ppie

User Manual - Paper Image Extraction

Turn a scanned paper crash report into clean, located field readings - every box read, how sure it is, and exactly where on the page it was found - ready for your crash system to review.

Version: September 30, 2026
Bad Marine LLC

Contents

1. **Welcome to ppie**
   What it does, how it fits with your crash system, who uses what, a scan start to finish
2. **Getting around**
   The three screens, the top bar, conventions used everywhere
3. **Extracting a scan**
   The Extract screen, reading one scan, the results list and the page, warnings
4. **Reading the results**
   Confidence, blank versus not read, dates and multi-part fields, person rows, what ppie leaves to people
5. **Forms ppie knows**
   The built-in forms, multi-page packets, the Forms screen, export and import
6. **Training a form**
   Starting from a blank or an existing form, the box editor, moving and zooming, repeating a box, check-box groups, saving and testing
7. **Sending scans to your crash system**
   The watched folder, a one-time batch, a dry run, what each result line means
8. **Running ppie (IT and administrators)**
   The folder watcher, the engine service and its endpoints, connecting to your crash system, air-gapped operation
9. **When something looks wrong**
   Common problems and what to do
10. **Quick reference**
    Labels, messages, glossary
11. **Known limits**
    What ppie doesn't do yet
12. **Trying the online demo**
    The sample scan, what the demo limits, your privacy

1. Welcome to ppie

1.1 What ppie does

ppie (Paper Image Extraction) reads scanned paper crash reports. Most crash reports now arrive electronically, but some still come in on paper - mailed in, faxed, or printed from a driver's own portal. ppie turns each of those scans into the same kind of data an electronic report carries, so they can be reviewed the same way.

For every scanned report, ppie:

- **recognizes the form** from its own printed title - you don't have to say which form it is;
- **straightens the scan** and lines it up against a blank copy of the same form, even when the page was fed in crooked or the scanner stretched it a little;
- **subtracts the printed form**, so only what the officer or driver wrote is left;
- **reads every box** - typed and handwritten text, code boxes, and check marks;
- **reports each reading with a confidence and a location** - what it read, how sure it is (0 to 1), and exactly where on the page it was found.

The result is one **envelope** per report: every field's reading, its confidence, and its place on the page, plus the page images' fingerprints so the same scan is never filed twice.

> **What ppie doesn't do:** It never decides what a value means. If the box says 04, ppie reports 04 - turning 04 into "Cloudy", filling in "Not Stated", or checking the report against the rules is your crash system's job, where a person reviews it.

1.2 How it fits with your crash system

ppie sits in front of your crash system. It reads the paper; your crash system (the Refinery, or any system that accepts ppie's envelope) does everything after that:

ppie	Your crash system
Recognizes the form, straightens the scan, reads every box, reports confidence and location.	Takes the envelope in, files it with its case, interprets the codes, and puts every value in front of a reviewer - with the scan beside it, and a click that jumps to the box ppie read.

ppie runs **on your own machine**. The scans never leave it except to go to your own crash system. It needs no internet connection: the reading engine is built in, and nothing is looked up online.

ppie is also a **standalone piece**: it can be pointed at any system that accepts its envelope, not only the Refinery, and new forms can be taught to it without programming (chapter 6).

1.3 Who uses what

Who	What they use
Records staff / intake clerk	Nothing directly - scans dropped into a watched folder are read and sent on automatically (chapter 7).

Who	What they use
Reviewer	Your crash system's review screens, where ppie's readings show up with their confidence and location.
Form trainer	The **Train** screen - teaches ppie a new form, or adjusts one it already knows (chapter 6).
Anyone checking a scan	The **Extract** screen - drop a scan in and see exactly what ppie reads (chapters 3 and 4).
IT / system administrator	Installs ppie, runs the engine service and the folder watcher, and connects it to the crash system (chapters 7 and 8).
Someone evaluating ppie	The online demo - no install needed (chapter 12).

1.4 A scanned report, start to finish

1. A paper report is scanned to a PDF - one file per report - into ppie's **incoming folder**.
2. ppie picks it up, recognizes the form from its title and, for a multi-page packet, works out what each page is.
3. It straightens each page, lines it up with the blank form, removes the printing and reads every box.
4. It sends the envelope and the scan to your crash system, and moves the file to **_processed** (or **_failed**, with the reason, if something went wrong).
5. A reviewer works the report in the crash system, confirming each value against the scan.

2. Getting around

2.1 The screens and the top bar

ppie's screens open in your web browser on the machine ppie runs on. The top bar is on every screen:



The top bar

Screen	What it's for
Extract	Drop in a scan and see what ppie reads - every value, its confidence and where it was found (chapter 3).
Train	Teach ppie a new form, or adjust the boxes of one it already knows (chapter 6).
Forms	Every form ppie can read: the built-in ones and any you've trained. Export, import, delete (chapter 5).

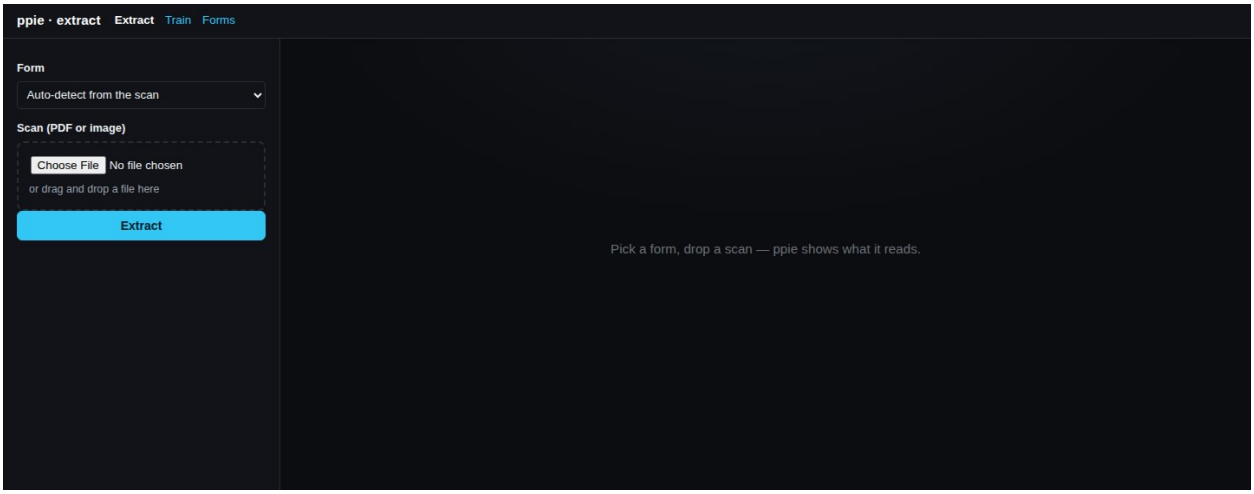
> **Local by default:** These screens only answer on the machine ppie runs on. Your administrator decides whether anyone else can reach them.

2.2 Conventions used everywhere

- **Confidence colors** - every box ppie reads is outlined in green (0.80 or higher), amber (0.50 to 0.80) or red (under 0.50, or blank). The same key sits above every page.
- **Hover a box** on the page to see its field, value, confidence and how it was read.
- **Pages show at their real size.** Scroll the page area to move around it; each side panel scrolls on its own.
- **? buttons** open the instructions - on the Train screen, **? How it works** in the top bar; on the Extract screen, the **?** beside a warning.
- **Esc** closes any open panel or window.
- **Forms are named by their printed title** - for example "Form 40 (Investigator's Motor Vehicle Crash Report)".

3. Extracting a scan

The Extract screen reads one scan and shows you everything ppie found, right on the page. Use it to check a scan, to see how well a form reads, or to test a form you've just trained. Nothing is sent anywhere from this screen.



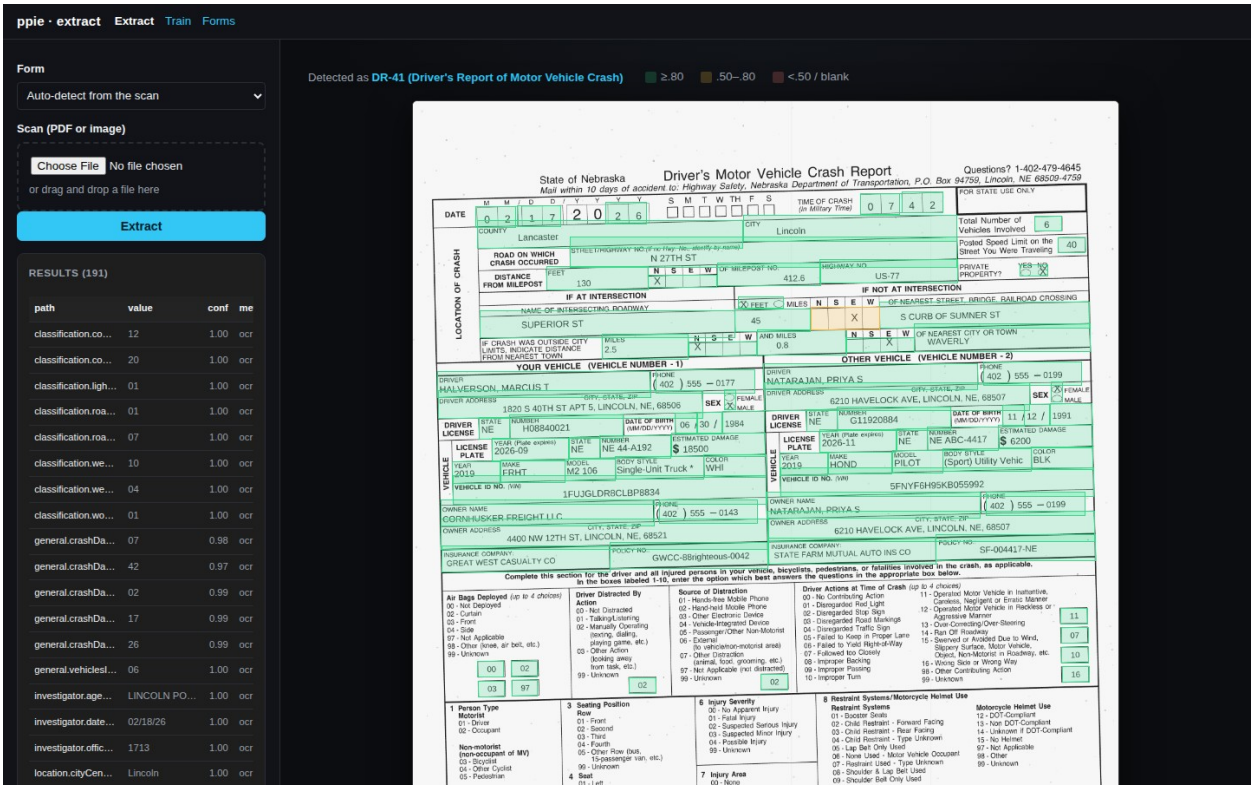
The Extract screen, before a scan

3.1 Reading a scan

1. Choose the **Form**, or leave it on **Auto-detect from the scan** (the usual choice).
    - Auto-detect reads the form's printed title on the first page.
    - Pick a form yourself only for a scan whose title is missing or unreadable.
2. Choose the **Scan** - click the box, or drag the file onto it. A PDF (any number of pages), or an image of one page (PNG, JPG, TIFF, BMP).
3. Click **Extract**. A single-page form takes a few seconds; a full multi-page packet can take a minute or more.

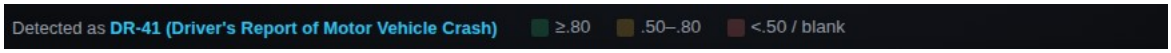
> **Picked the wrong form?** ppie checks the scan's own title anyway. If it reads as a different form, ppie reads it as the form it really is and tells you so in a warning (section 3.4).

3.2 The results



A scanned driver's report after Extract

The left side lists every field ppie read; the right side shows the scan with a box drawn around each one.



Above the page: the form it was read as, and the color key

Above the page is the form ppie read it as - **Detected as** when it recognized the form itself, **Read as** when you picked it - and the confidence color key.

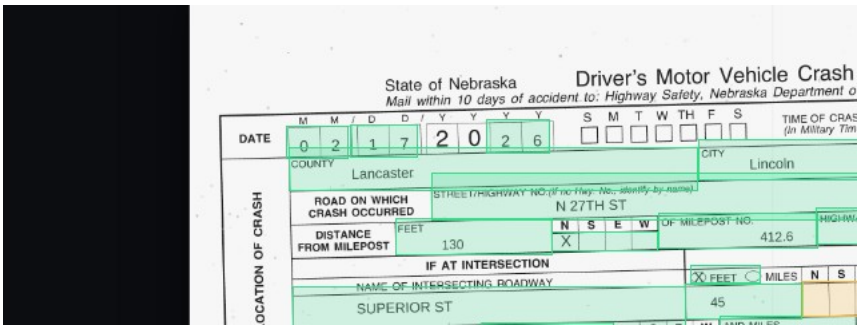
3.3 The results list and the page



The results list

Column	What it shows
path	Which field this is - the name your crash system uses for it, for example `location.countyId` or `vehicles[0].make`. Hover a cut-off name to read it all.
value	Exactly what the box says. (blank) means ppie looked and the box is empty.
conf	How sure ppie is, from 0.00 to 1.00 (chapter 4).
method	How it was read: **ocr** (text), **mark** (a check box or circled option), **blank** (empty).
pg	The page of the scan it's on.

On the page, every box is outlined in its confidence color. Hover any box to see its field and value.



Hover a box to see what was read there

A dashed grey box on the page (**unlabelled** in the key) is a box on a trained form that hasn't been given a field name yet - it's shown so you can see what's still unmapped, but nothing is read from it until it has a name (section 6.4).

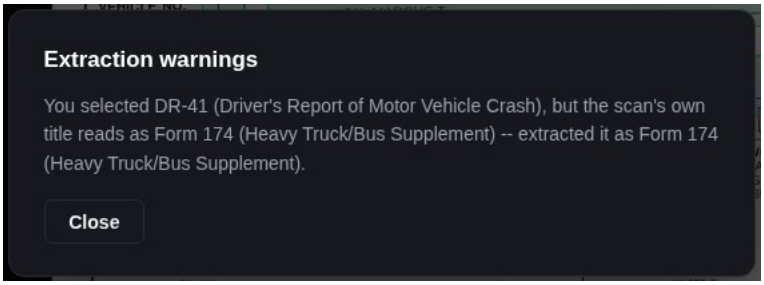
3.4 Warnings

If anything about the scan needs a second look, an amber line appears above the page:



A warning line

Click **?** for the details:



The warning details

Warnings don't stop the extraction - the results are still there. The common ones are explained in the Quick reference (chapter 10).

4. Reading the results

4.1 Confidence

Every reading carries a confidence from 0 to 1 - how sure ppie is that it read the box correctly.

Confidence	Color	What it means
0.80 and up	Green	A clean read. Still worth a glance, but usually right.
0.50 to 0.80	Amber	Probably right; check it against the scan.
Under 0.50	Red	Low confidence - check it. Handwriting that's faint, crowded or crosses the box lines lands here.

Confidence is about the reading itself. How well the whole page lined up with the blank form is reported separately: a page that couldn't be lined up well gets a **registration weak** warning (chapter 10), and every reading on that page deserves a closer look - a box can land a little off its field.

> **Confidence is a guide, not a verdict:** a low score says "look at this one"; a high score is not a promise. That's why every value is still confirmed by a person in your crash system.

4.2 Blank versus not read

You see	It means
(blank), method blank	ppie looked at the box and it's empty. Your crash system can treat that as "not stated".
No row at all for a field	ppie didn't read it - the form has no box for it, or the box couldn't be located. Your crash system treats it as unknown, not empty.

4.3 Dates, phone numbers and multi-part fields

- A field printed as several boxes - a date's month, day and year; a phone number's parts; a name split into last, first, middle - comes back as **one row per box**. Your crash system puts the parts back together.
- A field that allows **"up to N choices"** comes back with one row per choice box, in order.
- A **check box group** (a Y/N pair, or a list of numbered options) comes back as one row: the code of the option that's marked. Which code each option stands for is set on the form (section 6.8).

4.4 Rows of people

Some forms list people in a grid of rows - the Form 40 "All Drivers & Occupants" pages, or the injured-persons rows at the bottom of a driver's report.

- **Form 40's grid** has a Vehicle No. and Occupant No. on every row. ppie reads them and files each row under the right vehicle and person. A row whose numbers can't be read is skipped with a warning - never guessed, because a wrong guess would put one person's injuries on someone else.

- **A driver's report's injured-persons rows** don't say whose they are. ppie sends each row as it's printed (row 1, row 2, ...) and your crash system places each person using the person type written on the row.

4.5 What ppie leaves to people

- **Values are literal.** ppie reports what's written - it never turns a code into its meaning, corrects a spelling, or fills in a missing value.
- **It never decides a final value.** Every value is confirmed by a reviewer in your crash system.

> **Handwriting:** clean, typed or printed entries read very well. Real handwriting reads less reliably today - that's what the confidence score and the reviewer are for (see Known limits, chapter 11). Expect more amber and red on hand-filled forms.

5. Forms ppie knows

5.1 The built-in forms

ppie comes with six crash forms already mapped, box by box:

Form	Pages	Notes
DR-41 (Driver's Report of Motor Vehicle Crash)	2	The driver's own report.
Form 40 (Investigator's Motor Vehicle Crash Report)	8	The officer's report. Its vehicle, driver and occupant pages repeat for every vehicle and person - see 5.2.
Form 40a (Continuation Report)	1	More property damage and more narrative when Form 40 runs out of room.
Form 174 (Heavy Truck/Bus Supplement)	1	One per heavy truck or bus.
Form 178 (Non-Motorist Supplement)	1	One per pedestrian or cyclist.
Form 179 (Fatal Crash Report)	1	Added when someone died.

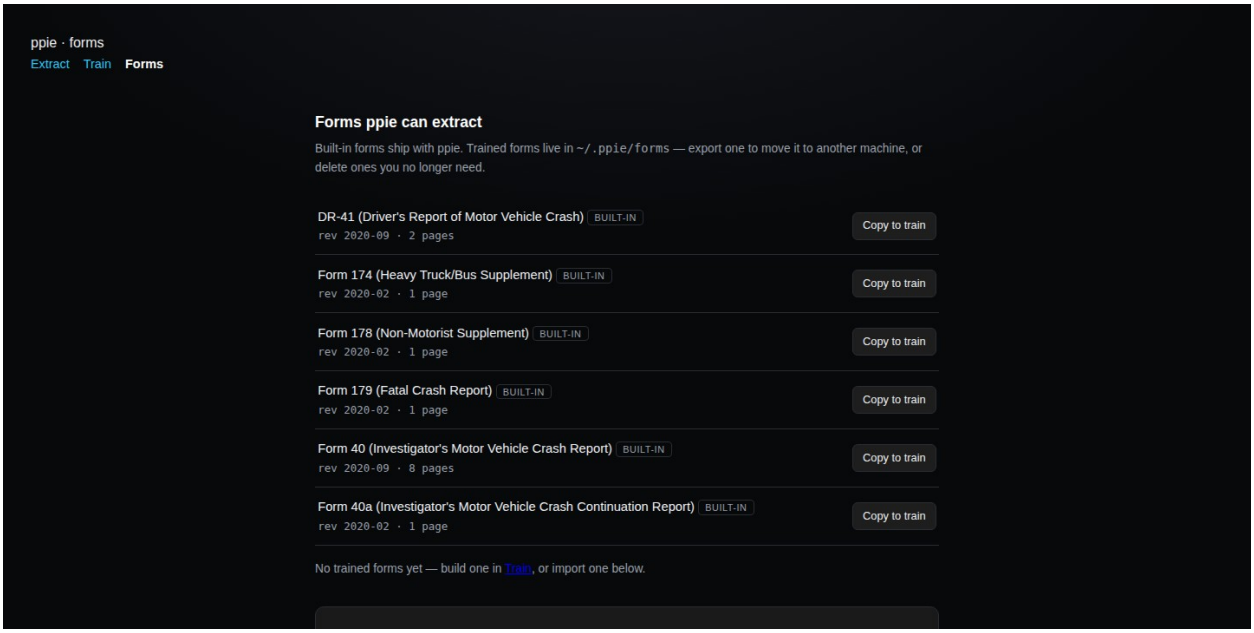
5.2 Multi-page packets

An officer's report usually arrives as one scanned **packet**: the Form 40 pages, repeated for each vehicle, plus any supplements, in whatever order they were stacked. Scan the whole packet as **one PDF** - ppie sorts it out:

- Each page is recognized by its **own printed title and footer**, not by its position in the file.
- Repeated vehicle pages are filed as vehicle 1, vehicle 2, and so on, in the order they appear.
- A Form 174 or 178 is filed under the vehicle or non-motorist number written in its own **Vehicle No.** / **Non-Motorist No.** box. If that box is blank or unreadable, it falls back to packet order and says so in a warning.
- Form 40a's extra narrative lines continue straight on from Form 40's.
- A page ppie doesn't recognize is left out, with a warning - the rest of the packet is still read.

> **One report per file:** put each crash's packet in its own PDF. Two crashes in one file would be read as one report.

5.3 The Forms screen



The Forms screen

Every form ppie can read, built-in first, then any you've trained:

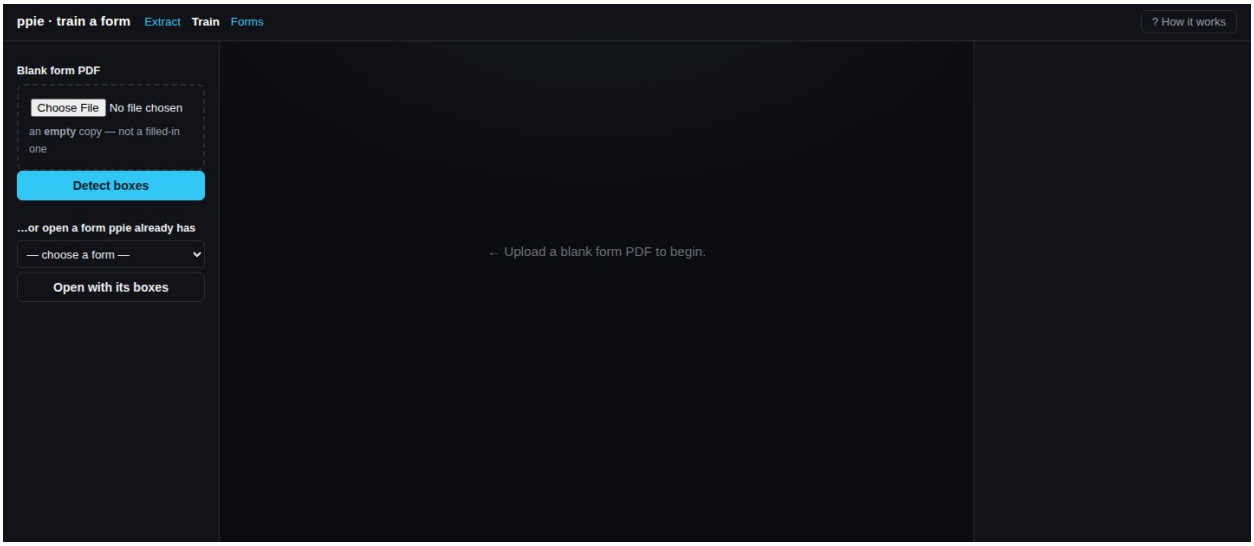
Button	What it does
Copy to train (built-in forms)	Opens the form in the Train screen with all its boxes, as the start of a new trained form. The built-in form itself never changes.
Edit (trained forms)	Opens the trained form in the Train screen to adjust it. Saving updates it in place.
Export (trained forms)	Downloads the form as a .zip - its blank form, its boxes and its check-box groups - to move it to another machine.
Delete (trained forms)	Removes it from this machine, after you confirm. It can't be undone - export it first if you might want it back.

5.4 Importing a form

At the bottom of the Forms screen, **Import a form** takes a .zip exported from another ppie. Choose the file and click **Import**. You can give it a different **Form key** (2-64 lowercase letters and digits) - for example when this machine already has a form with that key. Imported forms are ready to use straight away; there's nothing to restart.

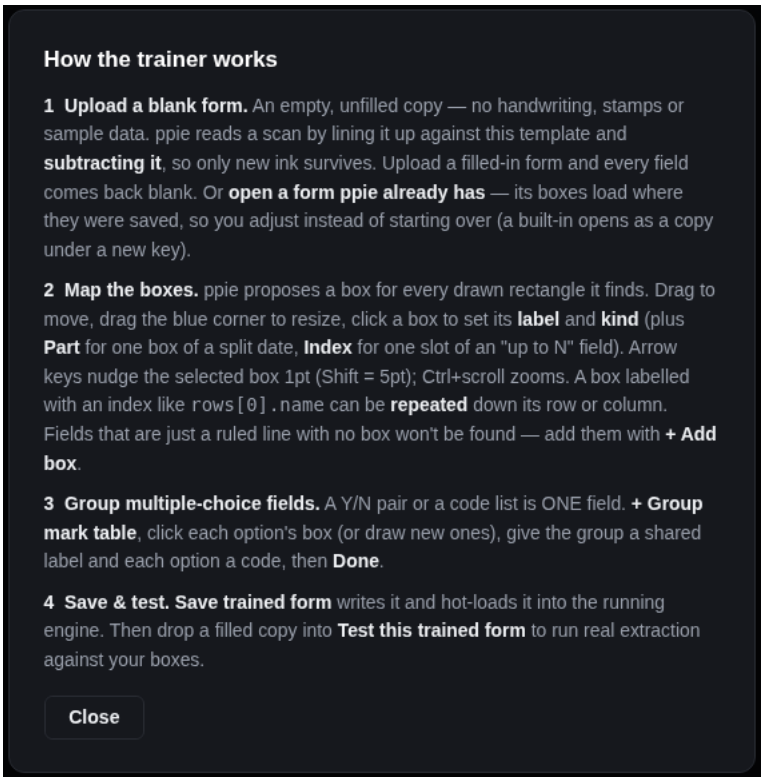
6. Training a form

The Train screen teaches ppie a form: you draw a box around each field, give it a name, and save. No programming, and no restart - a saved form can be used the moment it's saved. Use it for a form ppie doesn't know yet, or to adjust the boxes of one it does.



The Train screen

The screen has three parts: the left side to start and for tools, the page in the middle, and on the right the lists of boxes and check-box groups, with **Save** at the bottom. **? How it works** in the top bar has a four-step summary.

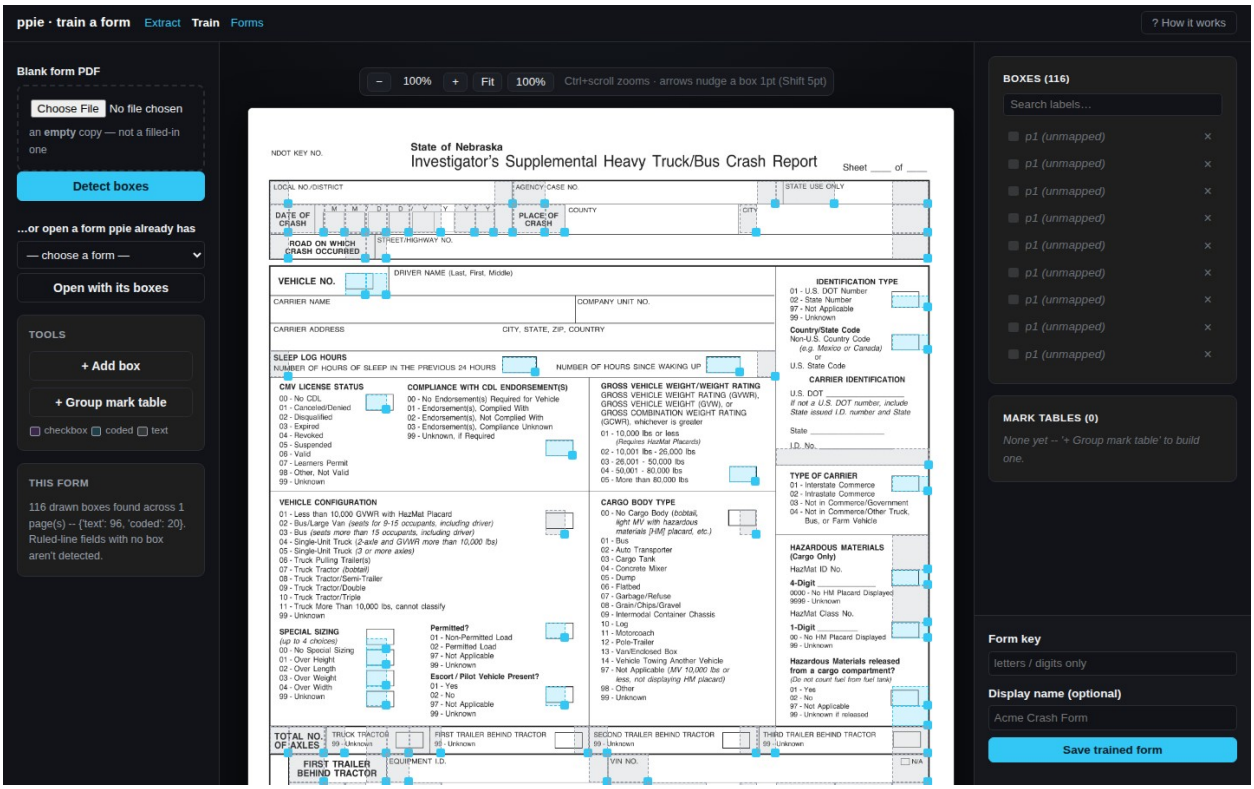


? How it works

6.1 Start from a blank form

> **Use an empty form:** ppie reads a scan by lining it up with this blank and subtracting it, so only new writing is left. Train on a filled-in copy and everything already written on it is subtracted away - every field on every scan would come back blank. Use the form's original PDF, or a scan of a clean, unused copy.

1. Under **Blank form PDF**, choose the empty form's PDF.
2. Click **Detect boxes**. ppie finds every drawn box on the form and places one for each.

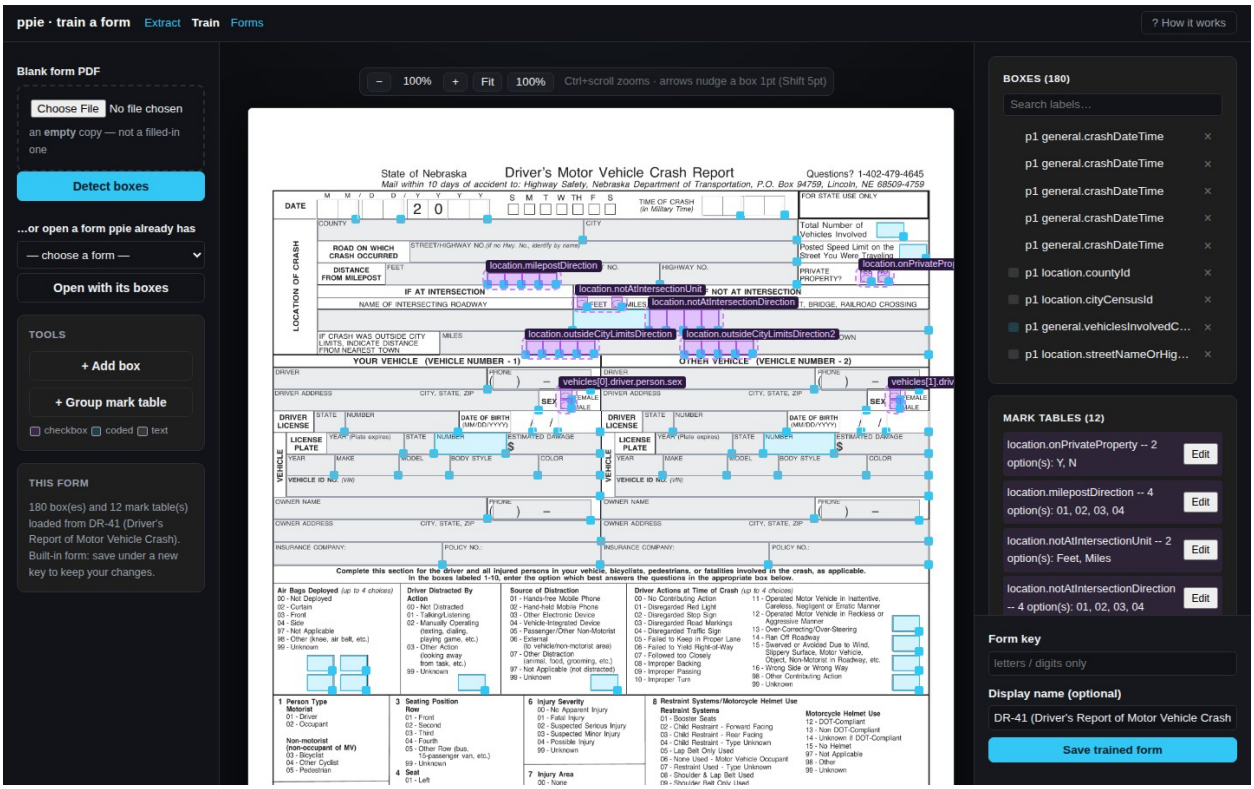


Boxes found on a blank form

Boxes are colored by what ppie thinks they are - **checkbox**, **coded** (a short code box) or **text** - and a dashed border means the box has no field name yet. Fields that are only a line to write on, with no box drawn around them, aren't found; add those yourself (6.5).

6.2 Or open a form ppie already has

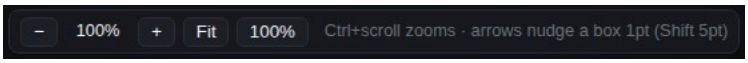
Under **...or open a form ppie already has**, choose a form and click **Open with its boxes** (or use Edit / Copy to train on the Forms screen). Every box and check-box group loads exactly where it was saved, so you adjust instead of starting over.



A form opened with its boxes

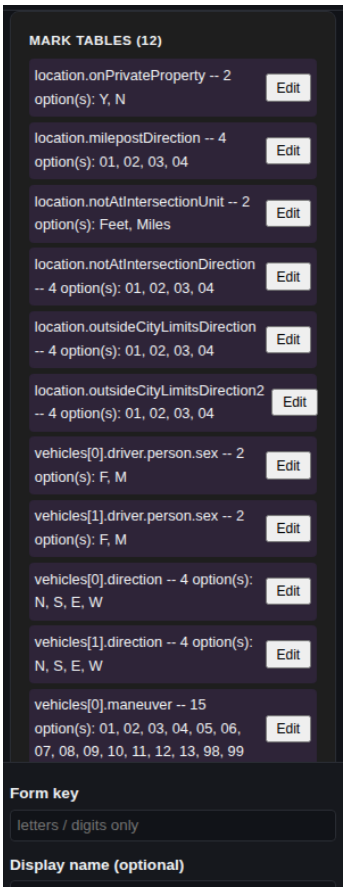
- A **trained** form opens under its own key - Save updates it.
- A **built-in** form opens as a copy: the Form key starts empty, and you save it under a new key. The built-in form never changes.

6.3 Moving around the page



The zoom bar

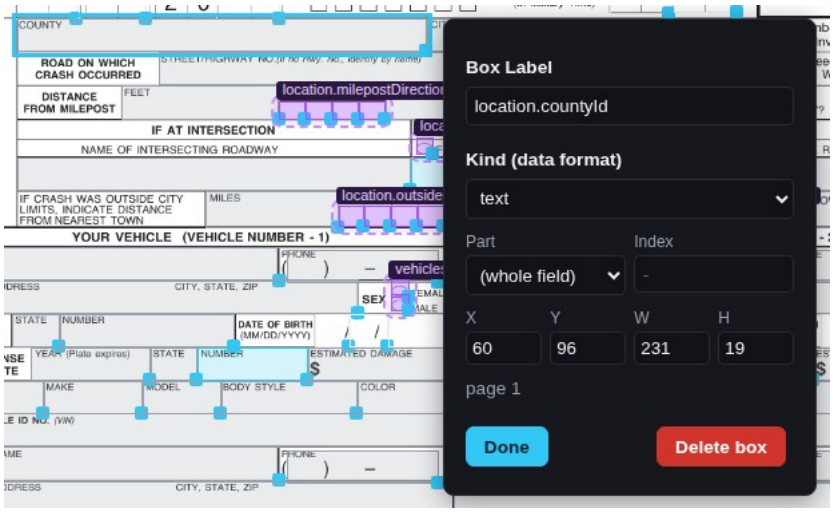
- **- / +** zoom out and in; **Fit** fits the page to the width of the middle area; **100%** is actual size. **Ctrl + scroll** zooms too.
- The **Boxes** list on the right has every box; type in **Search labels** to narrow it, and click a row to jump to that box.



The Boxes and Mark tables lists

6.4 The box editor

Click a box to open its editor beside it:



The box editor

Field	What to put
Box Label	The field's name, in the form your crash system uses - letters and digits, no spaces, a dot between parts, and [0], [1] ... for a repeated item. For example `crashDate`, `driver.licenseNumber`, `vehicles[0].make`. A red note appears while it isn't valid.
Kind	**text** for anything written in words or numbers, **coded** for a short code box (01, 99), **checkbox** for a single box that's marked or not.
Part	Only for one box of a split field: month, day, year, hour, minute, or a phone number's area code / prefix / line. Leave it on (whole field) otherwise.
Index	Only for one slot of an "up to N choices" field: 0 for the first box, 1 for the second...
X, Y, W, H	The box's position and size in points (1/72 inch), for exact placement.

- **Done** closes the editor. **Delete box** removes the box. The **x** on a row of the Boxes list deletes that box too.
- **Drag** a box to move it; drag its blue corner to resize it.
- **Arrow keys** move the selected box 1 point; hold **Shift** to move 5.

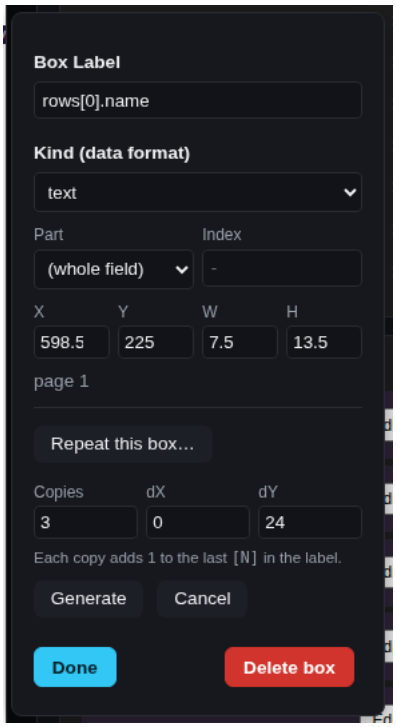
> **Unnamed boxes:** a box with no label is saved but never read. After a test extraction it shows as a dashed grey outline, so you can see what's left to name.

6.5 Adding a box

1. Click **+ Add box** (it turns green while it's on).
2. Drag out a rectangle on the page around the field - or just click, for a standard-size box you can then resize.
3. The editor opens for the new box: give it a label and kind.

6.6 Repeating a box

For a field that repeats down a row or across a column - one line per occupant, say - draw and label the first box, then let ppie draw the rest. The label must end in an index, like `occupants[0].name`.



Repeat this box

1. Select the first box and click **Repeat this box...**.
2. Enter **Copies** (how many more), and **dX** / **dY** - how far apart they are, in points (a positive dY goes down the page).
3. Click **Generate**. Each copy gets the next index: `[1]`, `[2]`, `[3]`...

Each copy is an ordinary box afterwards - move, resize or delete it on its own.

6.7 Check-box groups (mark tables)

A Y/N pair, or a list of numbered options where the officer marks one, is **one field** with several boxes. Make it a group:

1. Click **+ Group mark table**.
2. Click each option's box on the page (or drag out new ones). Each is numbered #1, #2 ... in the panel.
3. Give the group its **Shared path** - the field's name - and, optionally, a **Note**.
4. Type each option's **code** - the value to report when that box is marked (for example Y and N, or 01, 02...). Hover a row in the panel to light up its box on the page.
5. Click **Done**.

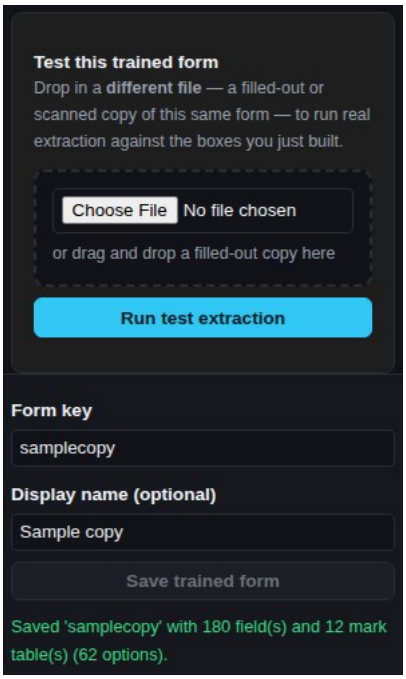


A check-box group being edited

- The group is outlined on the page and listed under **Mark tables**; **Edit** there, or click any of its boxes, reopens it.
- The **x** on an option row removes that option. **Ungroup all** turns the boxes back into ordinary boxes.
- Every option needs a code before Done will accept it.

6.8 Saving and testing

1. At the bottom right, enter the **Form key** - 2 to 64 lowercase letters and digits, no spaces or dashes (for example `acmecrash`) - and, optionally, a **Display name**.
2. Click **Save trained form**. The form is ready to use at once. The button greys out until you change something again.
3. Under **Test this trained form**, drop in a **different** file - a filled-in or scanned copy of the same form - and click **Run test extraction**. It opens the Extract screen's results for that scan (chapter 3).



After saving: test the form

> **Everything came back blank?** the blank you trained on probably wasn't empty (see 6.1). Train again from an unused copy.

7. Sending scans to your crash system

In daily use nobody opens ppie's screens: scans land in a folder, and ppie reads each one and sends it to your crash system. Your administrator sets this up once (the commands are in chapter 8).

7.1 The watched folder

1. Scan each paper report - one crash per file, as a PDF - into the **incoming folder**.
2. ppie checks the folder every 30 seconds, reads each new PDF, and sends its envelope and the scan to your crash system.
3. When it's done with a file, it moves it:
   - into **_processed** - sent, or already in your crash system;
   - into **_failed** - something needs a person (the reason is in ppie's log).

Both folders sit inside the incoming folder, so a file is only ever handled once and nothing is lost.

> **Scanned the same report twice?** no harm done. Your crash system recognizes the same pages and doesn't file them again; ppie counts the scan as done and moves it to _processed.

7.2 One-time batches and dry runs

- A **batch** reads every PDF in a folder once and stops - useful for a backlog.
- A **dry run** reads every scan and reports what it found, but sends nothing. Use it to check a folder of scans, or a new setup, before going live.

7.3 What each result line means

For every file, ppie writes one line, then a total:

Line	Meaning
OK ... **ingested (recordId=...)**	Sent and filed in your crash system as a new report.
OK ... **already in the Refinery: ...**	Your crash system already has this report (a repeat scan). Nothing new filed; nothing to do.
OK ... **dry-run OK - N fields, form=...**	Dry run only: read successfully, not sent.
SKIP ... **extraction failed: ...**	ppie couldn't read the file - not a PDF, damaged, or not a form it knows. Check the scan.
SKIP ... **Refinery did not ingest it: ...**	Your crash system received it but didn't file it - the reason follows (for example "no crash data").
SKIP ... **(connection or HTTP error)**	Your crash system couldn't be reached, or refused the key. Tell your administrator.
N/M done (K already in the Refinery)	The total at the end.

A file that fails never stops the others - the rest of the folder is still processed.

8. Running ppie (IT and administrators)

ppie runs on one of your own machines, as a command-line program and an optional web service. It needs no internet connection and keeps no database: nothing about a scan is kept after it's been read and sent.

8.1 The folder watcher and batches

Command	What it does
`ppie watch <folder> --refinery-url <url> --refinery-key <key>`	Watches the folder for good: reads and sends each new PDF, then moves it to `_processed` or `_failed` inside the folder. `--poll-seconds` changes the 30-second check.
`ppie batch <folder> --refinery-url <url> --refinery-key <key>`	Reads and sends every PDF in the folder once. Add `--move-done` / `--move-failed` to move files afterwards.
`ppie batch <folder> --dry-run`	Reads every PDF and reports, without sending anything.

- The address and key can also come from the settings `PPIE_REFINERY_URL` and `PPIE_REFINERY_API_KEY`, so the key isn't typed on a command line.
- The address is your crash system's OCR intake - for the Refinery, `/api/intake/ocr`. The key is the intake key your crash system's administrator sets.
- `--form <key>` reads every file as one form; leave it out to let ppie recognize each scan.

8.2 The engine service

`ppie serve` (or the supplied container image) runs ppie as a web service, for a system that would rather send scans to ppie than drop them in a folder. It listens on port 8080 of the local machine by default (`--host`, `--port` to change).

Endpoint	What it does
`GET /health`	Is it running, and which reading engine is it using.
`GET /api/v1/forms`	Every form it can read: key, name, revision, pages.
`POST /api/v1/extract`	Send a scan (file field `scan`, optional `form`); the envelope comes back. If `PPIE_EGRESS_URL` is set, the envelope is also sent there - your crash system's own intake.
`POST /api/v1/preview`	The same as extract plus each page's image, for a screen that shows the result. Never forwarded anywhere.
`GET /api/v1/stats`	Counts since it started: scans read, failures, average confidence and time. Never a field value.

> **No sign-in on the service:** the engine service has no passwords or limits of its own by design - it's meant to sit on a private network, behind whatever front door you already use. Don't expose it directly to the internet.

8.3 Where things live

- **Built-in forms** ship inside ppie and change only with a new ppie release.
- **Trained forms** are kept in a forms folder (`PPIE_FORMS_DIR`, or `.ppie/forms` in the user's home folder) - one folder per form. Back it up with your other configuration.
- **The reading engine** is included in ppie and checked every time it starts, so it never has to be downloaded.

9. When something looks wrong

What you see	Likely cause	What to do
Every field blank on a trained form	The form was trained on a filled-in copy, so everything already written was subtracted away.	Train again from an unused blank (section 6.1).
Couldn't identify which form this is	The first page's title is missing, cut off, or too crooked to read - or it isn't a form ppie knows.	Pick the form yourself on the Extract screen, or rescan the first page straight. A new form needs training (chapter 6).
Registration weak warning	The page couldn't be lined up well with the blank form - badly skewed, stretched, cropped, or a different edition of the form.	Rescan the page flat and straight, including all four edges. Check readings on that page closely.
Readings are the printed labels, not what was written	The page lined up only roughly, so the printed text wasn't fully removed.	As above - a straighter, fuller scan.
Lots of amber and red on a hand-filled form	Real handwriting is harder to read than print - faint, crowded, or across the box lines.	Expected today; the reviewer confirms these (chapter 11). A darker pen and writing inside the boxes help.
An occupant-grid row skipped warning	That row's Vehicle No. or Occupant No. couldn't be read, so ppie wouldn't guess whose row it was.	The reviewer enters that person from the scan.
A scan in **_failed**	See the reason in ppie's log (chapter 7).	Fix and drop it in the incoming folder again.

> **Scan settings that help:** 300 dpi, black and white or grayscale, the whole page including its edges, one crash per PDF.

10. Quick reference

10.1 Labels you'll see

Label	Where	Meaning
Green / amber / red outline	Extract page	Confidence 0.80 and up / 0.50-0.80 / under 0.50 or blank.
Dashed grey outline	Extract page	A trained box with no field name yet - not read.
Detected as / **Read as**	Extract page	The form ppie recognized / the form you picked.
ocr / **mark** / **blank**	Results list, method	Read as text / a marked box or option / an empty box.
checkbox / **coded** / **text**	Train screen colors	What kind of box it is.
Dashed border	Train screen	A box with no label yet.
built-in / **trained**	Forms screen	Ships with ppie / taught on this machine.

10.2 Messages

Message	What to do
You selected ... but the scan's own title reads as ...	Nothing - ppie read it as the form it really is.
page N registration weak	Rescan that page straighter; check its readings (chapter 9).
occupant-grid row (slot N) has an unreadable or ambiguous Vehicle No./Occupant No. - skipped, not guessed	The reviewer enters that row from the scan.
... routed by packet order ... which may be misrouted	A supplement's Vehicle No. / Non-Motorist No. box couldn't be read; check which vehicle or person it belongs to.
Couldn't identify which form this scan is from its first page	Pick the form, or rescan the first page.
Unsupported file type	Use a PDF, or a PNG, JPG, TIFF or BMP image.
Form key must be 2-64 lowercase letters/digits only	No spaces, dashes or capitals in a form key.
... is a built-in form key - choose another	Save your copy under a new key.
Box path ... isn't valid	Fix the label: letters and digits, a dot between parts, [N] for a repeated item.
Every option needs a code	Type a code for each option in the group.
No boxes to save	A form needs at least one ordinary (non-group) box.

10.3 Glossary

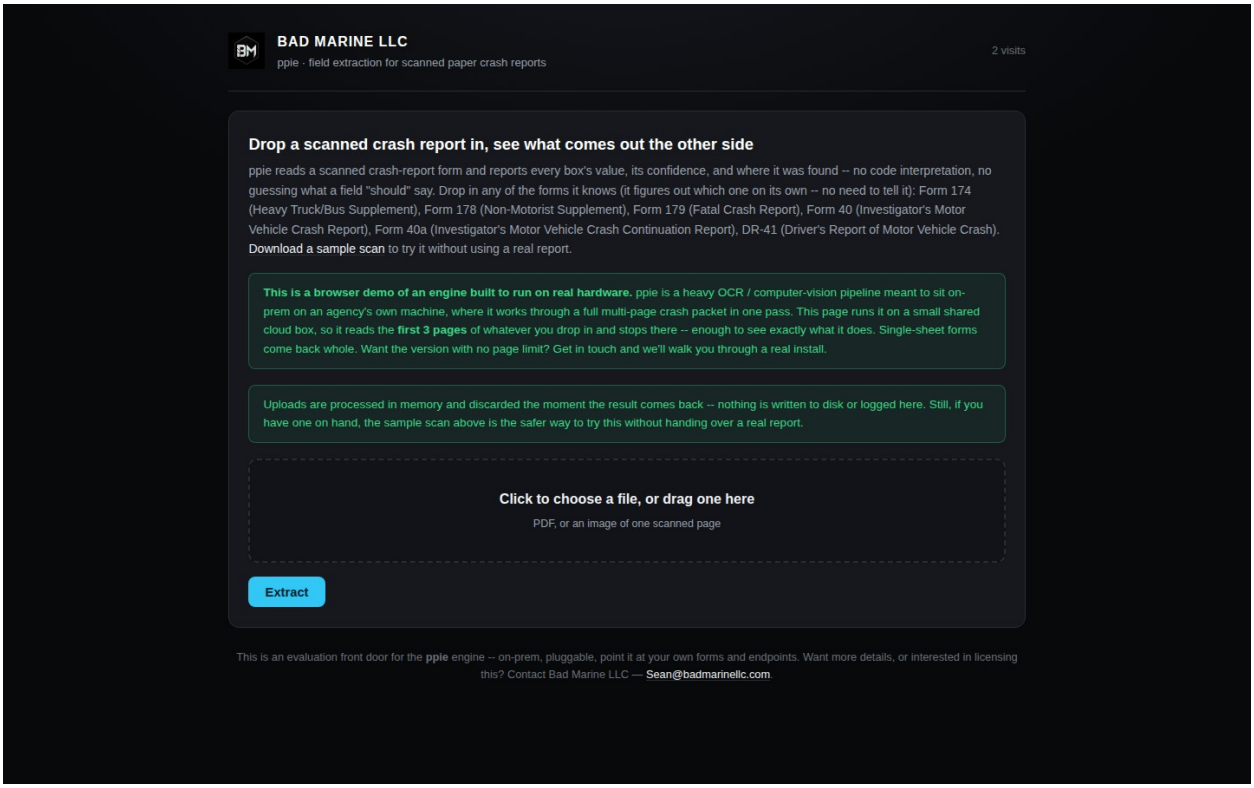
Term	Meaning
Envelope	What ppie produces for one report: every field's reading, confidence and location, and the page fingerprints.
Confidence	How sure ppie is of a reading, 0 to 1.
Blank form	An unused copy of a form - what ppie lines each scan up against.
Registration / lining up	Matching a scan to the blank form so every box lands on its field.
Packet	All the pages of one crash report - the main form plus any supplements - in one file.
Path / label	A field's name, as your crash system knows it (for example `vehicles[0].make`).
Mark table / check-box group	Several boxes that together answer one field; the marked one gives the value.
Form key	A form's short internal name - lowercase letters and digits.
Crash system	The system that takes ppie's envelopes in and puts them in front of reviewers (for example the Refinery).

11. Known limits

- **Handwriting** reads less reliably than print today. The reading engine is general-purpose and hasn't yet been trained on crash-form handwriting; training it on a state's own reviewer-confirmed scans is how that improves, and that training step isn't built into ppie yet. Meanwhile every value carries a confidence, and every value is confirmed by a reviewer.
- **Form 40's occupant grid**: a row whose Vehicle No. or Occupant No. can't be read is skipped rather than guessed.

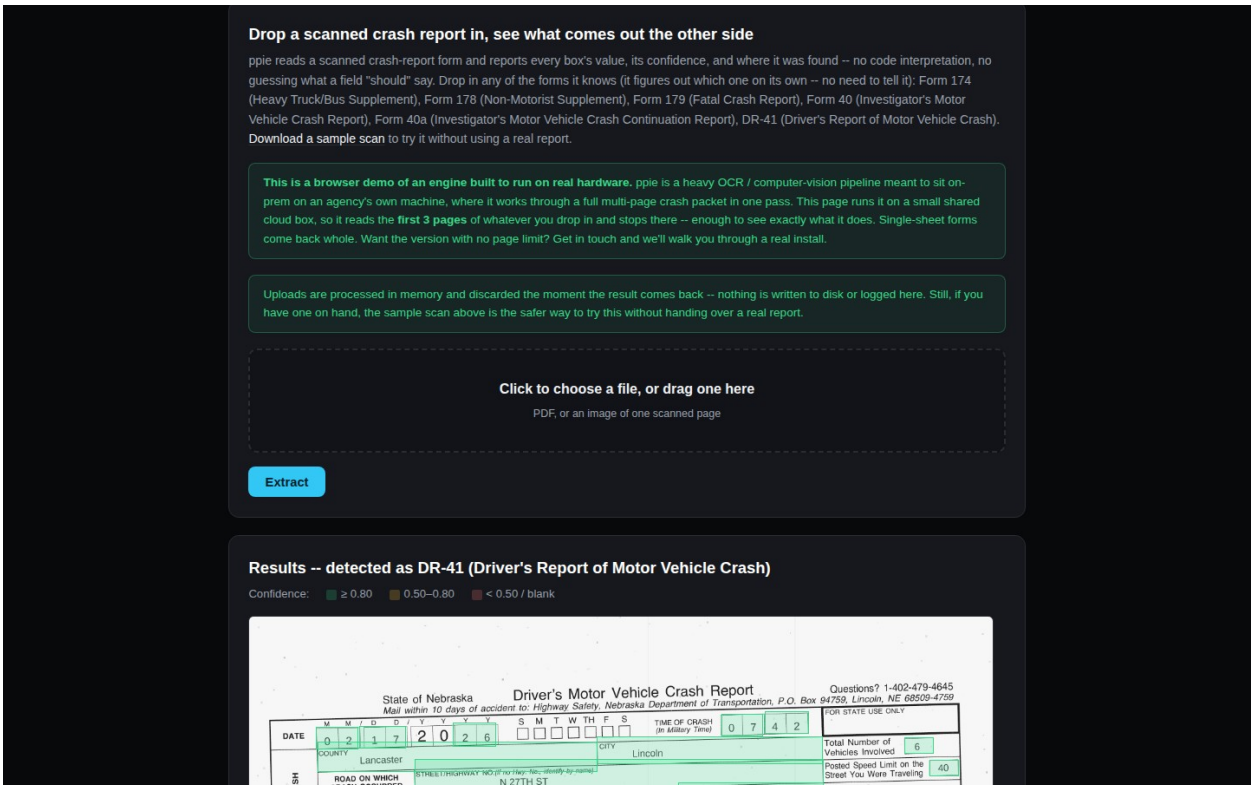
12. Trying the online demo

The online demo lets anyone try ppie in a web browser, with nothing to install. It runs the same reading engine as an installed system, on a small shared cloud machine.



The online demo

- **Download a sample scan** - a made-up, filled-in driver's report - if you don't have a scan of your own to try.
- **Drop a scan in** (a PDF, or an image of one page) and click **Extract**. There's no need to say which form it is.
- The result shows the form it recognized, each page with every box outlined in its confidence color, and the full list of readings.



A result on the online demo

The demo's limits, which an installed system doesn't have:

- **The first 3 pages** of a file are read, then it stops - enough to see what ppie does. A single-sheet form comes back whole; a full packet is cut short, and the result says so.
- **5 free extractions** per visit. An unlock code from Bad Marine LLC removes the limit.
- It can take longer than an installed system, because the machine is small and shared.

> **Your scans:** uploads are read in memory and discarded as soon as the result comes back - nothing is saved or logged. The sample scan is still the safer way to try it than a real report.