

Data for what you're building.

Tell us what your model needs.
We'll work with our data partners to find it.



Data we can help source

• Off-the-shelf & bespoke requests

Code & software

Private repositories, COBOL, commit history, pull requests and code metadata.

Images & video

Objects, scenes, activities, visual inspection and annotated media.

Documents & OCR

Scans, handwriting, forms, invoices, receipts and document layouts.

Finance & economics

Prices, filings, economic indicators and aggregated transactions.

Industry, IoT & telecom

Equipment telemetry, time series, defects and network performance.

Cybersecurity & fraud

Security events, threat indicators, fraud patterns and labelled detection data.

Text & language

Multilingual corpora, conversations, Q&A, domain text and translations.

Egocentric & human action

First-person video, hand interactions, daily activities and motion capture.

Healthcare & life sciences

De-identified clinical records, medical imaging, dictation and research data.

Retail & commerce

Product catalogues, pricing, reviews, inventory and shopping trends.

Climate, energy & agriculture

Weather, emissions, energy demand, crop imagery and soil measurements.

Synthetic & multimodal

Generated text, images, scenes and tabular data; paired audio, video and text.

Audio & speech

Speech, accents, dialogue, transcripts, music and ambient sound.

Robotics, 3D & autonomy

Demonstrations, trajectories, depth, LiDAR, point clouds and sensor streams.

Geospatial & earth observation

Satellite imagery, maps, points of interest and land-use data.

Business & market intelligence

Company profiles, market research and aggregate audience insights.

Transport, property & travel

Mobility patterns, logistics, property listings and travel demand.

AI training & evaluation

Instruction tuning, preferences, agent tasks and expert evaluations.

More domains on request: education, legal, science, public sector, demographics, sports, media and gaming.

Start with your requirement.

Share the data type, intended use, volume, geography, language and timeline.
We'll coordinate catalogue options, sample enquiries and quotes.

Talk to Forge ↗

Categories describe sourcing scope through our partners. Availability, samples, provenance, permitted use and delivery are confirmed per request. Exclusive or non-exclusive licensing is subject to the dataset and agreed terms.