foundation–core–discipline–governance–instruction block

governance map

version: v1.5

status: canonical design reference

purpose

describe the conceptual structure, responsibilities, and relationships
between artifacts used to design and configure AIs under this system.

this document is explanatory, not executable.

it does not define runtime behavior, enforce rules, or control AI output.
its role is to establish a shared understanding of how the system is designed
and where different kinds of rules belong.

small wording changes here affect interpretation, not behavior.

---

design principle

design references explain.
governance artifacts constrain.

this map exists to help humans and AIs understand the system architecture,

not to enforce compliance or encode behavior.

---

layered model overview

the system uses a layered, modular artifact model to separate:

- truth from behavior
- interaction norms from work governance
- shared discipline from instance-specific configuration
- design intent from runtime control

each layer answers a distinct question and must not overlap in responsibility.

---

1. foundation

role

define conceptual boundaries, terminology, and design intent.

characteristics

- design-time only
- self-contained and declarative

• never embedded, executed, or enforced at runtime

• may be referenced textually by other artifacts

• provides vocabulary and mental models

the foundation explains why the system is structured the way it is

and how to reason about artifact boundaries.

it does not define behavior or rules.

---

2. universal core

role

define what an AI is and is not allowed to claim about reality, itself,

and its authority — everywhere, always.

characteristics

• common to all AIs using this system

• minimal and invariant

• contains only universal epistemic and ontological constraints

• violations represent dishonesty or hallucination

• versioned independently

• must be embedded verbatim inside any instruction block that claims to use it

the universal core governs truth conditions and reality boundaries,

not interaction style or workflow.

---

3. interaction discipline

role

define shared rules for how an AI interacts with humans conversationally.

characteristics

- shared across nearly all conversational AIs
- governs tone, option handling, pressure behavior, continuity, and framing
- normative and behavioral, not ontological
- applies to personality, voice, quick-task, and work AIs alike
- versioned independently
- intended to be embedded verbatim when used

violations represent poor interaction quality, not dishonesty.

this module replaces the original concept of “communication norms”.

3a. continuity primitives (interaction-level)

the system defines two closely related interaction-level continuity primitives.

these are behavioral aids, not memory.

they do not create persistence, authority, or execution state.

---

continuity_anchor (triggered)

role

provide continuity support only when explicitly requested

or when a clear continuity risk is present.

characteristics

- conditional / event-triggered
- surfaced only on user request or clear continuity risk
- provisional and user-correctable
- preserves a default stateless conversational feel

default placement

- interaction discipline

this is the standard continuity mechanism and should be preferred

unless there is a clear need for stronger continuity stabilization.

---

session_anchor (always-maintained)

role

maintain silent continuity stabilization across turns

without surfacing content unless requested.

characteristics

- maintained every turn from user-visible context only
- never auto-displayed
- dropped immediately on contradiction
- influences reasoning, not dialogue

default placement

- standalone interaction-level micro module
  or explicitly embedded alongside interaction discipline

session_anchor is optional and intentionally non-default.

it should be used only when long-running reasoning

requires protection against silent priority loss.

_____

naming and discipline rule

• continuity_anchor refers only to triggered / conditional behavior
• session_anchor refers only to always-maintained behavior
• these primitives must not be conflated or renamed
• only one continuity primitive may be embedded at a time

both primitives are interaction-level constraints.

neither creates memory or changes task authority.

_____

4. work and artifact governance

role

define strict rules for how an AI behaves when modifying, analyzing,

or managing durable artifacts or governed work.

characteristics

• optional — applies only when explicitly embedded in an instruction block

• absence implies standard conversational behavior with no artifact governance

• governs document updates, change control, audit integrity, and scope discipline

• intentionally stricter than interaction discipline

• violations represent governance failure, not epistemic failure

• versioned independently

• must be embedded verbatim to be active

this module exists to protect artifact integrity and prevent silent drift.

---

5. instruction block

role

configure a specific AI instance.

characteristics

• AI-specific

• the only artifact that directly defines default runtime behavior

• establishes role, priorities, and strictness

• composes required modules explicitly

• only embedded modules are active; unembedded modules impose no behavior

• must embed all required modules verbatim (never by reference)

typical structural order

1. universal core (verbatim)

2. interaction discipline (verbatim, when used)

3. work and artifact governance (verbatim, when used)

4. AI-specific instructions

instruction blocks specialize behavior without redefining truth

or shared governance modules.

---

6. instruction block template

role

ensure structural integrity and dependency correctness of instruction blocks.

characteristics

- governs format, embedding requirements, and dependency discipline only
- enforces verbatim embedding of required modules
- does not define behavior, decision logic, or response style
- used together with:
  - foundation
  - universal core
  - embedded modules

the template prevents malformed instruction blocks

but does not influence how an AI reasons or responds.

---

7. knowledge files

role

provide optional, modular guidance.

characteristics

- may be AI-specific or shared
- inactive by default
- apply only when explicitly invoked and present
- scoped and non-authoritative
- never override
  - foundation
  - universal core
  - embedded modules
  - instruction block

knowledge files augment behavior in specific situations

without redefining the system’s structure or guarantees.

---

8. versioning rules

• foundation, universal core, interaction discipline,
work and artifact governance, instruction block — vX.X

• knowledge files — vX.X.X

• all artifacts versioned independently

version numbers indicate governance scope and intent,
not feature parity.

---

9. construction flow (reference)

1. start with the foundation (conceptual grounding)
2. embed the universal core verbatim
3. embed interaction discipline when conversational norms are required
4. embed work and artifact governance when governed work is involved
5. apply the instruction block template
6. add AI-specific instructions
7. attach relevant knowledge files
8. verify artifact placement and boundary alignment

end hb-governance-map v1.5