DOTCHECK

Technical Report

How DotCheck scores images, text, video, and audio — architecture, pipelines, and measured gates

VERSION	Docs v2026.7
DATE	26 July 2026
ENGINES	`Vermeer · Valla · Muybridge · Helmholtz`

The figures in this document come from held-out test sets. They estimate how AI-like content looks under our models — they are not proof of authenticity.

Petr Jaroch (trading as DotCheck), IČO 03330389, Jenikovice 2, 535 01 Prelouc, Czech Republic

dotcheck.ai/docs

AT A GLANCE

Vermeer (images, `inhouse@12`): real average **0.000** · AI average **0.938** · balanced accuracy **0.9950**
Valla (text English, `inhouse-text@9`): human average **0.045** · AI average **0.909** · balanced accuracy **0.980**
Valla also for Spanish, Portuguese, French, Italian, German, and Dutch — each head clears the same pass floors as English; results in the language table below
Muybridge (video frames, `inhouse-video@2`): bag p90 real **0.004** · AI **0.940** · balanced accuracy **0.985**
Helmholtz (audio, `inhouse-audio@2`): real average **0.004** · AI average **0.984** · balanced accuracy **0.986**

1. What DotCheck is

DotCheck helps you answer a simple question: does this look AI-made? It is a Chrome extension and a website. As you browse, or when you upload a file, you get a readable score for images, videos (frames, and soundtrack when present), standalone audio on Dashboard, and text in supported languages.
Everyday scoring always uses DotCheck's own models. Free, Premium, and Pro share the same models — there is no hidden "better model" behind a picker. Commercial detectors (Sightengine for media, Winston for text) run only when a Pro user explicitly asks for a second opinion — never as a silent substitute for Free or Premium.
This report explains how the system is wired, how each scoring pipeline works, how we measure the models, and where the limits are. It is for a technical reader. It is not a training cookbook and not a legal authenticity opinion.

2. What a score means

Each score is a probability between 0 and 1. Higher means the content looks more AI-generated under the model for that media type. The interface may group scores into soft bands such as "likely human" or "likely AI." Those bands are only a helpful reading of the number.
A score is an estimate on the kinds of content we tested. It is not a certificate of authenticity, not a courtroom deepfake verdict, and not enough on its own for high-stakes automated decisions about a person.

3. System architecture

Your browser talks to our API (Express). The API scores content with the in-house inference service, can return a cached result when the same content comes back, and enforces fair-use limits. Sign-in and billing use Firebase and Stripe. Optional Pro confirms call Sightengine or Winston only after you ask.

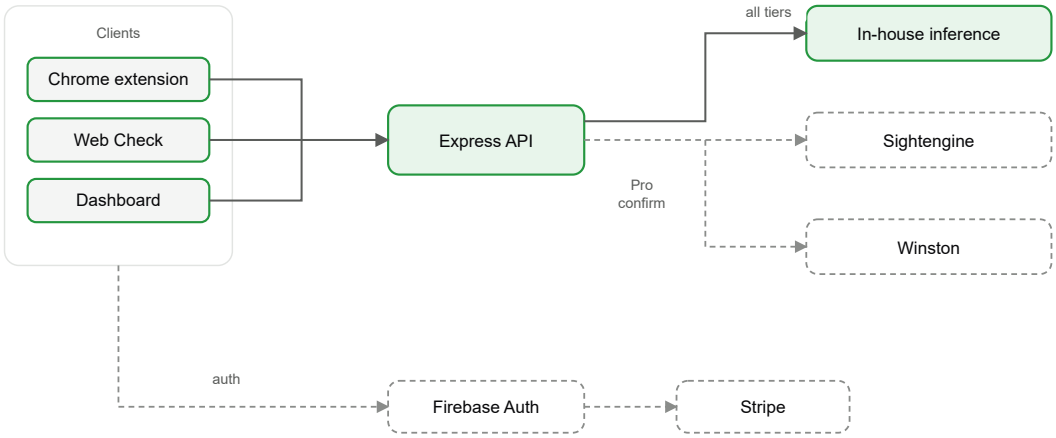

Figure 1 — Runtime path: clients → Express → in-house inference; Pro confirm vendors are explicit-only.

You meet DotCheck in three places:

- **Chrome extension** — dots on images (and video frames where we can score them), underlines on supported-language text, hover for a percent.
- **Check on the web** — upload an image or short MP4 and get the same kind of score; optionally build a short Share clip from that real score (never from a number you type in).
- **Dashboard** — a signed-in workspace for media, text, documents, history (Premium+), Share, and Pro confirm.

Where models run. Everyday scoring uses DotCheck-operated inference hosts. The extension and website do not download model weights as the normal product path. Live heads are also published under Apache-2.0 on Hugging Face. If inference is unavailable, the API returns a clear error (typically HTTP 503). It does not quietly send Free or Premium users to a paid vendor.

Two backend planes. Scoring, hash cache, and fair-use counters live on Express with MongoDB and the inference service. Identity and billing live on Firebase Auth, Firestore tier fields, and Cloud Functions that talk to Stripe.

3.1. Packages and responsibilities

PACKAGE	OWNS
Chrome extension	DOM scan, dots/underlines, page aggregate, local cache keys; talks to Express
Express API	Routes, tier policy, FUP, hash cache, Share jobs, Pro confirm gating
Inference service	FastAPI scorers; loads npz heads; CPU serve; clear 503 if a required head is missing
Website	Check, Dashboard, Share UI, docs PDFs, marketing pages
Cloud Functions	Stripe checkout, webhook, portal, subscription status

Tier wiring uses `subscriptionTier`: `free` | `premium` | `pro`. Express boots only when `TWO_TIER_ENABLED=true`. Legacy Firestore/Stripe ids such as `proplus` / `unlimited` map to `pro`.

3.2. Quotas that affect scoring

COUNTER	LIMIT	NOTES
Extension free / UTC day	30 signed out · 300 Google Free	Surface `extension`; soft wall when exhausted
Check free / UTC day	6 analyses · 3 Shares	Surface `check`; guest quota by public IP
Dashboard free / UTC day	30 analyses incl. 1 doc (3 pages)	Surface `dashboard`; signed-in account
Premium/Pro analyses / UTC day	Premium 3 000 / Pro 30 000 fair-use FUP	Same three models; tall fair-use wall
Share clips / UTC day	Free 3 · Premium/Pro unlimited	Separate from analyze FUP; never invents a score
Pro confirms / UTC month	300	Sightengine + Winston pooled; media confirm = 1 unit even if multiple frames

Batch caps: free up to 10 items per request; Premium/Pro up to 30. Text outside the supported set returns `unsupported_language` without burning free daily analyses.

3.3. Primary HTTP surfaces

ROUTE	ROLE
`POST /analyze-image`	In-house image score; Pro may set confirm retest
`POST /analyze-text`	Language-gated text score (7 langs); Winston confirm on Pro ask
`POST /analyze-video` / frames	Up to three frames; video-profile head
`POST /analyze-document`	Premium+ PDF/DOCX; in-house text + capped images
`POST /upload-and-analyze(-video)`	Website Check upload path
`POST /share` + status/download	Share job from a real in-house score
`GET /usage`	Remaining analyses, Shares, Pro confirms
`POST/GET/DELETE /history`	Dashboard library metadata; Premium+

Health checks expose whether inference is reachable. Operators should treat `inference.ok=false` as a hard stop for in-house scoring — not as a cue to route Free traffic to paid vendors.

4. Scoring pipelines

This section walks the request path for each modality. All tiers use the same in-house models; only Pro can branch into an optional commercial confirm after an in-house score already exists.

4.1. Image path

An image request hashes the cleaned bytes. On a cache hit for that `(clean_hash, engine)` pair, Express returns the prior score without calling inference again. On a miss, inference embeds the image with SigLIP2,

applies the trained linear head stored in the npz artifact, applies stored temperature calibration (sigmoid of logit divided by temperature), and returns P(AI). The npz also stores the decision threshold used for soft UI bands and gate evaluation.

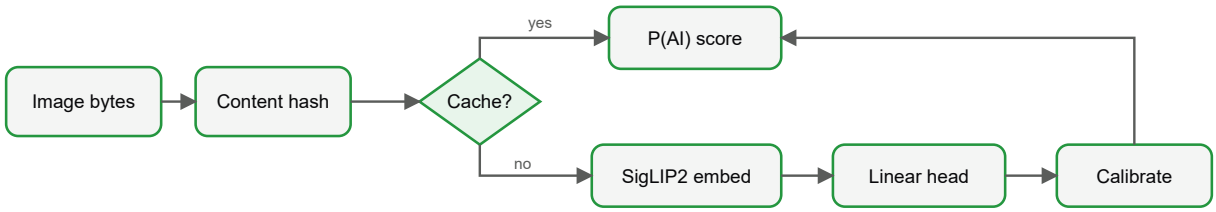

Figure 2 — Image pipeline: hash/cache, SigLIP2 embedding, linear head, calibrated score.

4.2. Text path

Text scoring supports English, Spanish, Portuguese, French, Italian, German, and Dutch. Express detects the language first. Unsupported input receives `unsupported_language` and does not burn the free daily analysis allowance. Supported text shares one TMR + Fakespot feature pair on the inference host; a **language-specific** logistic head maps those features to P(AI). Portuguese (`pt`) uses the Brazilian-trained text model.

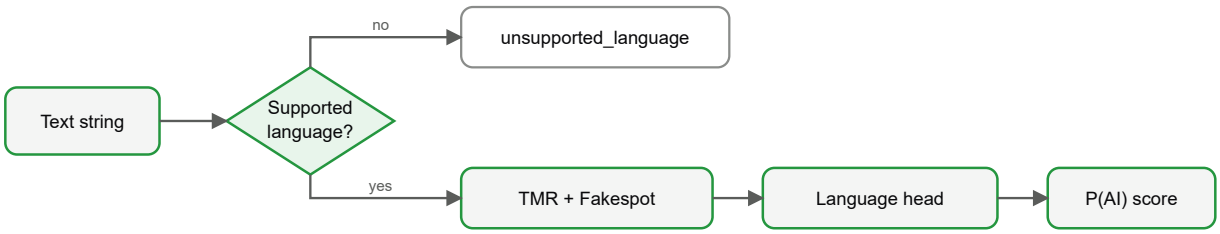

Figure 3 — Text pipeline: supported-language gate, shared TMR+Fakespot features, language logistic head.

4.3. Video path

Short video is scored as frames (not a full temporal deepfake timeline). The product typically samples frames (VS3 densify when duration warrants), scores each with Muybridge (SigLIP2 head under `profile=video`, wire id `inhouse-video@2`), and aggregates with a bag p90. When a soundtrack is present, Express may also score short audio windows with Helmholtz and fuse `max(frame, audio)`. Silent clips stay on frames. If the video head is missing, scoring stops with a clear error — we do not silently reuse the still-image head.

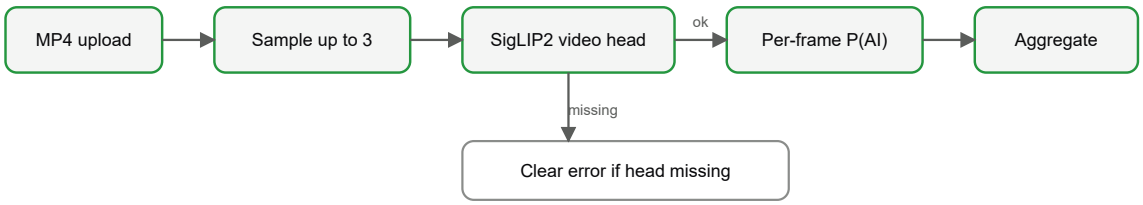

Figure 4 — Video pipeline: frame sample, video-profile head, aggregate.

Uploads for Check and Share are capped (for example MP4 up to 60 seconds and 25 MB). Share clips are built only from a real in-house score.

4.4. Audio path

Standalone audio uploads (Dashboard, Check, or the extension popup) hash the bytes and call inference `/v1/analyze-audio` (Helmholtz / `inhouse-audio@2`). Cache key is `(sha256(raw), inhouse-audio@2)`. On a miss, Dasheng-Base embeds protocol windows and a logistic head returns P(AI); multi-window uses max. Missing audio head ⟶ HTTP 503. Video soundtrack windows reuse the same `audio_analysis` cache when the wav hash matches.

4.5. Documents

Premium and Pro can score PDF/DOCX documents. Document scoring inherits the text and image models: extracted text goes through the text path; a small number of page images (capped) go through the image path. Page and length limits apply (`page_limit`, `text_too_short`, and related clear errors).

4.6. Pro confirm path

A confirm is a retest the user chooses after they already have an in-house score. It spends monthly confirm quota (300 per UTC calendar month for Pro) before the vendor call, and it does not use the free daily analysis allowance. Free and Premium never call Sightengine or Winston.

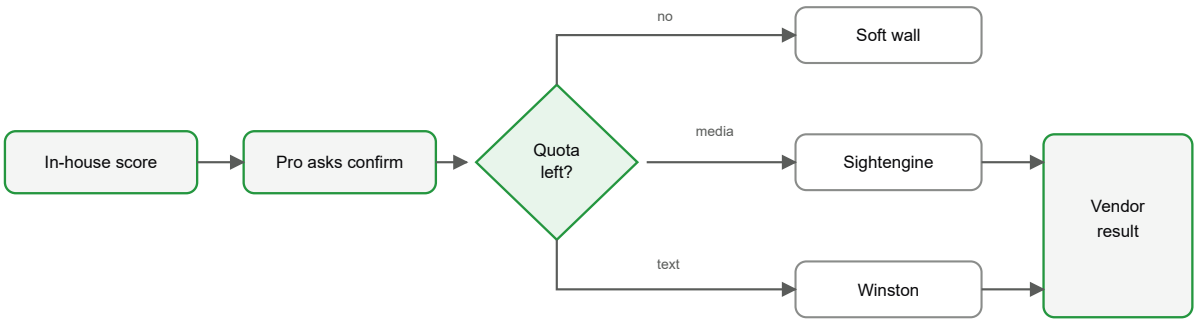

Figure 5 — Pro confirm: explicit retest after in-house score; vendor results are separate from the gate tables below.

Media confirms may include generator hints from Sightengine; text confirms return a Winston score. Those vendor results are commercial second opinions — they are not part of the in-house gate tables in this report.

5. Engine matrix

Public names (Vermeer, Valla, Muybridge, Helmholtz) are marketing labels; wire ids below are the API and cache keys.

SLOT	ENGINE	ARTIFACT / STACK
Image	Vermeer (`inhouse@12`)	Adapted SigLIP2-base + `siglip2_…_linear_head_v12lora.npz`
Text EN	Valla (`inhouse-text@9`)	Shared TMR + Fakespot + `text_stack_head_v9.npz`
Text ES	Valla (`inhouse-text-es_v2`)	`text_stack_head_es_v2.npz`

SLOT	ENGINE	ARTIFACT / STACK
Text PT	Valla (`inhouse-text-pt_BR_v1`)	`text_stack_head_pt_BR_v1.npz` (`pt` → this head)
Text FR	Valla (`inhouse-text-fr_v2`)	`text_stack_head_fr_v2.npz`
Text IT	Valla (`inhouse-text-it_v1`)	`text_stack_head_it_v1.npz`
Text DE	Valla (`inhouse-text-de_v1`)	`text_stack_head_de_v1.npz`
Text NL	Valla (`inhouse-text-nl_v1`)	`text_stack_head_nl_v1.npz`
Video frames	Muybridge (`inhouse-video@2`)	SigLIP2 + `…_linear_head_video_v2.npz` (`profile=video`)
Audio	Helmholtz (`inhouse-audio@2`)	Dasheng-Base + `dasheng_…_audio_v2.npz`

There is one live in-house **image** id, one live **video** id, and one live **audio** id. Text uses seven heads that share the same two open feature models in-process (Droplet RAM). Secondary models, Blend modes, and a Scoring-method picker were removed from the product path. Rollback heads may exist under an archive directory for operators; they are not loaded by the live serve entrypoint.

Express selects models through configuration and the language gate (`ENGINE_IMAGE_PRIMARY`, per-lang text models, `ENGINE_IMAGE_VIDEO`). Cache and sync reject retired tags (for example forensics-only or openai-detector style ids).

6. Image model — Vermeer

The image model (Vermeer) wire id is `inhouse@12`. Artifact: `siglip2_base_patch16_224_linear_head_v12lora.npz`. Backbone: SigLIP2-base (Apache-2.0), registered so serve reads backbone id and embedding dimension from the npz. Head: a linear classifier we trained and temperature-calibrated. Serve runs on CPU.

6.0.1. What the head stores

The npz carries the weight vector (or matrix) for the linear head, the decision threshold, and a temperature used at inference time. Inference computes a logit from the SigLIP2 embedding, divides by temperature, and applies a sigmoid to obtain P(AI). Soft product labels read that probability; they do not invent a second model.

6.0.2. Training and holdout

We train on commercial-clean self-generated AI images (Stable Diffusion family) and a diversified mix of real photographs, with ordinary JPEG and resize stress. Fit construction is generator-grouped so one pipeline load yields many samples. Before a model becomes the product default, we check it on a held-out test set that was not used for training: real photos in a Commons-style collection, and AI images from Kandinsky 2.2 (a different generator family). Roughly 200 images in each group.

6.0.3. Measured gates

WHAT WE MEASURED	TARGET	RESULT	PASS
Average score on real photos	≤ 0.12	**0.000**	YES
Average score on AI images	≥ 0.88	**0.938**	YES

WHAT WE MEASURED	TARGET	RESULT	PASS
Gap between AI and real averages	≥ 0.55	**0.970**	YES
Balanced accuracy at our threshold	≥ 0.92	**0.9950**	YES

Decision threshold is about 0.67; temperature about 0.19. For comparison, the previous Vermeer head (`inhouse@12`, CLIP ViT-B/32) scored about 0.101 on reals, 0.906 on AI, and 0.920 balanced accuracy on a comparable held-out check — showing the separation gain of the SigLIP2 `@5` head.

7. Text models — Valla

Everyday text scoring (Valla) uses **one shared feature stack** and **seven language-specific logistic heads**. Feature sources: TMR (MIT) and Fakespot (Apache-2.0). On the inference host those two models load once; each language head is a small npz of coefficients, threshold, and temperature. Scoring supports English, Spanish, Portuguese, French, Italian, German, and Dutch. Other languages receive `unsupported_language` and do not use up the free daily allowance.

7.0.1. English — Valla (`inhouse-text@9`)

Artifact: `text_stack_head_v9.npz`. This is the public headline text claim.

7.0.1.1. Training and holdout

Human samples are real human writing: a mix of Wikipedia, WikiText, and Gutenberg prose, with harder Wikipedia negatives weighted more heavily in the fit mix. **AI samples** are generated with open models (Qwen2.5-7B and Mistral-7B at several decode temperatures). The held-out test set uses AI text from Qwen2.5-1.5B and an equal-thirds human mix of the same three sources, so evaluation is not limited to a single generator or a single human corpus. About 400 human and 400 AI samples.

We have not published a holdout built from live ChatGPT or Claude scrapes. Generalization to those closed models is therefore unmeasured here; the open-model gates below remain the public claim.

7.0.1.2. Measured gates

WHAT WE MEASURED	TARGET	RESULT	PASS
Average score on human text	≤ 0.12	**0.045**	YES
Average score on AI text	≥ 0.85	**0.909**	YES
Balanced accuracy at our threshold	≥ 0.90	**0.980**	YES

Decision threshold is about 0.72. On the same v9 holdout, the prior Valla head (`inhouse-text@9`) scored about 0.045 human / 0.853 AI / 0.966 balanced accuracy — showing the `@9` lift in AI mean and balanced accuracy while keeping human mean low.

7.0.2. Other supported languages

The Spanish, Portuguese, French, Italian, German, and Dutch Valla heads clear the same absolute floors as English (human mean ≤ 0.12, AI mean ≥ 0.85, balanced accuracy ≥ 0.90).

Human samples are lead prose from Wikipedia articles in each language, filtered to continuous encyclopedic paragraphs. **AI samples** are generated with open models (the same panel as English) using native-language prompts that ask for encyclopedic prose, so both classes sit in a comparable domain. Portuguese requests (`pt`) use the Brazilian-trained text model (`inhouse-text-pt_BR_v1`). Each language holdout has about 200 human and 200 AI samples.

LANG	ENGINE	HUMAN MEAN	AI MEAN	BAL. ACC.
Spanish	Valla (`inhouse-text-es_v2`)	0.088	0.953	0.972
Portuguese	Valla (`inhouse-text-pt_BR_v1`)	0.103	0.945	0.928
French	Valla (`inhouse-text-fr_v2`)	0.082	0.976	0.975
Italian	Valla (`inhouse-text-it_v1`)	0.069	0.986	0.972
German	Valla (`inhouse-text-de_v1`)	0.064	0.968	0.958
Dutch	Valla (`inhouse-text-nl_v1`)	0.101	0.978	0.933

All six passed `TEXT_GATES_OK` on their language holdouts. Cite these measured numbers only — do not invent a single "multilingual accuracy" figure that averages across languages.

7.0.3. Why this stack

Open RoBERTa-class detectors are practical on CPU. The product path stacks two commercially clean feature sources and learns thin logistic heads on top, rather than shipping a second heavyweight generative model per language. New languages stay behind the gate until their own holdouts clear the same floors.

8. Video model — Muybridge

The video model (Muybridge) wire id is `inhouse-video@2`. Artifact: `siglip2_base_patch16_224_linear_head_video_v2.npz`. Same SigLIP2 backbone as still images; a separate head trained under `profile=video`.

8.0.1. Why a separate head

Still-image and video-frame distributions differ (motion blur, compression, generator fingerprints). The product therefore trains and serves a video-profile head. If that file is absent, inference refuses the video profile with a clear error instead of silently scoring frames with the still-image head.

8.0.2. Training and holdout

We train on open video generators that fit our license and hardware constraints (CogVideoX-2b and Wan2.1, both Apache-2.0). The held-out test set uses Wan2.2-TI2V-5B (Diffusers) clips kept out of training, plus real Commons video frames sampled with the same VS1 stamps the product uses. Fit reals also mix a fraction of commercial-clean still photos. Roughly 100 real and 100 AI frame rows, scored as bags of three. The product aggregates each bag with a p90 over frame probabilities. The table below is that bag-p90 contract (secondary single-frame floors also pass; see MODEL-CHOICE for detail).

8.0.3. Measured gates

WHAT WE MEASURED	TARGET	RESULT	PASS
Bag p90 average on real frames	≤ 0.12	**0.004**	YES
Bag p90 average on AI frames	≥ 0.85	**0.940**	YES
Balanced accuracy at our threshold	≥ 0.90	**0.985**	YES

Decision threshold is 0.5. On the same bags, the prior Muybridge head (`inhouse-video@2`) scored about 0.716 balanced accuracy — showing the `@2` lift.

9. Audio model — Helmholtz

The audio model (Helmholtz) wire id is `inhouse-audio@2`. Artifact: `dasheng_base_linear_head_audio_v2.npz`. Frozen Dasheng-Base masked audio encoder + logistic head.

9.0.1. Why this stack

Open audio encoders on CPU are practical for short windows. Product scoring uses protocol windows (typically 4 s) and a max aggregate; serve may single-pass prewindowed clips. Missing head → `/v1/analyze-audio` returns 503.

9.0.2. Training and holdout

Fit and holdout use commercial-clean speech/music/general corpora (CodecFake, DFADD, AudioGen-class generators as listed in `Data.json` / `MODEL-CHOICE.md`). Public claim floats are the standalone holdout means rounded for product docs — not a fused-video claim table.

9.0.3. Measured gates

WHAT WE MEASURED	TARGET	RESULT	PASS
Average score on real audio	≤ 0.12	**0.004**	YES
Average score on AI audio	≥ 0.85	**0.984**	YES
Balanced accuracy at our threshold	≥ 0.90	**0.986**	YES

10. How we evaluate

For each kind of media we keep a test set aside and do not train on it. When we describe performance, we report:

- **Average score on reals / humans** — should stay low so natural content is not over-flagged.
- **Average score on AI test samples** — should stay high on the generators we reserved for testing.
- **Gap** — AI average minus real/human average (reported where the gate set includes it).
- **Balanced accuracy** — accuracy averaged across the two classes at the stored decision threshold.

A model only becomes the product default after those checks pass (`IMAGE_GATES_OK`, `TEXT_GATES_OK`, `VIDEO_GATES_OK`, `AUDIO_GATES_OK`). Candidates that fail stay out of the product.

When the same file or text comes back again, a hash cache keyed by `(clean_hash, engine)` can return the prior score. That saves compute and protects free daily limits. How often the cache hits is an operations detail — not a quality claim.

On a typical mid-range CPU host, an image score is often on the order of about 200 milliseconds; exact time depends on the machine and load.

10.0.1. What “pass” means

Gates are pass/fail against fixed targets chosen before looking at the candidate’s final numbers. A head that misses any target stays out of the product path, even if one metric looks strong in isolation. When quoting DotCheck performance, use the measured gates for the engine ids that are actually wired — not earlier same-domain experiments or incomplete holdouts.

10.0.2. What this report measures

The gates in this document are for the current models: Vermeer (`inhouse@12`, images), Valla (`inhouse-text@9` English plus the six language heads above), Muybridge (`inhouse-video@2`, video frames), and Helmholtz (`inhouse-audio@2`, audio). Each table is the held-out result for that wire id. Earlier heads are historical comparison only; they are not the product default described here. Public video floats remain the 3-frame bag holdout; fused frame+audio has no separate public claim row yet.

11. Leviathan (product scoring path)

Leviathan is the product name for everyday scoring: the named models above, plus a generator-surface prior when the Chrome extension scores assistant output on major AI chat and image studios, plus a shared content-hash cache that can reuse a prior product score when the same text or media appears again. More extension users on those studios can densify the shared memory; early on the memory is thinner and models still carry the day. The memory is keyed to content, not a trail of where it was first seen or who scored it.

Generator context can sharpen coverage and consistency. It does not replace the measured model gates in this report. The public AI index records **model** probability only so provenance blending cannot inflate the published mean. Check and Dashboard ignore generator context on the request but still benefit from cache hits. Pro Confirm (Sightengine / Winston on ask) stays outside Leviathan.

12. Cache, Share, and Dashboard notes

Hash cache. Successful in-house scores may be stored under `(clean_hash, engine)`. Unique indexes prevent duplicate rows for the same pair. Cache hits are operational — they do not change the scientific meaning of the gate tables. When a generator-blended product score is stored, a later hit can return that product probability without re-blending; pure-model writes do not erase an existing blended product score.

Share. Website Check can enqueue a short “IS IT AI?” MP4 built from a real in-house probability. The client cannot supply an arbitrary percent. Share quota is separate from analyze FUP. Render may be disabled with a clear error when the Share worker is off.

Dashboard history. Premium and Pro can keep metadata rows (kind, percent, label, date) without storing the media itself. Reopening history does not trigger a Pro confirm; confirm always requires an explicit action on a live score session.

Extension local cache. The Chrome extension may store recent scores under versioned local keys. Those keys must match the free primary model ids; stale tags from retired models are not treated as authoritative product scores.

13. Failure modes and clear errors

DotCheck prefers an explicit failure over a silent wrong path:

- **Inference down or missing video head** — HTTP 503 / clear in-house unavailable errors. No silent Sightengine/Winston substitution for Free or Premium.
- **Unsupported language** — text is not scored; free daily allowance is not burned.
- **Document too short / too many pages** — explicit `text_too_short` / `page_limit` style errors.

- **Share busy or disabled** — `409 share_busy` / `503` when render is off; Share never fabricates a percent.
- **Confirm quota exhausted** — soft wall for Pro; Free/Premium cannot confirm.

These behaviors are part of the product contract. They exist so a degraded backend cannot look like a successful commercial recheck.

14. Licensing choices for the stack

We avoid non-commercial training weights and known-encumbered datasets for the models we ship. That is why the image path uses SigLIP2 (Apache-2.0), and why the text path uses the TMR + Fakespot stack for everyday scoring. Video generators used for fit (CogVideoX-2b, Wan2.1) are Apache-2.0. We do not present closed commercial generators such as Midjourney as in-house test sets; on Pro you can still ask Sightengine for a second opinion.

15. Plans and optional Pro confirm

PLAN	PRICE	WHAT YOU GET
Free	$0	Same three models; surface caps (Ext 30/300, Check 6, Dash 30+1 doc); 3 Shares/day; no commercial confirms
Premium	$7.99/mo or $79/yr	High-volume in-house (3k/day FUP); unlimited Shares; documents and history; no commercial confirms
Pro	$24.99/mo or $249/yr	Unlimited in-house under fair use (30k/day FUP); everything in Premium, plus up to 300 explicit confirms/month

Confirm never auto-fires because two models disagree or because a score sits in a middle band. Free and Premium soft-wall if they attempt confirm. That policy is enforced server-side. Vendor results remain separate from the in-house tables in this report (see Pro confirm path above).

16. Privacy in short

We use the media or text you check in order to give you a score — that is the purpose. We are not building a diary of full pages you browse. Temporary uploads are cleaned up after scoring. Models stay on our servers rather than shipping as a mandatory download to every device. The Privacy Policy at dotcheck.ai/privacy-policy has the full legal notice, controller identity, and processor list.

Mitigations that matter for scoring: short-lived temps, hash caches instead of full browsing diaries, Free/Premium never hitting paid vendors, and a language gate that avoids scoring text outside the evaluated set.

Controller: Petr Jaroch (trading as DotCheck), IČO 03330389, Jenikovice 2, 535 01 Prelouc, Czech Republic.

17. Limitations

- New generators, heavy edits, screenshots-of-screenshots, and unusual compression can move scores.

- Text supports English, Spanish, Portuguese, French, Italian, German, and Dutch. We have not published holdouts from live ChatGPT or Claude scrapes; generalization to those closed models is unmeasured here.
- Video uses a few frames, not full temporal or audio forensics; video holdout is reserved Wan clips, not a third unseen generator brand.
- Document scoring (Premium+) inherits text and image limits and page caps.
- Optional Pro rechecks from Sightengine or Winston are separate commercial opinions.
- Please do not rely on DotCheck alone for employment, academic misconduct, credit, or similar automated decisions about a person.

18. Related reading

- **Model Card** — shorter summary on the same docs page
- **Privacy Policy** and **Terms** — legal detail on the DotCheck website
- https://dotcheck.ai/docs — PDF downloads

Docs v2026.7 · 26 July 2026

19. Upstream models (lockstep)

Commercial-clean serve backbones and train/holdout generators. SSOT: `Data.json` → `models`.

MODEL	ROLE	LICENSE
`google/siglip2-base-patch16-224`	image_video	Apache-2.0
`Oxidane/tmr-ai-text-detector`	text	MIT
`fakespot-ai/roberta-base-ai-text-detection-v1`	text	Apache-2.0
`mispeech/dasheng-base`	audio	Apache-2.0
`stabilityai/sd-turbo`	image fit	MIT
`segmind/tiny-sd`	image fit	Apache-2.0
`stabilityai/sdxl-turbo`	image fit	MIT
`warp-ai/wuerstchen`	image fit	MIT
`Qwen/Qwen2.5-7B-Instruct`	text fit	Apache-2.0
`mistralai/Mistral-7B-Instruct-v0.3`	text fit	Apache-2.0
`zai-org/CogVideoX-2b`	video fit	Apache-2.0
`Wan-AI/Wan2.1-T2V-1.3B-Diffusers`	video fit	Apache-2.0
`kandinsky-community/kandinsky-2-2-decoder`	image holdout	Apache-2.0
`Qwen/Qwen2.5-1.5B-Instruct`	text holdout	Apache-2.0
`Wan-AI/Wan2.1-T2V-1.3B-Diffusers`	video holdout	Apache-2.0
`Wan-AI/Wan2.2-TI2V-5B-Diffusers`	video holdout	Apache-2.0

Hub cards: https://huggingface.co/DotCheck · https://huggingface.co/collections/DotCheck/dotcheck-engines-6a659e57f7934cb510677785