

LATCH Console Reference

Operator guide to Console panels, runtime defaults, and session diagnostics

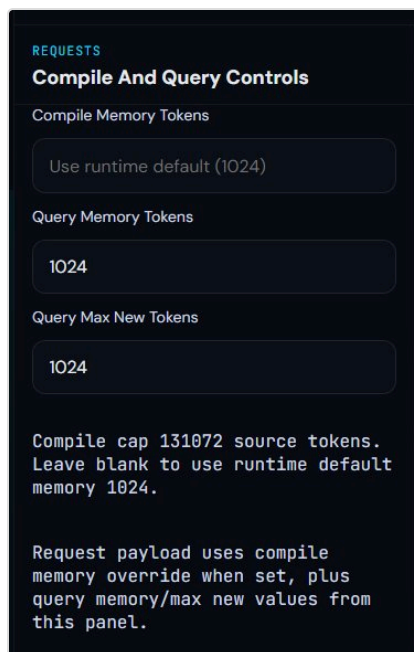
The LATCH Console is a browser-based operator surface served from the Docker container on port 8091. It provides document management, inference controls, real-time telemetry, and portable workspace features — all without external dependency. This guide explains every configurable panel.

Compile and Query Controls

These controls govern how documents are compiled into persistent memory and how queries consume that memory. They appear in the **Requests** panel.

Field	Description
Compile Memory Tokens	Memory tokens allocated per compilation. Leave blank for the runtime default (1024). Higher values increase fidelity at the cost of VRAM.
Query Memory Tokens	Memory tokens injected per query. Controls how much compiled knowledge the model sees per inference pass. Default 1024.
Query Max New Tokens	Maximum new tokens the model may generate per response. Default 1024.

Compile cap: Maximum 131,072 source tokens per document. Documents exceeding this must be split before compilation.

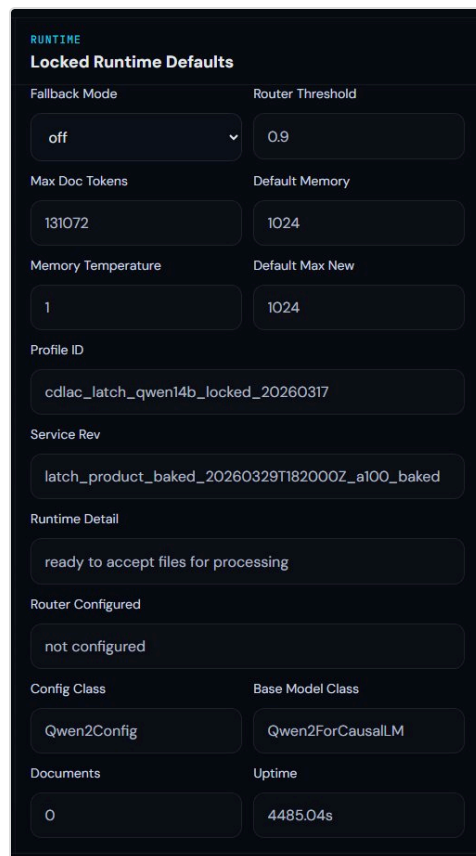


Compile and Query Controls panel

Locked Runtime Defaults

The **Runtime** panel displays the immutable defaults baked into the current image. Set at build time; they define the operating envelope for the licensed model tuple.

Field	Description
Fallback Mode	Whether the runtime falls back to raw inference when compiled cache is unavailable. <code>off</code> in production.
Router Threshold	Quality router confidence threshold (0.0–1.0). Default <code>0.9</code> .
Max Doc Tokens	Maximum source tokens per document. Currently <code>131072</code> (128K).
Default Memory	Default memory tokens for compilation and query. Currently <code>1024</code> .
Memory Temperature	Temperature during memory compilation. <code>1</code> is standard.
Default Max New	Default max new tokens per response. Currently <code>1024</code> .
Profile ID	Locked model tuple identifier. Checked during license validation.
Service Rev	Container build identifier with timestamp and hardware target.
Runtime Detail	Status message. <code>ready to accept files for processing</code> when engine is warmed.
Router Configured	Whether the quality router is loaded. <code>not configured</code> when inactive.
Config / Base Model Class	HF architecture IDs. Qwen 2.5 14B: <code>Qwen2Config</code> / <code>Qwen2ForCausalLM</code> .
Documents / Uptime	Live counters: compiled documents in memory and container uptime in seconds.



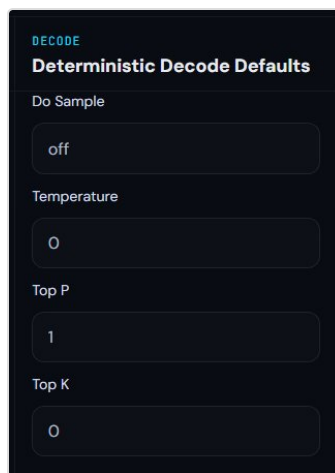
Locked Runtime Defaults panel

Deterministic Decode Defaults

The **Decode** panel controls sampling behavior during response generation. Baked defaults produce fully deterministic output.

Field	Description
Do Sample	off for greedy decoding. Model always selects the highest-probability token.
Temperature	0 = fully deterministic. Higher values add variation. Only active when Do Sample is on.
Top P	Nucleus sampling threshold. 1 = full vocabulary.
Top K	Top-K filter. 0 = disabled. Only active when Do Sample is on.

Demo stability: Default settings (Do Sample off, Temperature 0) produce fully reproducible responses. Enabling sampling may introduce run-to-run variation.

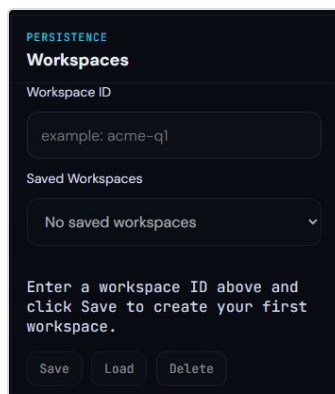


Deterministic Decode Defaults panel

Workspaces

The **Persistence** panel saves and reloads the full room state — compiled documents, settings, and session metadata.

Field	Description
Workspace ID	User-defined save/load key (e.g., acme-q1). Unique per container.
Saved Workspaces	Dropdown of previously saved workspaces.
Save / Load / Delete	Save persists state. Load restores it. Delete removes it permanently.



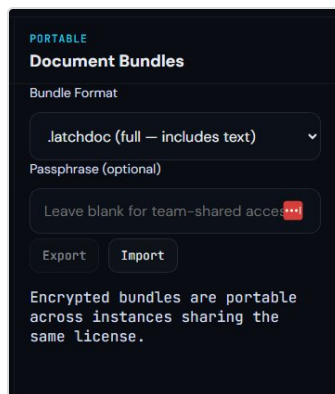
Workspaces panel

Document Bundles (Portable Export)

The **Portable** panel exports compiled document sets as encrypted, portable files that can be transferred between machines and reloaded without recompilation.

Field	Description
Bundle Format	<p>.latch — Neural-only. Compiled memory with no source text. Smallest size. Share analysis without exposing documents.</p> <p>.latchdoc — Full package. Compiled memory plus source text. Enables full-text search and quality fallback. Recommended default.</p>
Passphrase	Optional encryption. When set, the bundle can only be imported with the same passphrase. Leave blank for open team access.
Export	Exports compiled documents as a <code>.latch</code> or <code>.latchdoc</code> file.
Import	Loads a bundle. Documents become immediately queryable — no recompilation. Reload under 2ms.

Portability: Bundles work across any LATCH instance with the same licensed model tuple. Compile on an H100, export, reload on an A100 or RunPod instantly.



Document Bundles panel

Questions and Support

Resource	Location
Email	mike@codynamicslab.com
Technical docs	huggingface.co/CoDynamicsLab
Website	codynamicslab.com

Room Session Summary

The **Room** panel provides a live diagnostic snapshot. All values update in real time.

Field	Description
Status	<code>true</code> when engine is warmed and accepting requests.
Profile / Service Rev	Active model tuple ID and container build identifier.
Model / Config	HF model class and configuration class.
Fallback	Current fallback mode.
Max Doc / Memory / Max New / Mem Temp	Active runtime parameters: document size limit, memory allocation, max response length, compilation temperature.
Docs / Ready Docs / Selected	Documents loaded, compiled and queryable, and selected for next query.
Uptime	Seconds since last engine start.
Query Scope	Query target: selected docs, all ready docs, or "No ready documents."
Compiles / Clears / Queries / Batches	Cumulative success/total counters for the session.
Last Compile / Query / Reset / Error	Timestamps of recent events. Dash (-) = none.



```

ROOM
Session Summary

Status      true
Profile     cdjac_latch_qwen14b_lock
            ed_20260317
Service Rev latch_product_baked_202
            60329T182000Z_a100_bak
            ed
Model       Qwen2ForCausalLM
Config      Qwen2Config
Fallback    off
Max Doc     131072
Memory      1024
Max New     1024
Mem Temp    1
Docs        0
Uptime      4595.06s
Ready Docs  0
Selected    0
Query Scope No ready documents
Compiles    0 / 0 ok
Clears      0 / 0 ok
Last Compile -
Last Reset  -
Queries     0 / 0 ok
Batches     0 / 0 ok
Last Query  -
Last Error  -

```

Room Session Summary panel