anyone wanted the thing, whether the AI could do the job on its data, or who inside the business would own it. So the route below is organised around decisions, not deliverables. It is the four-stage method Ferrous Labs uses, The Science of SaaS, and every stage ends with a gate at which stopping is a normal outcome.

The governing rule is that you only ever commit to the stage in front of you. After each one the remaining forecast moves, up or down, and you decide again.

8.1 Stage 1: Product Hypothesis

The question: which opportunity, for whom, and is what you think you have actually there?

What happens: a Data Exploration Sprint, five days. The strongest sign from Part 1 becomes two or three candidate ideas, scored the way Part 4 scores them. Then the foundations audit from Part 5 gets done for real: the data profiled rather than assumed (how many records, how consistent, do they contain decisions or only inputs, what is the match rate between systems), the process written down, the client contracts read, the owner named. The output is a written hypothesis: the one-sentence offer, the user, the value, the shape from Part 3, the risks in order, and the specific things that have to be true.

What it costs: £8,000.

The gate: is there a candidate worth testing, is there enough in the data to test it with, and is there an owner for it? This is the cheapest place in the whole guide to stop, and the most common. A week one answer of "the signal is not there" has cost £8,000 rather than the £70,000 the programme would have run to.

8.2 Stage 2: Product Experiment

The question: will anyone pay for it, and can the AI actually do the job to the standard your experts set?

Those are two separate risks and they need two separate tests. Run them in parallel, because a product that works but nobody wants and a product that sells but does not work fail in exactly the same way.

The demand test. A Research and Design Sprint, five days, £8,000. It answers the commercial question rather than the technical one: what exactly is the product, who buys it, at what price, how does it reach them, and what does it do to your existing service. It is where the twenty conversations in Part 6 turn into a written product definition that either side can walk away from.

The technical test. A Model Training Round, five days, £12,000, aimed at the thing most likely to kill the product. Usually that is accuracy: can the model do this job, on your data, at or below the human error rate you measured in Part 9.2? Sometimes it is extraction quality from awkward documents, or whether the historical records hold enough decisions to learn from. The output is a number and a test set, not an opinion. Products with genuine judgement in them typically need two rounds; clean, well-separated data needs one.

What it costs: £20,000 to £32,000 for the stage, depending on how many training rounds the data demands. Clean separation means fewer rounds and the bottom of the range.

The gate: did demand show up, and did the technical test pass? If demand is there but the technology is not yet, park it with a date; the cost curve in Part 7.2 moves in your favour every year. If the technology works but demand did not show up, do not build it, however impressive the result.

8.3 Stage 3: Product Formulation

The question: what exactly gets built, and does it hit the standard?

What happens: Engineering Sprints, ten days each at £15,000. A tightly scoped first product can be a single sprint; most take two or three. The sprint structure exists so scope is reviewed against evidence every fortnight rather than fixed in a specification written before anyone used anything. Each sprint ends with something working that your people can look at, and the acceptance criteria from Part 9.3 are the standard it is measured against, not "does it feel right".

What it costs: £15,000 for a single-sprint product, £30,000 to £45,000 for most. Larger where the product also has to be packaged for people outside your business, for example as an API other companies call. If the Experiment stage has done its job, the scope here is smaller than you expected at the start.

The gate: is the error rate at or below the human baseline, and are the pilot customers actually using it? Part 12 sets out the conditions under which the answer is no.

8.4 Stage 4: Product Execution

The question: does it keep working, and is it earning?

What happens: the product is live, so it needs application maintenance at £1,500 a day and, where there is a trained model, model maintenance at £2,000 a day, covering drift, data refresh, service upkeep, retraining and the model retirements described in Part 9.5. Two days a month is typical for one application and one model; unused days roll over. Cloud infrastructure is passed through at cost. The numbers you watch are in Part 11.4.

8.5 How the commercial side can work

The stages describe the work. There are three ways to pay for it, and the right one depends on how much cash you have and how much of the upside you want to keep. The work is identical in all three; what changes is what you pay with.

Delivery partnership. You pay the rate card. You own the product, the code, the models, the derived IP and all of the revenue, outright. This is the cleanest thing to explain to an investor, a client or an insurer, and there is nothing left to negotiate later.

Revenue partnership. Reduced fees in exchange for a share of what the product earns, for an agreed period. You still own everything; this changes how you pay, not what you get. It suits a business that believes in the opportunity but would rather keep cash in the service while the product finds its feet, and it aligns the builder with the product selling rather than merely shipping.

Equity partnership. The largest reduction in fees, in exchange for a share of the product. Two shapes: a stake in your own company settled against the balance of the fees, or a joint venture where the product becomes its own company that both sides own and both sides sell. A joint venture only makes sense from Formulation onwards, for the simple reason that nobody can jointly own a product neither party can yet describe. It is the lightest route on cash and the heaviest on shared upside: you are funding part of something you co-own rather than buying a discount, and what you give up is sole ownership of what it earns.

Reduced-fee routes carry meaningful discounts against the rate card, which compound across a programme. The trade is always the same one: equity is the more expensive currency if you succeed, and cheaper than cash if you are pre-revenue.

The pattern to avoid, whichever route you choose, is the one Part 2 describes: a fixed-price build quoted from a specification, delivered to a business that has not tested demand, with nobody inside the business who owns it afterwards.

8.6 The alternatives, and when they make sense

Doing it yourself with no-code or AI-assisted builders. Genuinely useful for Shape 1 internal tools and for the concierge version in Part 6. Cost from free to a few hundred pounds a month [54][55], speed measured in days. What to know before relying on it: some platforms do not let you export your code, so the product lives on their pricing and their roadmap [54]; usage-based pricing can move sharply as you grow [56]; and security defaults are yours to own. In 2025 more than 170 apps built with one AI-assisted builder were found exposing personal data, payment details and credentials because database access rules had not been set (CVE-2025-48757) [57]. There is also no evaluation discipline built in (Part 9): a tool that seems to work on ten cases is not a product. Use this route to prove a workflow to yourself, and be honest about the point at which it stops being the right answer.

A generalist software agency. Fine for conventional software. The question that separates a software builder from an AI product builder is how the AI parts will be tested: ask how accuracy will be measured on your own cases, what will be shown to prove it, and what happens when the model is retired. If the answer is "we'll try it and see", you are buying build hours, not a product. Remember the evidence in Part 2: failures cluster around problem definition, data and ownership rather than code, and externally built tools reached deployment roughly twice as often as internal builds [8].

Hiring in-house. Sensible once the product has customers and a roadmap. Rarely sensible before, because the first year needs a wide range of skills for a short time each: product, data science, AI engineering, design, data protection, evaluation. The transition point is usually when maintenance

and improvement become steady work rather than bursts.

8.7 Questions to ask anyone who wants to build it for you

- What have you built for a business like ours, and can we speak to them?
- What do you need from us, in days per week, and from whom?
- How will you test that the AI parts are accurate enough, and what will you show us to prove it?
- What happens when the model you have used is retired?
- Who owns the code, the prompts, the models, the evaluation data and the infrastructure accounts?
- What does it cost to keep running in year two if we change nothing?
- What is the smallest piece of work that would tell us whether to carry on, and what does that cost?
- What would make you tell us to stop?

The last two matter most. A partner who cannot price a small first step, or who has no conditions under which they would advise you to stop, is selling a build rather than a product.

Part 9: What AI can and can't do reliably, and how to prove it

You do not need to understand how the models work. You do need to be able to hold a build partner to account. This part gives you the five things to insist on.

9.1 Models make things up, at a measurable rate

On a standard test of summarising a document faithfully, the best models invent content in around 2% to 3% of cases and many capable models in 5% to 10% [61]. In a professional setting the rates are higher: Stanford found the leading legal research assistants gave wrong or fabricated answers in 17% to 33% of queries, against 58% to 82% for general chatbots [62]. OpenAI's own explanation is that standard training rewards a confident guess over "I don't know" [63].

Implication: a product that claims expertise cannot ship on "it seems to work". It ships on a measured error rate on your own cases, compared with the human rate on the same task.

9.2 "Good enough" means "better than the human baseline you measured"

The usual ceiling for acceptable manual data entry is around 1% of fields; for billing-critical fields, well below that [64]. For judgement tasks the human error rate is higher and usually unmeasured. So the first job is to measure it: give your experts fifty past cases and see how often two of them disagree. That number is the bar.

A warning from the research on checking: a person "eyeballing" already-entered data caught almost no errors compared with structured double entry [64a]. A human reviewer skimming AI output is not a control. Structured review of a sample, against a rubric, is.

9.3 Evaluations: the thing to insist on

The practitioners who build these products say the same thing: systematic evaluation is "the most important part of the whole system" [65], and generic benchmarks "barely correlate with application-specific performance" [66]. In practice, for an SME product:

1. Build a test set from your own history. Fifty to a few hundred real cases (anonymised where needed, Part 5.4) with the correct answer attached by your experts. This is the single most valuable thing the business contributes to the build, and it comes from Sign 5.
2. Score every version of the product against it. Pass or fail per case, with a note from an expert on why it failed. Look at the failures; three failure types typically account for most of the errors [65a].
3. Separate precision from recall for anything that classifies or extracts [66]. "95% accurate" is meaningless if the 5% are all the cases that matter.
4. Re-run the whole set every time the prompt, the model or the data changes.

Acceptance criteria worksheet (fill in before the build starts):

ITEM	YOUR ANSWER
Task the product performs	
Human error rate on 50 past cases	____ %
Required product error rate (at or below)	____ %
Cases the product must refuse or escalate	
Test set size and who labels it	____ cases, labelled by ________
Who signs off the score before launch	

9.4 Keep a human where it matters

Two triggers for human escalation, from OpenAI's own guidance for building agents: when the system exceeds a failure threshold (for example, cannot classify after repeated attempts), and for any action that is "sensitive, irreversible, or high stakes" until confidence is earned [67]. Anthropic's advice is to find "the simplest solution possible" and add complexity only when needed; a single well-prompted model call with retrieval and examples "is usually enough" [68]. For an expertise product: route low-confidence and high-stakes cases to a person, log what the person changed, and feed that back into the test set.

9.5 The model will change under you

Model providers retire models on notice: at least 60 days for one major provider [69], at least six months for another's general models, with several widely used models scheduled for shutdown in late 2026 [70]. Research on one provider's model between March and June 2023 found accuracy on a simple maths task fell from 84% to 51% between versions [71]. Plan for a re-test and possible re-tune roughly once a year, and budget for it in Part 7.2.

What a build partner should already be doing about it: routing model calls through an abstraction layer so a provider swap is a configuration change; versioning prompts per model; keeping the test set so a swap can be scored before it goes live; keeping an open-weight model as a fallback so that no provider outage or price rise is total [72].

9.6 What to buy, what to build

Buy: the model, hosting, authentication, payments, document parsing. Build: the workflow, the prompts and rules, the test set, the escalation logic, and the integration with how your clients actually work. The first list is a commodity. The second list is the product.

Part 10: Regulation as a product signal. The UK list

The checklist's "Signals to watch" page says a new compliance regime can force an entire industry to perform the same structured task. Here is the current UK list, with dates, so you can find your own sector. Correct as at September 2026; several of these are moving.

REGIME	WHO AND WHEN	THE REPEATABLE TASK	PRODUCT ANGLE
Making Tax Digital for Income Tax [73]	Sole traders and landlords: income over £50k from 6 April 2026; over £30k from April 2027; over £20k from April 2028	Digital records, quarterly updates through recognised software, software-only year-end	Getting a specific trade's records quarterly-ready with less accountant time
Companies House reform (Economic Crime and Corporate Transparency Act) [74]	Identity verification mandatory for new directors and PSCs from November 2025, existing ones within 12 months; software-only accounts filing and mandatory P\&L for small companies from 1 April 2028	Verification through an authorised agent; structured accounts filing	Agents and accountants processing verification and filings at volume
Failure to prevent fraud offence [74a]	Large organisations (two of: 250+ staff, £36m turnover, £18m balance sheet) from 1 September 2025; their suppliers by extension	Documented "reasonable procedures"	Risk assessment and evidence tooling for mid-market suppliers who must show large customers they comply
Procurement Act 2023 [75]	Every supplier to the public sector, since 24 February 2025	Registration on the Central Digital Platform, share codes per bid, supplier data kept current; 30-day payment terms down the chain	Bid-readiness and supplier-data maintenance for SMEs; notice monitoring
FCA Consumer Duty [76]	Tens of thousands of FCA-regulated businesses, in force since July 2023 with annual board reporting; a June 2026 consultation proposes narrowing scope, closing 18 September 2026	Outcomes monitoring, fair value assessments, annual board report	Evidence collection and reporting for small regulated businesses and the compliance consultants who serve them
Building Safety Act: higher-risk buildings [77]	Around 12,500 registered buildings; Building Safety Regulator standalone from January 2026; second-staircase rule from 30 September 2026; Building Safety Levy from 1 October 2026	Digital "golden thread", safety case reports, occurrence reporting, building assessment certificate applications (73% of decided applications refused in 2025/26)	Golden-thread structuring, safety-case drafting support, gateway readiness. High refusal rates are a demand signal

REGIME	WHO AND WHEN	THE REPEATABLE TASK	PRODUCT ANGLE
Employment Rights Act 2025 [78]	Every employer. April 2026: day-one sick pay and paternity leave. October 2026: six-month tribunal limit, "all reasonable steps" harassment duty. January 2027: unfair dismissal qualifying period cut to six months, compensation cap removed	Policy updates, documented procedures, probation management	HR consultancies serving businesses under 50 staff (Sign 2 territory)
Data (Use and Access) Act 2025 [79]	Every organisation processing personal data; main provisions from February 2026; marketing fines raised to £17.5m or 4% of turnover	Electronic complaints procedure with 30-day acknowledgement, subject-access handling, automated-decision safeguards	Complaints and rights-request handling for small organisations
Cyber Security and Resilience Bill [80]	Managed service providers, data centres, designated critical suppliers; at Lords report stage in September 2026, not yet law	24-hour initial and 72-hour full incident reporting; registration	Watch, do not build yet
Martyn's Law (Terrorism (Protection of Premises) Act 2025) [81]	Venues with 200+ capacity (standard) and 800+ (enhanced); guidance published April 2026; commencement not before April 2027	Notify the regulator; document evacuation, invacuation, lockdown and communication procedures; enhanced tier submits measures	Procedure templates and evidence for standard-tier premises. The Home Office says venues "do not need to spend money on consultants"; a product that makes compliance cheap fits that message
Biodiversity Net Gain [82]	Developers and planning consultants; mandatory since February 2024; small-sites exemption up to 0.2 hectares from August 2026; nationally significant projects from November 2026	Metric calculations, habitat plans, off-site unit sourcing	Calculation and evidence tooling for smaller developers and their consultants
Awaab's Law [83]	Social landlords. Phase 1 since 27 October 2025 (emergency hazards within 24 hours; damp and mould investigated within 10 working days). Phase 2 from 30 November 2026 adds cold, heat, falls, structural, fire,	Triage, deadlines, written summaries, evidence	Hazard triage and deadline tracking for smaller housing providers and their contractors

REGIME	WHO AND WHEN	THE REPEATABLE TASK	PRODUCT ANGLE
	electrical and hygiene hazards. Private-sector extension enabled, undated		
Renters' Rights Act 2025 [84]	Private landlords and letting agents. From 1 May 2026: section 21 abolished, periodic tenancies, one rent rise a year; landlord database late 2026; ombudsman around 2028	Compliant tenancy documentation, rent-increase process, database registration	Compliance workflow for small letting agents
Private rented sector EPC C [85]	All private landlords by 1 October 2030, with a £10,000 per property cost cap; new EPC metrics expected from 2027	Property-by-property retrofit assessment and evidence	Assessment and planning tools for portfolio landlords and energy assessors
UK Sustainability Reporting Standards [86] [87]	Nobody yet. Standards published February 2026 for voluntary use; FCA proposes mandatory climate reporting for around 515 listed companies from January 2027. EU supply-chain demand on UK SMEs weakened after the EU narrowed its rules in February 2026	Structured climate and sustainability disclosure	Not yet for SMEs. Watch

On AI regulation itself: there was no AI bill in the May 2026 King's Speech, and the FCA confirmed in June 2026 that it will not write AI-specific rules [88]. Existing law (data protection, Consumer Duty, professional-body standards) governs AI products in the UK for now.

How to use the list: find your sector. Write down the structured task the regulation forces. Count how many of your clients have asked you about it in the last month. If it is more than once a week, take it into Part 4 as a candidate idea and score it. Compulsory spend (criterion 2) and repeatability (criterion 3) will already be 2s.

Part 11: What changes inside the business

11.1 Ownership and money

One senior owner, not a committee; a ring-fenced budget released in stages against evidence; a small squad of builders plus your experts [47]. A chief operating officer quoted in the same research said of product versus service: "Selling at scale is just so different." Expect that.

Cash flow changes shape. Service fees arrive in lumps; subscriptions arrive in trickles [48]. Model the first eighteen months of product revenue against the service revenue it displaces (Part 7.4) before committing.

11.2 Professional-body rules

Check before launch, not after. Current positions:

- **Solicitors (SRA).** A law practice may own a separate unregulated business (the SRA's own example is "a separate company that provides online legal services") but must obtain informed consent for referrals, tell clients the separate business lacks Legal Ombudsman cover, compensation fund and compulsory insurance, and keep confidential information and IT separate [89]. The SRA's warning notice on misuse of AI covers verification and client data in tools; it does not address selling tools [89a].
- **Surveyors (RICS).** The RICS standard on responsible use of AI became mandatory on 9 March 2026 and covers "responsible development" as well as use: risk registers, a responsible-use policy, client disclosure and procurement due diligence [90].
- **Engineers.** The Engineering Council and Royal Academy of Engineering 2026 ethical principles require engineers to "consider fully the risks and ethical implications of emerging and fast-moving technologies, such as artificial intelligence" [91].
- **Regulated financial businesses.** No AI-specific FCA rules; Consumer Duty and the Senior Managers regime apply to any AI product used with customers [88].
- **Accountants (ICAEW).** No specific guidance found on member businesses selling software; new professional indemnity regulations applied from September 2024 [92]. Ask your institute.

11.3 Insurance

Standard consultancy professional indemnity policies often exclude or fail to describe software products, and typically exclude cyber, data breach and IP infringement. Technology errors-and-omissions cover is written for software. Early-stage products are typically insured at £1m to £2m; public-sector frameworks demand at least £1m [93]. Cover cannot be arranged retrospectively. Tell your insurer before the first paying customer, not after.

11.4 Numbers to watch in year one

Early-stage B2B software benchmarks (median, under $1m revenue): gross revenue retention 92%, net revenue retention 100%, customer acquisition payback around five months, gross margin 74% [43]. AI-native products as a group do far worse on retention (median gross retention around 40%) [44], which is the cheap-product churn effect from Part 7.3. Most software companies take two to five years from first revenue to reach $1m of annual recurring revenue [94].

Four numbers for the monthly dashboard:

1. Paying customers (paid for a second period, Part 6).
2. Gross revenue retention (of last year's revenue, how much is still paying).
3. Cost to serve per customer (Part 7.3 worksheet).
4. Product error rate on the test set (Part 9.3).

Part 12: Kill criteria. How to know when to stop

Write these down before starting, with the owner's signature, so that stopping is a decision made in advance rather than an argument made under sunk cost.

Stop after Part 5 if: nobody can be named as owner with two days a week; the data turns out to be inputs without decisions and the idea depends on decisions; client contracts prohibit the use and clients will not licence it.

Stop after Part 6 if: fewer than five of twenty people reached "share data" or higher on the commitment ladder; the Van Westendorp range is below your cost to serve; nobody described the same problem twice, unprompted; nobody would pay for the concierge version.

Stop during the build if: the product error rate cannot be brought to or below the human baseline on the test set after two iterations; pilot customers have not used it for two consecutive weeks; the build has passed 150% of the fixed-price budget.

Stop after launch if: gross revenue retention is under 70% at twelve months (the "under $250 a month" churn band) [44]; cost to serve exceeds price after six months; senior time freed by the product is not being redeployed to revenue.

Parking is a valid outcome at every gate. Record the date, the reason and the condition that would reopen it.

Part 13: A worked example, start to finish

An anonymised composite, so you can see the whole path on one page. Every step uses the tools above.

The business. A twelve-person structural engineering consultancy in the Midlands, turnover around £1.4m. Three chartered seniors sign everything. Most revenue is design work for developers; a growing slice is reviewing other people's designs for local authorities and insurers.

The score (Part 1). Sign 1: 3 (seniors are the bottleneck; a three-week turnaround is mostly queue). Sign 2: 2 (they decline small domestic reviews under £1,500). Sign 3: 1. Sign 4: 2 (two software products have appeared in structural review). Sign 5: 3, doubled to 6 (eleven years of review reports, each with findings, severity and the accept/reject decision, in a document management system). Total 14 of 18.

The ideas (Part 4). Three, all from Signs 1 and 5. A: a first-pass review assistant that reads a submitted design pack, checks it against the practice's review checklist, drafts the standard comments and flags what needs a senior. Scored 13. B: a self-serve pre-submission checker sold to small contractors so their packs arrive complete. Scored 9. C: a benchmark database of common design faults sold to insurers. Scored 7. Idea A wins. It is Shape 1 with an obvious path to Shape 2: offering faster, cheaper reviews to the small work they currently decline.

The audit (Part 5). Process: the review checklist exists and is used; seniors agree on what needs them. Data: 2,100 review reports with decisions; consistent template since 2018; almost no personal data. Contracts: the local authority framework assigns deliverables to the client but carves out "methodologies, tools and know-how"; the insurer terms are silent. A short licence request to both for anonymised historical use is granted within a fortnight. People: the associate who runs the review team takes ownership at two days a week, backfilled by a contractor.

The validation (Part 6). Twenty conversations: eight local authority and insurer clients, seven contractors previously turned away, five strangers found through the regional ICE branch. The clients want turnaround, not price. The contractors want to know what will fail before they submit, and would pay for that. Eleven of twenty reached "share data" or higher; four contractors sign letters of intent for a pre-submission check at £150 to £400 per pack (Van Westendorp crossing point £220). Decision: build the internal assistant with a product partner on a fixed-price prototype, and run a concierge version of the contractor check by hand for eight weeks using it.

The money (Part 7). A Data Exploration Sprint at £8,000 profiles the eleven years of reports and answers the first question: are the findings and decisions consistent enough to learn from, and do the scanned drawings match the digital ones. They do, at a 91% match rate. A Research and Design Sprint (£8,000) defines the contractor-facing product and its price; one Model Training Round (£12,000) answers whether the model can find what the seniors find. Two Engineering Sprints build the internal tool, £30,000. Data protection advice and legal £4,000. Running costs of £2,000 a month from go-live for two maintenance days plus infrastructure. Programme total about £62,000, of which £8,000 was spent before the decision to carry on. R\&D relief claimed on the exploration

sprint and the training round. Expected return: senior review time per pack down from about four hours to about ninety minutes on the 80% of packs the tool handles; capacity for roughly 60% more reviews with the same three seniors; concierge contractor checks bringing £3,000 to £5,000 a month by month six.

The proof (Part 9). Test set: 200 past packs labelled by the seniors with the findings that should have been raised. Human baseline: seniors disagreed with each other on 6% of findings. Acceptance: the tool must raise at least 94% of the findings the seniors agreed on, with a false-alarm rate a senior can dismiss in under a minute per pack, and must escalate anything involving transfer structures or basements. First version scored 81% recall; error analysis showed most misses were in scanned hand-annotated drawings; a better document-parsing step took it to 93%, then 95%.

The change (Part 11). Insurer notified; PI policy endorsed for the internal tool at no extra premium, with a quote for technology cover ready for when the contractor product goes live. The associate's targets move from billable hours to reviews completed and tool error rate.

The gate (Part 12). Twelve months on: if the contractor concierge has fewer than fifteen repeat customers, it stays a service. If it has more, it becomes the Shape 2 build, priced at £220 per pack or £600 a month unlimited, with the test set as the launch gate.

What the example shows: nothing in the year required a decision about "AI" as such. Every decision was about a queue, a checklist, a contract, a price and an error rate. That is the point of the checklist's closing line.

Part 14: Where to go next

Everything in this guide maps onto the four stages Ferrous Labs uses to take an expertise business from idea to product.

STAGE	WHAT IT ANSWERS	PARTS OF THIS GUIDE
Product Hypothesis	Is there an opportunity worth pursuing, which one, and who owns it?	Parts 1, 4, 5 and 8.1
Product Experiment	Will people pay, and can the AI do the job to the standard your experts set?	Parts 6, 9 and 8.2
Product Formulation	What exactly gets built, for how much, against what acceptance criteria?	Parts 3, 7, 8.3 and 12
Product Execution	Running it, measuring it and growing it	Parts 7.2, 8.4, 11 and 12

If the checklist and this guide have surfaced a process, a dataset or an area of expertise that could become a product, the next step is still not to build it. It is to find out whether it is worth building, and that is a four-week piece of work, not a four-month one.

The Opportunity Finder at ferrouslabs.co.uk takes about ten minutes and turns your answers to Parts 1 and 4 into a first view of where the opportunity sits and what it would take to test it. No AI brief required. Start with the business opportunity.

Sources

UK ADOPTION AND CONTEXT

- [1] ONS, Artificial intelligence in UK businesses: 2023 to 2026, 20 July 2026.
- [2] ONS, Business insights and impact on the UK economy, 2 October 2025.
- [3] DSIT, AI Adoption Research (IFF Research and Technopolis), published January 2026, updated 13 February 2026.
- [4] British Chambers of Commerce with Intuit, 22 September 2025.
- [5] British Chambers of Commerce, Future of Work: AI in the Workplace, 18 March 2026.
- [6] Enterprise Nation Tech Hub research, June 2026 (1,320 UK small businesses).
- [7] Simply Business SME Insights Report, 8 September 2026.
- [8] MIT NANDA, The GenAI Divide: State of AI in Business 2025, August 2025.
- [8a] Futuriom, "Why we don't believe MIT NANDA's AI study", 26 August 2025.
- [9] RAND, The Root Causes of Failure for Artificial Intelligence Projects, 13 August 2024.
- [10] Gartner press releases, 29 July 2024 and 25 June 2025.
- [11] BCG, Where's the Value in AI?, 24 October 2024.
- [12] Nambisan, "Why Service Businesses Are Not Product Businesses", MIT Sloan Management Review, Summer 2001.
- [13] Kuhn, "The $30M Lesson: When Consulting Firms Try to Build Products", Medium, 26 June 2025.
- [14] Elayat (Euclid Ventures), "AI Services Disagreement", 16 April 2026.

FRAMEWORKS, MOATS AND EXAMPLES

- [15] a16z, "The Empty Promise of Data Moats", 2019, updated 2024.
- [16] a16z, "'AI Inside' Opens New Markets for Vertical SaaS", 2 December 2024.
- [17] Bessemer Venture Partners, "Building Vertical AI: an early-stage playbook", January 2026; "Ten principles for building strong vertical AI businesses", 2025.
- [18] Euclid Ventures interview with David Haber (a16z), "Market Structure Is the New TAM", 6 May 2026.
- [19] Better Tomorrow Ventures, "Services Won't Become Software", 26 February 2026.
- [20] VC Cafe, "Vertical AI in 2026", 7 January 2026, and "AI-Native Services", 6 May 2026; Marilyn Wo, "Productizing Consulting", 9 March 2026.
- [40] Sacra, Luminance company profile (revenue estimate), 2025; Luminance Series C announcement, February 2025.
- [41] Construction Management (CIOB), "MCIOB develops own AI software for SMEs using AI", 17 April 2026.
- [42] EU-Startups, Scope raises €17.3m led by Index Ventures, 22 May 2026 (company-reported metrics).

- [45] GS Ventures and TechCrunch reporting on 11x, March 2025.

COSTS, HOSTING, MODELS AND MAINTENANCE

- [21] TulipTech, MVP development cost UK, 2 June 2026.
- [22] Coderacle, How much does an MVP cost in the UK, 12 May 2026.
- [23] rsnake.co.uk, UK software development cost 2026, 9 February 2026 (public-sector rate cards).
- [24] UX Continuum, SaaS hosting costs breakdown 2026, 14 August 2026.
- [24a] DesignRevision, How much does it cost to build a SaaS, 16 February 2026.
- [24b] Uvik, AI development cost, updated 3 September 2026 (vendor).
- [24c] Lientz and Swanson maintenance estimate via Notkin (University of Washington); Wood, "The 60/60 rule", 97 Things Every Project Manager Should Know (O'Reilly).
- [24d] Lighter Capital, 2025 B2B SaaS startup benchmarks, updated 6 July 2026.
- [25] AI Cost Check, AI OCR document processing costs 2026, 17 April 2026; provider price pages (OpenAI, Anthropic, Google), September 2026.
- [25a] a16z, "LLMflation", 12 November 2024.
- [25b] Stanford HAI, AI Index Report 2025.
- [26] Vellum, Document data extraction: LLMs vs OCR, 3 December 2025.

VALIDATION

- [27] NSF I-Corps programme requirements (100 interviews in seven weeks).
- [27a] Steve Blank, Lean LaunchPad course posts, 2025 to 2026.
- [27b] Rob Fitzpatrick, The Mom Test, 2013.
- [28] Skimle, Customer discovery interviews: a practical guide, 3 May 2026.
- [29] Learning Loop, Letter of intent play.
- [29a] Van Westendorp Price Sensitivity Meter, 1976.
- [29b] CRV, "What is an MVP?" (concierge and Wizard of Oz examples).
- [29c] Headway, How to run a B2B software pilot programme, 2023.

DATA PROTECTION, IP AND CONTRACTS

- [30] ICO, Guidance on AI and data protection: lawfulness, updated 28 October 2024.
- [31] Skadden, summary of ICO generative AI consultation outcomes, January 2025.
- [32] ICO, The Data (Use and Access) Act 2025: what does it mean for organisations, 19 June 2026; Clifford Chance, February 2026.
- [32a] Arnold and Porter, The ICO's new statutory duty to produce an AI code of practice, August 2026.
- [33] Taylor Wessing and Bird and Bird on ICO anonymisation and pseudonymisation guidance (final 28 March 2025); A\&O Shearman on ICO recognised legitimate interests guidance, 23 March 2026.
- [34] ICO, When do we need to do a DPIA?, updated 13 October 2025.

- [35] Sprintlaw UK, IP and independent contractors (CDPA ss.11 and 90).
- [36] Bind Legal, IP ownership clause guide, 2026.
- [37] Pinsent Masons Out-Law, Database rights: the basics.
- [38] HSF Kramer, UK government report on copyright and AI, 18 March 2026.
- [39] Hogan Lovells, Copyright and AI statement of progress, December 2025.

PRICING AND BENCHMARKS

- [43] High Alpha, 2025 SaaS Benchmarks Report, November 2025.
- [43a] Growth Unhinged, 2026 State of B2B SaaS and AI Monetization, 13 May 2026.
- [43b] Futurum, 1H 2026 Enterprise Software Decision Makers survey, 12 May 2026 (methodology not published).
- [43c] The Register, Salesforce reverting to seat-based AI pricing, 12 December 2025.
- [44] Growth Unhinged, "The AI churn wave", 10 December 2025 (3,500 companies).
- [44a] SME Today and Fasthosts, UK SMEs wasting up to £10k a year on unused SaaS, 15 April 2026 (methodology not disclosed).
- [46] Vecteris, 2025 Productization Benchmarking Survey.
- [47] Vecteris, Organizational design and operating models for product-led services firms, 30 June 2026.
- [48] Eisha Armstrong (Vecteris) on the Consulting Success podcast, updated August 2025.
- [48a] Vecteris, The 3 Cs of Cannibalization.
- [48b] Harvard Law School Center on the Legal Profession, Couture, 24 February 2025.
- [94] ChartMogul, "Against the Odds", 15 October 2025 (6,525 companies).

FUNDING

- [49] HMRC, Corporate Intangibles Research and Development Manual (CIRD115000, CIRD123000, CIRD127000); Uplift Tax, merged scheme explained, August 2026.
- [50] UKRI, Smart innovation funding guidance (updated 26 August 2026); Innovate UK strategy, 19 March 2026.
- [50a] Innovate UK Business Connect, Business Growth; Myriad Associates, 16 February 2026.
- [50b] AI Opportunities Action Plan delivery page, delivery.ai.gov.uk.
- [50c] Made Smarter national rollout announcement.
- [50d] Growth Hub Info; Gloucestershire County Council, March 2026.
- [50e] Help to Grow: Management, September 2026 cohort notices.
- [51] British Business Bank, Growth Guarantee Scheme (updated July 2026).
- [52] startuploans.co.uk, September 2026.
- [53] GOV.UK, Apply for the Enterprise Investment Scheme (updated 6 April 2026); The Carry, SEIS and EIS limits 2026.

BUILD ROUTES

- [54] bubble.io/pricing, September 2026.
- [55] eesel, Lovable pricing, 9 June 2026.
- [56] BuildMVPFast, Should I use no-code for my MVP (vendor estimate).
- [57] Superblocks, Lovable vulnerabilities (CVE-2025-48757), 3 September 2025.
- [59] Max Pog, Big Startup Studios Research 2023 (407 studios).
- [60] Sprintlaw UK, Sweat equity agreements, updated 2026.

AI RELIABILITY AND ENGINEERING

- [61] Vectara, Hallucination leaderboard, November 2025 and May 2026 updates.
- [62] Stanford RegLab, Hallucination-free? Assessing the reliability of leading AI legal research tools, May 2024.
- [63] OpenAI, Why language models hallucinate, 5 September 2025.
- [64] Conexiom, Data entry error rate benchmarks, 16 June 2026 (citing Panko).
- [64a] Barchard and Pace, Computers in Human Behavior, 2011.
- [65] Hamel Husain, "Your AI Product Needs Evals", 29 March 2024.
- [65a] Hamel Husain, "A Field Guide to Rapidly Improving AI Products", 24 March 2025.
- [66] Eugene Yan, Task-specific LLM evals, 31 March 2024.
- [67] OpenAI, A practical guide to building agents, 2025.
- [68] Anthropic, Building effective agents, 19 December 2024.
- [69] Anthropic, Model deprecations, 2026.
- [70] OpenAI, Deprecations, 2026.
- [71] Chen, Zaharia and Zou, "How is ChatGPT's behavior changing over time?", arXiv 2307.09009, July 2023.
- [72] Maxim AI, How to avoid LLM vendor lock-in, 3 September 2026.

REGULATION

- [73] GOV.UK, Check if you're eligible for Making Tax Digital for Income Tax, 2026; ICAS, 7 August 2025.
- [74] GOV.UK, Changes to UK company law campaign (identity verification; changes to accounts).
- [74a] CPS, Organisations must prepare now for new fraud prevention law, 18 August 2025.
- [75] GOV.UK, Transforming public procurement collection; Procurement Act 2023 Central Digital Platform factsheet.
- [76] FCA, Consumer Duty information for firms; FCA CP26/23, 29 June 2026.
- [77] Griffiths and Armour on the golden thread; RWK Goodman, 20 February 2026; Fire Warden Training compilation of Building Safety Regulator statistics to May 2026 (secondary).
- [78] Acas, Employment Rights Act 2025; GOV.UK unfair dismissal factsheet; VinciWorks timeline, updated 28 July 2026.

- [79] DLA Piper Privacy Matters, February 2026; SI 2026/82; ICO DUAA page; Clifford Chance, February 2026.
- [80] House of Commons Library CBP-10442, updated 8 July 2026; bills.parliament.uk/bills/4035 (status 8 September 2026).
- [81] GOV.UK, Terrorism (Protection of Premises) Act 2025 factsheets; SIA section 12 guidance consultation, April 2026; ProtectUK.
- [82] GOV.UK, Biodiversity net gain guidance; Defra blog, 20 April 2026.
- [83] GOV.UK collection, Awaab's Law in the social rented sector.
- [84] GOV.UK, Renters' Rights Act 2025 implementation roadmap.
- [85] Home Energy Model Guide, MEES EPC C 2030 confirmed, 14 February 2026 (secondary; check the DESNZ response).
- [86] uksrs.org.uk timeline; KPMG summary of FCA CP26/5, February 2026.
- [87] EU Council press release on simplification of sustainability reporting, 24 February 2026.
- [88] IAPP, King's Speech signals no AI bill, 14 May 2026; FCA AI Lab; summary of FCA position, 9 June 2026.

PROFESSIONAL BODIES AND INSURANCE

- [89] SRA, Separate business guidance, 8 March 2022.
- [89a] SRA, Warning notice on misuse of AI.
- [90] RICS, Responsible use of AI in surveying practice, mandatory from 9 March 2026.
- [91] Engineering Council and Royal Academy of Engineering, Statement of Ethical Principles, 2026 revision.
- [92] ICAEW, New PII regulations from 1 September 2024.
- [93] SimplifyStream, Professional indemnity for software companies UK (broker content).