

An outside AI agent matched Mercury's 1099 page to the cent. Mercury's agent tools still leave out three things the job needs.

THE JOB	Send a 1099 to each contractor paid \$2,000+ in 2026; each needs a W-9.
THE TEST	Can an outside AI agent answer "Who needs a 1099, and what's missing?" through Mercury's CLI and MCP?
WHY THIS JOB	A hard deadline, and Mercury's 1099 page grades every answer.
WHAT I FOUND	Totals matched to the cent. Agents can't see payee IDs in MCP's payment list, a W-9's type or status, or every 1099-page payee.
STILL OPEN	Do customers feel this? Five conversations next. Production may differ from the sandbox.

Crispin Gathuru

Built with Claude Code in Mercury's sandbox (fake company, fake money, read-only)

Independent; not endorsed by Mercury

Pages 1–9: one finding each

A1–A10: evidence

Mercury's surfaces don't agree on what an agent can see. A parity test on one job shows where.

	Dashboard	API	CLI	MCP text	MCP structured
Payee ID, payment list	○	✓	✓	✗	✗
Payee ID, one payment	○	○	○	✓	✗
W-9 file	✓	✓	✓	✓	✗
W-9 form type	○ label only	✗	✗	✗	✗
W-9 status	✓	✗	✗	✗	✗
Every 1099-page name	✓	✗	✗	✗	✗
Dashboard: 3 of 6 · API: 2 of 6 · CLI: 2 of 6 · MCP text: 2 of 6 · MCP structured: 0 of 6					
✓ yes ✗ no ○ untested					
SANDBOX DATA					

- 01 Put the payee ID in MCP's payment list
- 02 Return each W-9's form type and status
- 03 List every 1099-page payee for agents

✓ **5 of 5** payees with payments match to the cent; the rest match at \$0.1 more with payments isn't compared (p. 3)

✗ **0/100** MCP list rows carry a payee ID

Appendix A1 (p. 10)

A January tax deadline is a fair test of outside agents on Mercury

WHAT

Who was paid \$2,000+ in 2026; which W-9s are missing?

WHY

Mercury's 1099 page grades every answer to the cent.

WHO

Businesses paying contractors, their bookkeepers, and their agents.

WHEN

By Feb 1, 2027; W-9s come first.

WHERE

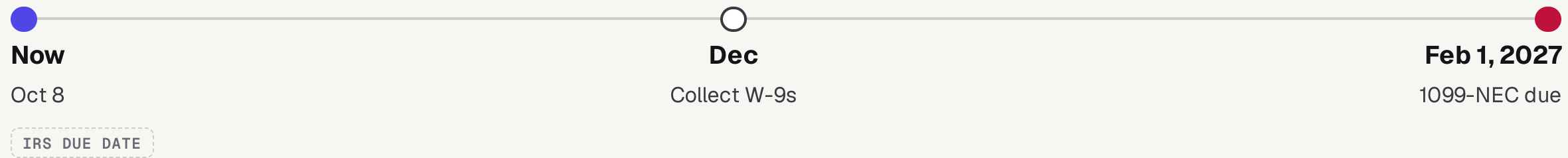
Outside agents via MCP, CLI and API. Command untested.

HOW

I built a read-only agent tool with Claude Code.

HOW MUCH

Unknown until customer interviews (K5).



Appendix A2 (p. 11)

Through the CLI, an agent's totals match Mercury's 1099 page to the cent

4
need action

1

2

3

4

10
W-9 problems (test files)

79
under \$2,000

✓ 5 of 5 payees with payments match to the cent; the rest match at \$0. Not compared: Currency Cloud Recipient (\$653); it isn't on Mercury's 1099 list.

✓ yes ✗ no ○ untested

SANDBOX DATA

1

Alex Rivera

Needs a 1099 · \$47,500

✓

2

Wire Person

Needs a valid W-9 · \$24,005

✓

3

IRS Tax Payment Recipient

Needs review · \$5,000

✓

4

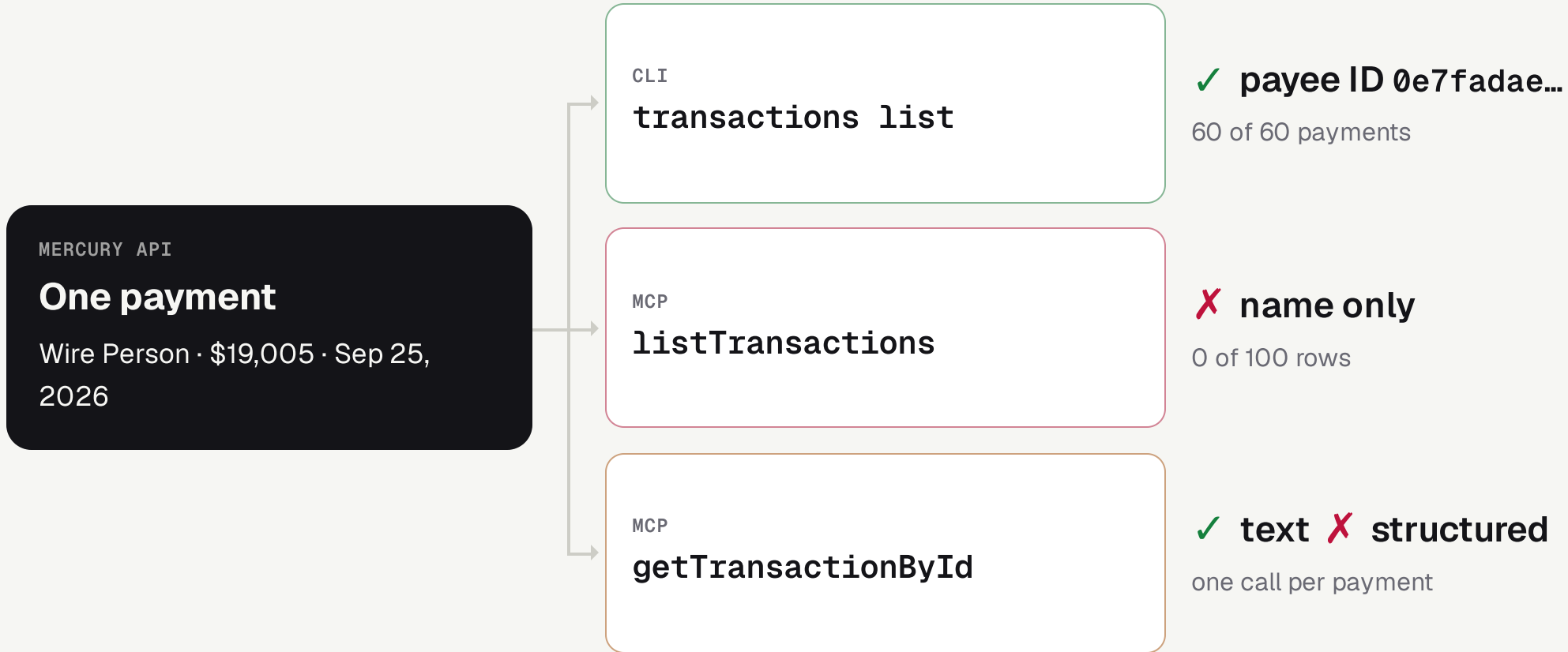
Patient Payee

Needs review · \$5,000

✓

Appendix A3 (p. 12) · A4 (p. 13)

The same payment reaches an agent with its payee ID through the CLI, and without one through MCP's payment list



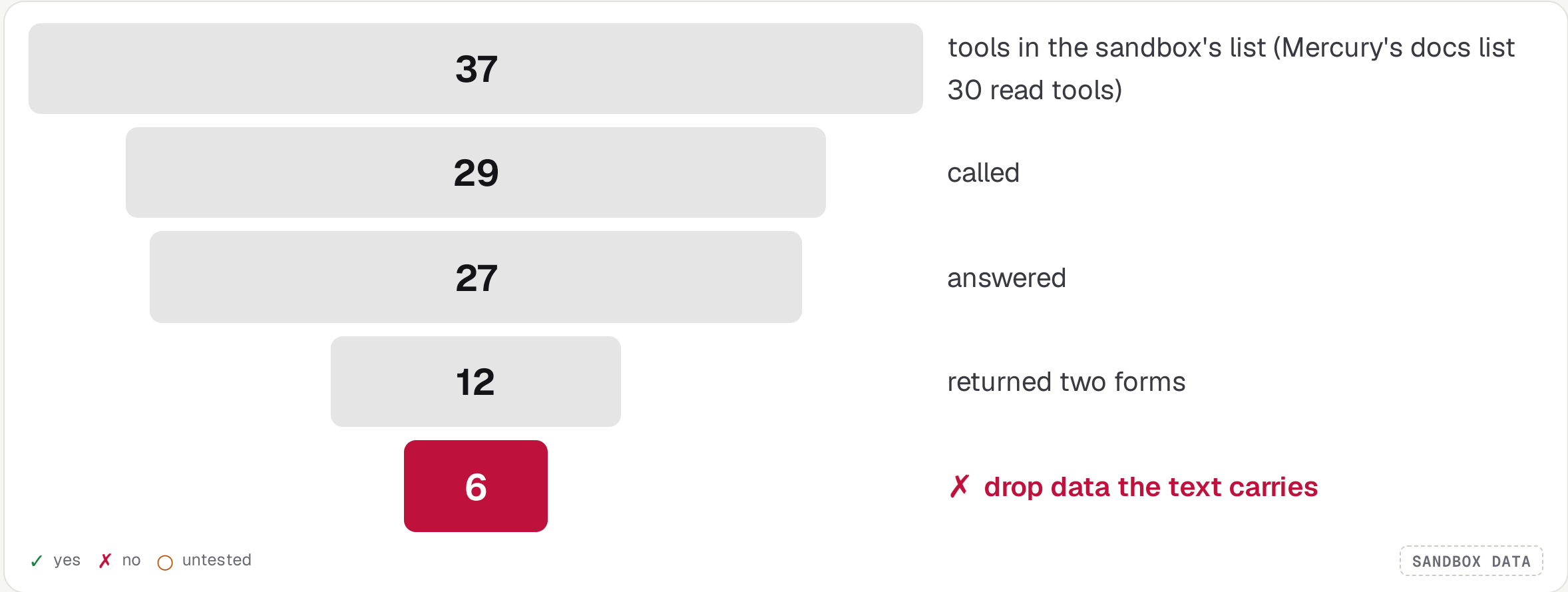
✓ yes ✗ no ○ untested

SANDBOX DATA

3 payees here are named “Banned Recipient”: a name join merges them.

[Appendix A5 \(p. 14\)](#)

In 6 of Mercury's MCP tools, the structured answer drops data the text carries, including the payment list



The six: getAccounts, listCards, getRecipients, listTransactions, getAccount, getTransactionById. **Dropped:** payee ID, payee files, next-page cursor.

Appendix A6 (p. 15)

A person can see when a W-9 fails. In the sandbox, an agent can't.

In the sandbox, no reading reached the API, CLI or MCP.

What a person sees

Details

W-9 Status

Processing Error

Email

—

Address

1 Test Street, Portland, OR 97201

Paid in 2026

\$0.00

View Payments >

W-9

h2h-riley-park.fillable.pdf

> Parsed tax details ⓘ

Mercury's dashboard

What an agent gets

id

2002c8f8...

recipientId

786f10c8...

fileName

h2h-riley-park.fillable.pdf

url

<temporary download link>

uploadedAt

2026-10-02T18:50:23.672277Z

X

No form type

X

No W-9 status

✓ yes X no ○ untested

SANDBOX DATA

Appendix A7 (p. 16)

2 of 5 pre-set tests pass; the two that decide whether this matters are still open

K1**Decisive**

Mercury's
reader vs mine

☐ **Open**

Not answerable in the
sandbox: all 11
uploads failed.

K2

My W-9 reader

✓ **Pass**
(synthetic files only)

25 of 25 typed W-9s
exact

K3

Registry
name match

☐ **Open**

Not built yet

K4

Totals match
Mercury's

✓ **Pass**

5 of 5 payees with
payments match to
the cent; the rest
match at \$0.1 more
with payments isn't
compared (A8)

K5**Decisive**

A business tries it

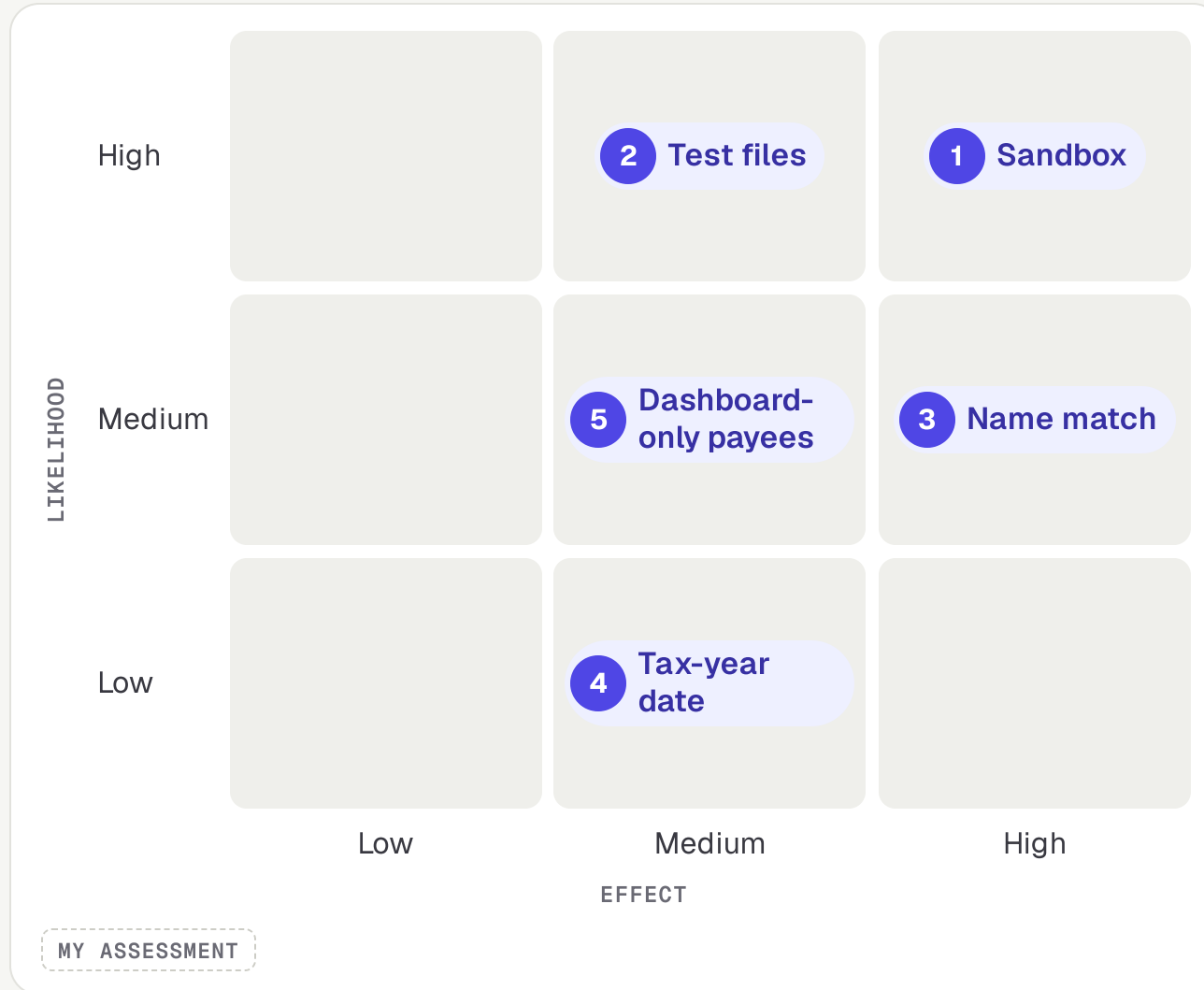
☐ **Open**

Not started: five
conversations
planned

SANDBOX AND TEST DATA

[Appendix A8 \(p. 17\)](#)

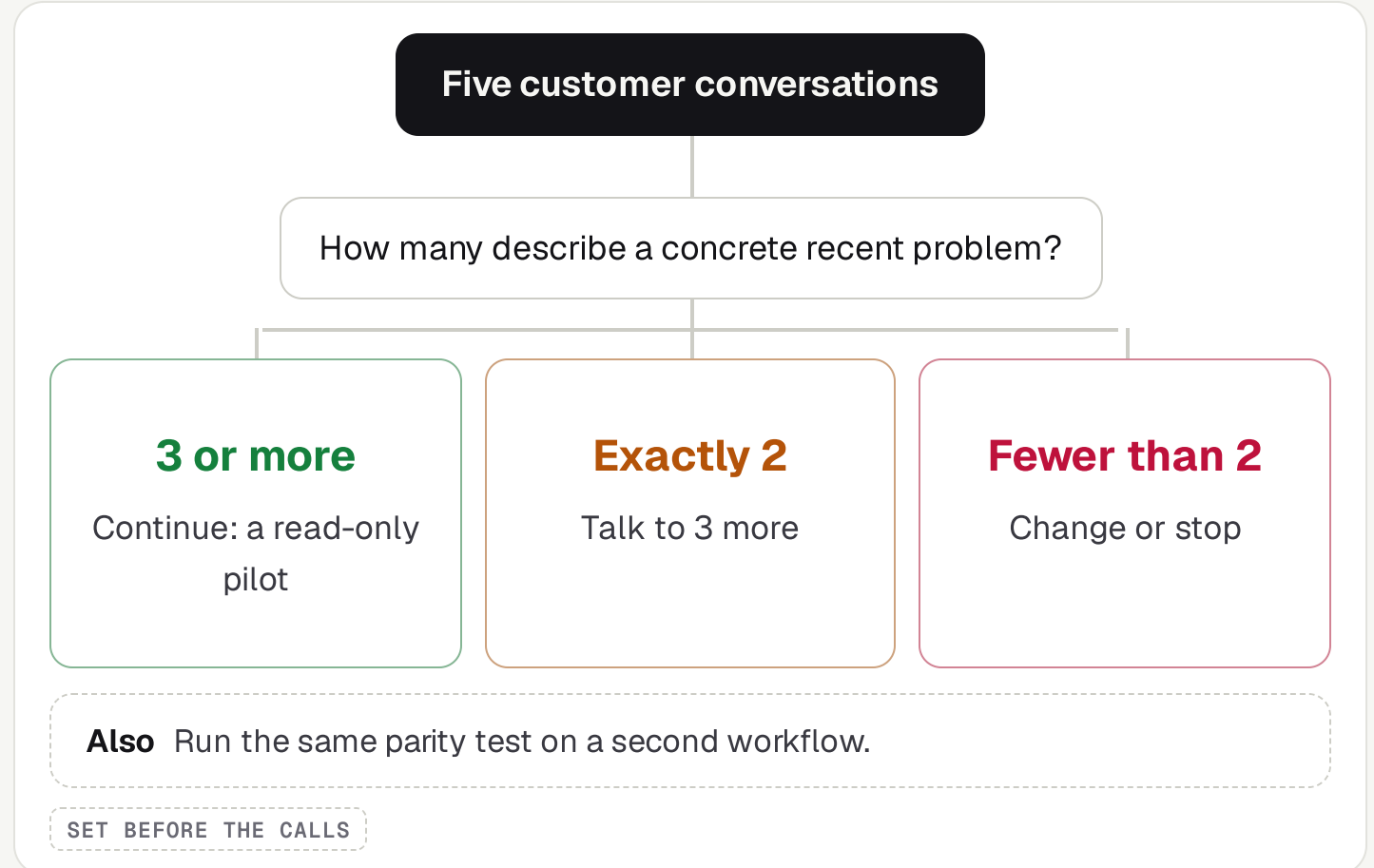
Five things could still make this wrong, and here's how each is checked



- 1 Run it read-only on real accounts, with consent
- 2 Rerun K2 on real W-9s, shared with consent
- 3 Name-only matches go to review; a person confirms each
- 4 Test a year-end payment on a real account
- 5 9 of 44 names checked on the 1099 list reach no agent tool, all \$0 in 2026; record type unknown. Change 03 asks Mercury to list them

Recommendation: three changes for Mercury. My next step: five customer conversations, then continue, change or stop.

- 01 Put the payee ID in MCP's payment list
- 02 Return each W-9's form type and status
- 03 List every 1099-page payee for agents



Appendix A10 (p. 19) · A1 (p. 10)

Source CUSTOMER_QUESTIONS.md (decision rule) · FINDINGS 3, 6, 21, 33, 34, 37, 39 · REQUIREMENTS.md

Crispin Gathuru · Portland, OR · readiness-1099.pages.dev · Also built: [Get-Pruuf](https://get-pruuf.com) (live Oregon contractor records)

✓ yes ✗ no ○ untested

The coverage grid, cell by cell, and the three changes in full

FACT	DASHBOARD	API	CLI	MCP TEXT	MCP STRUCTURED
Payee ID, payment list	○ the dashboard's payment list wasn't checked	✓ 60 of 60 payments 074057Z-read/009	✓ 60 of 60 payments 045355Z-demo/transactions	✗ 0 of 100 rows 150844Z-mcp/022	✗ no counterpartyId 150844Z-mcp/022
Payee ID, one payment	○ the dashboard wasn't checked	○ no single-payment call saved	○ no single-payment call saved	✓ counterpartyId in the text 150844Z-mcp/035	✗ no counterpartyId 150844Z-mcp/035
W-9 file	✓ shows the file screenshot (FINDINGS 39)	✓ 17 files 194232Z-fields	✓ 17 files 045355Z-demo/attachments	✓ 9 files 150844Z-mcp/016	✗ no structured result lists a file 150844Z-mcp
W-9 form type	○ label only: the file list calls every file “W9 Document”, even one that says it is not a W-9 screenshot (FINDINGS 2b)	✗ no taxFormType or formType 194232Z-fields	✗ no taxFormType or formType 045355Z-demo/attachments	✗ no taxFormType or formType 150844Z-mcp/016	✗ no files at all 150844Z-mcp
W-9 status	✓ “Processing Error” screenshot (FINDINGS 35, 39)	✗ no status field 194232Z-fields	✗ no status field 045355Z-demo/attachments	✗ no status field 150844Z-mcp/016	✗ no files at all 150844Z-mcp
Every 1099-page name	✓ the 1099 page itself screenshot (FINDINGS 16, 21)	✗ 9 of 44 missing 194232Z-fields/001	✗ 9 of 44 missing 045355Z-demo/recipients	✗ 9 of 40 missing; 4 created after this run 150844Z-mcp/016	✗ 9 of 40 missing; 4 created after this run 150844Z-mcp/016

○ untested: no saved evidence either way, or only a label (the dashboard's W-9 form type, finding 2b); not a gap. A9 (p. 18) lists the 9 names on the 1099 list that agents can't see.

01

Put the payee ID in MCP's payment list

Add counterpartyId to listTransactions rows, and keep it and the next-page cursor in structuredContent, with an outputSchema. Today an agent joins payments by name. 3 payees here are named “Banned Recipient”.

Findings 3, 6, 34

02

Return each W-9's form type and status

Return the form type and the reader's status (read, problems found, failed) on each payee file, through the API, CLI and MCP. The dashboard works out the status; its form type is only a label (finding 2b).

Findings 2b, 33, 37, 39

03

List every 1099-page payee for agents

List the ACH and wire pull counterparties that the 1099 page includes, with their payments, so an agent's list can match the page. No CLI command or MCP tool lists them today.

Findings 21

The rules behind every decision

Threshold	\$2,000 or more paid in 2026 needs a 1099 (\$600 before 2026).
Due date	Form 1099-NEC for 2026 is due Feb 1, 2027. Jan 31, 2027 is a Sunday, so the date moves to the next business day.
Corporations	Generally exempt, except medical and health care payments (1099-MISC box 6) and attorneys' fees (1099-NEC box 1).
Governments	Payments to the United States, a state or DC are not reportable.
Foreign payees	Form 1042-S, not a 1099; ask for a W-8BEN.
Missing tax ID	Get a W-9 with a TIN before the next payment. Until then, backup withholding applies: withhold 24% of each further reportable payment, deposit it, and report it on Form 945 and in box 4 of the 1099 (IRS Publication 1281).
Mercury's FAQ	Still describes 1099-NEC as for payments of \$600 or more (finding 42). Mercury's 1099 filing does not report backup withholding, so Form 945 and box 4 are filed outside Mercury (Mercury's 1099 Filing FAQ).
Unresolved warnings	A 1099 or exempt decision with a warning nobody has cleared goes to review, never straight through.
Not tested	Mercury Command, the in-app AI. This prototype is about outside agents.

Source IRS i1099mec (Rev. 12/2026) · support.mercury.com/hc/en-us/articles/41316677927060-1099-Filing-FAQs (Mercury's 1099 Filing FAQs, read 2026-10-03) · readiness/rules.py · FINDINGS 42, 43

✓ yes ✗ no ○ untested

The four decisions, rule by rule

PAYEE · DECISION	PAID	MERCURY'S PAGE	W-9 ON FILE	NEXT STEP · RULES
1 · Alex Rivera • Needs a 1099	\$47,500.00 50 payments	✓ same	Alex Rivera · individual · SSN ••2233 · signed	File a 1099-NEC (box 1) by Feb 1, 2027. R-REPORT
2 · Wire Person • Needs a valid W-9	\$24,005.00 2 payments	✓ same	Wire Person · individual · no TIN · signed	The W-9 has no tax ID. Get a new one; until then, withhold 24% of each further payment. R-NO-TIN
3 · IRS Tax Payment Recipient • Needs review	\$5,000.00 1 payment	✓ same	No W-9 on file	Exempt only if this was a payment to the government. The name is the only evidence, so confirm it. R-GOVERNMENT · R-REVIEW · W-GOV-NAME
4 · Patient Payee • Needs review	\$5,000.00 1 payment	✓ same	Patient Payee Medical Group Inc · C corp · EIN ••2223 · signed	Reportable if medical: the W-9 is for “Patient Payee Medical Group Inc”, and medical payments to corporations go on 1099-MISC box 6. Two more checks. R-CORP · R-REVIEW · W-NAME · W-INTL-WIRE · W-CORP-SERVICES

R-REPORT	\$47,500.00 of nonemployee compensation to an individual payee meets the \$2,000 threshold; report it on Form 1099-NEC box 1	R-NO-TIN	the W-9 has no taxpayer identification number
R-GOVERNMENT	payments to the United States, a state or DC are not reportable	R-REVIEW	exempt only if a person confirms what its warnings flag
W-GOV-NAME	exempt only because the name looks like a government payee; confirm the money went to the United States, a state or DC (a tax payment, say), not a business	R-CORP	payments to C or S corporations are generally not reportable
W-NAME	the W-9 on file is for 'Patient Payee Medical Group Inc', not 'Patient Payee'; confirm it belongs to this payee	W-INTL-WIRE	paid by international wire, yet a W-9 (US person) is on file; confirm the payee's status
W-CORP-SERVICES	corporations are generally exempt, but medical or health care payments to them are reportable on Form 1099-MISC box 6 at \$2,000 or more, and attorneys' fees on 1099-NEC box 1; confirm what was paid for		

All 93 payees

■ Needs a 1099 1

Alex Rivera	\$47,500.00
-------------	-------------

■ Needs a valid W-9 1

Wire Person	\$24,005.00
-------------	-------------

■ Needs review 2

IRS Tax Payment Recipient	\$5,000.00
---------------------------	------------

Patient Payee	\$5,000.00
---------------	------------

□ Under \$2,000, W-9 problem 1

RecipName	\$14.00
-----------	---------

■ Under \$2,000 1

Currency Cloud Recipient	\$653.00
--------------------------	----------

87 payees at \$0, 9 of them with a W-9 problem: listed in a drawer on the site.

Source `raw/20261008T045355Z-demo/packet.json` · an outlined label marks a W-9 problem under \$2,000

One payment, as each surface returns it

✓ Mercury CLI

● ● ● terminal

```
$ mercury --environment sandbox --format json
transactions list --max-items -1
```

● ● ● the payment, as the CLI returns it

```
{
  "id": "1c9dbe6c-be32-11f1-9e10-dfca487ec7e1",
  "amount": -19005.0,
  "status": "sent",
  "kind": "outgoingPayment",
  "counterpartyName": "Wire Person",
  "counterpartyId": "0e7fadae-be32-11f1-9e10-77d840416fa9",
  "createdAt": "2026-09-25T07:22:41.922878Z"
}
```

✗ MCP listTransactions

● ● ● tools/call

```
listTransactions({"limit": 100})
```

● ● ● the same payment, as MCP returns it (text)

```
{
  "id": "1c9dbe6c-be32-11f1-9e10-dfca487ec7e1",
  "amount": -19005,
  "status": "sent",
  "kind": "outgoingPayment",
  "counterpartyName": "Wire Person",
  "postedAt": "2026-10-02T07:22:41.922878Z"
  // no counterpartyId: the payee is only a name
}
```

✓ ✗ MCP getTransactionById

● ● ● text output · the one saved lookup: a payment to Patient Payee

```
{
  "id": "1d56171e-be32-11f1-9e10-0f1ab53c5fa7",
  "amount": -5000,
  "counterpartyName": "Patient Payee",
  "counterpartyId": "1d502ca0-be32-11f1-9e10-c71693c64413"
}
```

● ● ● structuredContent.transaction · same call

```
{
  "id": "1d56171e-be32-11f1-9e10-0f1ab53c5fa7",
  "amount": -5000,
  "counterpartyName": "Patient Payee"
  // no counterpartyId
}
```

Source raw/20261008T045355Z-demo/transactions.json · raw/20261002T150844Z-mcp/022_call_listTransactions.json · raw/20261002T150844Z-mcp/035_call_getTransactionById.json · FINDINGS 3, 23, 34

MCP's two forms, tool by tool

✓ yes ✗ no ○ untested

TOOL	FIELD	TEXT OUTPUT	STRUCTUREDCONTENT
getRecipients	payee files (attachments)	✓ 9 files	✗ missing
getTransactionById	payee ID (counterpartyId)	✓ present	✗ missing
listTransactions	next-page cursor (page.nextPage)	✓ present	✗ missing; says truncatedAt: null
listTransactions	payee ID (counterpartyId)	✗ missing	✗ missing

✗ No tool declares an outputSchema (37 tools listed).

✗ 2 failed calls (getAttachment, getTransaction) came back as plain text with no isError flag (finding 41).

The MCP spec (2025-06-18) says a tool that returns structured content should also return the same JSON, serialized, as text.

TOOLS THAT RETURN BOTH FORMS	FILLED FIELDS THE STRUCTURED FORM LEAVES OUT · OR THE TOOLS
✗ getAccounts	accountNumber, canSendRealTimePayments, createdAt, currentBalance, legalBusinessName, nickname, routingNumber
✗ listCards	accountId, budgets, createdAt, isAgentCard, name, remainingAmountCents, spendLimitType, updatedAt, userId
✗ getRecipients	accountNumber, accountType, attachments, bankCityState, bankCode, bankCountry, bankDetails, brazil, canada, checkInfo, chile, colombia, countrySpecific, defaultAddress, domesticWireRoutingInfo, dominicanRepublic, electronicAccountType, electronicRoutingInfo, fileName, honduras, iban, internationalWireRoutingInfo, kazakhstan, legalId, legalIdType, pakistan, paraguay, phoneNumber, realTimePaymentRoutingInfo, routingNumber, swiftCode, transitNumber, uploadedAt, url
✗ listTransactions	accountId, bankDescription, cardId, counterpartyNickname, externalMemo, failedAt, glAllocations, glCodeName, nextPage, page, postedAt
✗ getAccount	accountNumber, canSendRealTimePayments, legalBusinessName, name, routingNumber
✗ getTransactionById	accountId, address1, bankCityState, bankCountry, bankDetails, bankName, city, compliantWithReceiptPolicy, counterpartyId, country, createdAt, details, hasGeneratedReceipt, iban, internationalWireRoutingInfo, postalCode, postedAt, region, swiftCode
✓ Differ, nothing filled left out	listInvoices, listSendMoneyApprovalRequests, getUsers, getAccountCards, getAccountStatements
✓ Identical	getCashflowSummary

The W-9 test, upload by upload

✓ yes ✗ no ○ untested

PAYEE	WHAT IT TESTS	UPLOAD	MERCURY
Riley Park	Baseline: individual, SSN, signed	API	✗ Processing Error
Quinn Studio LLC	Payee named after line 2; line 1 is Morgan Quinn	API	✗ Processing Error
Granite Peak Corp	C corporation, exempt payee code 5	API	✗ Processing Error
Taylor Reed	No TIN	API	✗ Processing Error
Casey Morgan	Not signed	API	✗ Processing Error
Ironwood Supply LLC	Scan, no text layer; LLC taxed as partnership	API	✗ Processing Error
Control 1 IRS Blank Form	Can Mercury's reader handle the official file at all?	API	✗ Processing Error
Control 2 Riley NoGC	Did our generator's save step cause the error?	API	✗ Processing Error
Retry 1 Riley Park	Did the first job just die? (Mercury's tooltip says "try uploading again")	API	✗ Processing Error
Retry 1 Taylor Reed	Same question, second sample	API	✗ Processing Error
PROBE-2026-10-02 A Individual	Are dashboard uploads read when API uploads aren't?	dashboard	✗ Processing Error

```
● ● ● the dashboard-uploaded file, as the CLI returns it

{
  "id": "2002c8f8-be92-11f1-93e0-07ff7dde3bce",
  "recipientId": "786f10c8-be33-11f1-ab7e-ebcde1b4da30",
  "fileName": "h2h-riley-park.fillable.pdf",
  "url": "<temporary download link>",
  "uploadedAt": "2026-10-02T18:50:23.672277Z"
  // no tax-form type (taxFormType / formType)
  // no W-9 status: nothing says "Processing Error"
}
```

My reader, same file: Riley Park, individual, tax ID present, signed; it flags that the W-9 names Riley Park, not this payee.

Across the test, my reader read all 5 typed files and handed the 1 scan on. Comparing it with Mercury's needs production (K1).

Kill conditions K1 to K5, in full

TEST	KILL IF	THEN	RESULT	STATUS	RERUN
K1 Mercury's reader vs mine	Mercury's parsed W-9 details (finding 2b) already include name, classification and TIN, and that reaches the API or MCP	Drop w9_check	Not answerable in the sandbox: all 11 uploads failed.	Open	HEAD_TO_HEAD.md
K2 My W-9 reader	W-9 extraction accuracy is below 95% per field on the labeled set	Stop and write it up	25 of 25 typed W-9s exact	Pass (synthetic files only)	python -m readiness.eval_w9
K3 Registry name match	registry name-match precision is below 90% on a labeled sample	Drop registry_check	Not built yet	Open	REQUIREMENTS.md
K4 Totals match Mercury's	any payee total differs from Mercury's "Paid this year"	Fix it before anything else. If it can't be fixed, stop	5 of 5 payees with payments match to the cent; the rest match at \$0. Not compared: Currency Cloud Recipient (\$653); it isn't on Mercury's 1099 list.	Pass	python -m readiness.sandbox_report
K5 A business tries it	no real business will try it by the end of week 2	Ship it as a field report plus demo, not a product claim	Not started: five conversations planned	Open	CUSTOMER_QUESTIONS.md

✓ Also checked: 45 of 45 IRS-rule tests pass, rerun on every build (python -m unittest tests.test_rules); the agent won't guess between the 3 payees named "Banned Recipient" (smoke test, Oct 3, 2026: python agent/smoke_test.py).

The risk register

#	RISK · WHAT I KNOW	LIKELIHOOD	EFFECT	IF IT HAPPENS	HOW IT'S CHECKED	FINDINGS
1	Sandbox, not production Every number here comes from Mercury's sandbox (a fake company, fake money). Mercury's W-9 reader failed on all 11 of my uploads there.	High	High	Findings may not hold for real accounts; Mercury's W-9 reader may work there	Run it read-only on real accounts, with consent	37, 39
2	Circular W-9 test files My reader was graded on 25 typed W-9s I generated, with the same form fields it reads (finding 28).	High	Medium	A perfect score overstates real-world accuracy	Rerun K2 on real W-9s, shared with consent	28
3	Exemptions decided by name IRS Tax Payment Recipient counts as government only by its name.	Medium	High	A wrongly exempt payee gets no 1099	Name-only matches go to review; a person confirms each	7
4	Which date sets the tax year My totals use createdAt; MCP's list shows only postedAt (finding 44). All 57 payments here are dated Sep 25 – Oct 2, 2026, so no year boundary was tested.	Low	Medium	A Dec 31 payment posted in January lands in the wrong year	Test a year-end payment on a real account	44
5	Payees only the dashboard lists 9 names on Mercury's 1099 list that no agent tool returns: ABC Consulting Group, Deleted Payee, Freelance Designer, John Smith, Legal Services Corp, Merged Entity, Office Space Rentals, Patent Holder Inc, Tech Contractor LLC; all show \$0 in 2026 here, so they don't change this answer, but an agent can't reconcile against the list. John Smith's panel: W-9 status Missing, no email or address; View Payments shows no transactions this year. The record type is still unknown; Mercury's FAQ says the list also holds ACH and wire pull counterparties.	Medium	Medium	An agent's list looks complete when it isn't	9 of 44 names checked on the 1099 list reach no agent tool, all \$0 in 2026; record type unknown. Change 03 asks Mercury to list them	21

My assessment: likelihood and effect are judgments, low, medium or high, never probabilities. The facts in each row come from saved data.

Source risks.json · FINDINGS 7, 21, 28, 37, 39, 44 · HEAD_TO_HEAD.md · raw/screenshots/2026-10-08_1099-add-recipients_*.png, 2026-10-08_transactions_john-smith-view-payments.webp · raw/20261008T045355Z-demo/transactions.json

The next step, in full

Who	3 small-business owners or operators and 2 bookkeepers who pay contractors; Mercury customers first.
How	11 questions about the last time they prepared 1099s. No pitch, no tax IDs, no client names.
Rule, set before the calls	After 5 conversations: continue if at least 3 describe a concrete recent problem (a missing or wrong W-9, a mismatch notice, or hours of manual work); change or stop if fewer than 2 do; otherwise talk to 3 more.
If it continues	Pilot it read-only with real accounts, and settle K1 on a production account.
Also	Run the same parity test on a second workflow. List the facts that job needs, then check each one on the dashboard, API, CLI and MCP, as page 1 does.
What this is	An independent prototype on Mercury's sandbox (a fake company, fake money). Not built or endorsed by Mercury. Not tax advice: the rules follow IRS i1099mec (Rev. 12/2026).

Crispin Gathuru · Portland, OR · [readiness-1099.pages.dev](#) · Also built: [Get-Pruuf](#) (live Oregon contractor records)

Source CUSTOMER_QUESTIONS.md · REQUIREMENTS.md · demo/about.json