LAION-fMRI stimulus — Privacy notice

Version 2026-05-12

1. What we collect

When you submit the access form (in a browser or via the `laion_fmri` Python package) we record:

- full name;
- institutional email;
- institution / affiliation;
- PI / supervisor (optional);
- a free-text research-purpose description;
- the version of the Data Use Agreement you accepted and a timestamp of acceptance;
- a pseudonymised access-request identifier (HMAC of the raw identifier; the raw identifier is shown to you once and stored locally by the package, never persisted on our server).

For each download we record:

- which request id requested which files at what time;
- when the resulting download links expire.

The application database does **not** store IP addresses or user-agent strings. Rate limiting uses short-lived, in-process counters that are dropped after the rate-limit window. Separately, the webserver retains short-lived access logs for security monitoring as described in §3.

Providing the form data is not a statutory requirement, but it is a **contractual prerequisite** for receiving the stimuli: without a recorded acceptance of the Terms of Use we cannot give you the files. Decline simply means no access.

The website does not use cookies, local storage, browser fingerprinting, third-party analytics, or any tracking technology in scope of § 25 TDDDG.

2. Why we collect it (lawful basis)

The legal basis for accepting the Terms of Use and providing the stimuli to you is **Art. 6(1)(b) GDPR** (performance of a contract / pre-contractual steps at your request). The legal basis for the controlled-access audit log (download records, governance trail) is **Art. 6(1)(f) GDPR** — legitimate interests in maintaining the integrity of a controlled-access research dataset, responding to rights-holder

takedowns, and detecting abuse. The Art. 89(1) GDPR safeguards for scientific research apply, alongside § 27 BDSG.

We do not rely on consent (Art. 6(1)(a)) for this processing because acceptance of the Terms of Use is a precondition for the service rather than freely revocable consent.

No automated decision-making in the sense of Art. 22 GDPR takes place.

3. Recipients and international transfers

Your form-submission data is stored on the project's server hosted in Germany. The underlying virtual server is provided by **STRATO AG**, Pascalstraße 10, 10587 Berlin, acting as a processor under Art. 28 GDPR on the basis of STRATO's standard Auftrags verarbeitungsvertrag (AVV) — available at strato.de/datenschutz. The application database is accessible only to authorised project maintainers.

The webserver retains short-lived nginx access logs (request line, response status, timestamp, IP address, and, depending on server log format, user-agent string) for **14 days** for the sole purpose of security monitoring and abuse detection (Art. 6(1)(f) GDPR). These logs are not joined with the application database and are rotated automatically.

The stimulus files themselves are stored in Amazon S3 in the AWS **us-west-2** region (Oregon, USA), operated by Amazon Web Services Inc.; the EEA contracting entity for the account is Amazon Web Services EMEA SARL (Luxembourg). When you download a file, your IP address is processed by AWS as part of the HTTP request; we do not enable S3 access logging, CloudFront, or AWS WAF, so AWS only receives the operational telemetry covered by its standard service-level processing. AWS does not receive your name, email, or any other form-submission data.

The transfer to the USA is covered by the EU-US Data Privacy Framework adequacy decision under Art. 45 GDPR (Amazon Web Services is certified under the DPF), with the Standard Contractual Clauses in AWS's standard Data Processing Addendum as a fallback safeguard (Art. 46 GDPR). A copy of AWS's DPA is available at aws.amazon.com/compliance/data-protection.

4. Retention

Submission records and the download audit log are retained for **up to ten years from your last activity**, under the Art. 89(1) GDPR safeguards for scientific-research data processing. The 10-year window reflects the typical lifetime of a research dataset release plus a reasonable period for post-publication queries

and rights-holder takedowns; we will often delete or pseudonymise earlier in practice (for example when you revoke your access, or via the periodic cleanup job that drops identifying fields from long-inactive records).

Revoked requests are pseudonymised immediately: the identifying fields (name, email, institution, supervisor, research-purpose) are dropped, leaving an audit row tied only to the pseudonymous request-id hash. After the 10-year window the audit row itself is deleted.

You can request earlier deletion at any time — see §5.

5. Your rights

You have the rights under Arts. 15-21 GDPR, namely:

- access to your stored personal data (Art. 15);
- rectification of inaccurate data (Art. 16);
- erasure (Art. 17); we typically fulfil this through pseudonymisation of the identifying fields while keeping a minimal audit record under Art. 17(3)(d) GDPR (necessary for scientific-research archiving and for the controller's legal-defence interest);
- restriction of processing (Art. 18);
- data portability (Art. 20): on request we will export your submission record (name, email, institution, supervisor, research purpose, accepted terms version, timestamp) as a machine-readable JSON file;
- objection to processing carried out under Art. 6(1)(f) (Art. 21);
- revocation of your access to the stimuli at any time.

Send requests to the controller at the address listed at the foot of this document. We respond within one month (extendable by two months for complex requests, Art. 12(3) GDPR).

You also have the right to lodge a complaint with a supervisory authority. For us this is **Der Hessische Beauftragte für Datenschutz und Informationsfreiheit** (HBDI) in Wiesbaden, datenschutz.hessen.de. Users in other EU / EEA states may also complain to their local supervisory authority.

6. Free-text research-purpose field

Please use the research-purpose description on the access form to describe your own research plan only. Do not include personal data of third parties beyond your supervisor's name — no patient names, collaborator emails, or other identifying information about people other than yourself.

7. Security

Connections to the access service are encrypted with TLS (Let's Encrypt). Access-request identifiers are stored only as HMAC hashes keyed with a server-side secret; the raw identifier is shown to you once and held only by the loader package on your machine. Server logs are short-retention (see §3) and do not record sensitive fields.

8. Controller

The data controller for this service is **Prof. Dr. Martin Hebart** (acting as a natural person under Art. 4(7) GDPR for the LAION-fMRI dataset release). No Data Protection Officer is mandatory under Art. 37 GDPR for a private-person controller running a single research dataset release of this kind.

Controller: Prof. Dr. Martin Hebart. Contact: martin.hebart@uni-giessen.de. Imprint / Impressum: hebartlab.com/impress.