Fictional exchange. Companies, people, operational inputs, interviews and reviews are invented. Updated 15 September 2026. Listed attachments are rendered pages in this PDF, not native downloadable files.

Recommendation: stop the autonomous Copilot and ship the controlled workflow

FROM IQ256 | iq256.com/contact

TO Mara Vellstrand, HalcyonTrellis Systems GmbH <mara.vellstrand@halcyontrellis.example>

DATE 18 September 2025, 16:18 CEST

Hi Mara,

Our recommendation is decisive: choose option D. Stop the autonomous product as currently scoped. In 90 days, ship Guided Resolution for the eight recurring exception types. AI may draft, but a person approves every external action.

The proposed workflow targets the eight types representing 73% of the fictional case volume, with a plan about a year shorter than internal model development and an expected budget about 14% of its first-year cash requirement. It is intended to create better context, outcome, paid-adoption and error evidence; those results have not been achieved in this sample.

Do not buy VerityRelay, sign SableQuarry exclusivity, or build a proprietary decision model now. The attachments contain the decision, strongest dissent, thresholds, and plan. IQ256

ATTACHMENTS (3)

- **IQ256_Product_AI_Decision_Review.pdf** — HalcyonTrellis Resolution Copilot | 18 September 2025 — 01
- **HalcyonTrellis_90_Day_Plan.pdf** — Guided Resolution delivery and decision gates — 02
- **Decision_Record_and_Thresholds.xlsx** — Decision ID HTS-PAI-2025-09 | Owner Mara Vellstrand — 03

Product and AI decision review

HalcyonTrellis Resolution Copilot | 18 September 2025

Fictional worked example: scenario inputs, calculations and proposed actions.

DECISION

Choose option D. Stop the autonomous Copilot. Build Guided Resolution.

Approve a 90-day program capped at EUR 220k. Launch guided resolution for the eight recurring exception types. AI may draft explanations and messages, but a user approves every external action. Do not buy VerityRelay, sign SableQuarry exclusivity, or fund a proprietary decision model now.

Why this is the least expensive route that solves the real problem

1. Customers consistently ask for faster resolution and control, not unattended AI.
2. Eight recurring types represent 73% of volume and already have explicit policies.
3. The difficult tail is where context and labels are weakest.
4. No AI route meets the agreed safety threshold.
5. Seven willing-to-pay customers do not finance a EUR 600k-EUR 1.35M bet.
6. The controlled route preserves build, buy, and partner options while creating better evidence.

BOARD RESOLUTION

Fund controlled assistance because it captures most near-term customer value, limits downside, and produces the evidence required for any later autonomy decision.

The hidden decision

The apparent question is which AI route to choose. The actual question is whether HalcyonTrellis should automate judgment before it has standardized the work and captured the necessary context. The answer is no.

Apparent evidence	Counter-evidence	Resolution
11 of 18 RFPs mention AI	Only one reviewed loss was mainly caused by missing AI	Market signal exists; revenue urgency is overstated
Customers want automation	10 prohibit unattended messages; 17 expect it in base price	Buyers prefer effort reduction with control
Prototype reaches 94% on common cases	It falls to 51% on the long tail	Result supports workflow automation, not open judgment
Owning the model creates strategic IP	Context is complete in 54%; labels are ambiguous in 31%	The asset is context, policy, action, and outcome data

What HalcyonTrellis should own

- Customer policy model and exception taxonomy.
- Evidence-completeness checks and approved workflows.
- Outcome, correction, and reversal history.
- The evaluation set used to measure any future model or vendor.

Keep language models replaceable. They can draft readable text, but should not own the policy or final action.

Why the other routes lose now

Route	Decision	Strongest advantage	Decisive problem
Build	No	Maximum control and potential differentiation	EUR 1.35M before broad paid demand; 14-18 months; weak labels
Buy	No	Fastest AI-branded feature	57% exact action and a 24-month minimum
Partner	No	Domain experience	Selected-case evidence, 8-12 months, and exclusivity
Guided	Yes	Fast, controlled, reversible	No autonomous-agent story at launch

Economic check

Route	Two-year or first-year commitment	Evidence needed to recover it
Build	About EUR 1.77M over two years	About 158 paid customer-years at EUR 15k and the 75% gross-margin floor
Buy	About EUR 609k first year	About 55 paid customer-years at the same margin, before sales cost
Partner	About EUR 830k plus revenue share	Unverified performance and lost flexibility
Guided	EUR 184k expected; EUR 220k cap	About 17 paid customer-years at the same margin

Recovery counts are annualized gross-profit-equivalent payback proxies using the supplied 75% company gross-margin floor. They exclude sales and fixed operating costs.

Product and safety boundary

Capability	Launch?	Reason
Detect a covered exception	Yes	High volume and explicit definitions
Check required context	Yes	Missing context is the largest failure source
Show facts and policy	Yes	Needed for trust, audit, and correction
Recommend an in-policy action	Yes	Policy controls the action
Draft an explanation or email	Optional	Useful and reviewable
Send an external message	Only after approval	Current systems fail the severe-error threshold
Judge a long-tail case	No	Route it to a person and record the outcome

Technical release thresholds

- At least 98.5% exact recommended-action accuracy within covered workflows.
- Below 0.1% severe incorrect actions before unattended execution is reconsidered.
- At least 95% of generated explanations supported by displayed source fields.
- Required-context detection above 99% on the frozen evaluation set.
- No material performance gap above 3 points among pilot customers without a known cause.

Final commercial-release gates

- Median resolution time is 2.0 business days or less across every eligible covered case in the frozen release window.
- Manual touches are 1.8 or less across the same covered-case denominator.
- Paid use reaches at least 15% of all current customers eligible for the eight covered workflows; freeze eligibility before invitations. Or measured retention improves by at least 2 percentage points against the frozen baseline.
- Severe incorrect actions remain below 0.1% before any unattended external action.
- Company gross margin remains at or above 75% under Finance's standard definition.
- Committed ERP integrations are delayed by no more than six weeks.

Strongest dissent

> Simulated voice: Choose VerityRelay now. Competitors are shaping the category. Approval controls can contain harm while HalcyonTrellis gains an AI story in one quarter.

The objection is credible but does not win. VerityRelay returned the exact action in only 57% of HalcyonTrellis's blind cases, caused 2.6% severe errors, and requires a 24-month minimum. With mandatory approval, it becomes an expensive drafting assistant.

Interim continuation risks

The day-45 and day-90 measures below decide whether to keep testing. They do not replace the final commercial-release gates.

Risk	Early signal	Response
Buyers dismiss guided workflow	Three losses name missing autonomy as primary cause	Reopen a bounded buy evaluation
Review work does not fall	Touches remain above 2.5 at the interim review	Redesign approval flow before adding coverage
Pilots do not pay	Fewer than five commitments by day 45	Keep only retention-justified workflows
Policies resist encoding	More than 20% need interpretation	Narrow scope and improve policy capture
ERP roadmap slips	Forecast delay above six weeks	Reduce pilot scope

What would change the answer

- Reopen Build only with 25 paying customers, 80% context completeness, 25,000 reviewed outcomes, an 8-point proprietary advantage, and payback below 24 months.
- Reopen Buy only if a vendor reaches 90% exact action accuracy, costs below EUR 250k annually, and signs for 12 months or less.
- Reopen Partner only without exclusivity, with customer workflows owned by HalcyonTrellis, fixed external cost below EUR 400k, and at least 10 paid pilots.
- Pause the evidence program at day 90 if fewer than five customers are active, resolution time improves under 20%, or ERP delay exceeds six weeks. These are interim continuation gates, not commercial-release approval.

FINAL CALL

Approve Guided Resolution, the EUR 220k cap, the no-auto-send boundary, and the 90-day evidence plan.

Delivery record and remaining work

This fictional assignment is complete when the agreed review and documents are delivered and checked. Its business recommendations and future execution are separate from that completion.

Agreement	Boundary in this example
Delegated work	Compare build, buy, partner and guided workflow using supplied scenario data; recommend a route and prepare the budget, release gates and 90-day plan.
Work illustrated as delivered	A product review, strongest dissent, decision record and delivery plan. The fictional benchmarks and interviews were not conducted, and no Guided Resolution software was shipped.
Completion checks	Reconcile the EUR 184k plan/EUR 220k cap; distinguish interim continuation from final release gates; trace assumptions, payback proxies and reversal thresholds.
Authority retained by customer	Mara and the board retain investment and launch decisions. In this case, users approve external actions; IQ256 is not authorized to deploy, email suppliers or fund the program.
Remaining checks	Real customer/data validation, implementation, representative tests, paid-use/retention measurement, security review and production release approval.
Fee and timing	Agree the review scope, price and deadline before paid work. Implementation could be a separately scoped assignment; it is not included merely because the plan recommends building software.
Business budget	EUR 184k expected and EUR 220k cap are the fictional customer program budget, not the IQ256 review fee or a software implementation quote.

Where human input is agreed, IQ256 selects, briefs, coordinates and pays contributors. IQ256 integrates the agreed work and performs baseline completion checks. Extra verification answers a specified additional concern. External information packets require explicit customer approval; no approval is needed for every internal step within the agreed mandate.

If scope, price, deadline, sharing or customer effort changes materially, agree the next course before exceeding the mandate. A recommendation may name one option while the customer retains the decision; permission to act or commit money is agreed separately.

Discuss a defined assignment

https://iq256.com/contact

Start with a short, non-confidential description. Initial human reply normally within 2-48 hours, distinct from the agreed delivery deadline.

90-day plan

Guided Resolution delivery and decision gates

Fictional worked example: scenario inputs, calculations and proposed actions.

Period	Work	Decision gate
Days 0-15	Confirm eight types, sign pilots, freeze baselines, define context and policy	Continue with at least five paid pilots and no ERP delay above six weeks
Days 16-45	Build checks, policy engine, evidence display, approval, audit, and optional drafting	Release only at 98.5% exact action accuracy in scope
Days 46-70	Deploy to two design partners, then five additional pilots	Narrow any workflow above the severe-error limit
Days 71-90	Measure time, touches, conversion, burden, and customer variation	Continue evidence work only with five active users and at least 20% faster resolution

These are interim continuation gates. Commercial release still requires every final gate in the decision review, including 2.0-day resolution, 1.8 touches, paid-use or retention evidence, and company gross margin of at least 75%.

Delivery budget

Item	Budget
Product lead	EUR 34k
Engineering and platform	EUR 76k
Design and research	EUR 13k
Data review and QA	EUR 24k
Privacy, security, and legal	EUR 17k
Infrastructure and models	EUR 5k
Contingency	EUR 15k
Expected / board cap	EUR 184k / EUR 220k

Decision record and thresholds

Decision ID HTS-PAI-2025-09 | Owner Mara Vellstrand

Fictional worked example: scenario inputs, calculations and proposed actions.

Assumption	Confidence	If false	Next test
Eight types are about 73% of volume	High	Coverage falls	Verify each pilot before configuration
Buyers value assistance without autonomy	Medium	Commercial benefit weakens	Convert five pilots by day 45
Accuracy can rise from 97.4% to 98.5%	Medium	Release must narrow	Frozen blind test at day 45
EUR 15k average price works	Low-medium	Direct payback weakens	Test signed pilot terms

Residual uncertainty accepted

The controlled route may create less marketing impact than an AI-agent launch. The board accepts that risk because it reaches users sooner, costs materially less, and creates stronger evidence for the next decision.