



AI for logistics and transportation

Using AI for planning, documentation, exceptions,
customer updates and network learning.

WORKS WITH THE ASSISTANTS TEAMS ALREADY USE



EDITION 44 / EMBER & CLAY

EXECUTIVE WORKING EDITION - 08 OCT 2026

44

AT A GLANCE

Using AI for planning, documentation, exceptions, customer updates and network learning.



Logistics is a business of handoffs, documents and exceptions. Freight moves on paper and data created by many parties, and most of a dispatcher's or customer service rep's day goes to chasing, reconciling and explaining. AI can take on much of that reading, matching, drafting and summarizing, and can help planners spot patterns in network performance. It does not replace the operational judgment about a load, a driver or a customer, and it cannot fix poor master data. This playbook covers where AI fits in dispatch support, shipment documents, exception handling, customer updates and planning, how to run it, what to measure and how to avoid the common failures.

WHO IT IS FOR

Operations, dispatch, customer service, planning and technology leaders at carriers, brokers, forwarders, 3PLs and shipper logistics teams who are deciding where AI helps and how to run it.



TIME TO USE

70 minutes to read; two workshops to apply

What you will have at the end

- A ranked list of logistics workflows by fit, value and risk
- A map of which decisions stay with dispatchers, planners and customer owners
- A document and exception test set built from real shipments
- A metric set that links time saved to service, cost and claims
- A 90-day plan from one lane or customer to a decision to scale



Contents

01	What AI changes in logistics operations	03
02	Where AI fits in logistics, and where it does not	04
03	How it works, in plain language	06
04	Operating model, roles and decision rights	08
05	Risks, controls and human review	10
06	Measuring quality and value	12
07	Implementation guide	13
08	Pitfalls and a hypothetical scenario	15
09	Questions to ask before you buy or build	16
10	A first quarter that produces evidence	18
11	Terms used in this publication	19



EXECUTIVE BRIEF

What AI changes in logistics operations

01

II

AI helps most where logistics teams read, match, chase and explain, and helps least where the problem is missing data or a physical constraint.

A single shipment can generate a rate confirmation, a booking, a bill of lading, a commercial invoice, a packing list, proof of delivery, accessorial requests, appointment changes and a string of status emails. The documents arrive in different formats from shippers, carriers, warehouses, brokers and customs agents. Staff spend large parts of their day opening attachments, typing values into a transport management system, checking one document against another and writing the same type of update to a customer. That work is necessary and repetitive, which makes it a good fit for AI that reads, extracts and drafts.

The second area is exceptions. Late pickups, missed appointments, damaged freight, detention, weather and capacity shortfalls create a stream of decisions that need context from several systems and conversations. AI can assemble that context, propose options and draft the message, so the human responsible for the load decides faster. The third area is learning from the network: summarizing why lanes underperform, which customers generate the most exceptions and where dwell time accumulates.

The limits are equally concrete. A model does not know that a particular dock refuses early arrivals unless that fact is in your data. It cannot create capacity. It should not commit your company to a rate, a delivery promise or a liability position without a person. And if location, weight or address data is unreliable, AI will produce tidy summaries of unreliable facts.

- Start with documents and status communication, where volume is high and a person can check the result quickly
- Treat exception handling as decision support: the dispatcher or account owner still decides
- Use planning and optimization tools for routing and load building; use language AI for reading, writing and explaining
- Fix master data such as locations, appointment rules and customer instructions, because every use case depends on them
- Measure on-time performance, claims and cost per shipment alongside time saved



PLAIN TEST

If a new hire could do the task after reading the documents and following a written procedure, AI can probably assist. If the task depends on a relationship, a physical inspection or an unwritten site rule, assist only with preparation.

WORKING QUESTION

Which three document or message types does your team handle most often, and how many minutes does each take today?



WHERE IT FITS

Where AI fits in logistics, and where it does not

02



Match the technique to the problem: language AI for documents and messages, optimization for routes and loads, statistics for forecasts.

Logistics already uses mature mathematical tools. Routing engines, load planners and network optimizers solve constrained problems that language models are poor at. A language model asked to plan a multi-stop route may produce a confident answer that breaks a driver-hours limit or a time window. Keep the distinction clear: use dedicated optimization for planning math, and use language AI for the unstructured work around it, such as reading customer instructions, explaining a plan to a customer and summarizing why a route changed.



USE CASE	FIT	REASON AND CONDITION
Extracting fields from bills of lading, PODs and invoices	Strong	High volume, checkable against source; confirm before posting to system

USE CASE	FIT	REASON AND CONDITION
Matching documents to shipments and flagging mismatches	Strong	Reduces back-office reconciliation; exceptions go to a person
Drafting customer status updates from tracking and notes	Strong with review	Saves time; facts must come from system data, not guesses
Exception summaries and options for dispatchers	Strong	Pulls context together; dispatcher decides
Quote support from email requests	Moderate	Speeds data entry; pricing authority stays with sales or pricing
Route and load optimization	Use dedicated optimization	Constraints require solvers, not text models
Demand and capacity forecasting	Statistical methods, with AI for explanation	Forecast accuracy depends on data history and discipline
Claims and damage documentation	Moderate	Helps organize evidence; liability decisions stay with claims staff
Customs classification and compliance decisions	Cautious	Assist retrieval; qualified staff or brokers decide

Where it does not fit

- Making binding commitments on rates, delivery times or liability without human approval
- Safety-critical decisions on driver hours, vehicle condition or hazardous goods handling
- Replacing planning math with a chat interface
- Operating on addresses and location records that nobody maintains
- Handling customer disputes where the relationship needs a person

COMMON MISTAKE

Buying a customer-facing chatbot to answer where-is-my-shipment questions before your tracking data is reliable. If the data is wrong or late, the bot gives wrong answers faster.

WORKING QUESTION

For each planned use, is the core problem reading and writing, or constraint solving, and are you using the right tool for it?

HOW IT WORKS

How it works, in plain language

03



Read the document, check it against what the system knows, act on the clear cases, and route the unclear ones to a person with the evidence attached.

Document processing starts with turning an image or PDF into text, then identifying the document type and extracting values such as shipper, consignee, piece count, weight, reference numbers, charges and dates. A language model handles layouts it has not seen before better than template-based tools, but it can still misread a digit or merge two columns. Each extracted value should be compared with expected values from the order or booking. A mismatch, such as a weight 18 percent over booking, is a flag for a person to review rather than a value to quietly accept.



01 Ingest

Collect documents and emails from the mailbox, portal, EDI feed or driver app and classify them by type.

02 Extract and validate

Pull the fields and check them against the shipment record, reference formats and tolerance rules.



03 Decide the path

Clean matches post automatically within agreed limits. Mismatches and low-confidence values go to a work queue with the source image beside the extracted value.

04 Draft and notify

For status and exception messages, assemble facts from tracking, notes and appointment data, then draft the update for review or for pre-approved templates.

05 Learn

Store corrections and outcomes so the tests, rules and templates can be improved.

Exception support

For an exception such as a missed delivery appointment, the useful output is not a long narrative. It is a short card: what happened, which shipments and customers are affected, what the customer instructions say, what the options are, what each option costs in time and money, and a draft message. The dispatcher selects an option and edits the message. The AI is doing the gathering and first drafting, which are the slow parts, and the person keeps the judgment.

MODE	WHAT THE AI DOES	APPROPRIATE FOR
Assist	Extracts and drafts; person posts and sends	Rate-related messages, disputes, new customers
Auto with limits	Posts clean matches and sends template updates within rules	Routine PODs, standard milestones, low-value shipments
Escalate	Flags and routes with evidence	Mismatches, safety, high-value or sensitive freight

WORKING QUESTION

For your most common document, what tolerance would you accept for an automatic post, and what would send it to a person?

OPERATING MODEL

Operating model, roles and decision rights

04

Give each workflow an owner on the operations side, a clear list of what the system may do alone, and a feedback loop from the people who use it.

Logistics teams are used to service levels, escalation paths and handoffs between shifts. The same ideas apply to AI. A document queue has an owner, targets and a backlog. An AI-drafted customer message has an approver or an approved template. A change to a template or rule is made by a named person after a test. This keeps the work visible and avoids tools that nobody feels responsible for.

ROLE	RESPONSIBILITY	TYPICAL OWNER
Process owner	Sets scope, targets and rules for workflow	Operations or customer service manager
Dispatcher or account owner	Decides on exceptions, commitments and customer relationship matters	Dispatch, account management
Queue reviewer	Resolves flagged documents and records corrections	Back office, billing, customs staff
Master data owner	Maintains locations, appointment rules, customer instructions, rate references	Operations excellence, IT
AI product owner	Manages releases, templates, test sets and metrics	Digital or continuous improvement lead
Security and privacy owner	Controls access to shipment, driver and customer data	IT security, data protection
Customer success	Collects customer feedback on automated communication	Account management

Decision rights

- AI may post a document to a shipment only when values match the booking within written tolerances
- AI may send standard milestone updates from templates; any message that changes a promise, assigns blame or discusses cost needs a person
- Rates, accessorial approvals, claim settlements and credit decisions stay with authorized staff
- Safety, hours-of-service and hazardous goods decisions follow existing compliance procedures and are not delegated to a language model
- Any person may switch an automated flow to manual and report why



Shift handover is a hidden strength. Exception summaries that record what was tried, who was contacted and what was promised make it easier for the next shift to carry on without a call. Build that into the template from the start, and require people to confirm that the summary is accurate before the shift ends.



DRIVER AND STAFF DATA

Driver location, hours and performance data is personal. Decide in advance what is collected, who can see it and whether it feeds any performance review. Tell staff and drivers plainly, and check current employment and privacy requirements.

WORKING QUESTION

Who in your organization can approve a change to an automated customer message, and how quickly can they roll it back?



RISKS AND CONTROLS

Risks, controls and human review

05

The main risks are quiet data errors, overconfident customer messages, security exposure of commercial data and over-automation of commitments.

RISK	WHAT IT LOOKS LIKE	CONTROL
Misread value	Wrong weight, reference or charge posted to a shipment	Validate against booking; confidence thresholds; sample audit
Wrong promise to customer	Draft states a delivery time that the system does not support	Draft only from verified fields; block commitments language; human approval for changes
Duplicate or fraudulent documents	Altered invoice or payment redirection request	Verify against known contacts; require separate approval for bank detail changes
Data leakage	Rates or customer lists exposed to wrong users or tools	Role-based access, approved tools only, vendor terms on data use
Prompt injection through email	A message contains hidden instructions that steer the system	Treat inbound text as data; limit what the system can do; no autonomous payments or rate changes
Stale instructions	Customer site rules out of date	Owner and review date for each instruction
Customs or regulatory error	Misclassification or missing documentation	Qualified review; keep authority with licensed staff; check current requirements

Freight is attractive to fraudsters, and email is the common entry point. Changed bank details, fake carriers and doctored delivery proofs are known patterns. An AI that reads and acts on inbound email widens the surface, so apply a principle of least authority: the system reads and drafts, but cannot pay, change payment details, alter rates or release freight. Those actions need separate, authenticated approval by staff.

Minimum controls for automated posting

- Tolerance rules written and approved by the process owner
- Source image visible beside every extracted value
- Daily random sample reviewed by a second person
- Changes to bank details, rates and credit never executed by the system alone
- Audit log of inputs, outputs, corrections and who approved
- Weekly error review with the master data owner
- Documented fallback to manual processing

REGULATED FREIGHT

Hazardous goods, temperature-controlled products, customs filings and cabotage rules carry legal duties. Use AI to retrieve and organize requirements, but keep qualified people responsible and check current requirements in each jurisdiction. This playbook is not legal advice.

WORKING QUESTION

Which action in your workflow, if performed wrongly by an automated system, would cost you the most, and what stops it from happening without a person?

MEASURING VALUE

Measuring quality and value

06

Tie every efficiency measure to a service or cost measure, and capture a baseline before anything changes.

Logistics already has strong operational metrics: on-time pickup and delivery, claims ratio, cost per shipment, dwell time and days sales outstanding. Use them. Add a few that are specific to AI-assisted work, such as extraction accuracy, share of documents processed without touch, correction rate and message approval rate. Capture at least a month of baseline, preferably spanning a typical peak and a quiet week, because logistics volume is uneven and a month with a promotion or a storm can mislead.

METRIC	DEFINITION	WHY IT MATTERS
Touch time per document	Observed minutes from open to posted	Real labor effect
Straight-through rate	Share posted without human edit within tolerance	Shows automation extent
Field accuracy	Correct fields over audited fields, by field type	Flags weak spots such as weights or dates
Exception response time	Detection to customer notified	Service quality
Billing cycle time	Delivery to invoice sent	Cash impact
Claims and disputes	Count and value per thousand shipments	Shows hidden errors
Customer message edits	Share of drafts changed by staff	Template quality signal
On-time performance	Pickup and delivery within window	Ensures service did not slip

Illustrative example

These numbers are illustrative only. Suppose a billing team processes 6,000 proofs of delivery a month at an average of 3 minutes each, or 300 hours. If assisted extraction cut average touch time to 1.5 minutes while field accuracy on invoice-relevant fields remained at the

baseline level, the team would recover roughly 150 hours. If the faster process raised billing disputes, the saving would partly vanish. The value case should include the cost of review, licenses and maintenance, and should say what the recovered hours will be used for, such as faster invoicing or reduced overtime.

- Sample audited files evenly across customers, lanes and document types
- Separate the effect of AI from other changes such as a new carrier mix or a new TMS release
- Record override and correction reasons to guide fixes
- Report ranges and sample sizes, not single numbers
- Review value and quality monthly for the first quarter, then quarterly


WORKING QUESTION

Which existing metric would tell you first that automation was damaging service, and who watches it weekly?

II

IMPLEMENTATION

Implementation guide

07

K

Start with one customer, lane or document type, prove accuracy on real shipments, then widen.

01 Pick the first slice

Choose a document type and a customer or lane with enough volume to measure. Proof of delivery processing for one large customer is a common, low-risk start.

02 Clean the reference data

Check customer instructions, location records, reference formats and tolerance rules. Assign an owner for each and a review date.



03 Build the test set

Collect 200 to 300 real documents and messages, including poor scans, handwritten notes and odd formats. Have experienced staff mark the correct values.

04 Run in shadow

Let the system process documents alongside staff for two to four weeks. Compare outputs and examine every difference.

05 Release with limits

Allow automatic posting for clean matches within tolerance, and queue the rest. Review a daily sample.

06 Add exception and message support

After documents are stable, introduce exception cards and draft customer updates for one account team.

Integration and data

Most gains depend on how well the tool connects to your TMS, WMS, tracking feeds and mailboxes. Prefer read access first, then controlled write access for specific fields. Use unique identifiers, such as shipment and reference numbers, to match documents, and treat free-text matching as a fallback with lower confidence. Keep a plain copy of the original document with each posted record to support audits and disputes.

Readiness checklist

- Named process owner and queue reviewers
- Written tolerance and escalation rules
- Test set built from real, messy documents
- Master data owner assigned for locations and instructions
- Access controls and vendor data terms reviewed
- Drivers and staff informed of what data is used
- Manual fallback tested during a peak shift
- Weekly review meeting scheduled for the first 90 days

PEAK SEASON

Do not launch major changes just before peak. Start well ahead, run in parallel through at least one busy week, and keep the manual path ready.

WORKING QUESTION

What is the smallest slice of work, by customer and document type, that would give you a trustworthy result in four weeks?


PITFALLS

Pitfalls and a hypothetical scenario

08



Failures usually trace to bad reference data, unchecked automation and a missing owner for the exception queue.

- Automating before cleaning location and customer instruction data
- Setting tolerances too loose to avoid queue work, then absorbing errors in billing disputes
- Letting draft customer messages promise things the system cannot support
- Ignoring night and weekend shifts, where review capacity is thinnest
- Buying a point solution for each document type without a shared review and logging approach
- Counting hours saved without checking claims, disputes and on-time performance
- Leaving prompt and template changes uncontrolled


HYPOTHETICAL SCENARIO (ILLUSTRATIVE, NOT A REAL CASE)

A regional 3PL automates delivery status emails for one retail customer. Drafts use the latest scan time as the delivery time. A depot scans pallets late at night in batches, so customers receive emails showing deliveries hours after they occurred. The fix is to use driver proof-of-delivery time where available, label scan-based times as such, and add a weekly check comparing email times with signed POD times.



The lesson is that the model did exactly what the data told it. The failure was in the definition of delivery time, which no one had agreed. AI tends to expose such ambiguity in field definitions

that human staff quietly work around. Treat the discovery as a benefit and write down the definitions: what counts as delivered, late, damaged and complete, and which source wins when two disagree.



WARNING SIGN	LIKELY CAUSE	FIRST RESPONSE
Rising billing disputes	Loose tolerances or misread charges	Tighten tolerances, audit charge fields
Customers reply to correct automated messages	Wrong time or status definitions	Review field definitions and sources
Queue growing weekly	Too many low-confidence cases understaffed review	Inspect causes, adjust thresholds or staffing
Staff retype extracted data	Distrust or poor interface	Interview users, fix interface and accuracy



WORKING QUESTION

Where do two of your systems disagree about the same fact, and which one should the AI believe?

BUYING QUESTIONS

Questions to ask before you buy or build

09

Test products on your documents and your lanes, and ask how they behave when they are wrong.

Questions for any supplier

- Will you process a sample of our real documents, including poor-quality ones, and share field-level accuracy?
- How do values reach our TMS, and can a person see the source next to each value before posting?



- How does the product handle mismatches against bookings, and how are tolerances configured?
- What controls prevent the system from changing payment details, rates or credit decisions?
- How do you defend against instructions hidden in inbound emails and attachments?
- Where is data stored, who can access it, and is it used to train models for other customers?
- How are model changes announced, and can we pin versions and re-test?
- What logs and exports do we get for audits, disputes and customer inquiries?
- What is the full cost, including integration, review time, volume tiers and support?

OPTION	FITS WHEN	WATCH FOR
Capability inside your TMS or visibility platform	Your vendor already holds the data	Roadmap dependence, limited customization
Specialist document automation	Document volume is the main pain	Integration depth, handling of rare formats
General AI platform you can configure	Many varied workflows and internal skills	You own testing, security and maintenance
Custom build	Distinctive process and engineering team	Long-term support burden

Keep the assets that matter in your hands: your test set, your tolerance rules, your reference data and your logs. They make it possible to compare suppliers fairly and to switch without starting again. Halden & Company helps logistics teams choose the first slice, build the test set and design the review and control model. Free assessments are available at halden.company/assessments, and the team can be reached at halden.company/contact.

WORKING QUESTION

If your document automation vendor were unavailable for a week, how would you keep freight moving and billing on time?

30 / 60 / 90

A first quarter that produces evidence

10

Prove accuracy on one slice of real shipments, add controls while volume grows, then decide on scale using service and cost evidence.

First 30 days: ground it

- Choose one customer, lane or document type as the first slice
- Capture a baseline for touch time, accuracy, disputes and on-time performance
- Assign owners for customer instructions, location data and tolerance rules
- Build a 200 to 300 document test set with correct values from experienced staff
- Review data access, vendor terms and fallback to manual

Days 31 to 60: test it

- Run shadow mode and examine every difference from staff results
- Agree field definitions such as delivered, late and complete
- Release limited automatic posting with a daily second-person sample
- Introduce exception cards for one team
- Train reviewers and set weekly error reviews

Days 61 to 90: decide

- Compare results to baseline on cost, service, claims and disputes
- Decide to scale, change or stop, and record the reasons
- Add draft customer updates for one account team under template control
- Document stop conditions and rollback steps
- Choose the next slice using the same test and review pattern

GLOSSARY

Terms used in this publication

11



TERM	MEANING
Bill of lading	The document recording that goods were received for shipment, and the terms of carriage.
Proof of delivery	A record, often signed, that freight reached its destination, used for billing and disputes.
TMS	Transportation management system, used to plan, execute and settle freight movements.
Straight-through rate	The share of items processed with no human edit.
Tolerance	The allowed difference between an extracted value and the expected value before a person must review.
Accessorial	An extra charge beyond base freight, such as detention, liftgate or redelivery.
Dwell time	Time a vehicle or shipment waits at a facility.
Exception	Any shipment event that departs from plan, such as delay, damage or refusal.
Master data	Reference records such as locations, customers and rate tables that other processes rely on.
Prompt injection	Hidden or hostile instructions inside content that a language model reads, intended to change its behavior.
Shadow mode	Running AI beside normal work without affecting outcomes, so results can be compared.
Least authority	Giving a system only the permissions it needs, so mistakes cannot do wide harm.



A note on use

This playbook is general guidance, not legal advice. Transportation, customs and safety requirements vary by jurisdiction and change over time; check current requirements.



TAKE THE NEXT STEP

Take the free Halden assessment at halden.company/assessments, or talk to us at halden.company/contact about turning this into a plan for your organization.

