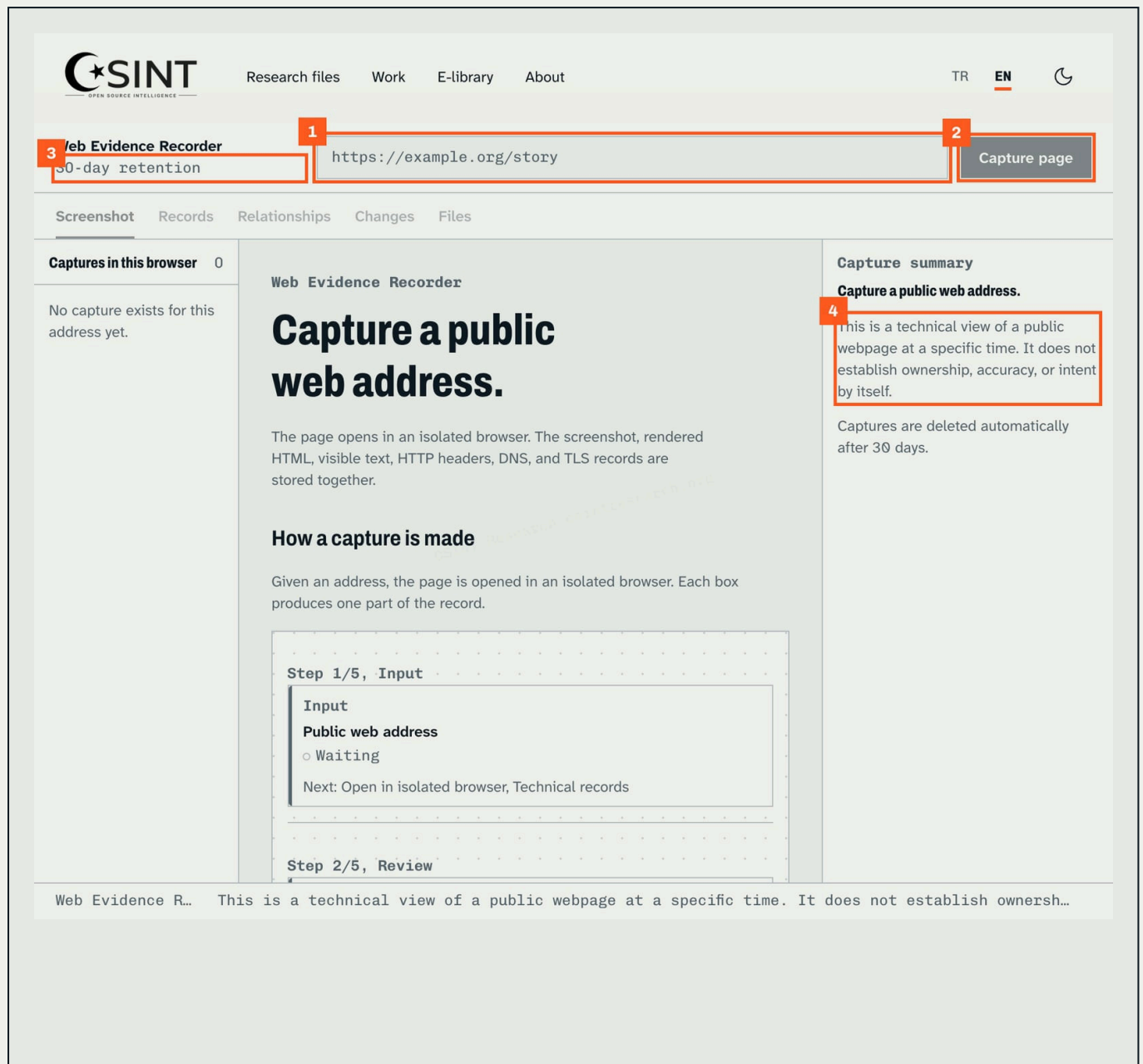




How the web evidence recorder is built

The web evidence recorder keeps a public web page as it was at one moment, with its source address, date and checkable digests. This handbook explains where the page is opened, what is kept where, and why I added each safeguard.

1 Why it needs a server

Most tools on the site run in your browser and your file never leaves your device. This one is different. To record a web page as evidence, the page has to open in a clean browser that is separate from yours. Otherwise your session, extensions and cookies would mix into the record.

The screenshots come from earlier captures. I did not create new captures while updating this handbook.

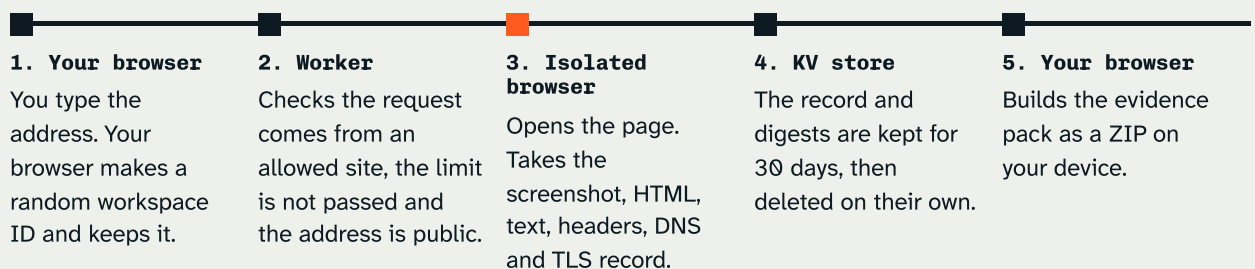
So a separate Worker running on Cloudflare opens the page. You give the address, the Worker opens the page in an isolated browser, collects the records and computes the SHA-256 digest of every file.

Step-by-step use of the tool is in the Web evidence recorder field guide. This one covers only the structure.

A few words

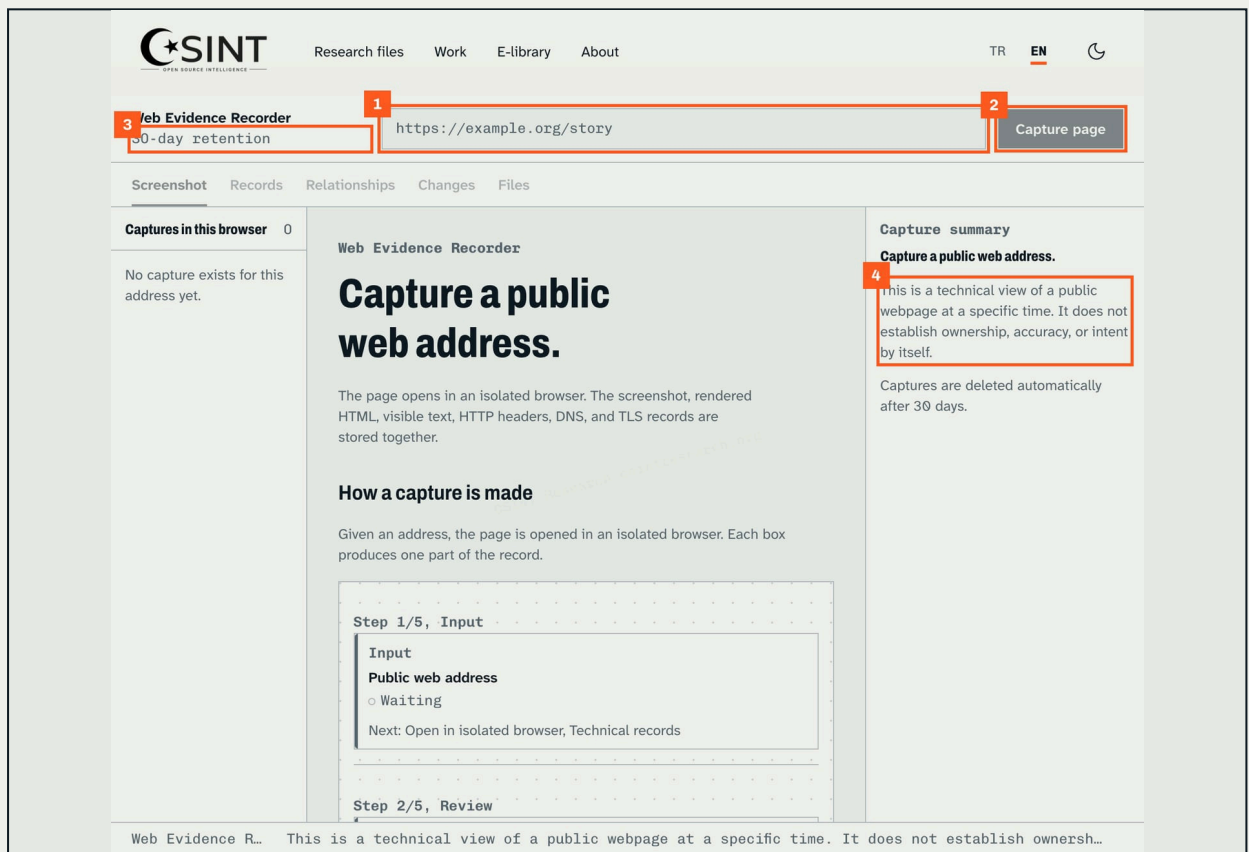
Worker	A small server program that runs on Cloudflare. The site is static, live work happens in Workers.
Isolated browser	A headless browser opened fresh for each capture and closed when it is done. The Cloudflare Browser Rendering service.
DNS	The system that turns a domain name into an IP address.
TLS	The site's encrypted connection and its certificate.
SSRF	An attack that tricks a server into sending requests to addresses inside its own network.

2 What runs where



Stays in the browser	The workspace ID. The Worker knows which captures are yours from this ID. No account.
Sent to the Worker	Only the address, the workspace ID and the language. The request body is at most 4096 bytes.
Stored on the server	Screenshot, rendered HTML, visible text and a manifest. The address and workspace ID are not written in plain form in the storage keys, their digests are used.
Time	Every capture is deleted after 30 days. At most 10 captures of history are kept per address.
Limit	6 captures a minute per workspace, 30 a minute overall.

3 The page frame



The web-evidence page before any capture. The screenshot, records, relationships, changes and files tabs open after the first capture.

1 Address

http or https only, up to 2048 characters. An address with a user name or password is not accepted.

2 Capture page

The moment you press it, the address goes to the Worker. There is no separate consent box. The button is the consent.

3 30-day retention

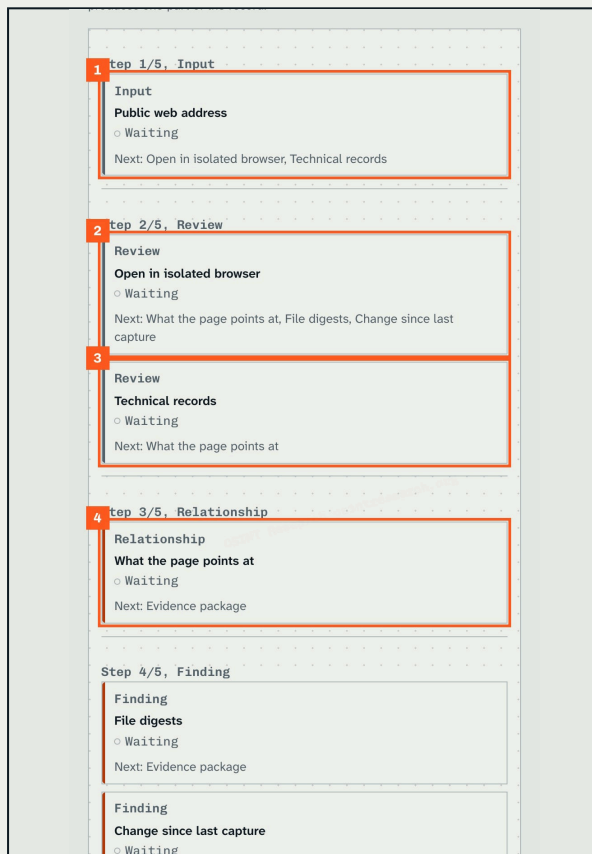
The retention period is written at the top. You can also delete a capture yourself before that.

4 What a capture is

A technical view at one moment. It does not establish ownership, accuracy or intent by itself.

4 A capture in five steps

The page shows how a capture is made with boxes. Each box produces one part of the record. When you select a box, it opens what it takes, what it does, what it produces and what it does not prove.



The method graph. Steps 1 to 3 and the start of step 4.

1 Input

The address and its DNS record are checked. Pages behind a login or closed pages are not captured.

2 Open in isolated browser

Opened at 1440 pixels wide, redirects are followed. If the page does not settle within 32 seconds, that is written as a note and the capture still goes ahead.

3 Technical records

HTTP headers, the redirect chain, DNS and the TLS certificate. Cookie and authorisation headers are not written.

4 Links on the page

Links, domains, emails and files are separated. Then come file digests, the difference from the previous capture and the evidence pack.

Review	Open in isolated browser
Takes	The checked address.
Does	Opens the page in an isolated browser, follows redirects, takes the image and the text.
Produces	1 Does not prove On very long pages the image stops at the first 12,000 pixels.
	Screenshot, rendered HTML, and visible text.

Details of the open in isolated browser box.

Locating

Finding File digests

Takes

The stored image, HTML, and text files.

Produces

One digest per file.

Does

Computes a SHA-256 for each evidence file.

2

Does not prove

A digest shows the file did not change, not that its content is true.

Details of the file digests box.

1 Limit of the image

On very long pages the screenshot stops at the first 12,000 pixels. HTML is capped at 1.5 MB and text at 750,000 characters.

2 Limit of the digest

SHA-256 shows the file has not changed since. It does not show the content is true.

5 Safeguards I put in the Worker

Letting a server open any address you give it is a risky feature. Someone could try to use it to reach Cloudflare's or someone else's internal network. So the Worker checks at every step.

- 1 Requests are accepted only from allowed sites.
- 2 Only ports 80 and 443. Internal names such as localhost, .local, .internal, .lan and names without a dot are refused.
- 3 The domain is resolved over DNS and a CNAME chain is followed for at most 5 hops. If even one resulting address is in a private or reserved range, nothing is captured.
- 4 Every sub-request made from inside the page goes through the same check. Requests other than GET, HEAD and OPTIONS are stopped.
- 5 Pop-up dialogs are dismissed. Each capture opens in its own browser session, which is closed when it is done.

6 Comparison and the pack

If the same address is captured again from the same workspace, the Worker compares the new capture with the previous one. Status code, title, canonical address and the three file digests are compared one by one. Added and removed links, domains, emails and files are listed. Text similarity is worked out over the first 10,000 words.

The evidence pack is built in your browser. It holds a screenshot, HTML, text, the capture record, a README and a list that computes each file's digest again. The pack is saved to your computer. Download it before the 30 days are up.

7 What it does not do

- ✗ It does not capture pages behind a login, a paywall or otherwise closed.
- ✗ It does not prove the content is true, complete or legally conclusive.
- ✗ A link on a page does not show ownership or cooperation.
- ✗ This page does not monitor on a schedule. Each capture is taken when you press the button.
- ✗ Captures are deleted after 30 days. Download the pack for any capture you want to keep longer.

Think of a capture like a witness statement. It says what was seen at that address at that moment. What it means needs separate research.