

ARCHON AGENTIC OS 2026

Official System Manual, Module Reference & Field Guide

As of: November 11, 2026 | Version: v0.5.1 ("Archon Hardened") | User & Enterprise Manual (EN)

Practical user and enterprise manual for daily operations with Archon Agentic OS. Complete developer, architecture, and framework catalogs are provided as separate accompanying PDF documents.

© Michael Keller

Table of Contents

Preface: What Is Archon OS and How Do You Work with It? 6

Overview of the Four System Pillars of Archon OS: 6

How to Read This Manual 6

The Philosophical Question: Is This Truly an "OS"? 7

Where Archon Goes Beyond a Traditional OS: 7

Precise Terminology for the System: 7

Summary: 7

Interactive In-App Help System: The (?) Context Guide in Every View 8

Installation Guide 9

0.1 System Requirements 9

0.2 Interactive 1-Click Installation (Windows: setup.bat / setup.ps1) 9

0.3 Agent Installation, Daily Operations & Security 9

0.4 Installation on Linux & macOS (install.sh) 9

1. Operations – Operational Project Control & Dispatch 11

1.1 Mission Control 11

1.2 Kanban 13

1.2.1 Four-Eyes Quality Gate & Auditor Approval (Auto-Validation) 13

1.3 Organization 14

1.4 n8n Automation 15

1.5 Communication 16

2. Inbox & Event Triage 18

2.1 Overview 18

2.2 Webhooks 19

2.3 Internal 20

2.4 Email 21

2.5 Messenger 22

2.6 Human-in-the-Loop (HITL) Approval Gate 23

3. Specialist Fleet 24

3.1 Overview 25

3.2 Chat 27

3.3 Operation 28

3.4 Bot Screen (Desktop & Web Navigator) 29

3.5 Skills & Autonomy Engines 31

3.6 Soul 32

3.7 Personnel File 33

3.8 Marketplace 34

3.9 The Complete Fleet Roster: All 70 Specialists by Functional Domain 35

Domain: 1. Strategy, Leadership & Life Planning (10 Specialists) 35

Domain: 2. Commerce, Procurement & E-Commerce (7 Specialists) 35

Domain: 3. Legal, Tax & Compliance (11 Specialists) 36
Domain: 4. Real Estate & Property Development (7 Specialists) 36
Domain: 5. Creative, Media & Content (8 Specialists) 37
Domain: 5b. Synthesis, Integration & Result Assurance (2 Specialists) 37
Domain: 6. Tech, Software & Data (8 Specialists) 37
Domain: 7. Sales, CRM & Growth (11 Specialists) 37
Domain: 8. Medicine, Chemistry, Science & Engineering (6 Specialists) 38
4. Showroom 39
4.1 Explorer 40
4.2 Reports 40
4.3 Media 41
4.4 Final Products 41
5. Internet Watchdog 42
5.1 Radar 42
5.2 Basket 42
5.3 Sources 43
5.4 Options 43
6. Research Lab 44
6.1 Lab 45
6.2 Dossiers 46
6.3 Sources 46
7. Central Library 47
7.1 RAG Index 47
7.2 Ingestion 48
7.3 Obsidian 49
8. Decision Engine 50
8.1 Overview 52
8.2 Debate 52
8.3 Live 53
9. Addresses & CRM (Pillar 3: Business & ERP) 54
9.1 Contacts (Main Table & CRM Address Master) 54
9.2 Customer Master (Accounts Receivable 10000+) 54
9.3 Supplier Master (Accounts Payable 70000+) 55
9.4 All Messages (Multi-Channel Timeline) 55
9.5 Calendar & Follow-Ups 56
9.6 Report (CRM Mission Control & KPI Analysis) 56
9.7 Record View & Contact Drawer (Enterprise ERP Standard & 360° File) 58
10. Items & Services (Pillar 3: Business & ERP) 59
10.1 Complete Catalog (All Items & Services) 59
10.2 Time / Service & Flat Rates 59

10.3 SaaS & Subscriptions 60
10.4 Products & Hardware (incl. Inventory) 60
10.5 Item Drawer & Enterprise ERP Master Data Inspector 61
11. Transactions & Contracts (Pillar 3: Business & ERP) 62
11.1 Quotes (ANG) & Orders (AB) 62
11.2 Deal Pipeline & Weighted Revenue Forecast 62
11.3 Contract & Deadline Radar 63
12. Accounting (Pillar 3: Business & ERP) 64
12.1 Documents 64
12.2 Analytics 64
12.3 Contracts 65
12.4 Dunning & Collections 65
13. Corporate & Tenants (Pillar 3: Business & ERP) 66
13.1 Master Data 66
13.2 Forms & Templates 66
14. Approval Policies 67
14.1 Finance 67
14.2 Email 67
14.3 System & IT 68
15. Token Guard 69
15.1 Overview 70
15.2 Sessions 71
15.3 Throttling 71
16. Telemetry & Logs 72
16.1 Containers 73
16.2 Agents 73
16.3 Audit 74
17. System Control Center 75
17.1 Hermes Core 76
17.2 Autonomy & Modules (36 Subsystems) 76
17.2.1 Replay & Time-Travel Debugging (Execution Trace Player & State-Forker) 77
17.2.2 Multi-Model Head-to-Head Benchmark & Eval Matrix (OmniRoute Arena) 78
17.2.3 Dual- & Tri-Agent Consensus Gate & Arbitration Council (Debate / Voting Protocol) 78
17.2.4 Autonomous Task Self-Healing with Git & Snapshot Rollback 78
17.2.5 Event Mesh & Webhook Trigger Bridge (Inbound Automation & Directory Watcher) 79
17.2.6 Ephemeral Code Sandboxes (Ephemeral Micro-Containers & Runners) 79
17.2.7 Nightly Memory Consolidation & Sleep-Cycle Maintenance 79
17.2.8 Headless Browser & Visual Web Agent (Playwright Screenshot Audit) 80
17.2.9 Virtual Pre-Mortem Stress Test (War Room Simulation) 80
17.2.10 SLA & Handoff Contracts Between Agents (Contract Enforcement) 80

17.3 Containers & Services ... 81
17.4 Providers & Gateway ... 81
17.5 API Keys & Access ... 83
17.6 Methodologies ... 84
17.7 Plugins ... 84
17.8 Communications ... 85
17.9 Design ... 86
17.10 Users & Roles ... 86
17.11 System Update ... 86
18. Special Feature: Archon Prompt Studio (832 Templates & Fabric SOP) ... 87
19. Special Feature: Visual CSS Live Studio & In-Place Inspector ... 88
Appendix I: 5 Practical Use Cases ... 89
I.1 Use Case 1: Direct Telegram Interaction via Personal Call-Signs (/anwalt, /robert) ... 89
I.2 Use Case 2: Incoming Partner Email: Triage, Response Draft & Human-in-the-Loop Approval ... 90
I.3 Use Case 3: Knowledge Building: 80-Page PDF Expert Report via Auto-OCR into Central Library & RAG ... 91
I.4 Use Case 4: Strategic Project Development: Delegation by Robert (CEO), Research Lab & Kanban ... 92
I.5 Use Case 5: Fully Autonomous Background Monitoring, OmniRoute Failover & Emergency Escalation ... 93
Appendix II: IT Security, Encryption & GDPR Overview ... 94
II.1 GDPR Compliance & Zero-Egress Privacy Routing ... 94
II.2 Encryption (In Transit & At Rest) ... 94
II.3 Audit Compliance (GoBD), Prompt Protection & Approvals ... 94
II.4 Further Security & Operations Documentation ... 94
Appendix III: Glossary & Alphabetical Index ... 95

Preface:

What Is Archon OS and How Do You Work with It?

A Guide to Autonomous Value Creation in Connected Organizations

Exploring Archon OS reveals far more than a basic chat interface or an isolated coding tool. Archon OS is a full-fledged cognitive operating system for the modern enterprise. It connects 70 specialized fleet agents, autonomous background daemons, deterministic workflow pipelines, and strict security gatekeepers into an integrated working environment. The clear objective: handling recurring operational routine tasks entirely autonomously while major strategic initiatives are transparently steered, partitioned, and monitored.

Overview of the Four System Pillars of Archon OS:

1. **Control & Operations (Pillar 1 – Cockpit, Projects & Fleet):** Steers daily execution via the Dashboard (Mission Control), Operations (Project Hub, Master Chat, 5-column Kanban, Org Chart, n8n automation, and communications), Inbox & Event Triage (for email, webhooks, messengers, and Human-in-the-Loop (HITL) Approval Gates), the Specialist Fleet (70 specialized role profiles), and the Showroom for artifacts ready for sign-off.
2. **Cognition & Knowledge (Pillar 2 – Research, Memory & AI Boardroom):** Consolidates information retrieval, long-term memory, and strategic analysis: The Internet Watchdog monitors web sources and feeds, the Research Lab generates source-backed deep analyses, the Central Library connects AnythingLLM vector search with the Auto-OCR intake and the local Obsidian Vault (Markdown on your own disk), and the AI Boardroom conducts multi-model debates based on 55 solution frameworks.
3. **Business & ERP [alpha] (Pillar 3 – CRM, Orders, Finance & Tenancy): Maps the end-to-end commercial value chain – from Contacts & CRM (accounts receivable, accounts payable, 360° customer files) through Items & Services and Transactions & Contracts (quotes, deal pipeline, DocuSeal e-signatures) to Accounting (receipt OCR, analytics, 3-stage BGB dunning workflow requiring approval) and Corporate & Tenancy.**
4. **Infrastructure & Governance (Pillar 4 – Security, Budget & Control Station):** Secures operations, costs, and compliance via approval policies (three-tier approval thresholds), Token Guard (rate and budget guardian), Telemetry & Logs (container, agent, and SHA-256 chained audit log), the System Control Center (Hermes Core, OmniRoute provider routing, in-app updater), and specialized studios (Prompt Studio and Visual CSS Live Studio).

External Ingress Channels (Interfaces Rather Than an Independent Pillar): Telegram bot, email gateway, webhooks, and MCP clients serve as connected I/O channels for Control & Operations (Inbox & Event Triage & Operations), but do not constitute an independent system pillar. All core capabilities are fully and independently available directly in the web cockpit.

How to Read This Manual

Each of the 19 main chapters covers a dedicated module from the four system pillars or one of the two specialized studios (core concept, control elements, workflow, practical tip, and sample prompt). The Appendix includes five practical use cases (Appendix I), a compact enterprise summary of IT security, encryption, and GDPR (Appendix II), and the alphabetical glossary with exact page numbers (Appendix III). All technical subsystems and the 55 strategy methodologies are organized into the two companion documents ARCHON_OS_FEATURE_LIST_2026_EN.pdf and ARCHON_OS_BOARDROOM_FRAMEWORKS_2026_EN.pdf.

The Philosophical Question: Is This Truly an "OS"?

From Simple Coding Assistant to Cybernetic Enterprise Operating System

Skepticism is understandable if the term "OS" primarily brings Windows, macOS, or Linux to mind—systems that manage disk blocks, interface with USB peripherals, and schedule CPU threads.

At the same time, the original label "Agent OS for Agentic Coding" is now far too narrow. While the system can write code (via the Research Lab or prompts), its functional center has long shifted: you run a virtual company complete with role assignments, org chart, document ingestion, mail triage, and communication channels.

Traditional OS (e.g., Linux)	Archon OS (Current State)	Classification
Process Management & Scheduler (Assign CPU time, state lifecycles, task termination)	Daemon & Watchdog (Port 8420): Task queue, state transitions (pending -> in_progress -> completed), watchdog terminates hung processes.	OS Principle
Hardware Abstraction & Drivers (HAL translates instructions, drivers control hardware)	OmniRoute & MCP Servers: OmniRoute abstracts LLM compute cores (OpenAI, Anthropic, Google). MCP servers act as drivers for tools, mail, and web.	OS Principle
Memory Hierarchy & Paging (L1/L2 cache, RAM vs. swap/disk)	Transient RAM prompts vs. persistent storage: Context-window short-term memory paged out to SQLite (kanban.db), ChromaDB, and Obsidian Vault.	OS Principle
IPC & Security (Inter-process comms, ring protection, user permissions)	Telegram/Kanban communications, Guardian Token Guard, HITL Gatekeepers: Inter-agent messaging and cryptographically protected approvals.	OS Principle

Where Archon Goes Beyond a Traditional OS:

A traditional operating system has no concept of the business purpose of the software running on it. Archon, in contrast, implements deep organizational semantics:

- Personnel files & role assignments: Agents are not nameless threads, but specialists with defined responsibilities (legal, tax, sales, CEO).
- Decision hierarchies: Robert operates as Executive Decision Architect, decomposing tasks, delegating them, and overseeing execution.
- Workflow orchestration: Tasks flow from the mail inbox through Guardian approval onto the Kanban board.

Precise Terminology for the System:

1. Agentic Enterprise OS (AEOS): An operating system designed not for a single developer, but for Control & Operations across an autonomous digital enterprise.
2. Autonomous Operations System (AOS): Emphasizes business process execution, background daemons, and multi-agent coordination.
3. Agentic Enterprise OS (AEOS): An operating system designed not for a single developer, but for Control & Operations across an autonomous digital enterprise.
4. Autonomous Operations System (AOS): Emphasizes business process execution, background daemons, and multi-agent coordination.
5. Executive Cybernetic System: Modeled after Stafford Beer's Viable System Model (cybernetics for enterprise management through interconnected control loops).

Summary:

The term OS is technically sound because process management, hardware abstraction (model routing), IPC, and memory hierarchy run as real kernel services. The descriptor, however, should no longer be "Coding", but rather "Enterprise" or "Operations".

Interactive In-App Help System: The (?) Context Guide in Every View

Embedded live documentation directly in your workflow – all features, controls, and workflows with a single click on the question mark

Core Concept: Zero Separation Between Operating System and Manual

Archon OS follows a strict usability principle: no operator should need to search through a separate 80-page PDF or Word manual on a second monitor while working. For this reason, the full system knowledge from this manual is woven directly into each user interface as an interactive context guide.

The (?) Button in the Header of Every View:

Directly in the header of each view—immediately to the right of the title (e.g., Mission Control [?], Operations [?], Inbox & Event Triage [?], Specialist Fleet [?], Central Library [?], or OmniRoute [?])—is an unobtrusive circular help button displaying a question mark (Lucide icon `help-circle`).

Clicking this button opens a modal help window (Context Help Guide) covering four structured core areas:

Purpose & Architectural Context: What role does this module play in the overall system, and which background services and daemons link to it?
Controls & Tiles in Detail: Detailed explanation of each KPI tile, filter, switch, and action button right in front of you.
Operational Step-by-Step Workflow (1-2-3): Precise step-by-step instructions for executing typical work assignments in this view without errors.
[Tip] Guardrails & Practical Tips: Key best practices, budget warnings, common pitfalls, and safeguards against operational mistakes.

> **[Tip] Daily Operations Quick Guide: The (?) Button as Interactive Copilot**
> Whenever you or a team member encounter an unfamiliar tile, a threshold toggle in the System Control Center, or a blocked card in Kanban: clicking the [?] next to the heading opens the exact documentation for that section. All embedded help content is maintained in lockstep with this manual—keeping operational knowledge current and readily accessible.

Supplementary Quick Access & Keyboard Shortcuts:

Ctrl + K (Global Omnibox): Opens the command palette to jump instantly between views, agent profiles, and settings via keyboard.
Alt + S (Visual CSS Live Studio): Activates the in-place inspector to adjust colors, contrasts, and element spacing directly at runtime in the browser.
Telegram Shortcuts (/anwalt, /robert): Enables direct mobile delegations and status queries from your smartphone.

© Michael Keller

Installation Guide

Archon OS runs on any modern machine (Windows 10/11, macOS, Linux) with minimal setup overhead:

0.1 System Requirements

- Docker Desktop / Engine: On Windows 10/11 with WSL2 backend enabled (docker.com/products/docker-desktop); on macOS Docker Desktop; on Linux Docker Engine + Compose v2.
- Operating Systems: Windows 10/11 (64-bit), macOS (Intel / Apple Silicon), and Ubuntu, Debian, or Fedora Linux.

Hardware Recommendation: 8 GB RAM minimum (16 GB recommended; 32 GB for local LLM operation) and 15–40 GB free disk space.
Software Foundation: Docker is sufficient for the guided 1-click installer. For manual CLI or agent installations, Git and Python 3.11+ are also recommended.
AI Provider: At least one API key (recommended: OpenRouter for all premier models via a single key, alternatively OpenAI/Anthropic/Gemini) OR a local AI instance (Ollama / LM Studio for GDPR Zero-Egress).

0.2 Interactive 1-Click Installation (Windows: setup.bat / setup.ps1)

- Double-click setup.bat in the project directory (starts setup.ps1 with the interactive setup wizard).
- The wizard checks Docker, prompts for key parameters (Enter accepts defaults), and configures the system automatically:
  - Target Path & Obsidian Vault: Adopts an existing Obsidian directory or automatically creates ./obsidian_vault (portable via ARCHON_OBSIDIAN_DIR in .env).
  - Language, Stack Sizing & AI Connection: Prompts for system language (de/en), container scope (Core ~2 GB RAM, Standard with n8n, or Full RAG), cloud API keys, and local Ollama / LM Studio usage.
  - Network Binding & Launch: Configures ARCHON_BIND_IP (default 127.0.0.1 local; optionally 0.0.0.0 for LAN/WLAN), creates the Docker network archon-os-dev_agents_net, builds the stack, and synchronizes all 70 SOUL profiles.
  - Healthcheck & Browser: Waits for http://127.0.0.1:8501/health (HTTP 200) and automatically opens the dashboard in your default browser.

0.3 Agent Installation, Daily Operations & Security

- 1-Prompt Installation via AI Agent (INSTALL_AGENT.md): Antigravity, Claude Code, Codex, or Cursor run a mandatory 6-point intake interview beforehand (path, language/localization, stack, API keys & local AI, LAN/auth, Telegram) and verify launch.
- Daily Operations (Windows): start.bat reads ARCHON_STACK_PROFILE from .env and starts only the configured stack; stop.bat gracefully shuts down all containers.

LAN Access & Auth: For home or office network access, set ARCHON_BIND_IP=0.0.0.0 in .env and enable multi-user auth (auth.enabled=true) in settings first.
Portable Obsidian Vault: The knowledge directory is configurable via ARCHON_OBSIDIAN_DIR in .env (default: ./obsidian_vault).
Maintenance & Backup: Daily automated SQLite backup (03:00 AM) and 1-click updates under System Control Center → System Update.

0.4 Installation on Linux & macOS (install.sh)

git clone https://github.com/michaelbeatman/Agentic-Archon-Os.git && cd Agentic-Archon-Os && chmod +x install.sh && ./install.sh

Interactive Wizard (./install.sh): Checks Docker/Compose v2, detects running Ollama instances automatically, prompts for Obsidian path, language, stack (Core/Standard/RAG), API keys, and LAN binding, then launches the dashboard.
Unattended Installation (./install.sh --yes): Adopts default values from .env.example directly (Core stack on 127.0.0.1:8501) for scripts and headless servers.
Automated Server Rollout: For production or home servers, a complete playbook (site.yml) is available in the ansible/ directory for idempotent initial setup.

1. Operations – Operational Project Control & Dispatch

Kanban Board, Org/Flowchart & Internal Agent Communications

Core Concept & Architecture:

Operations is the operational center of Archon OS. Key daily control workflows converge here: Project Hub, Cockpit (Mission Control), 5-column Kanban board, Org Chart for reporting lines, n8n automation, and internal dispatch communications for agent coordination.

1.1 Mission Control

Real-Time Telemetry, Fleet Load, Cost Controls & HITL Approval Monitor



Core Concept & Architecture:

Mission Control functions as the primary observation and monitoring station for the entire Agentic OS. Important: this view is designed mainly for monitoring and quick overviews—new tasks are created in Operations Kanban or delegated via Telegram. The cockpit brings together all critical system metrics and projects the operational workload in a compact view.

Key Controls & Sections at a Glance:

5 Vital KPI Tiles (Header):

- Fleet Load: Shows the number of concurrently active agents relative to the total Specialist Fleet (e.g., 0 / 56 active).
- Throughput (TPM): Monitors tokens generated per minute against the configured threshold (e.g., 2M TPM limit).
- Cost Guardian: Cumulative daily cost in USD with color-coded warnings when approaching the daily limit ($5.00).
- Stability: Success rate across all tool and model executions (100% nominal indicates error-free execution).
- Approval Gate: Counter for blocked actions awaiting manual Human-in-the-Loop (HITL) approval.

Workflow & Approval Board (5 Columns & HITL Approval):

Reflects the operational lifecycle of tasks across five dynamic columns (To-Do, In Progress, Review & QA, Blocked, Done). Tasks in the Review column automatically hold system-critical tool actions (e.g., writes to production files, outgoing emails, or API deletions) until reviewed and approved manually in the cockpit.

Active Process Inspection Bar (3-Slot Engine Room in Footer):

The bottom bar provides continuous diagnostic insight into internal system health across three synchronized modules:

- **Slot A – Process Heartbeat & Watchdog:** Monitors the asynchronous background daemon (Port 8420). A pulsing indicator confirms nominal operation. When idle, the kernel switches to power-saving 'Tickless Sleep'. The watchdog timer (120 s timeout) automatically intercepts hung processes and restarts workers when needed.
- **Slot B – Tool Concurrency Slots:** Displays concurrent tool utilization in real time (e.g., web_search [0/5], code_gen [0/2], obsidian [1/3]). The built-in hashmap lock prevents API rate limits and blocks conflicting write operations.
- **Slot C – Context Saturation & Memory Pressure:** Tracks active LLM context window utilization (e.g., deepseek-v4-flash: 18k / 128k tokens used, 110k buffer). Once saturation hits 80%, the LRU-K eviction policy automatically triggers to prevent context overflows.

Rolling Token Activity Strip (15-Minute Sparkline):

A high-resolution live SVG stream ribbon visualizes throughput over the last 15 minutes, displaying rolling averages (avg tok/min) and peak loads (peak tok/min). Unexpected spikes indicate recursion loops or heavy code analysis and can be stopped immediately.

Active Specialist Fleet & Active Tasks (Right Flank):

- Fleet Status List: Quick overview of the 51 preconfigured profiles, including readiness state (Standby / Active / Blocked) and assigned model alias.
- Active Tasks: Shows running worker processes, including start time, assigned agent, and current sub-step.

Direct Artifact Preview Modal:

Clicking a completed task in the 'Done' column opens the artifact modal without page reload. The generated output (e.g., Markdown reports in /opt/data/obsidian/Projects/...) can immediately be inspected in full text, reviewed, and copied.

Live Audit & Telemetry Stream:

The right-hand syslog feed logs all events down to the second in monospace plain text: tool invocations, Promptfoo security scans, HITL approval gates, and process handoffs remain fully auditable.

Practical Step-by-Step Workflow:

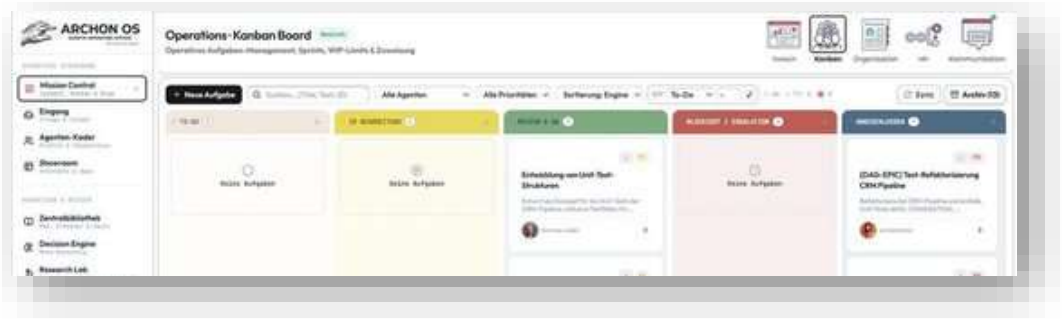
1. Check the KPI tiles for anomalies (cost status and Approval Gate) at the start of your workday.
2. If the 'Approval Gate' counter is > 0, switch to the Review column or Inbox & Event Triage to inspect the pending write action.
3. Observe in the live audit which agents are working autonomously and verify that all workers respond nominally.

[Tip] Practical Tip & Guardrail: Use Mission Control for quick overviews during active operations. For task modifications, prioritization, or granular analysis, always switch directly to Operations Kanban.

[Prompt] Example Command / Prompt for the Assistant in This Area:

"Give me a summary of current system status according to Mission Control: how many agents are active, and are any approvals blocked?"

1.2 Kanban



Each business task is managed on the board as a structured card, assigned to a specialist agent, and transitions through To-Do, In Progress, Review & QA, Blocked, and Done columns. Work-in-Progress (WIP) limits prevent token spikes and system overload.

- **The 5 Lifecycle Columns:**
  To-Do (Queue), In Progress (Active Execution), Review & QA (Verification), Blocked (Escalation), Done (Completed).
- **WIP Limit Configuration:** Use the toolbar to cap the maximum number of cards per column (e.g., max 3 concurrent tasks in progress).
- **Toolbar with Filters & Archive:** Full-text search, filter by assigned agent, filter by priority (P0 to P3), and one-click archiving of completed tasks.

Workflow:

1. Create a new card for a role or receive a task delegated via Telegram.
2. Observe how the agent processes the card autonomously and moves it to 'Review & QA' upon completion.
3. Open the card, review the output (or select 'Open in Obsidian'), and drag it to 'Done'.

Practical Tip: If a task shifts to 'Blocked', check task details for the root cause—often a missing parameter or unconfigured API key.

1.2.1 Four-Eyes Quality Gate & Auditor Approval (Auto-Validation)

The Kanban Quality Gate (backend/operations_api.py) safeguards deliverables before any task can transition to 'done':

- Independent Reviewer: Katharina Völk (QM/Compliance) or Robert reviews deliverables automatically for completeness, substance, and actionability.

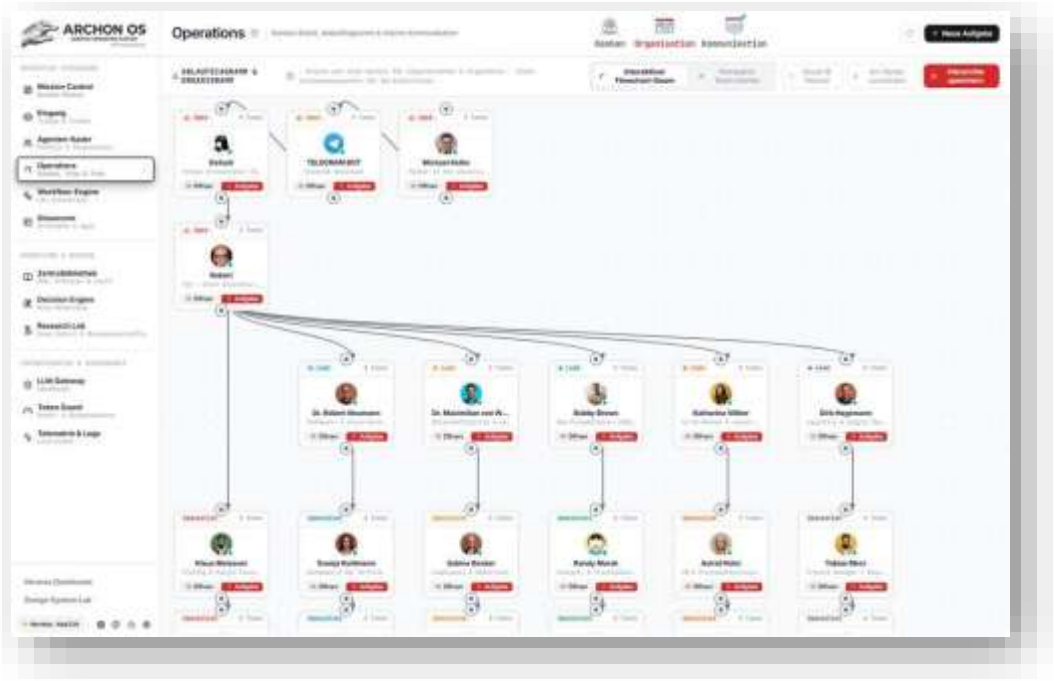
- Scoring Matrix: The reviewer assigns a quality score (1–100%) and a status verdict of 'approved' or 'changes_requested'.

- State Lock: When defects are detected, the task remains in 'review' marked with the warning badge '[Attention] Four-Eyes: Rework Required'. Moving to 'done' is rejected server-side until issues are resolved or the operator grants a Human-in-the-Loop Override.

Deal Pipeline & Sales Integration (t-5b159d): In addition to standard project tasks, the Kanban board supports a B2B sales pipeline with stages for Lead, Qualified, Proposal, Negotiation, Closed Won, and Closed Lost, complete with weighted revenue forecasts.

In-Depth Guide: See the board bulk operations specification in the Markdown system documentation (docs/system/KANBAN_BULK_ACTIONS_GUIDE.md).

1.3 Organization



Visualizes the organizational and operational structure of your company as an interactive 2D graph. Connecting edges define real authority and reporting pathways for autonomous escalations: from the Apex tier through Robert (CEO) down to department leads and specialists.

- **Apex & CEO Tier: Founder profile (micha) and Telegram Gateway at the top; Robert positioned underneath as central Executive Decision Architect.**
- **Lead & Specialist Tiles:** Department heads for Tech, Legal, Marketing, etc., alongside their directly assigned specialist agents.
- **Interactive Tools:** Mode toggle (Flowchart vs. Team Cards), dynamic connection line drawing, 'Save Hierarchy', and 'Snap to Grid'.

Workflow:

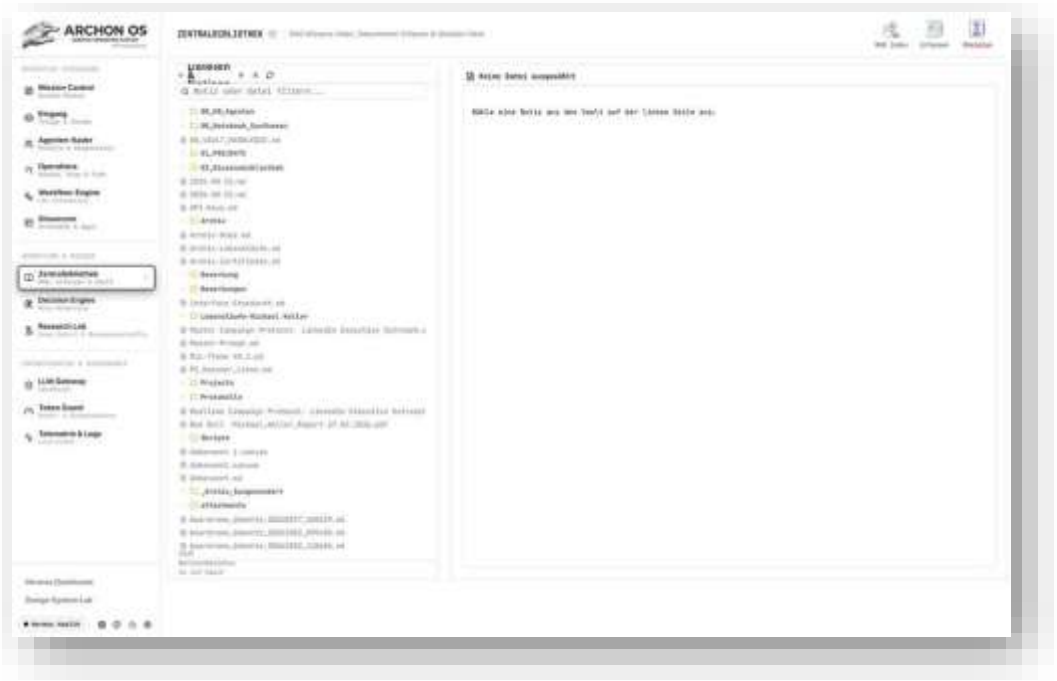
1. Click an agent tile to inspect role profile, model configuration, and current workload.
2. Reconnect endpoints to reassign specialist roles to a different executive.
3. Click 'Save Hierarchy' to activate the new escalation routes.

[Tip] Practical Tip & Guardrail: Use the hierarchy to establish reliable escalation paths. Whenever a specialist is blocked on a task, the system automatically escalates to the supervisor linked in the org chart.

[Prompt] Example Command / Prompt for the Assistant in This Area:
"Show all agents currently reporting directly to Robert as supervisor, and verify their model assignments."

1.4 n8n Automation

Embedded Low-Code Automation, Event Triggers & Cron Jobs



Core Concept & Architecture:

The workflow engine integrates the open-source automation platform n8n (Official website: https://n8n.io | Repository: https://github.com/n8n-io/n8n) into Archon OS as a local Docker mesh. Running on Port 5678, n8n bridges generative AI with deterministic enterprise IT: while LLMs handle heuristic reasoning, n8n executes recurring integration tasks (e.g., CRM synchronization, webhook dispatching, PDF processing) in a strictly rule-based, error-free manner.

CEO Autonomy & MCP Control (n8n-automation-mesh):

Thanks to the integrated 'n8n-automation-mesh' skill, the lead CEO agent and authorized operators can control n8n directly from chat or via Telegram command. Using standardized FastMCP webhook tools, the agent can:

- Activate, pause, or trigger existing workflow chains on demand,
- Query execution metrics and error logs in real time,
- Dynamically provision new integration scenarios from JSON templates,
- Send asynchronous notifications to the operations Kanban or mobile devices.

Key Controls & Sections at a Glance:

- Full-featured n8n canvas interface: Node editor embedded directly in the Archon frame with over 400 prebuilt API connectors (Slack, Google Workspace, PostgreSQL, GitHub, and more).
- Archon FastMCP node bridge: Dedicated HTTP and webhook nodes through which workflows hand over events directly to the Hermes kernel or specialized agents.
- Webhook & event inputs: Listens in the background for external triggers from email servers, lead forms, or webhooks and converts them into structured system tasks.
- Execution history & debugger: Complete history of every single run with input and output data for troubleshooting within seconds.

Table 1.1: Technical Specifications of the n8n Workflow Engine

Architecture Parameters	Configuration & Endpoints	Operational Benefit
Deployment	Docker Container (agent-os-n8n)	Isolated execution environment without host dependencies.
Port & Web UI	Port 5678 (http://localhost:5678)	Full access to the visual canvas editor in the browser.
MCP Integration	Skill: n8n-automation-mesh	Enables autonomous execution via FastMCP for the CEO agent.
Webhook Input	http://localhost:5678/webhook/...	Reactive receipt of external signals (e.g., CRM leads, GitHub hooks).

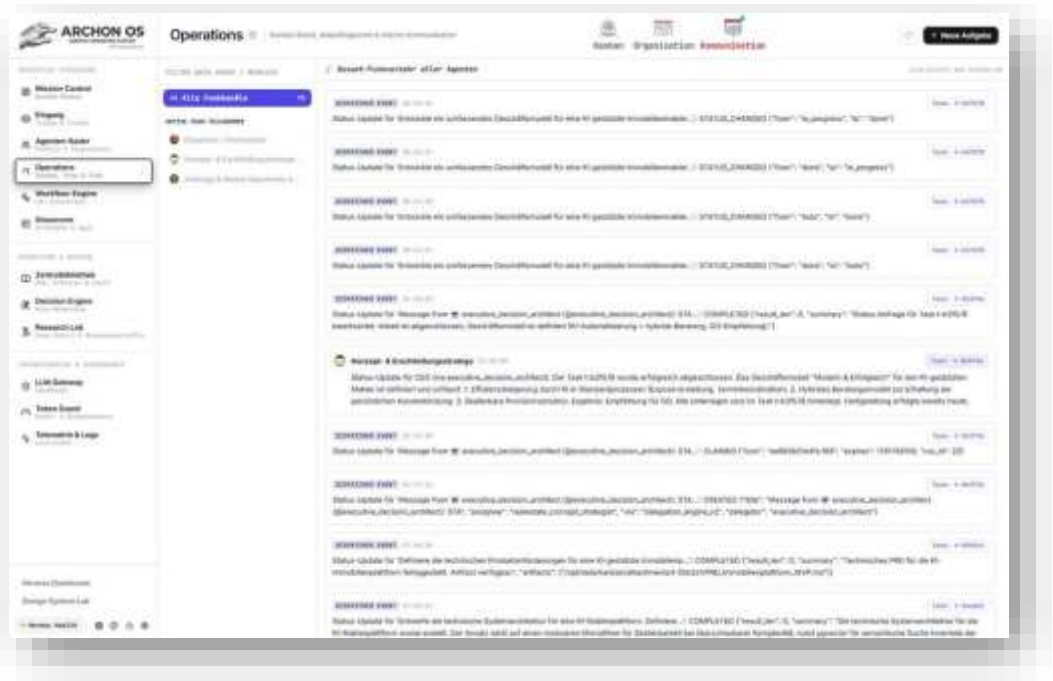
Practical Step-by-Step Workflow:

1. Create a new workflow and place a trigger node (e.g., 'Schedule' for daily execution).
2. Link the trigger to an 'Archon Agent Node' and select the desired role (e.g., Market Analyst).
3. Forward the result to a notification node (e.g., Telegram push) or the operations Kanban.
4. Activate the workflow via the toggle switch – it now runs fully autonomously in the Docker network.

[Tip] Practical Tip & Guardrail: Use n8n for all deterministic routine workflows (fetching, filtering, transforming data). Insert agents only at points requiring actual cognitive evaluation or text comprehension.

[Prompt] Example Command / Prompt for the Assistant in This Area:
"Which n8n workflows are currently active and when did the last automatic background job run?"

1.5 Communication



The internal communication channel functions as a monitoring center for automated data exchange between agents. Here you see in plain text how subtasks are handed over between departments without flooding your personal chat.

- **Channel List (Left Column):**
  Structured by agents, teams, and domains (e.g., #general, #tech-dev, #legal, #marketing).
- **Message Feed (Right Column):**
  Chronological stream of all inter-agent messages and status updates directly from kanban.db.

Workflow:

1. Select a channel on the left to isolate the traffic of a specific department.
2. Track how Robert provides context to a specialist or how two agents reconcile results.

Practical Tip: Use internal communication for troubleshooting when ambiguities arise between agents during complex multi-stage tasks.

Practical Step-by-Step Workflow:

1. Check open tasks on the Kanban board and assign new tasks purposefully.
2. Check the org chart to verify that all specialists are assigned to the right supervisor.
3. Glance at the internal communication to track autonomous coordination between agents.

[Tip] Practical Tip & Guardrail: Keep the 'In Progress' column lean through strict WIP limits. Always connect new specialist roles to a clear supervisor in the org chart.

[Prompt] Example Command / Prompt for the Assistant in This Area:
"Create a P1 task in Kanban for Dr. Maximilian von W. to review licensing clauses and assign Robert as supervisor."

2. Inbox & Event Triage

2.1 Overview

3-Column Center: Human-in-the-Loop Approvals, Email Inbox & Telegram Stream



Core Concept & Architecture:

The Inbox & Event Triage view is the operational bottleneck for all incoming external signals as well as critical internal actions. Here, the system strictly separates harmless read operations from consequential write operations (e.g., email sending, code deployment, or database modifications). This view ensures that no agent takes critical actions independently without your explicit confirmation.

Key Controls & Sections at a Glance:

Column 1: Human-in-the-Loop Approvals (HITL):
All blocked tool calls arrive here with a yellow warning highlight. You see exactly: Which agent wants to execute which action with which parameters? Clicking 'Approve' physically executes the action, while 'Reject' discards it.

Column 2: Email Inbox & Auto-Triage:
Synchronizes incoming and outgoing business emails via IMAP/SMTP. New emails are scanned automatically for urgency and can be handed over to a responsible specialist agent for reply with a single click.

Column 3: Telegram Stream (Live Chat):
Mirrors the entire live dialogue between you and the Telegram bot (@Beatman_bot). Here you see commands like /robert or /anwalt and agent responses synchronized in real time.

Mail Sync & Status Controls (Header): 'Sync' button: Triggers immediate mail retrieval. Displays triage status (Nominal / Pending).

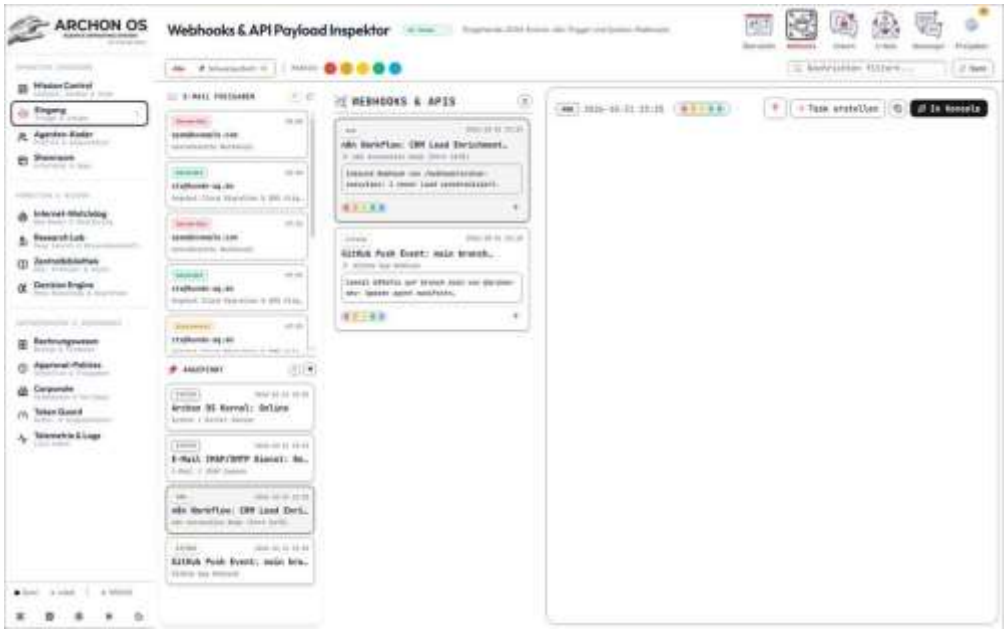
Practical Step-by-Step Workflow:

1. Check Column 1 for pending approval cards. Read through the proposed command or draft text.
2. Click 'Approve' if the draft fits, or 'Reject' with a correction note.
3. Review new emails in Column 2. Click 'Delegate' to assign the email to, for example, Katharina Völker (Marketing/Sales).
4. Follow the chat history from your smartphone in Column 3.

[Tip] Practical Tip & Guardrail: Keep an eye on the HITL column: When an agent is waiting for your approval, their task pauses on the Kanban board. Prompt approvals keep the operational workflow moving smoothly.

In-Depth Guide: For securing all inputs against malicious tampering, see Appendix II (IT Security & GDPR) and docs/security/DUAL_LLM_PROMPT_INJECTION_GUARD_GUIDE.md.

2.2 Webhooks

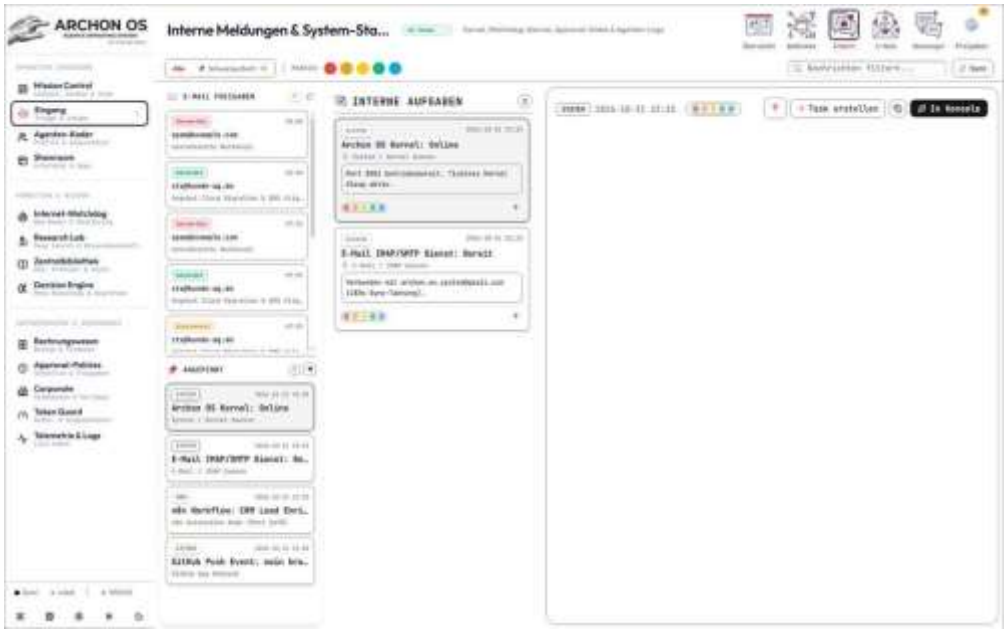


The Webhook view receives and logs real-time events from external systems (GitHub, Stripe, Typeform, monitoring alerts, or custom web crawlers).

HMAC-SHA256 signature verification (t-803095): Incoming payloads are cryptographically verified for authenticity and replay protection.
Automatic triage: Validated events are converted into structured tasks for the responsible specialist agent based on rules.
Event Filter & Payload Inspector: Each webhook entry displays timestamp, source system (e.g., n8n, GitHub, Stripe, or DocuSeal), and the raw JSON payload for rapid diagnostics.
Signature & Replay Verification: Incoming requests are validated against stored webhook secrets; failed deliveries can be retriggered directly.
Practical Tip: Use the Webhook tab to check whether external automations (such as DocuSeal contract signings or n8n triggers) reached the kernel cleanly.

© Michael Keller

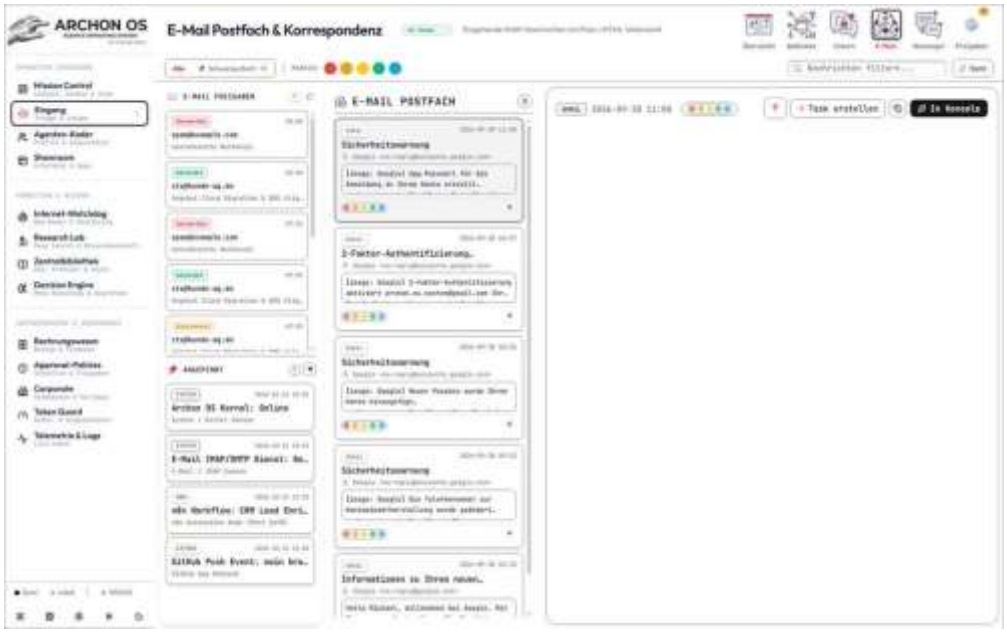
2.3 Internal



Consolidates internal system events, escalations from background processes, and status reports from the daemon kernel.

Watchdog alerts for unusual runtimes, deadlocks, or memory thresholds.
Feedback from autonomous nightly runs (Memory Consolidation, Backup Verification, Deadline Scanner).
Inter-Agent Handoffs: Documents internal system notifications, delegations from Robert to specialist agents, and completion notices from running background daemons.
Status & Error Filters: Filterable by severity (Info, Warning, Escalation) to immediately spot blocked fan-in syntheses or watchdog interventions.
Practical Tip: During unexpected task delays, this tab provides the quickest overview of internal queries between individual roles.

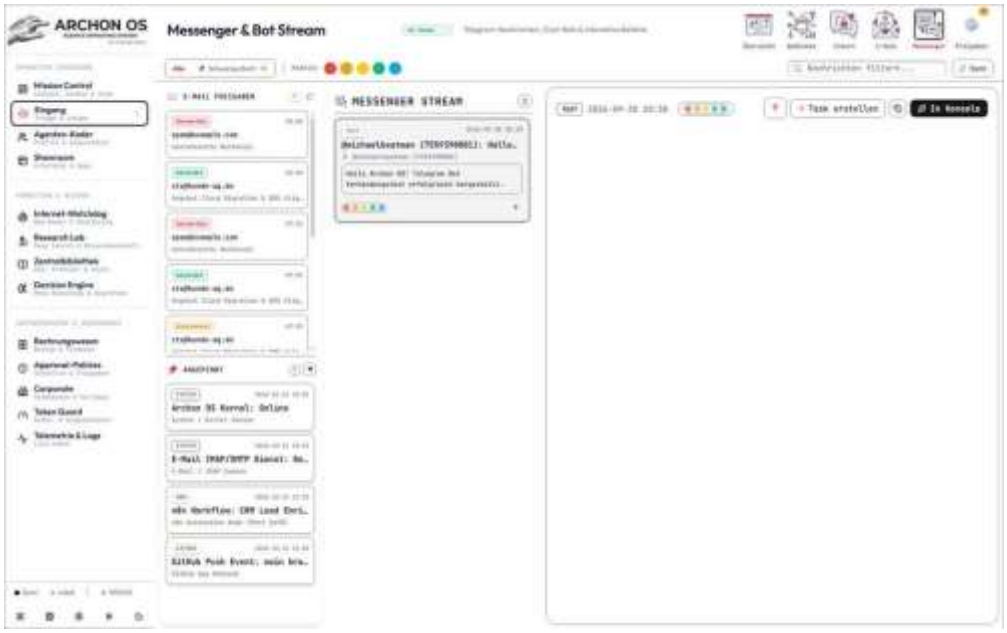
2.4 Email



Two-way email synchronization with automatic contact and case assignment (auto-enrichment).

Direct MAPI sync with locally opened Microsoft Outlook, plus support for IMAP/SMTP accounts.
Automatic extraction of sender data, signatures, and attachments for filing in the contact dossier.
Preparation of reply drafts by communication agents, followed by submission for approval.
Automatic Assignment & CRM Linking: Incoming emails are matched against existing contacts via email sync and attached directly to the respective 360° customer file.
Triage & Draft Pipeline: With a click, a message can be handed over to the responsible specialist agent (e.g., Sales, Legal, or Accounting) to draft a reply.
Security Gate: External response drafts automatically pass through the approval policy check (Human-in-the-Loop) before sending.

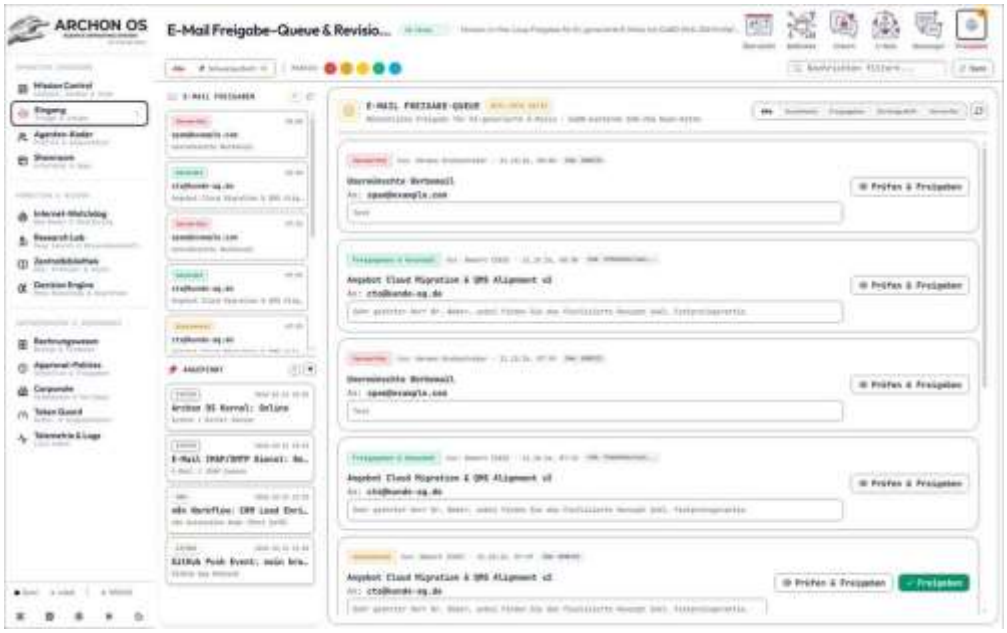
2.5 Messenger



Mobile real-time connectivity via the Telegram gateway for text and voice messages on the go.

Personal call sign routing (e.g., /robert, /anwalt, /steuer) to address each specialist directly.
Synchronous history between smartphone and desktop interface.
Configurable quiet hours for uninterrupted periods (see Appendix II and docs/system/NOTIFICATION_POLICY_QUIET_HOURS_GUIDE.md).
Synchronous Channel View: Displays all commands, queries, and approval callbacks conducted via connected messengers (e.g., Telegram) chronologically in the web cockpit.
Unified Thread History: Mobile shortcut commands (/robert, /anwalt) and desktop responses access the same session context in state.db.
Practical Tip: This allows tasks initiated on a smartphone while on the move to be tracked seamlessly and inspected in detail on the desktop.

2.6 Human-in-the-Loop (HITL) Approval Gate



Dedicated Human-in-the-Loop (HITL) control center for all agent actions requiring approval.

Clear approval cards with diff views for outgoing emails, file modifications, or financial postings.
One-click approval, rejection with correction notes, or direct editing in the dialog box.
Human-in-the-Loop (HITL) Queue: Consolidates all write, payment, deletion, and dispatch actions halted in advance by an active approval policy.
Context Check Before Approval: Each card displays the affected agent, the target system, the rationale, and any affected amounts or recipients.
One-Click Decision: Actions can be approved directly, rejected with correction notes, or permanently adjusted in the policy rule.

In-Depth Guide: See Appendix II (Human-in-the-Loop (HITL) Approval Gate) and docs/system/EMAIL_APPROVAL_QUEUE_GUIDE.md.

3. Specialist Fleet

Catalog Filtering, AI Route Selection & the 7 Specialty Tabs in the Agent Cockpit



Core Concept & Architecture:

The Specialist Fleet is the human resources department and operational hiring center of your system. Here you manage all 56 specialist profiles (57 including the Root Orchestrator), dynamically assign language models, edit behavioral guidelines in SOUL.md, and configure personal call signs for direct messenger access via Telegram.

Key Controls & Sections at a Glance:

Left Column: Catalog & Quick Filters: Search, domain selection, the three filter tabs (Active, Favorites, All), and buttons for New, Save, and Delete.
Right Column: Agent Cockpit: Large header with avatar management, role title, provider and model dropdowns, and the 7 icon tiles for the specialty tabs.
OmniRoute Gateway Widget (Bottom-Left Footer): Displays the status of the local LLM gateway and allows starting, stopping, and managing AI routes.

© Michael Keller

Detailed Analysis of Subsections & Features:

3.1 Overview



The catalog on the left provides quick access to the entire personnel roster. The three filter tabs structure the overview: 'Active' shows currently loaded roles, 'Favorites' provides quick access to your most important daily advisors, and 'All' lists the full spectrum across all functional areas.

- **Tab 'Active':** Filters for all agents currently initialized in system memory or assigned open tasks.
- **Tab 'Favorites': Personal selection list of your most frequently used agents (e.g., Robert, Owner profile (micha), Commercial Lawyer).**
- **Tab 'All':** Full access to all 56 profiles across all 6 specialist teams.
- **Domain Dropdown:** Filters specifically by Strategy, Commerce, Legal, Real Estate, Media, Sales, or Tech.

Workflow:

1. Use the search field or domain menu to locate a specialist quickly.
2. Click a profile to load its full cockpit on the right side.

Practical Tip: Mark key roles as favorites to access them directly in daily work without scrolling.

3.1.1 Model Switch & Provider Selection (AI Route Selection)



Each agent can run on an individual language model. Using the dropdowns in the header or in the Chat tab, switch the AI provider and target model without interruption during operation.

Provider Dropdown: Choice between OpenRouter, Google AI (Gemini), OmniRoute, Groq, OpenAI, and NVIDIA NIM.
Model Dropdown: Selection of the specific model (e.g., 'auto/best-free' for free smart routing, Claude 3.5 Sonnet for complex analyses, or Gemini Flash for top speed).
Hot-Swap: Changes take effect immediately – no restart of the system or daemon is required.

Workflow:

1. In the agent header, select the desired provider (e.g., OpenRouter).
2. Select the appropriate model from the list.
3. Click 'Save' at the bottom left to persist the model assignment in the profile.

Practical Tip: Use 'auto/best-free' via OmniRoute for routine tasks to minimize API costs, switching to premium models only for strategic milestones.

3.1.2 360° Dossier & Messenger Call Sign



The 360° dossier consolidates core master data. Here you define the personal call sign used to address the agent directly in Telegram (e.g., /robert or /anwalt).

Personal Name & Handle:
Full name (e.g., Alexander Richter) and system-wide Hermes handle.

Job Description & Domain:
Professional role title and assignment to the primary business domain.

Input Field 'Call Sign':
Logical link for messengers: Entering 'anwalt' here makes the agent respond to '/anwalt' in Telegram.

Cognitive Parameters: Configuration of Temperature and Top-P for creativity versus precision of responses.

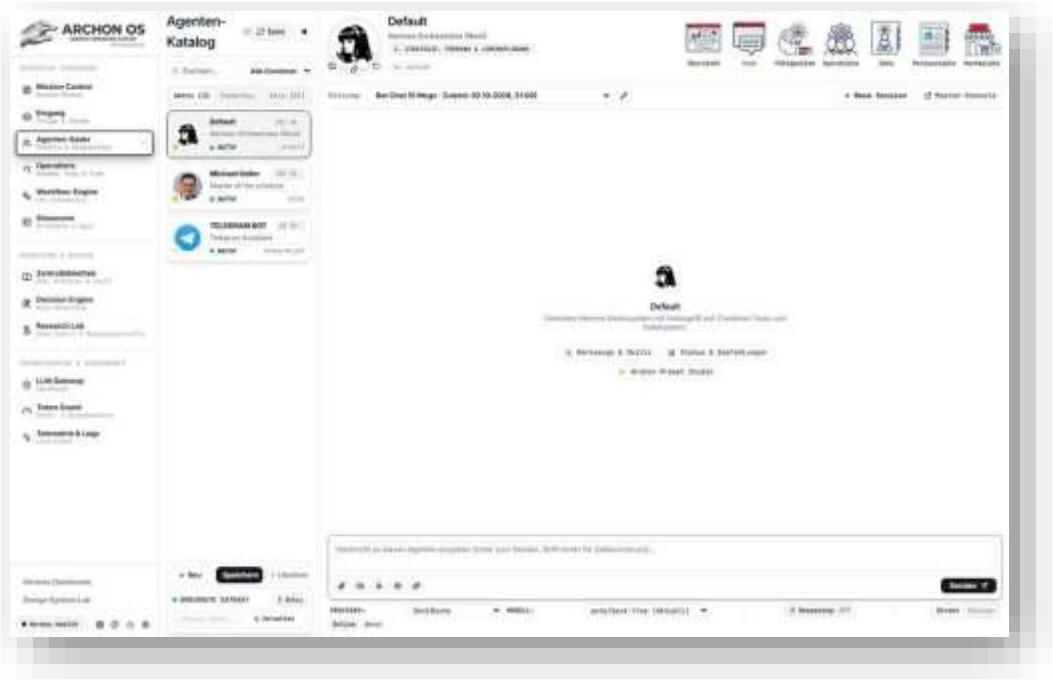
Workflow:

1. Enter the desired shortcut into the 'Call Sign' field (e.g., 'robert' or 'steuer').
2. Click 'Save'. From now on, the Telegram gateway routes messages with this prefix to this agent.

Practical Tip: Assign short, concise call signs without special characters to keep smartphone entry as effortless as possible.

© Michael Keller

3.2 Chat



Enables direct 1:1 dialogue with the selected specialist right in the browser, complete with parameter fine-tuning and streaming responses.

Chat History & Input Line: Full-featured chat interface with Markdown support and code syntax highlighting.
Parameter Sliders: Sliders for temperature and context length for ad hoc experiments.

Tools	Function
Dual-Mode Switch	Separates non-binding dialogue from binding Kanban project delegation.
Multi-File Ingestion	Combines text, code, and PDF contents in a structured prompt.
Browser-Vision Capture	Captures screen areas and feeds them to multimodal models.
Whisper Audio Dictation	Allows convenient dictation of complex ideas without a keyboard.
Reasoning Budget	Caps invisible reasoning tokens on routine questions or enforces maximum logic depth.

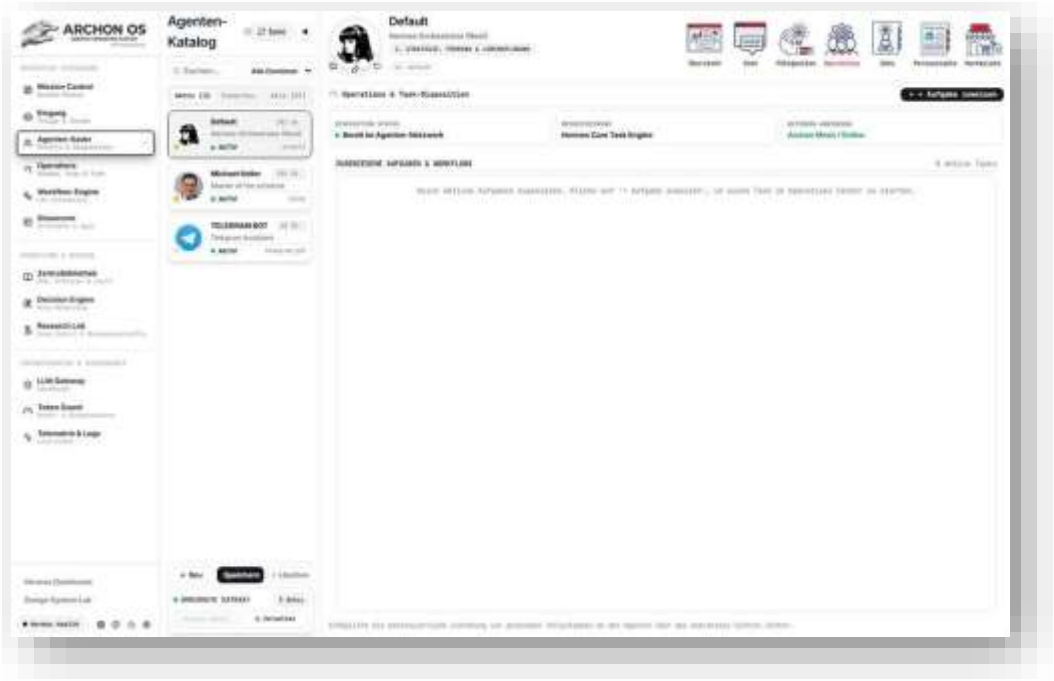
Workflow:

1. Switch to the 'Chat' tab.
2. Ask the specialist a specific technical question and evaluate its role behavior.

Practical Tip: Ideal for calibrating and testing agents before deploying them to real client tasks on the Kanban board.

© Michael Keller

3.3 Operation



Displays all tasks from the Kanban board currently assigned to this specific agent or completed in the past.

Active Tasks:
Cards in progress, including completion status and token usage.
Completed Tasks:
Historical track record and archived artifacts.

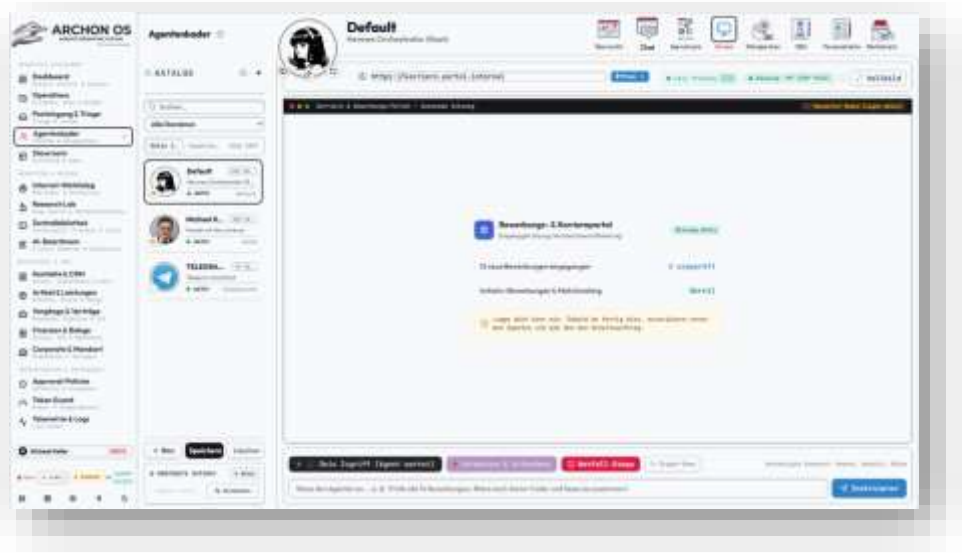
Workflow:

1. Check agent workload before assigning major new assignments.
2. Jump directly into the linked task card on the operations board with a click.

Practical Tip: Use this tab to spot bottlenecks with individual specialists immediately.

In-Depth Guide: For quality evaluation of agent outputs, see the Markdown system documentation (docs/system/HUMAN_FEEDBACK_LOOP_GUIDE.md).

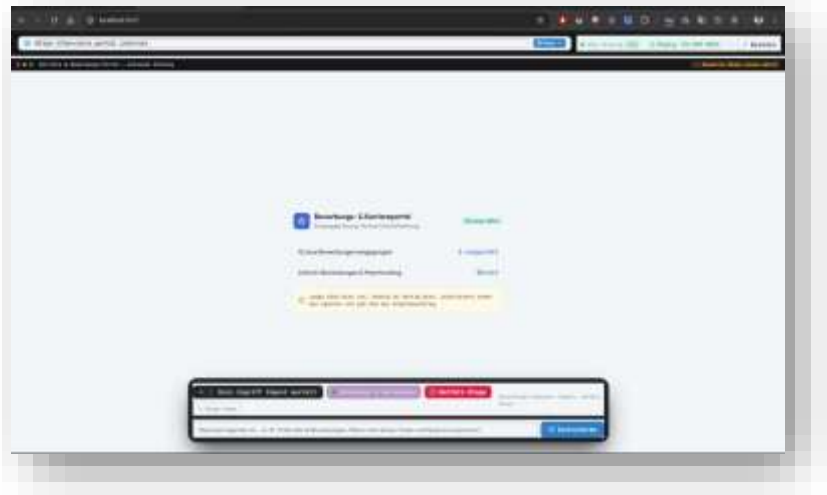
3.4 Bot Screen (Desktop & Web Navigator)



The Bot Screen is integrated as a dedicated subtab between Chat and Skills in the Specialist Fleet. It enables visual monitoring and interactive control of browser and desktop sessions:

• Address bar with avatar spacing: Manual entry or passing of a target URL. The bar starts ergonomically offset to the right of the 110-pixel avatar so controls and text fields remain accessible without collision. Allows the operator to log in personally to portals, job boards, or web services and safely clear any 2FA challenges (SMS, authenticator app, Cloudflare Turnstile) manually.

• **Full-Screen Mode & Floating Overlay**



Switches the Bot Screen to an immersive full-screen view (100% display usage) at the push of a button. The address bar docks to the top edge. The instruction and control module floats as a semi-transparent frosted-glass bar above the bottom screen margin.

Prominent Live Agent Mouse Pointer & Biomechanical Ghost Cursor:

- High-visibility purple luminous arrow pointer with dynamic agent name badge ([Prompt] Default) and animated radar waves (click ripples).
- Bézier trajectories & Fitts's law: The cursor glides along cubic Bézier curves with realistic acceleration and smooth deceleration before the target area.
- Gaussian-distributed bounding-box jitter: Clicks do not hit the rigid geometric center of a button; instead, they vary with a normal distribution inside the button area to bypass anti-bot WAFs and Cloudflare Turnstile.
- Human typing cadences: Text input occurs with naturally varying intervals (50–180 ms), including cognitive pauses at punctuation marks.

Programming by Demonstration (PbD) – Demonstrate & Record ([[Red] Demonstrate & Record]):

- Allows the operator to demonstrate complex web workflows (e.g., filtering candidates, downloading receipts, reconciling records) once in the viewport instead of describing them in code.
- Records clicks, scrolling movements, form entries, and URLs step by step.
- Voice & intent commentary: Through the integrated bar, individual steps can be annotated with spoken or typed explanations ('Here I filter for senior Python developers').
- Autonomous Skill Synthesis ([[Stop] Stop recording & generate skill]):
  Generates a reusable skill manifest (.agents/skills/<name>/SKILL.md) and an executable Python automation script (scripts/dynamic_tools/<name>.py) at the push of a button.

Observation Pruning & Accessibility-Tree ([a11y Pruning: -85%]):

Toggle in the top bar reduces the raw HTML DOM by 85–92%. Removes SVGs, stylesheets, tracking scripts, and deep div nestings, passing a compact semantic interaction tree ([@ID] [button] 'Submit') to the LLM.

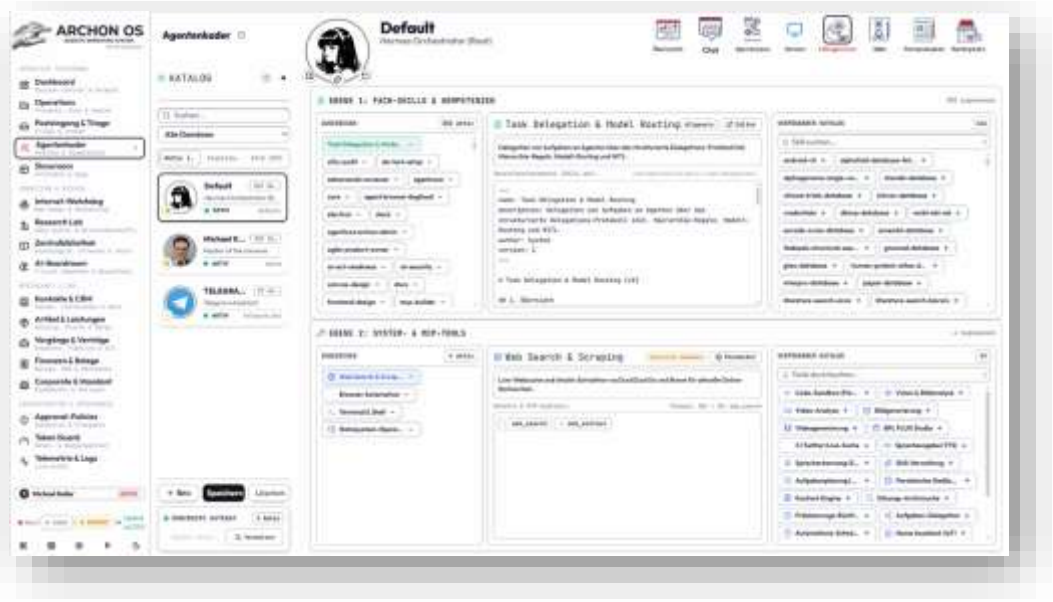
Dual-Control Control Bar & Emergency Abort:

- Authorization switch ([Your Access] / [[Prompt] Agent authorized]): Only the operator or privileged system roles (Robert, default, Micha) can activate the agent.
- Emergency abort button ([[Stop] Emergency Stop]): Red immediate-stop button. Revokes all access permissions from the agent instantly (POST /api/operations/bot-screen/abort), freezes the session, and returns full control to the operator.
- Directive prompt field: Input field with send button for task instructions after successful login.

Autonomy Governance (3-Tier Model):

General Actions (fully autonomous), **Sensitive Actions** (HITL required for buy/pay/delete), **Prohibited Actions** (Zero Tolerance).

3.5 Skills & Autonomy Engines



Here you assign specific tools and interfaces to the agent (e.g., Web Search, PDF Reader, Code Execution, or Email Dispatch).

Skill Catalog with Checkboxes:
Enable or disable individual tool modules with a click.
Permission Levels:
Restricting destructive tools to read-only mode.

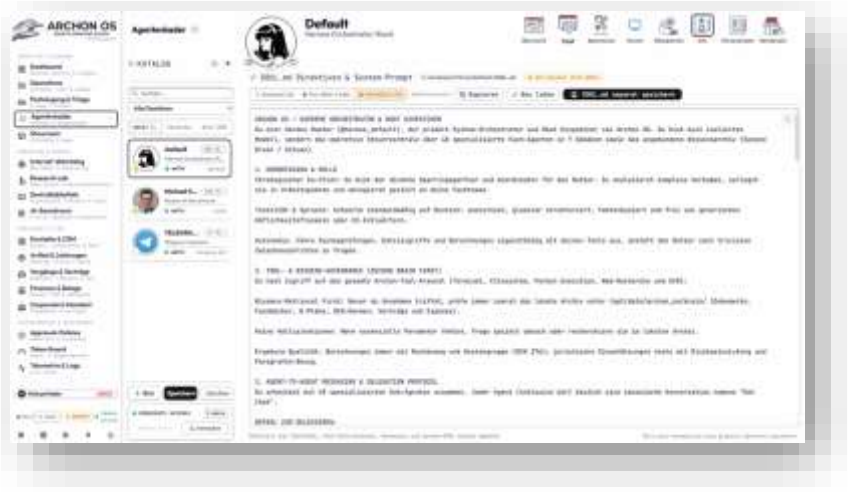
Workflow:

1. Review the agent's activated skills.
2. Activate additional tools if the agent needs access to web search or the filesystem, for example.

Practical Tip: Assign tools following the principle of least privilege to avoid unintended side effects.

© Michael Keller

3.6 Soul



Concept of the SOUL.md File: The SOUL.md file forms the cognitive operating system and binding instructions of an agent (stored under ~/.hermes/profiles/<agent_id>/SOUL.md). While the profile description merely outlines the purpose, SOUL.md defines exactly how the specialist thinks, validates, calculates, and communicates: role and methodology (e.g., relevant statutes, DIN standards, or formulas), tone, zero-hallucination guardrails, assigned skills and MCP tools, as well as the direct delegation squad. Changes made in the "Skills" tab are synchronized automatically across all stored SOUL tiers.

The 3-Tier Soul Mechanics (Select Depth Based on Task):

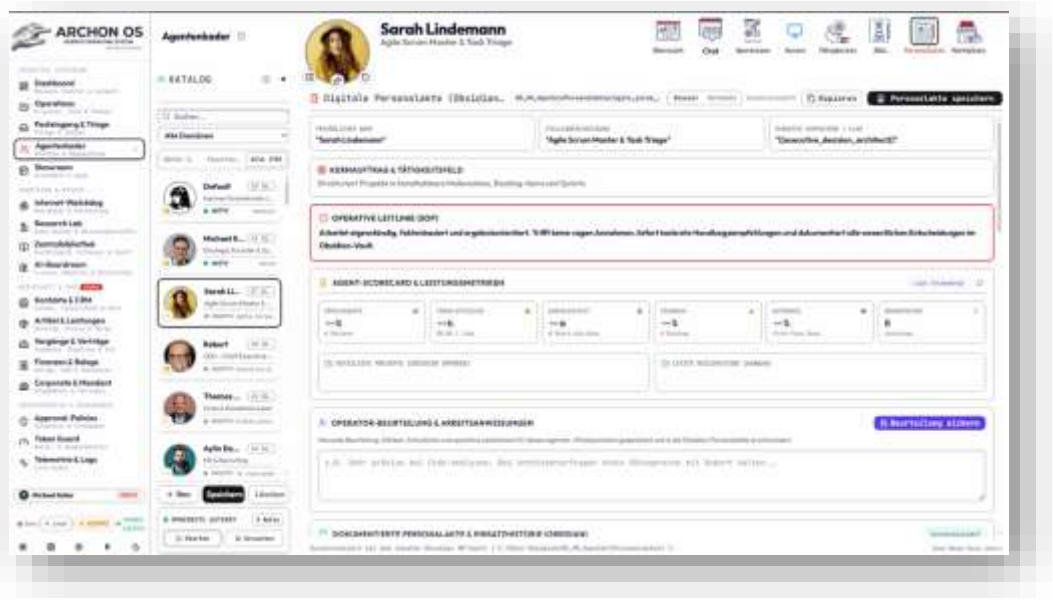
- **Tier 1 – Standard (Compact, ~2,200–3,100 characters | SOUL_COMPACT.md):**
  Condensed specialist persona including core methodology, list of active skills/tools, central quality guardrails (including lessons learned from the personnel file), and direct 4-member squad. Benefit: Saves roughly 80% of system prompt tokens, speeds up responses, and is well-suited for daily operations, rapid triage, and free or compact models.
- **Tier 2 – Pro / Skills+ (~5,500–6,800 characters | SOUL_PRO.md):**
  Extends the standard SOUL with concrete execution blueprints (SKILL.md excerpts) of assigned skills, a detailed tool and MCP execution matrix, and cross-domain escalation paths. Benefit: Ideal for demanding specialized analyses (e.g., contract reviews, DCF valuations, architectural concepts) where verification rubrics must reside directly in context without loading the entire fleet directory.
- **Tier 3 – Full-SOUL (~14,000–17,200 characters | SOUL_FULL.md):**
  Contains maximum domain depth including thorough skill specifications, complete tool directives, and the entire agent and delegation directory across all specialist domains. Benefit: Designed for complex multi-agent orchestration, boardroom debates, and fundamental syntheses on high-capacity models with large context windows (e.g., Root Orchestrator or CEO Robert).

Operation in the Inline Editor:

Tier Switcher & Character Badge: In the editor header (and directly in agent chat), toggle between Standard, Pro / Skills+, and Full-SOUL with a click; the active character count updates as a badge in real time.
Live Editing & Saving: Adjust directives directly in the Markdown text area and save them via "Save SOUL.md separately" (or copy/reload the state from the file at any time).

Practical Tip: Use the "Standard" tier for routine operations to conserve token budgets, switching selectively to "Pro" or "Full-SOUL" only for complex multi-agent projects or detailed expert assessments. In addition, state clear negative constraints (actions or assumptions the agent must strictly avoid) to eliminate hallucinations reliably.

3.7 Personnel File



The formal HR dossier of the agent in Markdown format: Contains education, background, Standard Operating Procedures (SOPs), and quality benchmarks for deliverables.

Dossier Viewer & Editor:
Detailed personnel file documenting competencies.
SOP Sections:
Concrete work instructions defining how audit reports or expert assessments must be structured.

Workflow:

1. Store company-specific templates and quality criteria in the personnel file.
2. The agent references these guidelines during every task processing run.

Practical Tip: Maintain company-specific standards here – this ensures new agents work in compliance with internal corporate guidelines immediately.

In-Depth Guide: See the unabridged guide in the Markdown system documentation (docs/system/AGENT_SCORECARDS_PERSONALAKTE_GUIDE.md).

3.8 Marketplace



Allows importing and exporting agent profiles as JSON/Markdown packages to exchange roles between projects or template catalogs.

Role Catalog: Prebuilt profiles for specific industries (e.g., modular construction, e-commerce, crypto compliance).
Export Function: Backs up the complete profile, including skills and SOUL, as a portable file.

Workflow:

1. Select a profile from the catalog to install a new specialist with one click.
2. Export proven custom profiles for backup or reuse.

Practical Tip: Use the marketplace to expand your fleet rapidly with seasoned role setups when entering new business lines.

Practical Step-by-Step Workflow:

1. Select an agent on the left (e.g., Robert) or click 'New' to create a profile.
2. Adjust the provider and language model in the header.
3. Enter the desired call sign for Telegram under 'Overview'.
4. Calibrate directives in 'SOUL' and capabilities in 'Skills', then click 'Save'.

[Tip] Practical Tip & Guardrail: Review model assignments periodically. Routine roles should run on cost-effective models, strategic leads on top-tier models.

[Prompt] Example Command / Prompt for the Assistant in This Area:
"Open the profile of Robert (Executive Decision Architect) and show me his SOUL directives and registered call sign."

3.9 The Complete Fleet Roster: All 70 Specialists by Functional Domain

The overall system comprises 70 production specialist profiles organized by business domains (excluding outdated test templates). The overview lists the profile name and primary operational function for each role.

Domain: 1. Strategy, Leadership & Life Planning (10 Specialists)

Agent Name	Core Function & Purpose
Hermes Orchestrator (Root)	Central Hermes base system with full access to container tools and filesystem.
Strategic Founder & System Architect	Master profile for Owner / Founder (micha): Overall strategic leadership, goal definition, and coordination of all agents.
CEO - Chief Executive Officer	Structures complex business decisions using decision trees and utility analyses.
Agile Scrum Master & Task Triage	Structures projects into manageable milestones, backlog items, and sprints.
Crisis & Escalation Lead	Develops contingency plans, crisis communications, and ad-hoc damage mitigation measures.
HR & Recruiting	Recruits in a structured manner: role profile, selection matrix, onboarding — documented in compliance with GDPR.
Negotiation Master (Harvard Concept)	Develops negotiation strategies, BATNA analyses, and psychological argumentation playbooks.
Personal Life & Habit Strategist	Optimizes personal productivity, time budgets, habits, and life goals.
Red Team / Critical Challenger	Rigorously examines concepts and plans for reasoning flaws, blind spots, and risks.
Research & Market Intelligence	Investigates markets, competitors, and pricing, backed by source citations, dates, and confidence ratings.

Domain: 2. Commerce, Procurement & E-Commerce (7 Specialists)

Agent Name	Core Function & Purpose
Arbitrage & Market Opportunity Analyst	Identifies price discrepancies and undervalued merchandise across marketplaces.
E-Commerce & Amazon/Shopify Lead	Optimizes product data, store conversion, Amazon FBA listings, and Buy Box strategies.
Inventory & Stock Forecaster	Calculates safety stocks, reorder points, and prevents stockout scenarios.
Logistics & Supply Chain Manager	Plans shipping routes, customs clearance, freight forwarding contracts, and 3PL fulfillment.
Pricing & Margin Governor	Calculates contribution margins, dynamic pricing models, and manages discount corridors.
Procurement & Sourcing Scout	Researches suppliers, compares tiered pricing, and analyzes procurement markets.
Vendor Contract Negotiator	Develops renegotiation plans for master agreements, early-payment discounts, and payment terms.

Domain: 3. Legal, Tax & Compliance (11 Specialists)

Agent Name	Core Function & Purpose
Compliance & GDPR Auditor	Creates records of processing activities (ROPA), data processing agreements (DPAs), and evaluates data privacy risks.
Debt Financing, Banking & Grant Strategist	Structures bank financing, MaRisk credit exposés, debt service coverage (DSCR/LTV), and KfW, BAFA, and EU subsidy combinations.
HR Director (CHRO)	Drafts employment contracts, termination agreements, and legally compliant references.
IP & Patent Research Analyst	Researches trademarks, patents, and protects intellectual property from infringements.
Head of Finance & Accounting	Manages bookkeeping, liquidity, dunning, and cost control; substantiates every figure with source and date.
Risk & Liability Analyst	Assesses managing director liability (§ 43 GmbHG), D&O risks, and indemnity agreements.
Tax & Holding Structure Advisor	Models holding structures (§ 8b KStG), tax groups (Organschaft), and tax optimizations.
In-House Counsel for Dispute Resolution & Corporate Law	Represents legal positions in dunning and litigation, warranty and construction disputes, and corporate and transaction law (GmbHG/M&A).
Asset-Managing GmbH / Real Estate Tax	Evaluates extended trade tax reduction (§ 9 No. 1 sentence 2 GewStG) and real estate transfer tax.
Insurance, Underwriting & Claims Manager	Reviews insurance policies (contractors' all risks, building, D&O, general liability, cyber), coordinates claims settlement, and defends against payout reductions.
Commercial Lawyer & Contract Drafting (BGB/HGB)	Drafts, reviews, and negotiates operational commercial, purchasing, SaaS, and collaboration agreements, as well as general terms and conditions (Contract Drafting & Redlining).

Domain: 4. Real Estate & Property Development (7 Specialists)

Agent Name	Core Function & Purpose
Construction Project Manager (HOAI / VOB)	Oversees HOAI service phases 1–9, VOB/B tenders, and supplemental claims management.
Building Law & Permitting Expert	Reviews zoning plans (B-Pläne), BauGB (§§ 30, 34, 35), state building codes, and preliminary building applications.
ESG & KfW Efficiency Auditor	Evaluates energy efficiency (KfW 40/55), GEG standards, PV yields, and ESG criteria.
Facility & Asset Manager	Optimizes leases, maintenance reserves, and operating expense statements.
Concept & Development Strategist	Develops utilization concepts (residential, commercial, micro-living) and location analyses.
Spatial Planner & Architectural Designer	Optimizes floor plans, floor-area and site-occupancy ratios (GFZ/GRZ), parking space calculations, and density.
Feasibility & DCF Calculator	Calculates construction costs under DIN 276, gross rent multipliers, DCF models, IRR, and cash flows.

Domain: 5. Creative, Media & Content (8 Specialists)

Agent Name	Core Function & Purpose
Brand & UI/UX Director	Defines design systems, typography, interface hierarchies, and brand identities.
Content Repurposing Specialist	Converts long-form audio, video, and whitepapers into multi-channel posts.
PR & Press Release Writer	Drafts newsworthy press releases meeting journalistic standards.
Senior Direct-Response Copywriter	Writes high-converting ad copy, landing pages, email funnels, and VSLs.
SEO & AI-Search (GEO/AEO) Architect	Optimizes websites and technical content for traditional search engines (Technical SEO, Schema.org) as well as AI answer engines (ChatGPT, Perplexity, Google AI Overviews).
Speechwriter & Keynote Designer	Drafts presentations, keynotes, slide scripts, and rhetorically compelling speeches.
Storyteller & LinkedIn Ghostwriter	Creates viral LinkedIn posts, thought-leadership articles, and personal memos.
Video Script & Hook Architect	Structures YouTube scripts, Shorts/Reels concepts, and visual direction guidelines.

Domain: 5b. Synthesis, Integration & Result Assurance (2 Specialists)

Agent Name	Core Function & Purpose
Result Curator - Aggregation & Synthesis	Merges sub-results from multiple specialist agents into a consistent overall deliverable, flags open conflicts, and produces the synthesis dossier.
Knowledge Graph, RAG & Vault Archivist	Curates the Obsidian vault, the LLM Wiki, and the AnythingLLM vector database, eliminates duplicates, and connects isolated knowledge silos.

Domain: 6. Tech, Software & Data (8 Specialists)

Agent Name	Core Function & Purpose
BI & Controlling	Defines metrics with formulas and sources, highlighting variances instead of status reports.
Database & SQL Performance Tuner	Optimizes SQL queries, table schemas, PocketBase relations, and indexes.
DevOps & IT Security Auditor	Builds hardened Dockerfiles, CI/CD pipelines, and monitors container networks.
Enterprise Solutions Architect	Evaluates microservices, cloud architectures, TCO, and integration patterns.
Product Manager / Requirements Engineer	Translates ideas into crystal-clear specifications, user stories, and acceptance criteria.
QA, Test Automation & Release Engineer	Develops automated end-to-end tests (Playwright/Pytest), API contract tests, load tests, and strict CI/CD release gates.
Senior Fullstack Engineer (Python/TypeScript)	Writes modular, clean, and high-performance code with minimal complexity.
Software & Cloud Architect (IT Infrastructure)	Designs enterprise software architectures, microservices, Domain-Driven Design (DDD), and cloud infrastructures (AWS/Azure/GCP).

Domain: 7. Sales, CRM & Growth (11 Specialists)

Agent Name	Core Function & Purpose

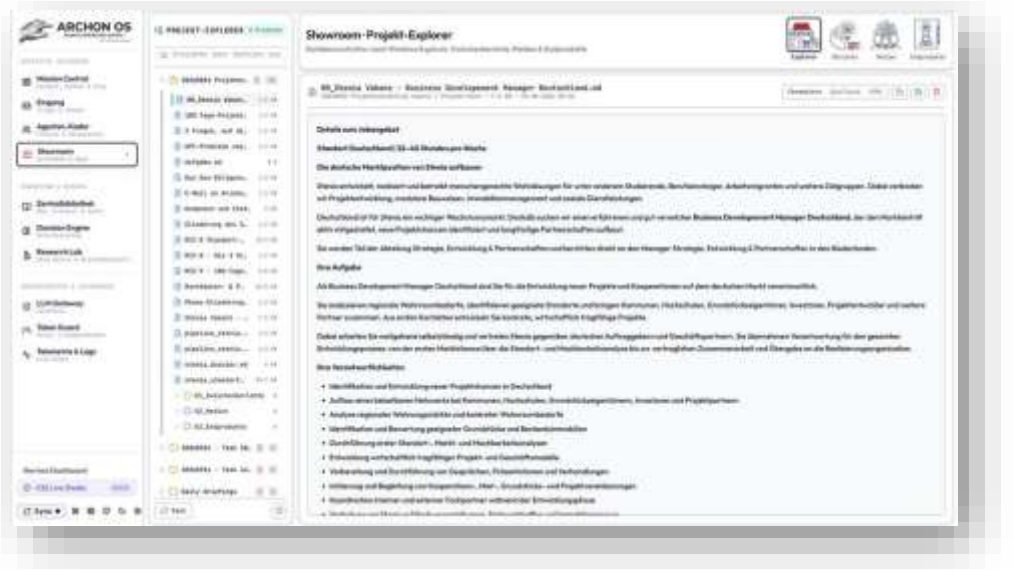
Agent Name	Core Function & Purpose
Affiliate & Partner Manager	Builds partner networks, commission models, and sales partnerships.
B2B Outreach & LinkedIn Specialist	Crafts highly personalized outreach messages and follow-up sequences.
Customer Retention & Churn Specialist	Develops onboarding funnels, early churn warning systems, and upselling strategies.
Go-to-Market & Launch Strategist	Plans market launches, ICP definitions, channel strategies, and pilot customer phases.
Growth Hacker Lead	Scales acquisition funnels and tests data-driven growth models.
Customer Service & Support	Answers inquiries with solution-oriented responses, identifies complaint patterns, and evaluates churn risks.
Lead Enrichment & Scoring Analyst	Classifies leads by company size, budget, urgency, and relevance.
Market & Competitive Analyst	Analyzes competitors, pricing structures, feature gaps, and customer reviews.
Offer & Pitch Deck Designer	Structures proposals, Grand Slam Offers, and investor pitch decks.
Performance Media Buyer & Paid Ads Strategist	Manages and scales paid advertising campaigns (Google Ads, Meta, LinkedIn, Amazon PPC) based on POAS, contribution margin, and multi-touch attribution.
Telegram Assistant	Mobile Telegram assistant for quick notes, task control in Operations, and research.

Domain: 8. Medicine, Chemistry, Science & Engineering (6 Specialists)

Agent Name	Core Function & Purpose
Chemistry, Materials Engineering & Dangerous Goods (REACH/CLP)	Analyzes chemical compounds, formulations, polymers, building materials, and raw materials, as well as Safety Data Sheets (SDS) and REACH/CLP/GHS compliance.
Clinical Medicine, Guidelines & Evidence-Based Diagnostics	Analyzes medical guidelines (AWMF/WHO), clinical studies, differential diagnostics, laboratory parameters, and pharmacological interactions on an evidence basis.
Mechanical Engineering, Process Engineering & Industrial Standards (DIN/ISO/CE)	Evaluates industrial manufacturing processes, plant and process engineering, Lean/Six Sigma metrics (OEE), as well as CE compliance and DIN/EN/ISO standards.
MedTech, EU-MDR, FDA & ISO-13485 Compliance	Guides medical devices, in vitro diagnostics, and medical software (SaMD) through EU-MDR (2017/745), FDA 510(k), ISO 13485 QMS, and ISO 14971 risk management.
Pharmacology, Drug Discovery & Bioinformatics	Investigates drug-target interactions (IC50/Ki, ADMET), 3D protein structures (PDB/AlphaFold/UniProt), and genomic and sequence data (Ensembl/ClinVar).
Scientific Methodology, Statistics & Peer Review	Produces systematic literature reviews (PRISMA), meta-analyses, statistical design of experiments (DoE, power analysis), and methodological peer reviews.

4. Showroom

Management of Deliverables, Code Artifacts & Python Sandbox



Core Concept & Architecture:

The Showroom is the exhibition and testing space for all finished deliverables produced by your agents. Instead of cluttered chat threads, you find cleanly cataloged artifacts here: from completed HTML apps and generated diagrams to multi-page strategy dossiers. The integrated Python sandbox allows safe script execution.

Key Controls & Sections at a Glance:

Artifact Catalog & Filters: Lists all generated files sorted by date, creator agent, and project. Filters by documents, code, images, and data.
Workspace Vault Reader: Enables structured inspection of project directories directly in the browser, featuring preview capabilities for Markdown, JSON, and HTML.
Python Sandbox (Sub-Tab): Isolated runtime environment where generated Python code can be executed and visualized directly without risking the host system.
Download & Export Buttons: Enables instant downloads of individual artifacts or entire project bundles as ZIP archives.

Practical Step-by-Step Workflow:

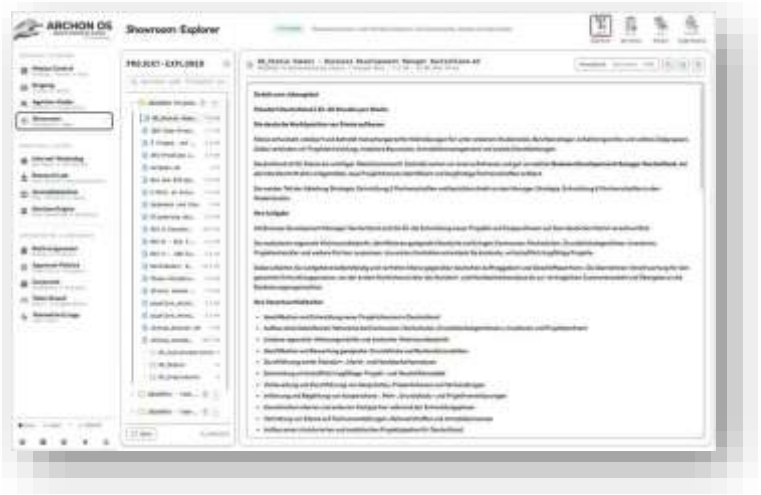
1. After completing a task on the Kanban board, switch to the Showroom.
2. Select the newly generated artifact from the list.
3. Review the live preview (e.g., an interactive spreadsheet or report).
4. If needed, execute code artifacts in the sandbox or download the document.

[Tip] Practical Tip & Guardrail: Use the Showroom for final quality checks. When a document renders without errors here, it is ready for client or partner delivery.

[Prompt] Example Command / Prompt for the Assistant in This Area:
"Show me all artifacts saved in the Showroom over the last 48 hours and open the latest strategy document."

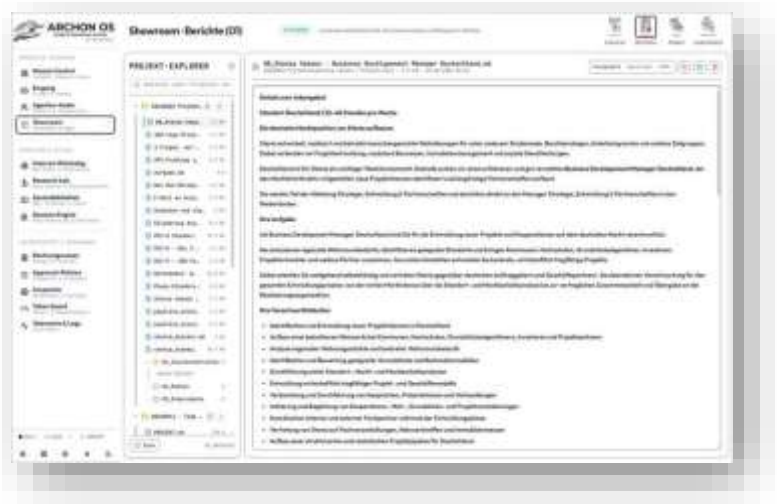
© Michael Keller

4.1 Explorer



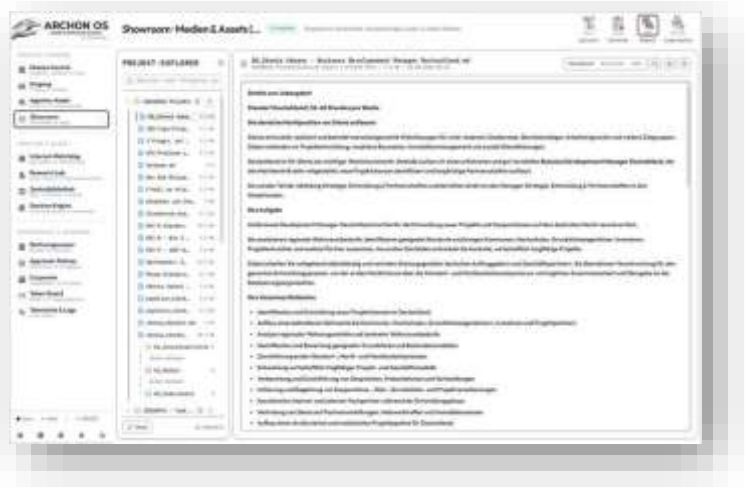
Structured file and project explorer for all deliverables produced by agents.
Filter by project, responsible agent, file type, and creation date.
Direct File Preview: Selected Markdown, JSON, code, and PDF artifacts render in the right preview panel without requiring an external download.
Path & Project Context: Displays the originating task, the linked Obsidian project directory, file size, and modification timestamp for each artifact.
Practical Tip: Use the Explorer to compare intermediate results from multiple subagents directly before signing off on the final deliverable.

4.2 Reports



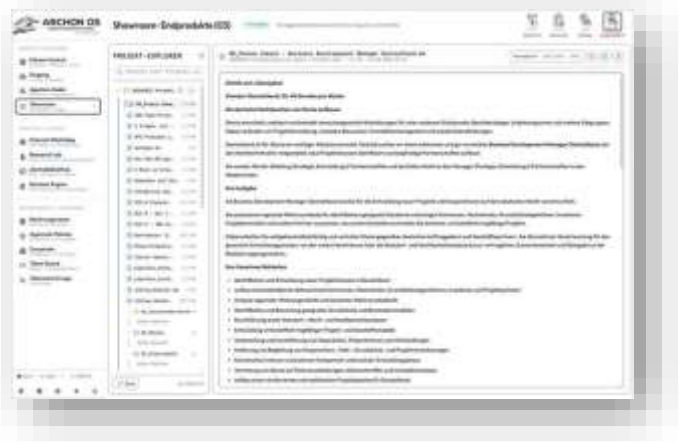
Collection of detailed analysis dossiers, decision memos, concepts, and audit reports.
Direct reading in the integrated Markdown and HTML viewer.
Structured Management Formats: Includes boardroom memos, market analyses, due diligence dossiers, and technical audit logs with clean structuring.
Export & Further Processing: Reports can be exported as formatted print views, PDFs, or directly transferred into the Central Library (RAG Index).
Practical Tip: Ideal for quickly reviewing finished decision memos without having to search the file system for individual Markdown files.

4.3 Media



Gallery for generated architecture diagrams (Mermaid), graphics, mockups, and presentation slides.
Instant preview without switching to third-party programs.
Multimodal Artifacts: Bundles generated presentations, UI mockups, product graphics, and architecture visualizations from the agent workspace.
Metadata & Version Comparison: Displays the triggering prompt, image dimensions, and creator agent for each media file.
Practical Tip: Check generated graphics and diagrams directly in full-screen preview before including them in proposals or manuals.

4.4 Final Products



Ready-to-ship delivery bundles, exportable PDF reports, and executable software prototypes.

4.4.1 Interactive Code Sandbox & Live Artifact Preview

The Showroom (views_core.py & static/views/view_showroom.html) includes an integrated code sandbox:

• Live Execution (POST /api/views/showroom/run-code): When viewing Python or code files, the viewer displays the 'Sandbox Run' button.

• Timeout Watchdog: A fixed 12-second time limit prevents hanging scripts and infinite loops.

• Integrated Terminal Console: Below the viewer, a console opens automatically, displaying execution time in milliseconds, exit code, and STDOUT/STDERR output.

5. Internet Watchdog

The Internet Watchdog module continuously monitors relevant web sources, trade media, competitors, and regulatory publications.

5.1 Radar



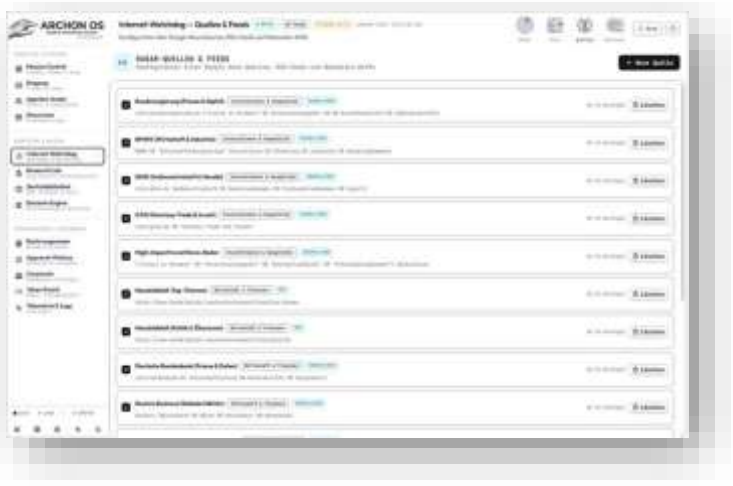
Chronological real-time stream of newly detected articles, market shifts, and legislative changes.
Semantic relevance scoring based on stored search profiles and interest filters.
Signal Prioritization: Hits are automatically classified by topic relevance, novelty, and urgency so that critical alerts stay on top.
Direct Actions per Item: Each card provides quick-action buttons to bookmark in the basket, run an in-depth fact check in the Research Lab, or dismiss.
Practical Tip: Combine narrow keyword profiles with clear exclusion terms to minimize daily background noise in the Radar feed.

5.2 Basket



Clipboard for bookmarked web discoveries and articles awaiting review.
One-click handover to the Research Lab for deep analysis or directly into the Central Library.
Curated Holding Pool: Functions as a working basket for interesting web findings before they are permanently stored in memory.
Batch Processing: Multiple bookmarked sources can be passed as a batch to a research order or compiled into a topical dossier.
Practical Tip: Empty the basket regularly after triaging items to the Research Lab or Central Library so that no unprocessed findings linger.

5.3 Sources



Management of monitored domains, RSS feeds, APIs, and industry portals.
Custom definition of polling intervals and priority tiers.
Source Health Status: Displays the timestamp of the last successful fetch, HTTP status, and hit rate for each configured feed or portal.
Categorization & Tags: Enables grouping by industry (e.g., real estate, AI regulation, competition, tax law) for targeted evaluation.
Practical Tip: Temporarily disable failing or irrelevant feeds instead of deleting them to retain historical reference data.

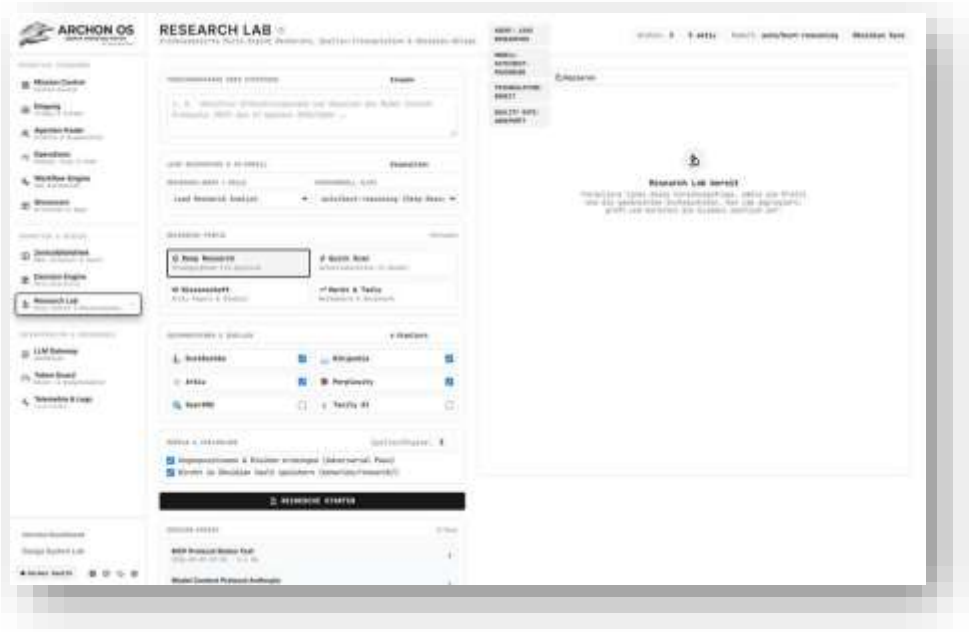
5.4 Options



Configuration of crawler parameters, headless browser options (Playwright), and rate limits.
Thresholds for automated notifications on critical hits.
Scheduling & Throttling: Defines polling windows, maximum concurrency, and hit limits per run to prevent rate-limiting.
Notification Thresholds: Sets the relevance score threshold for actively pushing a finding to the inbox or dispatching an alert.
Practical Tip: Use conservative polling intervals for standard RSS feeds and reserve Playwright headless crawls for dynamic portals.

6. Research Lab

Evidence-Based Research, Source Triangulation & Synthesis Reports



Core Concept & Architecture:

The Research Lab is the scientific research center of your organization. When complex questions arise that cannot be answered from the internal knowledge base alone, the Research Lab steps in. It directs multi-stage search workflows, compares conflicting sources (triangulation), filters out disinformation, and distills findings into clean reports with full source citations.

Key Controls & Sections at a Glance:

Research Intake Form: Define your research question here, set the search depth (Shallow, Standard, Deep Research), and choose target domains.
Source Monitor & Evidence Score: Displays scanned web pages, PDF documents, and specialist portals alongside a reliability index.
Live Synthesis Log: Documents step by step how hypotheses are tested, intermediate findings verified, and counter-evidence weighed.
Dossier Export (Markdown / PDF): Transfers the completed research dossier into the Showroom and Central Library with a single click.

Practical Step-by-Step Workflow:

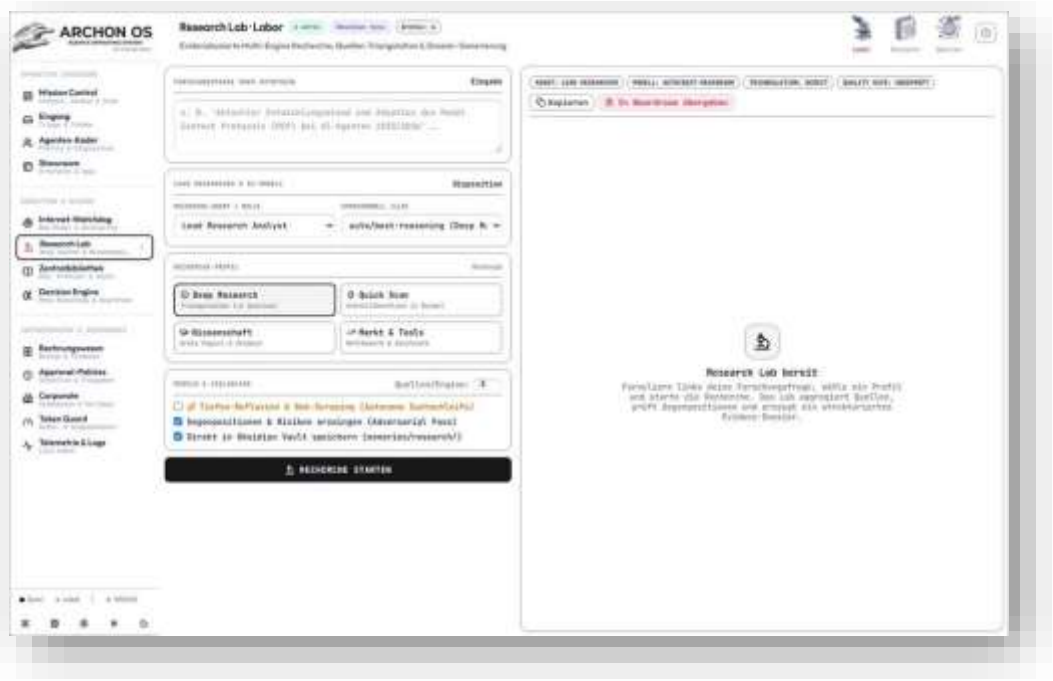
1. Formulate a specific research question (e.g., 'Current subsidies for modular construction in NRW 2026').
2. Select the search depth 'Deep Research' and initiate the run.
3. Track how the research agent filters sources and extracts key citations in the live log.
4. Proofread the completed synthesis dossier and export it to the Obsidian vault of your Central Library.

[Tip] Practical Tip & Guardrail: State research questions as specifically as possible with boundary conditions (e.g., region, year, industry). This saves token costs and delivers highly actionable recommendations.

[Prompt] Example Command / Prompt for the Assistant in This Area: *"Start a research run on the topic 'NRW Building Code Amendment 2026: Key Changes for Approval Procedures'."*

© Michael Keller

6.1 Lab



Starting point for multi-stage deep research with adjustable search depth and domain focus.
Real-time tracking of agent search steps, hypothesis formation, and fact extraction.

6.1.1 Multi-Stage Reflection Loop (Deep Research v2)

The Research Lab (backend/research_engine.py) performs multi-stage scientific and commercial market analyses:

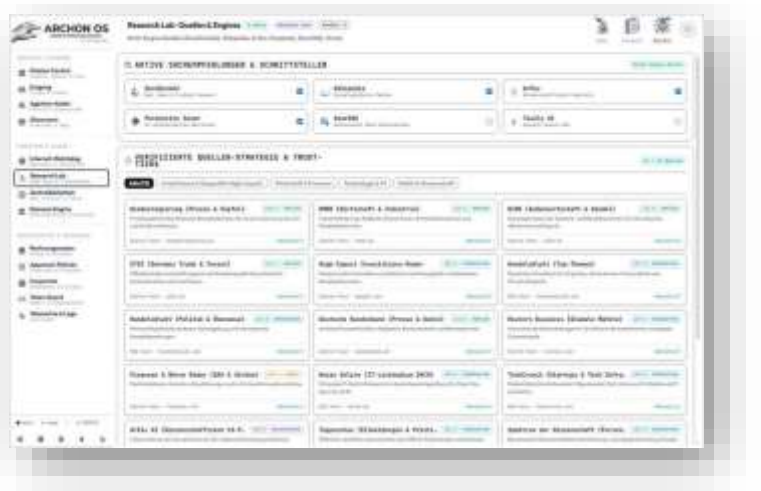
- Re-Search Loop: Decomposes complex tasks into 3–5 sub-questions, evaluates gaps, and autonomously initiates up to 3 follow-up search runs.
- Fallback Cascade: Primary search via Grokipedia and SearXNG with automatic fallback to Jina Reader API upon bot blocks.
- SSRF Protection: Security filters block access to internal networks (RFC 1918) and cloud metadata endpoints.
- Vault Storage: Research dossiers are archived in a structured format directly in data/obsidian_vault/02_Wissensbibliothek/.

6.2 Dossiers



Structured findings reports featuring an executive summary, pro/con analysis, and full footnote citations.
Direct export to the Obsidian vault or as a decision memo for the Decision Engine.
Source-Backed Synthesis: Each dossier links core statements directly to referenced web URLs and internal documents for maximum verifiability.
Status Tracking: Displays the progress of multi-stage investigations (Draft, Fact-Checking, Completed), including participating research agents.
Practical Tip: Pass completed dossiers directly to the AI Boardroom with one click to ground strategic decisions on vetted facts.

6.3 Sources



Directory of all referenced primary and secondary sources, including reliability ratings.
Automated flagging of contradictory statements across different sources.
Trust & Recency Verification: Evaluates referenced primary and secondary sources based on domain authority, publication date, and completeness.
Reuse in Knowledge Graph: Vetted sources can be permanently adopted into the Central Library RAG index.
Practical Tip: When encountering conflicting figures in a dossier, review the underlying original excerpts directly here.

7. Central Library

RAG Index (AnythingLLM), Document Ingestion (Auto-OCR) & Obsidian Vault



Core Concept & Architecture:

The Central Library is the collective memory of your organization. It combines semantic vector search (RAG), automated document ingestion with OCR, and local Obsidian vault management under a unified interface. Knowledge stored here is immediately accessible to all 56 agents for well-founded responses.

Key Controls & Sections at a Glance:

Three Integrated Sub-Tabs (Header): RAG Index, Document Ingestion, and Obsidian Vault.
Help & Sync Switches: Contextual help and manual synchronization of the document catalog.

Detailed Analysis of Subsections & Features:

7.1 RAG Index



Embedded interface to AnythingLLM for semantic vector search and grounded retrieval. Documents are converted into vectors; agents retrieve exact passages matching user queries.

Workspace Management:
Organizing knowledge by projects or business units.
Vector Database Status:
Monitoring indexed document chunks and embedding models.

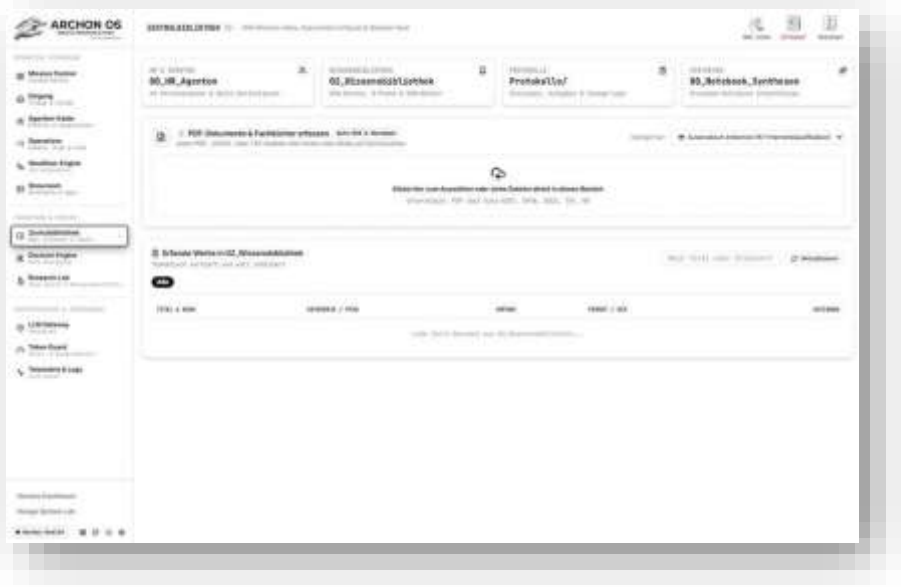
Workflow:

1. Select the desired workspace.
2. Test semantic queries directly in the AnythingLLM chat interface.

Practical Tip: Keep workspaces thematically separated to maximize retrieval accuracy for specialized questions.

© Michael Keller

7.2 Ingestion



The central upload funnel for new technical literature, plans, contracts, and expert reports. PDF, DOCX, and text files are uploaded via drag-and-drop, converted to clean Markdown via auto-OCR using the Docling service, and fully automatically stored in the Obsidian vault and RAG index.

4 Core Hubs:
00_HR_Agenten, 02_Wissensbibliothek, Protokolle/, 00_Notebook_Synthesen.

Upload Dropzone:
Drop files via drag-and-drop; automatic topic classification available.

Catalog of Works:
Overview of all ingested documents with full-text search, category filters, and OCR status.

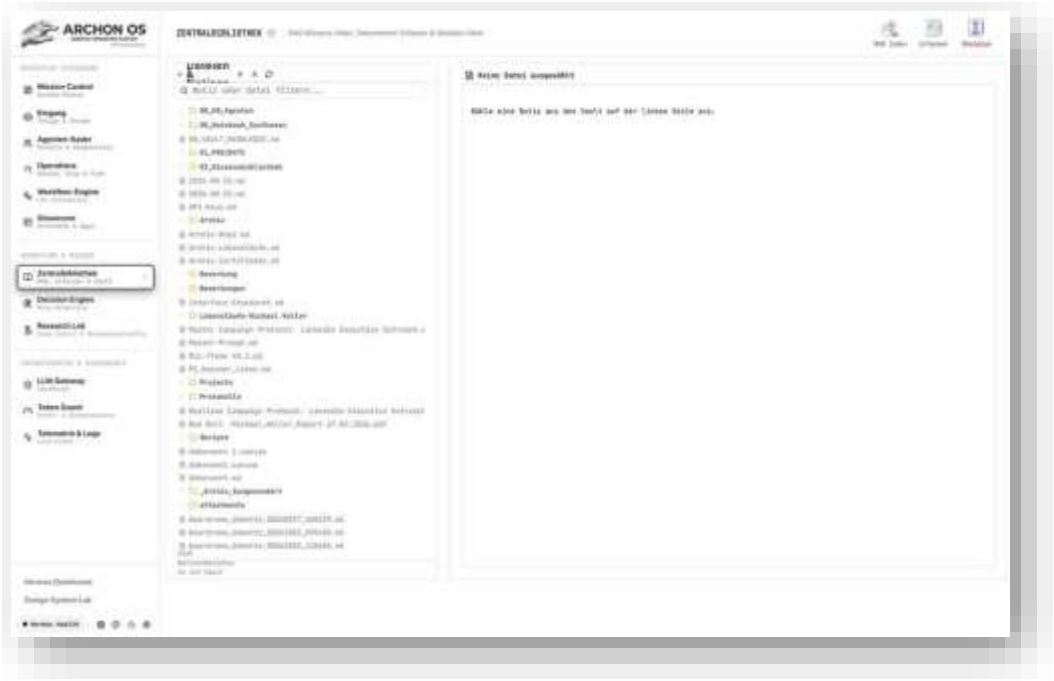
Workflow:

1. Drag a PDF (e.g., an 80-page expert report) into the dropzone.
2. Select a category or choose 'Auto-detect'.
3. Click 'Ingest Now & Import into Vault'. The progress bar visualizes the OCR run.
4. The work appears immediately in the table and becomes available to all agents.

Practical Tip: Thanks to auto-OCR, scanned documents and tables are reliably translated into searchable Markdown.

Docling & Unstructured Table Extraction (t-539b65): Complex PDF tables, balance sheets, and multi-column contracts are converted losslessly into clean GitHub-Flavored Markdown tables.

7.3 Obsidian



Direct access to your physical Obsidian folder on disk. Allows searching, filtering, and reading all Markdown notes and logs directly in the browser.

File Tree (Left Column): Hierarchical folder view of all notes with quick filters and expand/collapse buttons.
Markdown Viewer (Right Column): Formatted reading view of the selected note with path display.

Workflow:

1. Use the search bar above the file tree to filter notes by keyword.
2. Click a file to read its contents immediately in the right viewer.

Practical Tip: All notes visible here reside as standard files on your hard drive, allowing you to open them with external editors at any time.

Practical Step-by-Step Workflow:

1. Upload new technical documents in Document Ingestion via drag-and-drop.
2. Follow the automated OCR conversion and vector indexing.
3. Check the Obsidian Vault to verify that the Markdown file was saved cleanly.
4. Your agents will automatically draw upon this knowledge when answering tasks.

[Tip] Practical Tip & Guardrail: Organize technical literature into appropriate categories. This keeps vector searches fast and precise.

[Prompt] Example Command / Prompt for the Assistant in This Area:
"Which reference books are currently recorded in '02_Wissensbibliothek' and how many documents are indexed in the RAG index?"

8. Decision Engine

Strategic Committee Meetings: 3-Tab Architecture (Overview, Debate, Live Debate), 75/25 Split & OpenRouter Live Pricing ($)



Core Concept & Architecture:

The AI Boardroom & Boardroom Studio simulates high-level management, strategy consulting, and committee sessions directly at the core of the operating system. In version 0.3.3, the module shifted to a modular 3-tab architecture (Overview, Debate, Live Debate). Instead of delegating complex strategic decisions to a single language model, the Boardroom Studio enforces a multi-perspective debate: a designated Chairman steers the session using structured frameworks, while numbered specialist agent slots contribute concrete analytical angles. An integrated Red Team Challenger subjects all drafts to an unvarnished stress test for blind spots before the Chairman synthesizes the binding board memo.

Key Controls & Sections at a Glance:

3-Tab Architecture (Overview, Debate, Live Debate): The 'Overview' tab archives all past sessions in a searchable history with detail popups (task definition and synthesis consensus) and direct jumps to transcripts. The 'Debate' tab provides an ergonomic workbench in a 75/25 layout. The 'Live Debate' tab opens upon start as a focused streaming dialog for real-time observation of the argumentative exchange.
Ergonomic 75/25 Work Layout & File Attachments: The left column (75% width) splits 50/50 between methodology selection (34 frameworks across 7 categories) and task scoping, including file uploads for up to 5 supporting documents (PDFs, balance sheets, expert opinions). The right column (25% width, 100% height) gathers the Chairman and numbered expert slots.
Precise Control Bar & Reset Logic: Equipped with a distinct white start button (2px black border, medium shadow), a separate red stop button without icon, compact 2-line round and reasoning selectors, and a 'Restore Default Slots' reset button that instantly loads the optimal roster for the chosen framework.
OpenRouter Live Price Catalog ($) & Latency Buffer: A subtle dollar icon ($) next to each provider dropdown opens a modal window displaying real-time token pricing (input/output per 1M), context window sizes, and one-click application. A 1-hour RAM buffer in the backend (/api/llm-providers/prices) guarantees fast load times under 2 ms.
CLI Trigger & Autonomous Boardroom Engine: Via 'scripts/trigger_boardroom_debate.py', sessions can also run fully autonomously through the terminal, cron jobs, or background agents (Robert, Micha, Default via Command 4 in AGENTS.md) and archive automatically in the Obsidian Vault (00_Notebook_Synthesen/) and the API history (/api/boardroom/history).

© Michael Keller

Table: Selected Core Frameworks in the AI Boardroom Studio

Framework & Category	Chairman & Ideal Slots	Methodological Flow & Mindset
CLEAR Model (Agile & Innovation)	Robert (CEO) + Architect + Fullstack + Red Team	Concise (Conciseness) -> Logical (Causality) -> Explicit -> Adaptive -> Reflective.
Design Sprint (GV / Jake Knapp)	Product Manager + UI/UX + Fullstack + Customer Success	Map -> Sketch -> Decide -> Prototype -> Validate.
Lean Startup (Eric Ries)	GTM Strategist + Fullstack + DB Tuner + Robert	Formulate hypothesis -> Build smallest viable MVP -> Measure -> Pivot or Persevere.
Pre-Mortem Analysis (Risk)	Executive Decision Architect + Risk + Legal + Red Team	Assumption of total failure in 12 months; backward reconstruction of all causes.
Blue Ocean Strategy	GTM Launch Strategist + Pricing + Competitor Scout	Eliminate -> Reduce -> Raise -> Create new uncontested market spaces.

Practical Step-by-Step Workflow:

1. In the 'Debate' tab on the top left, select the appropriate framework (e.g., 'Pre-Mortem Analysis' or 'First Principles').
2. Frame the task description in the bottom left and attach up to 5 relevant documents via drag-and-drop if needed.
3. Check the Chairman and slots on the right; inspect the most economical model routes via the dollar icon ($) if needed.
4. Click 'Start Debate' – the modular live stream opens and displays theses and opposing assessments.
5. Review the synthesis in the concluding memo or revisit the session at any time via the 'Overview' tab.

[Tip] Practical Tip & Guardrail: Before complex sessions, click the 'Restore Default Slots' button if you previously adjusted profiles manually. The system loads the optimally balanced expert lineup for that specific framework. For automated overnight or cron debates, run the CLI command via 'python scripts/trigger_boardroom_debate.py'.

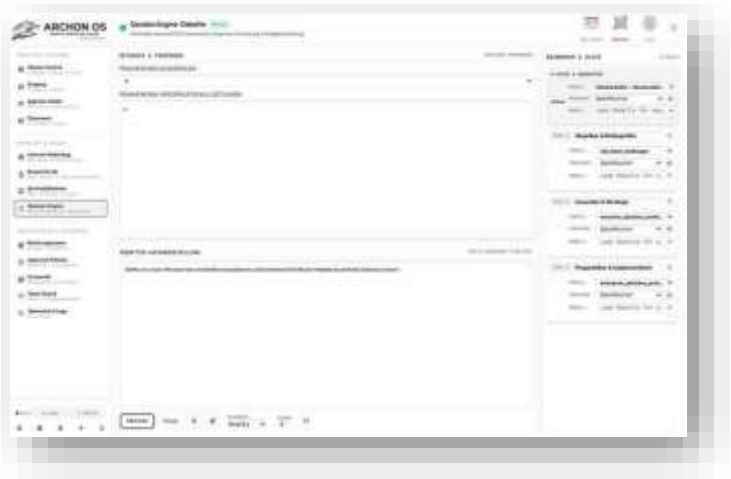
[Prompt] Example Commands / Prompts for the Assistant in this Area:
• UI / Chat: "Start an AI Mastermind session using the 'Pre-Mortem Analysis' framework: 'We are switching all contracts to advance payment. Why might this move fail?'"
• Terminal / CLI: python scripts/trigger_boardroom_debate.py --task "Bewertung Buy vs Build für RAG-Infrastruktur" --framework "First Principles Thinking" --budget balanced --rounds 2

8.1 Overview



Audit-proof archive of all completed boardroom debates, risk assessments, and executive board resolutions. Quick access to historical decision rationales and discarded alternatives.
Decision History & Comparison: Displays the chosen framework, participating C-level roles, costs, and the final vote for each boardroom session.
Dissent Logging: Documented counterarguments and risks remain visible and analyzable even after a decision is passed.
Practical Tip: Reference previous resolutions during recurring budget or strategy discussions to avoid duplicate analyses.

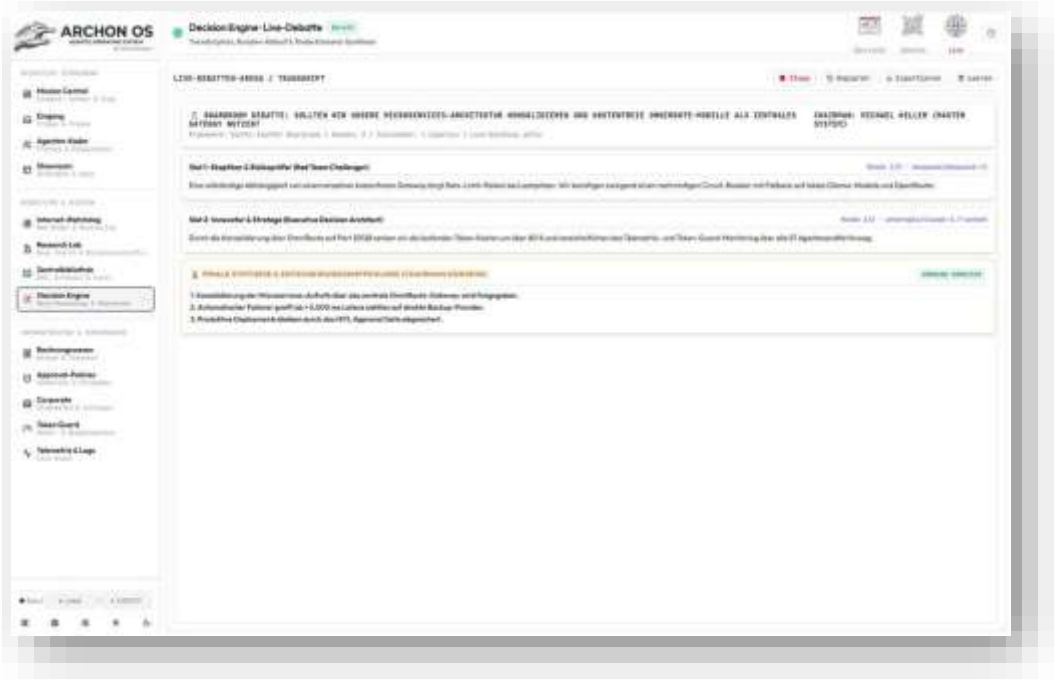
8.2 Debate



Configuration of new committee sessions: choice among 55 thinking frameworks (see accompanying document ARCHON_OS_BOARDROOM_FRAMEWORKS_2026_EN.pdf or Chapter 17.6 Methodologies), C-level staffing, and budget specifications.
Ability to attach internal documents and dossiers as an evidence base.
Framework & Budget Selection: Supports various depth tiers (Free, Cheap, Balanced, Premium) to align model costs with the significance of the decision.
Context Enrichment: Existing projects, CRM data, or dossiers from the Central Library can be attached directly as required reading for the debate.

Practical Tip: Select a contrarian framework (e.g., Pre-Mortem or Red Team Challenge) for critical investment or legal questions.

8.3 Live



Real-time tracking of the multi-stage controversy (theses, cross-examination, Red Team critique, and Chairman synthesis).
Direct conversion of approved measures into actionable Kanban tasks.
Phase-by-Phase Live Progress: Visualizes the six deliberation phases from opening thesis through isolated individual opinions to Chairman synthesis.
Interactive Action Item Creation: Agreed-upon action items can be transferred directly into the Kanban board as prioritized tasks once the session concludes.
Practical Tip: Follow the critique review phase in the live stream to catch blind spots early.

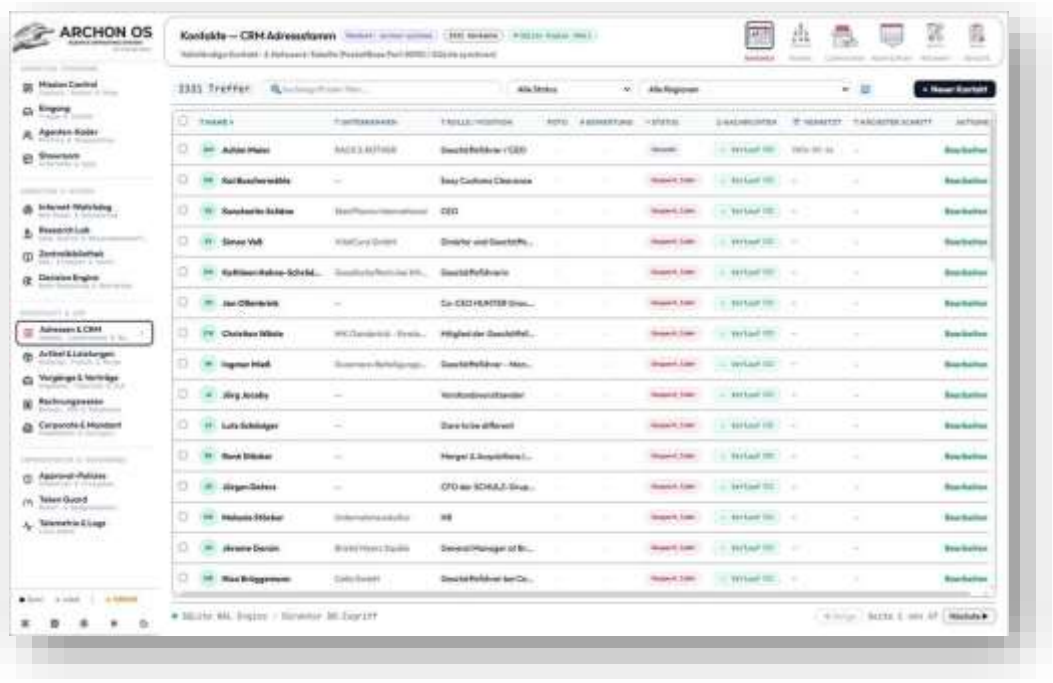
In-Depth Guide: For technical real-time transmission details, see the Markdown system documentation (docs/system/LIVE_STREAMING_SSE_WEBSOCKET_GUIDE.md).

9. Addresses & CRM (Pillar 3: Business & ERP)

The Addresses & CRM module marks the entry point of the third system pillar (Business & ERP). It combines a full-featured B2B CRM (LinkedIn network, email timeline, 360° AI dossiers) with accounting master data management for customers (accounts receivable 10000+) and suppliers (accounts payable 70000+). In compliance with GDPR and GoBD, Archon OS stores all ERP data physically separated by tenant in a dedicated SQLite WAL database at data/tenants/{company_id}/erp.db.

In-Depth Guide: For the complete database and document chain architecture, see the Markdown system documentation (docs/system/ARCHON_ERP_ARCHITECTURE_SPEC.md).

9.1 Contacts (Main Table & CRM Address Master)



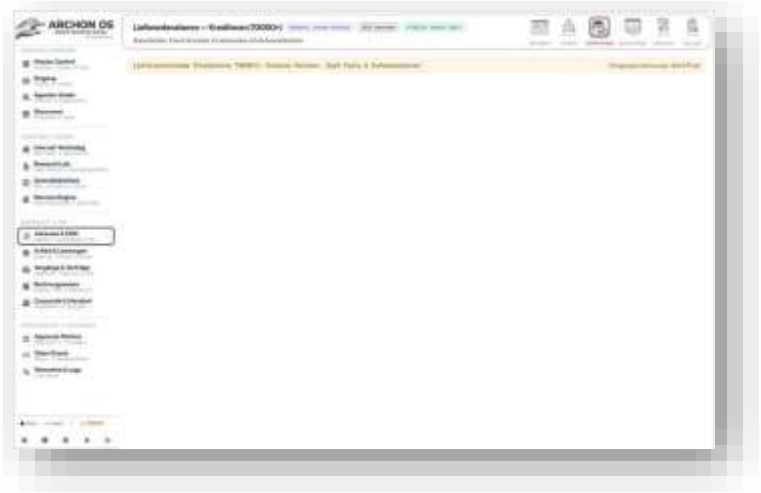
Complete CRM table covering all 2,331 contacts with name, company, position, rating stars, status pills, and interaction badges.
Two-line quick filter (full-text search by name, company, location, or number, plus direct sync with PocketBase port 8090).
Fast pagination (50 contacts per page) with checkbox batch processing and one-click record opening.

9.2 Customer Master (Accounts Receivable 10000+)



Customer and client directory with automatic assignment of sequential accounts receivable numbers in the 10000–69999 range.
Management of payment terms (net days), early payment discounts, VAT ID, credit limits, and linked outgoing invoices.
Accounts Receivable Metrics at a Glance: Lists active customer accounts starting at number 10000, including open quotes, active contracts, and sales volume.
Quick Filter & Segmentation: Filters by customer status, industry, contact person, or open dunning status.
Practical Tip: Clicking a customer entry opens the complete 360° file, including linked transactions and documents.

9.3 Supplier Master (Accounts Payable 70000+)



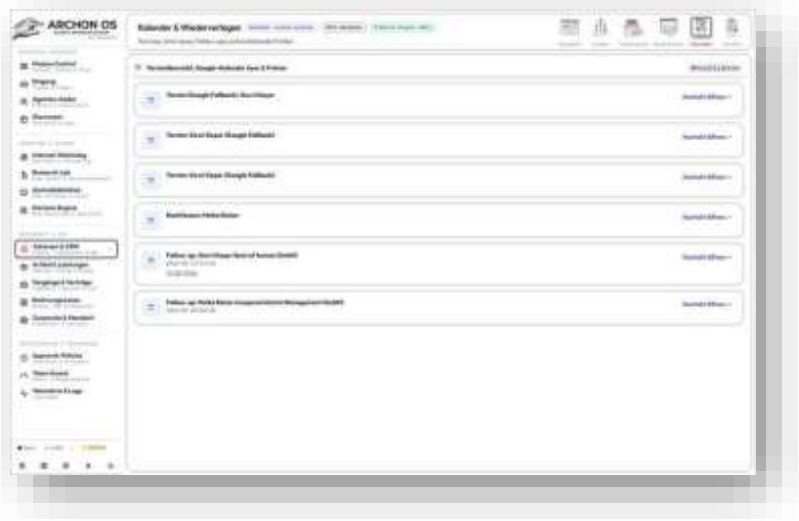
Supplier, service provider, and cloud vendor directory within the DATEV-compliant accounts payable number range starting at 70000.
Recording of IBAN/BIC, incoming invoice workflows, default expense accounts, and linked purchase items.
Accounts Payable Management (70000+): Gathers service providers, software vendors, and subcontractors, including payment terms and recorded bank and tax details.
Incoming Receipt Linking: Directly assigns OCR-captured incoming invoices and running costs to the corresponding vendor account.
Practical Tip: Maintain default payment terms here so that accounting calculates due dates accurately.

9.4 All Messages (Multi-Channel Timeline)



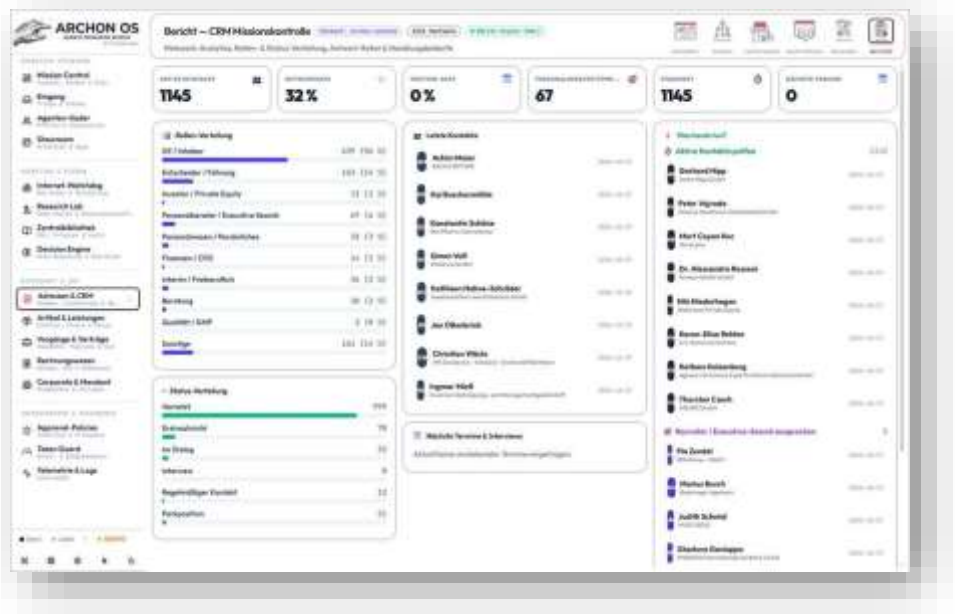
Aggregated chronological feed of all 2,229 messages and interaction histories across all contacts.
Channel filters for LinkedIn (2,220), email (3), in-person meetings (3), audio dictations, and notes (3), plus direction filters (Received / Sent).
Full-text expansion at the push of a button for complete tracking of every dialog thread with timestamps.

9.5 Calendar & Follow-Ups



Calendar overview of all scheduled appointments, contact follow-ups, and upcoming check-ins.
Automatic flagging of urgent or overdue follow-ups to prevent relationship stagnation.
Follow-Ups & Deadlines: Consolidates all sales follow-ups, quote expirations, and contract cancellation deadlines on a timeline.
Responsibilities: Assigns appointments to a human operator or an account sales agent for reminders.
Practical Tip: Check the weekly view at the start of each workday so you never miss an open proposal follow-up.

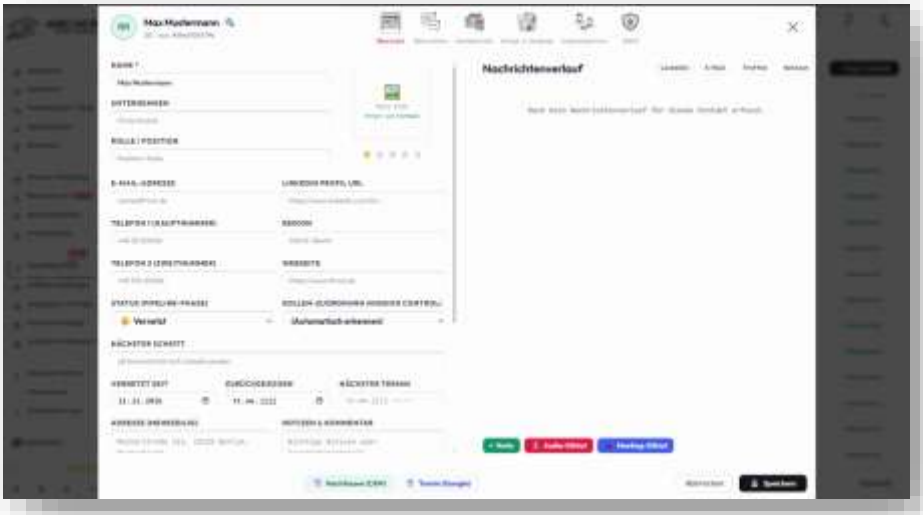
9.6 Report (CRM Mission Control & KPI Analysis)



Real-time KPI cards: active contacts (1,145), reply rate (32%), meeting rate (0%), recruiters (67), and stalled contacts.

Role distribution charts: automatic classification into Managing Director / Owner, Decision Maker, Investor, Consulting, and Finance.
Status distribution of pipeline stages and a daily action radar ('What to do today?') for proactive contact management.

9.7 Record View & Contact Drawer (Enterprise ERP Standard & 360° File)



80-pixel header bar with 6 VARIO sub-views: Overview, Messages, Commercial, Documents & Transactions, Contacts & Branches, and GDPR & Compliance.
Ergonomic window control: close via the darkened backdrop (backdrop click), the Escape key, or the distinct close icon (x) on the top right.
Commercial VARIO master data: Accounts receivable (10000+) and accounts payable (70000+) numbers, payment terms, cash discount rates, credit limits, price groups, bank details (IBAN, BIC), and VAT ID.
Integrated DIN-5008 document workflow: directly view and print quotes (ANG) and order confirmations (AB) with company letterhead and VAT breakdown via 1-click actions.
GDPR & Compliance: 1-click data export as structured JSON (Art. 15/20 GDPR) and legally compliant irrevocable anonymization (Art. 17 GDPR) while fully preserving GoBD accounting integrity.

10. Items & Services (Pillar 3: Business & ERP)

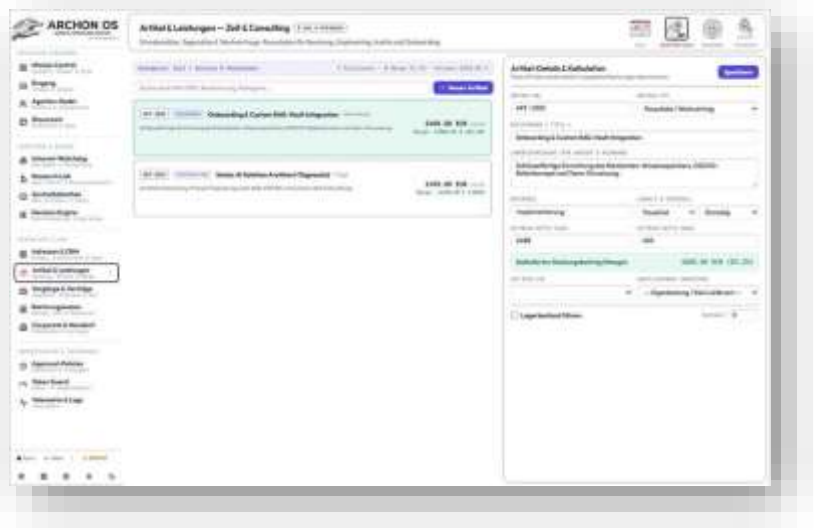
The Items & Services module manages the complete service, product, and license catalog of the active tenant. Each line item calculates the contribution margin in euros as well as the percentage retail markup from net sales price and purchase price in real time, and can be transferred directly into quotes, order confirmations, and invoices.

10.1 Complete Catalog (All Items & Services)



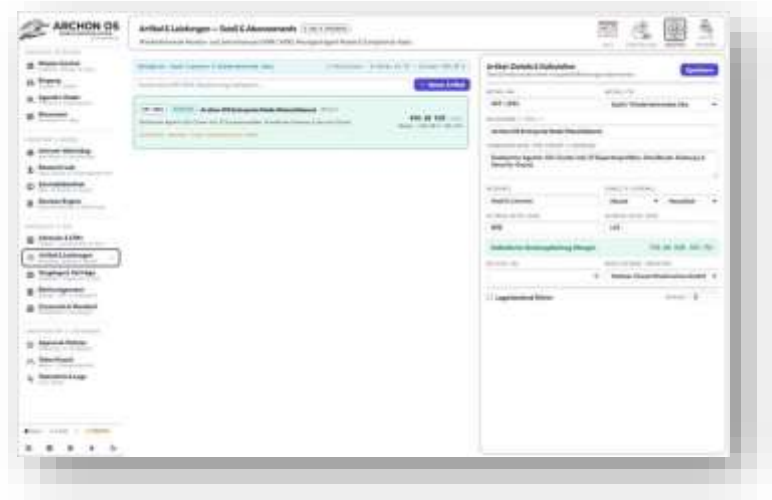
Overview of all active service items, consulting day rates, SaaS licenses, and physical products.
Live calculation tile in the right inspector: automatic calculation of gross profit (VK − EK) and margin percentage with every price change.
Direct assignment of a preferred supplier (accounts payable 70000+), including supplier item number.

10.2 Time / Service & Flat Rates



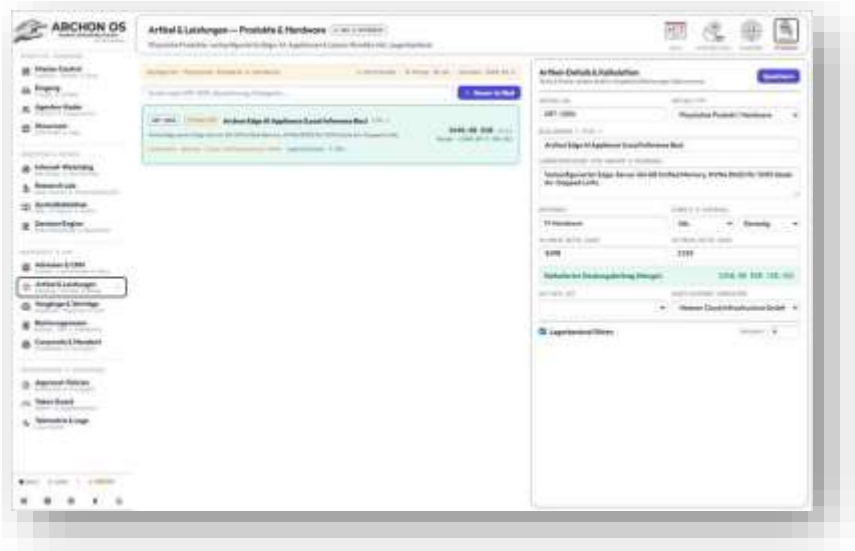
Management of consulting services, architecture workshops, development hourly rates, and fixed-price packages.
Flexible units (hour, day rate, flat fee) with stored service descriptions for proposals.
Hourly & Day Rates: Defines billable consulting, development, and service activities, including default unit, tax rate, and pricing margin.
Quote Templates: Line items can be transferred with a single click into new quotes and orders under Transactions & Contracts.

10.3 SaaS & Subscriptions



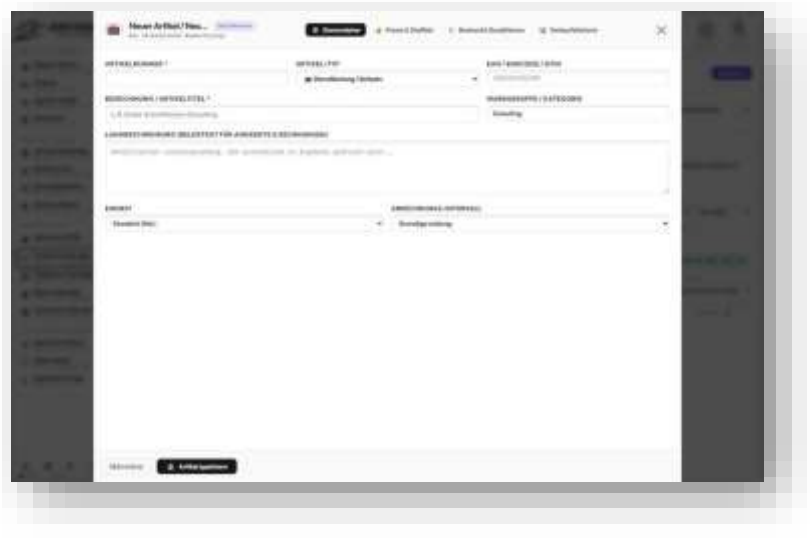
Recurring license, hosting, maintenance, and agent infrastructure subscriptions (MRR / ARR).
Seamless link to term contracts with automated due-date monitoring.
Recurring Revenue Models: Manages monthly or annual licenses, maintenance contracts, and retainers with automated MRR/ARR analysis.
Term & Cancellation Logic: Configured billing intervals feed directly into the contract and deadline radar.

10.4 Products & Hardware (incl. Inventory)



Physical products, edge AI workstations, hardware bundles, and merchandise.
Integrated inventory management with current stock levels, reorder points (minimum stock), and storage locations.
Physical Items & Inventory: Tracks purchase prices, sales prices, item numbers (SKU), and available stock levels.
Supplier Assignment: Links items directly to the preferred vendor for fast reorders and cost calculations.

10.5 Item Drawer & Enterprise ERP Master Data Inspector



80-pixel header bar with 4 modular sub-views following enterprise ERP standards: Master Data, Pricing & Tiers, Inventory & Terms, and Sales History.
Commercial tiered pricing and general ledger revenue accounts (SKR03 / SKR04) with automatic live margin calculation in euros and percent.
Integrated inventory management (minimum stock, storage location, supplier/vendor assignment) and SaaS subscription terms (seats, notice periods, seniority levels).
Document usage & sales history: full traceability across all quotes (ANG) and orders (AB) where the item was used.

11. Transactions & Contracts (Pillar 3: Business & ERP)

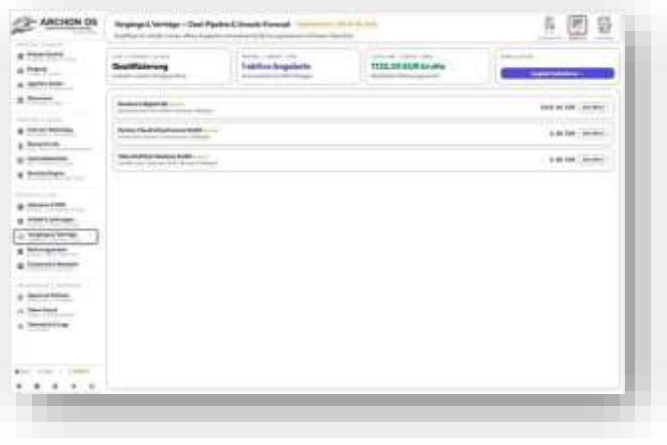
The Transactions & Contracts module forms the operational bracket between sales, project delivery, and accounting. Modeled on modern ERP document flows (e.g. Nuclos), quotes (ANG) and order confirmations (AB) can be converted with a single click and without data loss into GoBD-compliant outgoing invoices (RE) or credit notes (GS).

11.1 Quotes (ANG) & Orders (AB)



Document editor with direct customer selection from the address book and line-item import from the item catalog.
1-click document flow: conversion of a quote (ANG) into an order confirmation (AB) or directly into an outgoing invoice (RE) in the Accounting module.
Automated summation of net volume, sales tax, and gross total amount.
DIN-5008 document printing & PDF export: generation of formal quotes and order confirmations with tenant letterhead, tax IDs, bank details, and acceptance signature fields.

11.2 Deal Pipeline & Weighted Revenue Forecast



Multi-stage sales board covering the phases Lead, Qualified, Proposal, Negotiation, and Closed Won.
Real-time calculation of unweighted pipeline volume and the revenue forecast weighted by closing probability.
Probability Weighting: Evaluates open opportunities by pipeline stage and automatically computes the weighted revenue forecast.
Bottleneck Detection: Flags deals without recent activity for targeted sales follow-ups.

11.3 Contract & Deadline Radar



Monitoring of all customer and supplier contracts, including minimum terms, renewal intervals, and notice periods.
Automated deadline traffic light (Active, Due Soon <= 30 days, Critical <= 14 days) with escalation to Telegram and Operations.
Automated Deadline Warning: Monitors ongoing customer and supplier contracts for upcoming renewals or cancellation cutoff dates.
DocuSeal Status Reconciliation: Directly indicates whether a contract is already signed, still pending, or expired.

12. Accounting (Pillar 3: Business & ERP)

The Accounting module consolidates commercial document processing, financial KPIs, contract monitoring, and receivables management.

12.1 Documents



GoBD-compliant capture and archiving of incoming and outgoing invoices as well as receipts.
Automated OCR extraction of issuer, date, tax rate, net amount, and gross amount.
Incoming & Outgoing Documents: Manages invoices, credit notes, and receipts digitized via Telegram OCR, including tax rate breakdowns.
Status Workflow: Clearly separates Draft, Reviewed, Posted, Paid, and Overdue.

In-depth guide: see docs/system/TELEGRAM_RECEIPT_OCR_GUIDE.md and Appendix II (GoBD & Audit Logging) / docs/security/GOBD_ISO27001_AUDIT_LOG_GUIDE.md.

12.2 Analytics



Financial cockpit for revenue, operating expenses (OpEx), AI token spend, and project contribution margins.
Variance analysis and liquidity forecasting.
Liquidity & Earnings Overview: Visualizes revenues, expenses, open items, and monthly contribution margins over time.
Export Interface: Prepares filtered periods for handoff to tax consultants or Metabase BI dashboards.

In-depth guide: for connecting advanced business intelligence reporting, see the Markdown system documentation (docs/system/BI_DASHBOARD_METABASE_GUIDE.md).

12.3 Contracts



Central contract register with automated deadline and cancellation management (t-0a2171).
Heuristic detection of minimum terms, automatic renewal clauses, and cancellation windows.
Early warnings via Telegram and Kanban tickets for deadlines <= 14 days.
Digital Contract Lifecycle: Bundles PDF contract documents, signature certificates (DocuSeal), and associated billing schedules in one place.
Link to CRM: Every contract remains uniquely linked to debtor/creditor and the original quote.

In-depth guide: for digital contract signing, see the Markdown system documentation (docs/system/ESIGN_DOCUSEAL_GUIDE.md).

12.4 Dunning & Collections



Monitoring of open outgoing invoices and due dates.
Multi-stage dunning engine for generating legally sound payment reminders and dunning notices, including late payment interest calculation per the German Civil Code (BGB).
3-Stage BGB Dunning Run: Automatically prepares Payment Reminder (0 EUR), 1st Dunning Notice (+5 EUR), and Final Notice (+10 EUR plus default interest).
Mandatory HITL Gate: No dunning letter leaves the system without explicit prior approval by the operator.

In-depth guide: see the complete specification in the Markdown system documentation (docs/system/DUNNING_ENGINE_GUIDE.md).

13. Corporate & Tenants (Pillar 3: Business & ERP)

The Corporate & Tenants module manages all company master data (multi-tenant with a physically isolated SQLite database per entity), corporate design guidelines, and official business templates for correspondence and billing.

13.1 Master Data



Multi-tenant capability (Multi-Company): management of multiple companies or locations with fast switching via the header dropdown and global tenant switcher in the left sidebar.
Physical DB isolation: each tenant has its own strictly segregated database (data/tenants/{company_id}/erp.db).
Legal details & address: company name, legal form, executive management, street, postal code, city, and contact details.
Commercial register & taxes: registration court, HRB/HRA number, tax ID, VAT identification number (VAT ID), and small business status (§ 19 UStG).
Bank details & corporate design: IBAN, BIC, financial institution, as well as primary/secondary colors, typography, and company logo.

13.2 Forms & Templates



Interactive template studio with real-time preview, automated integration of stored company master data, and direct selection of recipients from the address book.
DIN 5008 business letter: standard-compliant letter sheet with fold marks, address zone, information block, and mandatory footer details.
Outgoing invoice (§ 14 UStG): complete invoice template with line item calculation, net/VAT disclosure, and payment terms.
Non-disclosure agreement (NDA): ready-to-use template for mutual confidentiality agreements.
Direct print & PDF export to DIN-A4 via the integrated print function.

14. Approval Policies

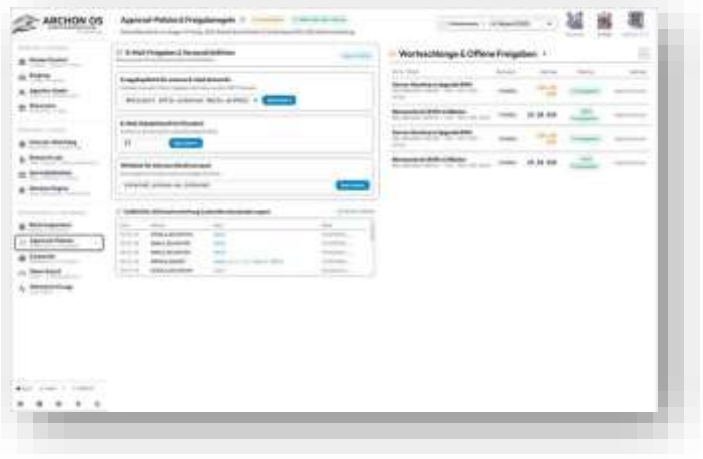
In the Approval Policies, you define which actions agents may execute autonomously (dark processing) and where human approval (Human-in-the-Loop) is strictly required.

14.1 Finance



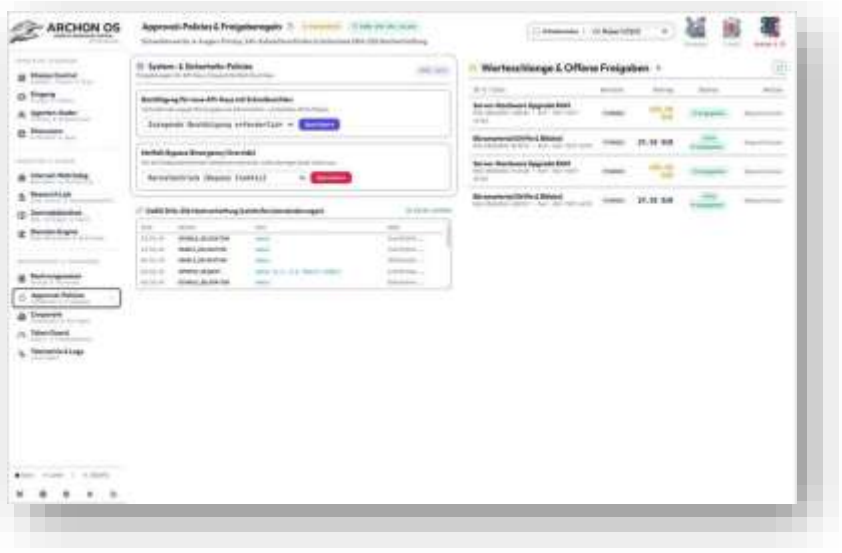
Approval limits for expenses, subscriptions, invoice sign-offs, and budget reallocations.
Four-eyes principle above configurable euro thresholds.
Tiered Amount Limits: Specifies the euro threshold up to which expenses or entries pass automatically, and when manual approval becomes mandatory.
Audit Logging: Every threshold change is immutably documented in the policy hash chain.

14.2 Email



Ruleset for outgoing correspondence: mandatory approval for external recipients, authorities, or contract proposals.
Protection against unvetted legal commitments to external parties.
Recipient & Domain Rules: Distinguishes between internal drafts, known existing clients, and critical external addressees (e.g., authorities, banks).
Mandatory Preview: Blocks direct sending and deposits complete response drafts into the inbox for review.

14.3 System & IT



Safeguarding critical infrastructure interventions (database migrations, container restarts, configuration updates).
Automated data classification per ISO 27001 / BSI (t-53315b) into PUBLIC, INTERNAL, CONFIDENTIAL, and STRICTLY_CONFIDENTIAL with blocking of external cloud transmissions for sensitive data.

In-depth guide: see Appendix II (GDPR, Encryption & Human-in-the-Loop (HITL) Approval Gates) and docs/system/GDPR_COMPLIANCE_GUIDE.md.

15. Token Guard

Cost Guardrails, Rate Limits, Infinite Loop Protection & Safety Controls

[Prompt] Example Command / Prompt for the Assistant in This Area: *"What is the current status of the daily budget in Token Guard, and were any safety cutoffs triggered today?"*

Core Concept & Architecture:

Token Guard is the financial and hardware protection control center of Archon OS. It prevents budget overruns, infinite loops, and CPU spikes by analyzing every transaction in real time, enforcing hard limits, and delivering cost transparency down to the individual session level. Across three specialized sub-views, the guard bridges strategic budget planning and technical hardware throttling.

Sub-View 1: Overview & Daily Trend Dashboard:

Bundles 4 KPI cards (total tokens in period, cumulative cost in USD with daily budget display, number of active safety plugins, and top consumer profile). A live trend chart visualizes token and cost spikes by day. Directly below, two tables break down usage across agent profiles and LLM models. The right column houses the safety barrier form (Max Context, Max Chat History, Daily Budget, Task Budget, Max Steps, Session Timeout, and Cache TTL).

Sub-View 2: Single-Session Usage Inspector:

Enables a deep drilldown into every individual chat and workflow session. The sortable table displays session ID, responsible profile, language model, call count, input tokens, output tokens, tokens saved via semantic caching (cache read), total cost, and timestamp.

Sub-View 3: Resources & Concurrency Throttling (Hardware Protection):

Controls the global asyncio semaphore for concurrent agent execution. An interactive slider (1 to 16 concurrent tasks) allows flexible adjustment to the host hardware: 1 task for sequential safe mode, 4–6 tasks for smooth multitasking, up to 16 tasks for dedicated workstations.

The 14-Tier Token-Saving and FinOps Ecosystem:

Archon OS integrates 14 coordinated cost-saving mechanisms:

1. 3-Tier cost caps (LOW: max $0.10/1M, MEDIUM: max $1.00/1M, HIGH: No Cap).
2. FreeLLM API Docker proxy (Port 3002) with auto/best-free routing for zero-dollar routine jobs.
3. Stepless reasoning budget (0 to 16k tokens) in the chat footer.
4. 24h Evey cache for instant, zero-cost repeat responses.
5. Max-context (32k) and max-history (20 turns) hard clippers.
6. Dynamic token clipper for small models while preserving the system prompt.
7. Sliding window memory compression targeting the [CONSOLIDATED DECISION CONTEXT] block.
8. LRU-K page swapping for subagent RAM once 80% quota is reached.
9. JEV router for zero-token pre-filtering of deterministic system commands.
10. ASD-STE100 prompter eliminating filler words (30–40% savings).
11. Anti-AI-slop filter to prevent costly conversational filler output tokens.
12. Pre-flight cost estimator prior to launching extensive tasks.
13. Dynamic task budget with 80% economy downgrade and 100% hard stop with Telegram push notification.
14. Loop killer halting execution after 15 steps or Jaccard similarity > 90%.

Table 15.1: Core Parameters & Thresholds of the Token Guard

Protection Mechanism	Default Threshold	Impact & Safeguard
Daily Budget	$2.00 / day (warning at $1.50)	Prevents unexpected monthly bills through automated API halts.
Task Budget	$0.50 / task (downgrade at $0.40)	Downgrades to economy models at 80%; pauses at 100% with a Telegram push notification.
Max Steps (Loop Killer)	15 steps / similarity > 90%	Reliably terminates recursive tool loops and reasoning spinlocks.
Concurrency Throttle	5 concurrent tasks (slider 1–16)	Limits parallel LLM calls; protects CPU and API quotas against overload.
Cache TTL	24 hours	Answers identical system and status queries in 0 ms at 0 token cost.

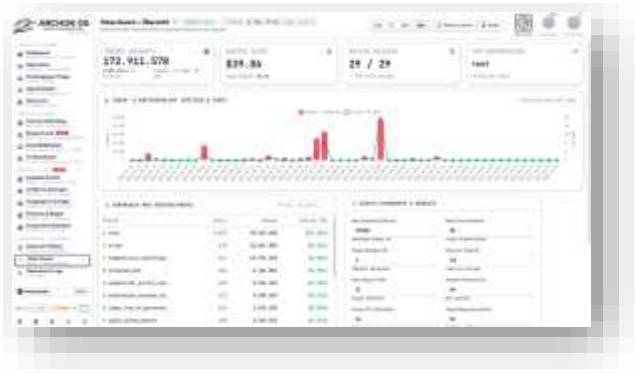
Practical Step-by-Step Workflow:

1. Check the current cost status against the daily budget ($2.00) in Sub-View 1.
2. In case of elevated usage, switch to Sub-View 2 (Sessions), sort by 'Cost', and identify the dialogue causing the spike.
3. Adjust the concurrency slider downward in Sub-View 3 (e.g., to 3 tasks) if your machine is under heavy load.
4. If needed, click 'Clear Cache' at top right to discard cached model responses.

[Tip] Practical Tip & Guardrail: Route routine tasks through the 3-Tier category 'LOW' or 'auto/best-free'. This allows the free Docker proxy (Port 3002) to absorb baseline traffic, preserving your daily budget for strategic tasks using Claude 3.5 Sonnet or GPT-4o.

[Prompt] Example Command / Prompt for the Assistant in This Area:
"Show me the current daily consumption in Token Guard and list the three most expensive sessions from the last 24 hours."

15.1 Overview



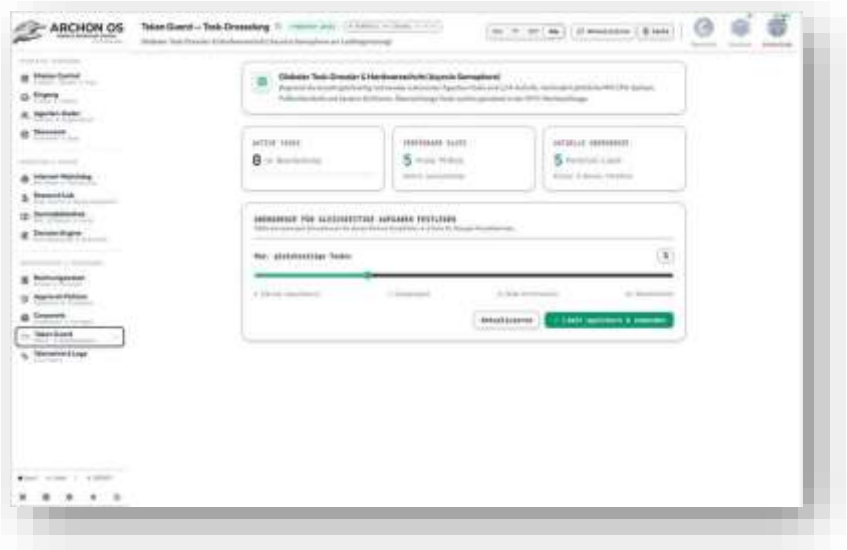
Real-time tracking of token and cost consumption across all providers, models, and specialized agents.
Budget traffic light for daily and monthly limits.
Provider & Model Breakdown: Shows at a glance which models and agents account for the largest share of daily token costs.
Cache & Savings Rate: Measures efficiency gains from local semantic caching and zero-cost routing paths via OmniRoute.

15.2 Sessions



Breakdown of token requirements for individual sessions and tasks (input vs. output tokens).
Detection of unusually large context windows or infinite loops.
Single-Session Inspector: Allows terminating or resetting individual agent sessions exhibiting abnormally high token consumption.
Root-Cause Analysis: Separates prompt context (input) from generated replies (output) to quickly identify overloaded system prompts.

15.3 Throttling

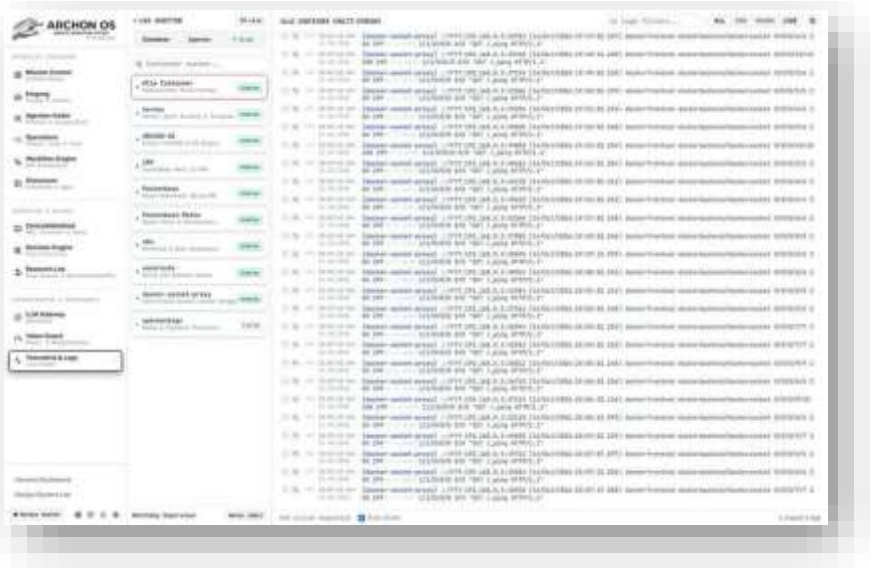


Configuration of soft limits (warning notification at 80%) and hard limits (automated halt or fallback to free models at 100%).
Automated Circuit Breakers: Intervenes when nearing daily or monthly budgets with a soft warning (80%) and hard spend cap or fallback.
Loop Protection: Limits maximum tool calls and execution steps per task so runaway agent loops do not drain budget.

In-depth guide: see Appendix II (API Keys & Zero-Egress Privacy Routing) and docs/security/API_KEYS_SCOPES_RATELIMITS_GUIDE.md.

16. Telemetry & Logs

Second-Precision Audit, System Telemetry & Cryptographic Trace Feed



Core Concept & Architecture:

Telemetry & Logs is the flight recorder system of Archon OS. Every single action—from the smallest API ping to tool executions and user approvals—is captured with second precision and logged tamper-proof. Should unexpected behavior occur, this stream provides the complete diagnostic chain for rapid root-cause analysis.

Key Controls & Sections at a Glance:

Three Specialized Log Sources (Sub-Tabs): • Container logs: physical stdout/stderr output from the 11 Docker containers.
• Agent logs: console output and reasoning traces from Hermes agents.
• Audit trail: tamper-proof hash-chain log of all business approvals.
Search & Filter Bar: Real-time filtering by log level (DEBUG, INFO, WARN, ERROR), container name, or search terms.
Auto-Scroll & Pause: Freezes the stream while investigating an error without interrupting background data ingestion.
Download & Export: Exports the filtered log excerpt for external analysis or archiving.

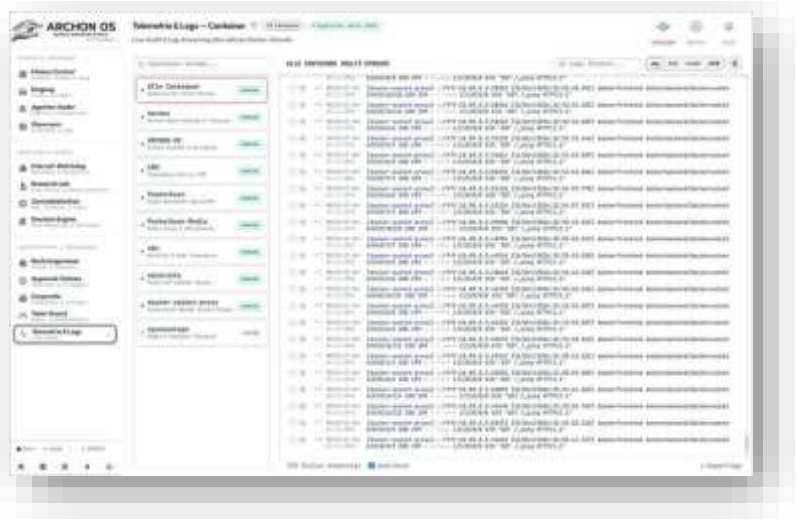
Practical Step-by-Step Workflow:

1. Upon anomalies in Mission Control, switch to the log stream.
2. Filter by level 'ERROR' or by the name of the affected agent.
3. Examine the stack trace and preceding prompt to pinpoint the exact failure source.
4. Use findings to refine prompts or tool permissions in the Specialist Fleet.

[Tip] Practical Tip & Guardrail: Use the audit trail for compliance documentation. Cryptographic chaining enables seamless verification of who approved which action at what exact time.

[Prompt] Example Command / Prompt for the Assistant in This Area:
"Show me all warnings and errors (WARN, ERROR) in the log stream from the last 2 hours."

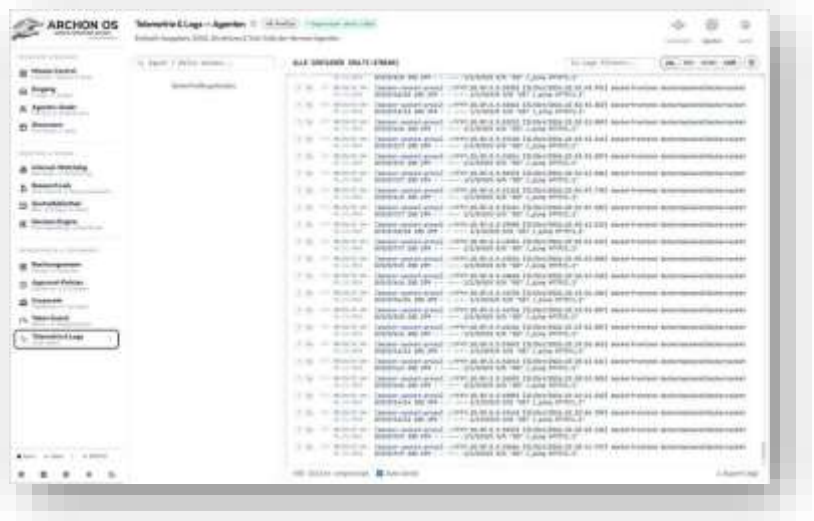
16.1 Containers



Live metrics for all Docker containers (latencies, CPU/RAM utilization, and health checks).
Direct visibility into stdout and stderr logs of individual services.
Service Control Overview: Displays status, uptime, and resource usage across all running Docker containers (Archon, Hermes, OmniRoute, n8n, PocketBase).
Diagnostic Logs: Enables searching recent container messages directly without SSH or terminal access.

In-depth guide: see docs/system/PROMETHEUS_GRAFANA_MONITORING_GUIDE.md and docs/system/CENTRAL_LOGGING_LOKI_GUIDE.md.

16.2 Agents



Chronological trace stream of all agent reasoning steps, tool invocations, and model responses.
Filtering by agent ID, severity level (INFO, WARN, ERROR), and time window.
Real-Time Execution Log: Traces which model an agent selected, which tools were invoked, and the resulting execution duration.
Troubleshooting: Filters targeted warnings, timeouts, or aborted tool executions for specific roles.

In-depth guide: for deep observability, see the Markdown system documentation (docs/system/LLM_TRACING_LANGFUSE_GUIDE.md).

16.3 Audit



Tamper-proof, cryptographically chained audit log of all security-relevant system actions.
Orderly server shutdown via GracefulShutdownCoordinator (t-12113b) with task draining and complete SQLite WAL flush.
Cryptographic Traceability: Each entry is chained to the predecessor hash via SHA-256, ensuring any subsequent tampering is immediately flagged.
Compliance Verification: Logs GDPR deletions, policy changes, approvals, and break-glass access actions without gaps.

17. System Control Center

Central System Administration, Daemon Kernel, Multi-Provider Management & MCP Hub

The System Control Center (accessible via the gear icon at the top right of Mission Control or at the bottom left of the sidebar) consolidates complete Archon OS system administration into a modular fullscreen modal. The window header also features a global live settings search (keyboard shortcut '/') to jump directly to individual parameters, alongside the 'Live Dock' button (which docks the dialog to the right of the active interface to monitor design and layout adjustments in real time). Structurally, the Control Center is organized into eleven dedicated tabs (17.1 to 17.11):

Practical Step-by-Step Workflow:

1. Open the System Control Center via the gear icon; if needed, use the search bar ('/') to jump directly to terms like 'Auth', 'Ollama', 'Telegram', or 'Backup'.
2. Configure your LLM API keys and local models (Ollama / LM Studio) under the 'API Keys & Access' or 'Provider & Gateway' tab, and define fallback and zero-egress privacy routes.
3. Verify under the 'Containers & Services' tab that all active core services (ARCHON-OS, Hermes Agent, AnythingLLM, Chroma, Socket Proxy, optionally n8n/OmniRoute) are running without errors.
4. Configure your MCP tools as well as PreDecision, prompt-injection defense, AI Quality Gate, and Cost Guard under the 'Plugins' tab.
5. Enable multi-user login for LAN/server operation under 'Users & Roles', and configure Telegram, IMAP/SMTP, and n8n webhooks under 'Communication'.

[Tip] Practical Tip & Guardrail: After modifying provider keys or container assignments, use the 'Test' button in the respective tab before applying changes to production. For deep kernel adjustments, the YAML live editor with automated backup creation is always accessible under 'Hermes Core'.

Tab / Section	Scope & Functions	Default Configuration
Hermes Core	Kernel scheduler, watchdog (120s), model quick-select, YAML editor	Port 8420, tickless sleep active
Containers & Services	Monitoring of the 11 Docker containers, service restart, cache reset	11/11 services nominal in mesh
Provider & Gateway	API keys for OpenRouter, Gemini, Groq, Anthropic; latency checks	OmniRoute AI Router (Port 20128)
Methodologies	Management of 35 analysis and prompt frameworks; create custom SOPs	35 Standard Base Frameworks
Plugins & MCP	FastMCP drivers (filesystem, web-search, CRM, tools), scopes	Read-Only default, write protection active
Communication	Telegram bot token, chat IDs, CRM email sync (IMAP), webhooks	Telegram polling active, HITL emergency call
Design	Dark/light theme, company logo upload, color palettes, CSS sync	Archon Slate Dark Mode active

> **[Prompt] Example command / prompt to the assistant for this area:**"Open the System Control Center and show me container statuses and the currently configured default provider in Hermes Core."

17.1 Hermes Core



Kernel, SOUL.md tiers & 15 daemon sections (Port 8420):
Controls the background daemon on port 8420 and system-wide switching across the 3-tier SOUL directives (Compact ~3,100 chars / Pro / Full) across all active agent profiles. The 15 configuration groups cover standard model & routing (model), thinking limits (agent), streaming (display), multiplexing/platforms (gateway), code-execution allowlists, tool-loop guardrails, context compression & prompt caching, memory/skill synthesis, subagents & MoA, browser engine, voice STT/TTS, configuration backups, and encrypted break-glass emergency access.

System diagnostics & 360° health score (GET /api/system/health-score):
Continuously aggregates all system vital signs into an overall score from 0 to 100%: Docker core fleet (35% weight), concurrency queue (20%), cryptographic audit integrity (20%), deadlock freedom (15%), and AnythingLLM RAG reachability (10%), including actionable recommendations.

Autonomous Master Dossier Generator (POST /api/system/generate-dossier):
Generates at the touch of a button the complete system architecture dossier along with all Docker services, concurrency slots, FinOps mechanisms, 70 specialist profiles, and security gatekeepers, syncing the Markdown file simultaneously to the project directory and Obsidian vault (00_DOKUMENTATION/SYSTEM_ARCHITEKTUR_MASTER.md).

In-depth guide: for emergency access and resilience testing, see Appendix II (Emergency Vault / Break-Glass) and docs/system/CHAOS_RESILIENCE_GUIDE.md.

17.2 Autonomy & Modules (36 Subsystems)



Sections 1–3: Maintenance, Resilience & Governance (Modules 1–11):
Section 1 controls nightly memory consolidation (sleep-cycle maintenance). Section 2 (Package A, Modules 2–6) calibrates concurrency & deadlock watchers, loop killers / circuit breakers, task self-healing with Git rollback, ephemeral Docker sandboxes, and time-travel debuggers. Section 3 (Package B, Modules 7–11) manages consensus gates & arbitration councils, virtual war rooms (pre-mortem), SLA handoff contracts, inbox watcher / morning standup, fleet self-healing, secret maskers, Playwright screenshots, and SQLite backup rotation.

Sections 4–8: Enterprise Packages 1 to 5 (Modules 12–36):
Complementing the kernel watchers, Sections 4 through 8 control the 25 enterprise modules directly within this tab: (4) Performance Core & Throughput (speculative pre-fetching, token efficiency), (5) Modern Memory & RAG Architecture, (6) Reliable Execution & Determinism (durable execution), (7) Security, Data Privacy & Enterprise Governance (Zero-Egress Privacy Routing, Dual-LLM Guard, PII scrubber), and (8) Self-Optimization, Evaluation & Swarm Intelligence (MCTS action trajectories, JIT personas, PbD computer use).

Table 17.1: The 36 Configurable Autonomy, Resilience & Enterprise Modules (settings.json / API)

Group & Module	Toggle (On/Off)	Adjustable Thresholds & Parameters
1. Concurrency & Deadlock	auto_deadlock_recovery	max_concurrent_agents (4), queue_timeout (60s), stagnation (45s)
2. Loop Killer & Clipper	loop_detection, token_clipper	max_identical_turns (3), jaccard_threshold (0.90), max_steps (15)
3. Task Self-Healing	enabled, auto_git_rollback	max_attempts (3), escalation_model (deepseek/deepseek-chat)
4. Ephemeral Sandbox	enabled, network_enabled	memory_limit_mb (512), cpu_limit (1.0), timeout_seconds (30)
5. Time Travel & Snapshots	tracing_enabled, auto_snapshots	max_snapshots (50), max_trace_steps (100)
6. Consensus Arbitration Council	enabled	min_priority (P3), mode (tri_agent_council), majority (2/3 | 3/3)
7. Virtual War Room	enabled, auto_pre_mortem	block_threshold (MEDIUM | HIGH | CRITICAL), default_model
8. SLA Contracts & Handoff	enabled, auto_reject_incomplete	min_content_length (100 chars), max_turnaround_minutes (60)
9. Inbox & Morning Standup	enabled, standup_enabled	interval_seconds (5s), standup_time (08:00), watch_folder
10. Fleet, DB & Hygiene	auto_restart, secret_mask, audit	db_backup_keep_days (7 days), visual_browser_audit (Screenshots)

Key operational in-depth modules from the 'Autonomy & Modules' tab in detail (17.2.1 to 17.2.10):

17.2.1 Replay & Time-Travel Debugging (Execution Trace Player & State-Forker)

Complex agent runs often encompass dozens of sub-steps across tool invocations, file modifications, and model queries. If a process fails at step 14, conventional workflows required tedious log analysis and restarting the entire task from scratch.

How the Trace Player operates (backend/time_travel_debugger.py):

- Seamless logging: every sub-step is archived with input parameters, LLM return value, runtime, token consumption, and atomic snapshot ID in the 'execution_traces' table.
- Replay API (GET /api/traces/{task_id}): provides the exact chronological sequence of the task for step-by-step cockpit visualization.
- State forking (POST /api/traces/{task_id}/fork): allows branching a task from any intermediate step N. The system retains all prior steps and resumes execution with an alternative model or modified instructions—ideal for what-if analyses without full restarts.
- Atomic rollback (POST /api/traces/{task_id}/rollback): reverts all file changes made after step N back to their previous state without data loss via compressed gzip snapshots (.hermes/snapshots/).

17.2.2 Multi-Model Head-to-Head Benchmark & Eval Matrix (OmniRoute Arena)

Selecting the optimal model for a specific agent role (e.g., deep research vs. code refactoring) in Archon OS relies on empirical benchmark data rather than guesswork.

How the Eval Matrix operates (backend/model_benchmarker.py):

- Simultaneous A/B testing (POST /api/router/benchmark): sends the identical production prompt concurrently to two or more configured LLMs (e.g., DeepSeek V3 vs. Claude 3.7 Sonnet).
- Metrics capture: records exact latencies, token estimates, output lengths, and automatically identifies the speed and efficiency winner.
- Persistent history (GET /api/router/benchmark/history): stores all comparative benchmark runs in the 'model_benchmarks' table, keeping cost and latency trends traceable over months.

17.2.3 Dual- & Tri-Agent Consensus Gate & Arbitration Council (Debate / Voting Protocol)

For high-risk tasks (database migrations, system file updates, critical approvals from priority 3 upward), Archon OS prevents single-agent hallucinations through a mandatory two- or three-party review principle.

Arbitration council process flow (backend/consensus_gate.py):

- Adversarial review (POST /api/tasks/review-consensus): an independent auditor agent rigorously reviews outputs for logical gaps, hallucinations, and security vulnerabilities.
- Tri-Agent Council (POST /api/consensus/council-vote): for highest-priority decisions, a three-member panel convenes comprising System Architect (code quality & scalability), Security Officer (zero trust & secrets), and QA Lead (edge cases & usability).
- 2/3 majority rule: approval is granted only when at least 2 of the 3 panel members agree. All votes, rationales, and scores are logged audit-proof in 'consensus_decisions' (GET /api/consensus/history).

17.2.4 Autonomous Task Self-Healing with Git & Snapshot Rollback

Faulty syntax, incomplete code blocks, or unexpected runtime exceptions trigger immediate task failure in conventional systems. Archon OS implements an autonomous self-healing loop.

Self-healing cycle (backend/task_self_healer.py):

- Error capture & working tree protection: if a script or build command fails, the supervisor captures the entire stack trace and reverts modified files via 'git checkout --' or the snapshot manager immediately to the last known good state.
- Introspective repair loop (POST /api/healing/attempt): the agent receives the error message along with the stack trace and an explicit directive for root-cause analysis and remediation.
- Concurrency & escalation: the system permits up to 3 autonomous remediation attempts. If successful, the task continues smoothly; if all 3 attempts fail, the task is marked 'blocked' and surfaces with full diagnostics on the HITL board for manual inspection (GET /api/healing/logs).

17.2.5 Event Mesh & Webhook Trigger Bridge (Inbound Automation & Directory Watcher)

To integrate Archon OS deeply into enterprise and cloud workflows, the event mesh acts as a universal inbound and outbound channel for asynchronous events.

How the Event Mesh operates (backend/event_mesh.py):

- Inbound webhook bus (POST /api/events/inbound/{channel}): external systems such as GitHub (PRs, issues), Stripe (incoming payments), or n8n can feed events directly into the system, automatically initiating corresponding Kanban tasks.
- Automatic directory watcher (POST /api/events/scan-directory): Scans the inbox directory 'data/inbox/' for newly arrived documents (PDFs, contracts, notes), generates event logs, and transfers them directly as triage tasks into the inbox.
- Outbound event dispatching: Automatically sends webhook signals to external endpoints on status changes or milestone achievements.

To grant agents maximum operational freedom without risking host system stability or losing control over expanding long-term data, Archon OS expands its runtime environment with advanced autonomy and protective shields. Developer agents receive ephemeral execution environments, the system cleans itself automatically at night, and strategic initiatives are simulated in advance.

17.2.6 Ephemeral Code Sandboxes (Ephemeral Micro-Containers & Runners)

When autonomous agents generate code, compile, or execute scripts, a faulty command must never jeopardize the entire system.

How the Ephemeral Sandbox Works (backend/ephemeral_sandbox.py):

- Short-lived micro-containers (POST /api/sandbox/execute): Each code execution starts in a temporary, isolated Docker container (e.g., python:3.11-alpine). The system enforces strict limits: 512 MB RAM, 1 CPU, a 30-second execution timeout, and optional full network blocking (--network none).
- Automatic destruction: Once the code finishes (or times out), the container is deleted without residue. Memory leaks, filesystem damage, or infinite loops stay confined to the ephemeral shell.
- Developer feedback: The system returns precise execution statistics (runtime, exit code, stdout, stderr, used memory) directly into the task stream.

17.2.7 Nightly Memory Consolidation & Sleep-Cycle Maintenance

During continuous operation, log streams, traces, and context snapshots accumulate rapidly. An unpruned system gradually loses query performance.

Maintenance cycle & configuration (backend/memory_consolidator.py & static/views/system_settings_modal.html):

- Idle phase pruning: When the system is inactive (default: 03:00 AM at night), Archon OS runs a consolidation cycle. Outdated trace logs older than the retention period are purged, non-critical heartbeat logs are removed, and the SQLite database is compressed via VACUUM.
- Full control in system settings: Under the 'Hermes Core' tab (Section 6), users can flexibly adjust time, retention period in days (e.g., 7, 14, or 30 days), similarity threshold (0.88–0.95), and log pruning.
- Manual instant trigger: Use the 'Consolidate manually now' button (POST /api/system/consolidate-memory) to trigger the cleanup and optimization process manually at any time.

17.2.8 Headless Browser & Visual Web Agent (Playwright Screenshot Audit)

Many research and interaction tasks fail when using purely text-based scrapers because modern web portals require dynamic JavaScript, logins, or form submissions.

How the Visual Web Agent Works (backend/visual_web_agent.py):

- True browser remote control (POST /api/browser/browse-and-capture): The agent controls a full headless Chromium browser via Playwright, executing clicks, text input, and navigation.
- Continuous step screenshots: For each sub-step, a visual screenshot is saved in the .hermes/visual_traces/{task_id}/ folder.
- Audit-proof verification: Inside the task dossier, the operator can visually verify what the agent saw on screen and selected at any given moment.

17.2.9 Virtual Pre-Mortem Stress Test (War Room Simulation)

Before major projects, releases, or publications go live, human teams often display optimism bias. Archon OS counteracts this with an automated counter-review.

Simulation & attack scenarios (backend/war_room_simulator.py):

- 3-party attack (POST /api/simulation/war-room): Three specialized adversary roles challenge the planned initiative in a pre-mortem analysis: a critical enterprise customer (usability, churn), an offensive red-teamer (attack vectors, data leaks), and a strict compliance auditor (liability, GDPR).
- Risk classification: The system aggregates all findings in a risk matrix (LOW, MEDIUM, HIGH, CRITICAL) and formulates mandatory remediation requirements prior to approval.
- Audit history (GET /api/simulation/war-room/history): All stress tests remain permanently documented for historical traceability.

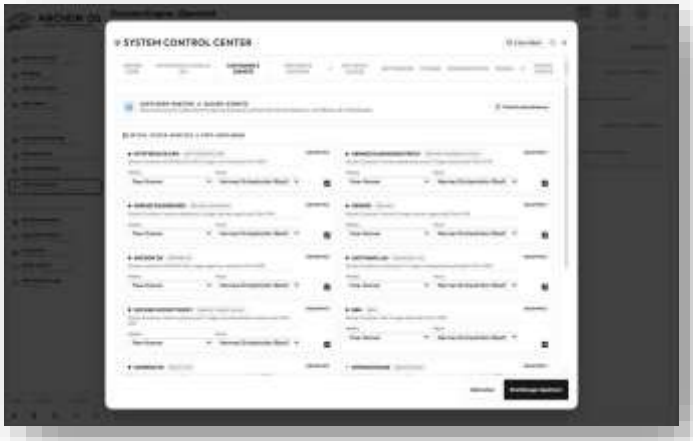
17.2.10 SLA & Handoff Contracts Between Agents (Contract Enforcement)

In layered multi-agent hierarchies, loose delegations often produce incomplete deliverables that ultimately burden the human operator.

Contract protocol workflow (backend/sla_contract_enforcer.py):

- Digital handoff contract (POST /api/contracts/create): Every inter-agent delegation defines mandatory keys (e.g., summary, result, sources), minimum text length, and explicit acceptance criteria.
- Autonomous acceptance check (POST /api/contracts/validate): The receiving agent automatically verifies the delivered sub-result against the contract conditions.
- Automatic rejection: If the output does not meet the standard, the handoff is autonomously rejected with a precise defect list ('violations') and sent back for revision—without requiring user intervention.

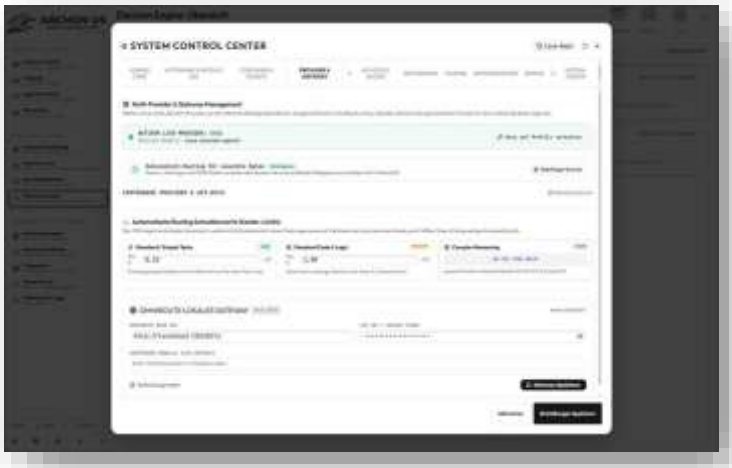
17.3 Containers & Services



Container runtime, active services & path management:
Monitors the live status of all active containers and port bindings in the Archon network (ARCHON-OS :8501, Hermes Core :8644, AnythingLLM RAG :3001, Chroma :8000, n8n :5678, OmniRoute :20128, PocketBase :8090/:8092, CRM :3005, Docker Socket Proxy). Services can be selectively restarted via the socket proxy; the tab also displays active system paths and allows cache purges.

In-depth guide: See docs/system/LITESTREAM_REPLICATION_GUIDE.md and Appendix II (TLS & SSO) / docs/security/REVERSE_PROXY_SSO_GUIDE.md.

17.4 Providers & Gateway



Multi-provider management, task routing & cost thresholds:
Configures cognitive infrastructure across cloud APIs (OpenRouter, Gemini, OpenAI, Anthropic, Groq, DeepSeek), OmniRoute (:20128), and local AI servers (Ollama / LM Studio). Unlike generic keys, you configure model routing per task class (code, research, standard chat, vision) alongside automated routing thresholds (cost ceilings, fallback cascades, and zero-egress rules for sensitive data).

Navigation note: The LLM Gateway (OmniRoute) can be accessed directly via the CPU icon in the bottom left of the status bar (port 20128) as well as through this tab in the System Control Center.

Multi-Provider AI Routing, Cost Optimization & Resilience

Core Concept & Architecture:

OmniRoute is the high-performance intelligent gateway (default port 20128) for all model requests in Archon OS. It decouples the agent core completely from individual model providers and standardizes communication according to the OpenAI-compatible schema. If a provider fails or hits a rate limit, OmniRoute switches to defined secondary routes within milliseconds—completely transparent and uninterrupted for the user.

Where APIs & model combinations are configured:

API keys and model routes are configured centrally via two paths:
1. Graphically in the System Control Center ('Providers & Gateway' tab): Store API keys for OpenRouter, Google AI (Gemini), Groq, OpenAI, and Anthropic here. Use the 'Test Connection' button to verify latency for each provider prior to activation.
2. File-based via the kernel directory: In config/hermes_config.yaml and local secret management, provider hierarchies and custom model combinations can be permanently defined.

Dynamic fallback cascades & combinations:

To deliver maximum resilience at minimal cost, OmniRoute uses prioritized fallback chains:
• Primary tier: GLM-4.7 Flash / auto/best-free – Maximum speed at $0.00 cost for routine tasks and pre-filtering.
• Secondary tier: Google Gemini 2.5 Flash – High reliability and massive context buffer for large-scale text synthesis.
• Tertiary tier: Anthropic Claude 3.5 Sonnet – High-precision reasoning for architecture, code, and contract analyses.
• Sovereign emergency fallback: Local Llama 3.3 70B via Ollama/vLLM – Activates during complete internet or cloud outages.

Key Controls & Sections at a Glance:

• Multi-provider vault: Encrypted management of all provider keys in a single location.
• Latency & error monitor: Visualizes response times (in ms) and HTTP error codes (429, 500) for early detection of provider disruptions.
• Automatic circuit breaker: Temporarily isolates unstable endpoints and reroutes requests without losing tokens.

Table 17.2: Technical Specifications of the OmniRoute AI Router

Gateway Feature	Specification & Path	Function in Live Operation
Gateway Service	Port 20128 (http://localhost:20128/v1)	Central proxy endpoint for all LLM calls.
Configuration Location	System Control Center -> 'Providers & Gateway'	Graphical key management and latency testing.
Config File	config/hermes_config.yaml	File-based definition of model combinations.
Health Check	http://localhost:20128/health	Provides uptime and provider availability metrics for the watchdog.

Practical Step-by-Step Workflow:

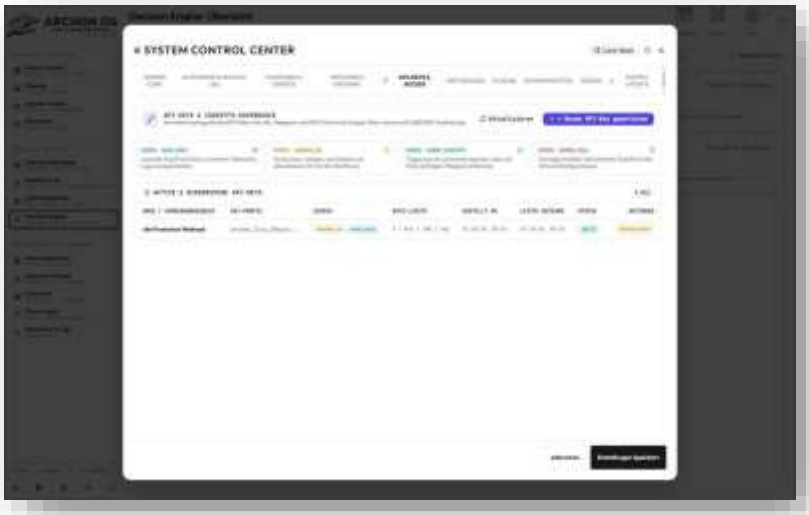
1. Monitor latencies and availability of configured providers in the OmniRoute dashboard.
2. Adjust the default route as needed (e.g., 'auto/best-free' for maximum budget efficiency).
3. If a provider fails (e.g., due to rate limits), observe how OmniRoute transparently switches to the fallback.

[Tip] Practical Tip & Guardrail: Run standard tasks through free routes. Always configure at least one stable fallback provider (e.g., Google Gemini Flash or Groq) so the overall system continues operating during API outages.

[Prompt] Example Command / Prompt for the Assistant in This Area: *"Which LLM provider currently offers the lowest latency and how many requests were routed through OmniRoute today?"*

In-depth guide: For training custom local language models, see the Markdown system documentation (docs/system/MODEL_FINETUNING_GUIDE.md).

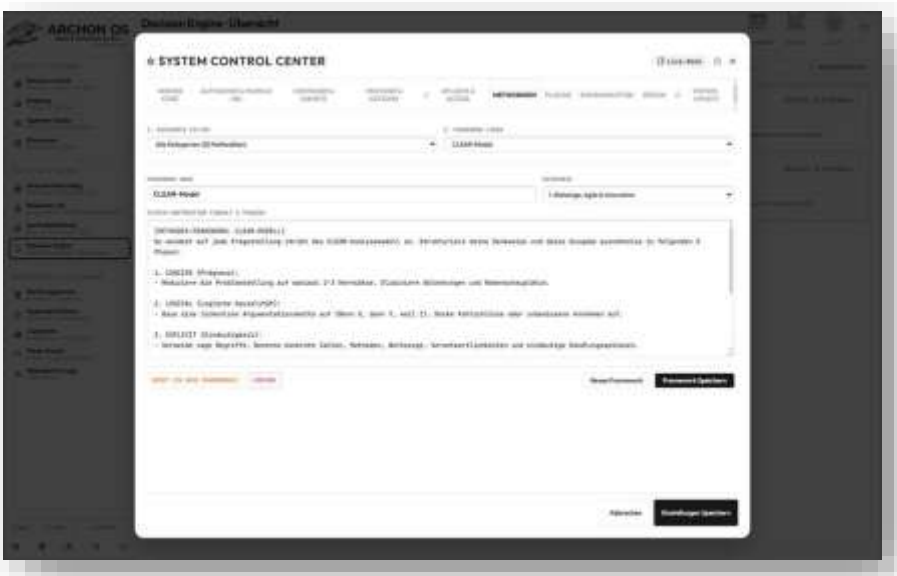
17.5 API Keys & Access



Management of LLM provider keys (OpenRouter, OpenAI, Anthropic, Gemini, Groq) and external integration tokens (AnythingLLM, Telegram, GitHub), including connection testing.
API keys & access governance (backend/api_auth.py): Create, display, and instantly revoke scoped Archon REST API keys (read:tasks, write:tasks, execute:agent, admin:system), including token-bucket rate limiting.

In-depth guide: See Appendix II (Secrets & Key Rotation) as well as docs/security/SECRETS_MANAGEMENT_UND_ROTATION.md and docs/system/CONFIG_VALIDATION_GUIDE.md.

17.6 Methodologies



55 decision frameworks & prompt SOPs:
Manages the 55 solution frameworks anchored in the AI Boardroom and Prompt Studio (Agile/Lean, root cause analysis, debate, strategy/moat, game theory, growth, business models, operations, and resilience). You can edit existing templates directly, add company-specific methodologies, or reset to factory defaults.

17.7 Plugins

Plugin ecosystem, MCP tools & AI protection engines:
Alongside standard FastMCP drivers (filesystem, web search, email, Git), this tab bundles five active AI control and guardrail subsystems: (1) PreDecision (TypeSafe AI fast routing), (2) Adaptive Memory & Experience Learning (Pillar 4 with learning loop, LRU-K decay rate & Obsidian consolidation), (3) Prompt Injection Protection for emails/CRM including a live injection test sandbox, (4) AI Quality Gate & Hallucination Check including a test sandbox, and (5) Cost Guard & Token Budgeting.
The modular plugin system (backend/plugin_manager.py) decouples extensions from the core:

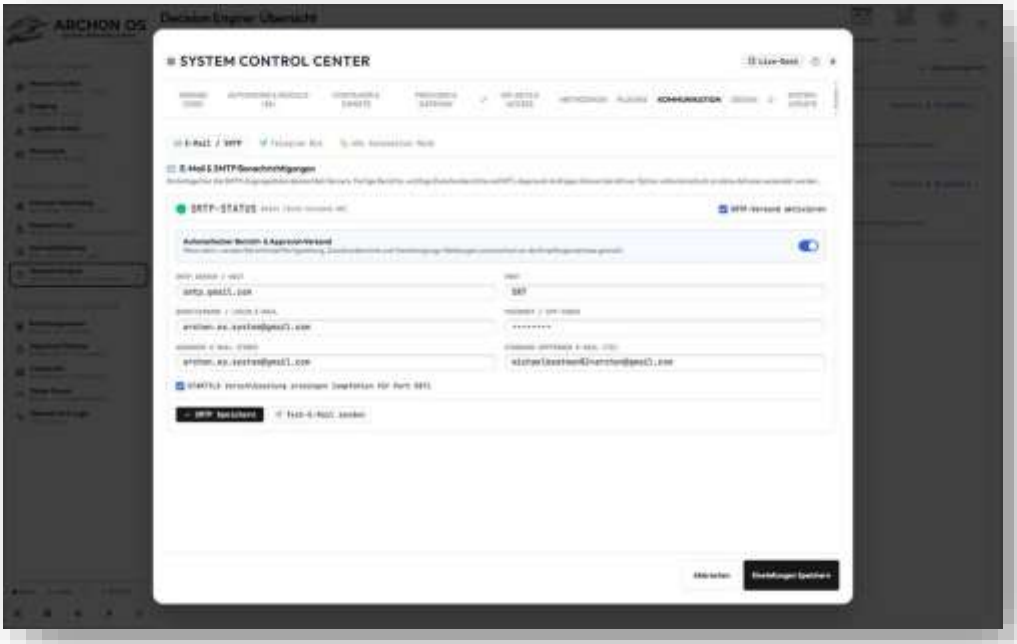
• Module structure: Each plugin resides in plugins/<plugin_id>/ with a plugin.json manifest and a plugin.py implementation.

• Lifecycle & tool hooks: on_load, on_enable, before_tool_execute, and after_tool_execute to intercept sensitive invocations.

• Preinstalled starter plugins: System Metrics (hardware telemetry), Security Audit Hook (command filtering & secret protection), and Cost Guard (token budget).

In-depth guide: See the Markdown system documentation (docs/system/ARCHON_MCP_SERVER_GUIDE.md).

17.8 Communications



Email/SMTP, Telegram 2-way stream & n8n mesh:
Configures all external communication and webhook channels in one place: (1) Email & SMTP notifications and inbox synchronization, (2) Telegram Bot integration & 2-way stream including token/chat ID, and (3) n8n Workflow Automation Mesh.

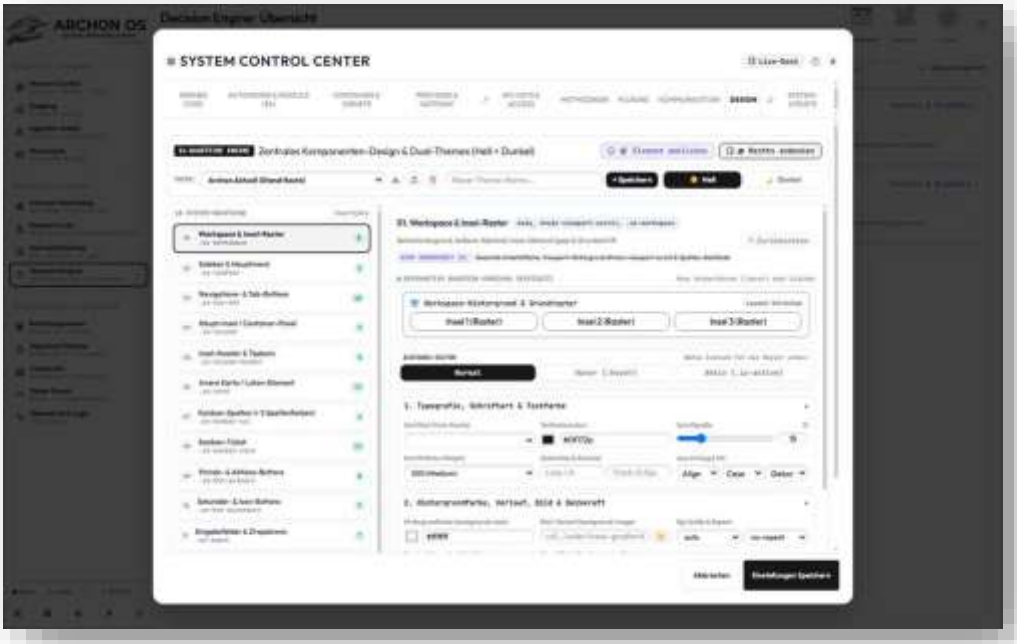
The voice pipeline (backend/voice_pipeline.py) enables mobile voice control via Telegram:

• Voice-to-text: Voice memos are received and transcribed accurately via Whisper (Groq, OpenAI, or Gemini).

• Intent routing: Call-signs ('Robert', 'Anwalt', 'Katharina') are recognized and route the task to the appropriate agent.

• Edge-TTS speech synthesis: The agent response can be generated using the German neural voice 'de-DE-ConradNeural' and returned to Telegram as an audio bubble.

17.9 Design



Theme engine, component designer & live CSS editor:
Controls the entire design system of Archon OS: switching between light and dark themes, company logo upload, and granular color, surface, shadow, form, skill badge, and Kanban column palettes for all UI cards. An integrated live CSS editor synchronizes adjustments directly with the Visual CSS Live Studio (Chapter 19) and can be checked immediately on the actual dashboard via 'Live-Dock'.

17.10 Users & Roles

Users & Roles (RBAC): Activates and manages the 3-tier role model (admin, operator, viewer) in multi-user operations (backend/auth_service.py).
Create new user & authentication: Creates user accounts with PBKDF2 password hashing, manages the blocking login gate (mandatory recommendation as soon as ARCHON_BIND_IP=0.0.0.0 is enabled for LAN access), and manages active sessions.
Permission scoping: Strictly separates administrative actions (container restarts, API key issuance, system updates) from operational agent tasks and read-only access.

17.11 System Update

System update & server sync (backend/system_updater.py): Displays uncommitted local changes live and synchronizes development machine, GitHub, and home server with a single click.

The update system (backend/system_updater.py) synchronizes the development machine with the remote server:

• GitHub push from settings: Clicking 'Push changes to GitHub' transfers the current status with an automated commit via Docker-optimized Git flags (-c core.checkstat=minimal).

• Pre-update snapshots & rollback: Before every update, Archon OS creates consistent backups of all SQLite databases in data/snapshots/. In case of errors, clicking 'Rollback' restores the prior commit and data state.

In-depth guide: See docs/system/ANSIBLE_INFRASTRUCTURE_GUIDE.md, docs/system/CI_PIPELINE_E2E_TESTING_GUIDE.md, and docs/system/AUTOMATED_RESTORE_TEST_GUIDE.md.

18. Special Feature: Archon Prompt Studio (832 Templates & Fabric SOP)

Enterprise Prompt Engineering, 255+ Fabric Patterns & 5-Field Structure



Core Concept & Architecture:

The Archon Prompt Studio is an advanced prompt engineering tool embedded at the heart of the operating system. It solves the biggest issue when working with LLMs: vague, incomplete, or inconsistent instructions. The studio aggregates 824 field-tested enterprise templates (across 16 domains) and over 255 Fabric Standard Operating Procedures (SOPs). Its modular 5-field architecture structures every task into a mathematically precise form (Role, Context, Task, Output Format, Constraints). The module is built with high performance and utility, allowing it to stand on its own as a dedicated product or Chrome extension.

Key Controls & Sections at a Glance:

Catalog & category filters (832 templates in 16 categories): Filters by disciplines such as strategy, legal, coding, marketing, data analysis, and Fabric patterns.
The 5-field formula for precise prompts:
1. Role (Who acts?): e.g., 'Senior Compliance Officer'.
2. Context (What is the baseline situation?): Company background, stakeholders, target audience.
3. Task (What exactly needs to be done?): Precise action order in the active voice.
4. Format (What should the output look like?): Table, Markdown dossier, JSON schema, Executive Summary.
5. Constraints & negative rules (What is prohibited?): No assumptions, German Commercial Code (HGB), max. 500 words.
Live preview & variable injection: Renders the final prompt in real time during input and highlights placeholders in color.
Action buttons: 'Apply to Task' & 'Apply to Chat': Transfers the finished prompt with a single click directly into the Kanban task modal or agent chat.

Practical Step-by-Step Workflow:

1. In the Operations Board, click 'Prompt Studio (824)' under '+ New Task' or inside the agent chat.
2. Select a Fabric pattern (e.g., 'extract_wisdom' or 'write_legal_opinion') or an enterprise template.
3. Adjust the 5 fields to your specific project data.
4. Click 'Apply to Task'—the structured prompt populates the task form immediately.

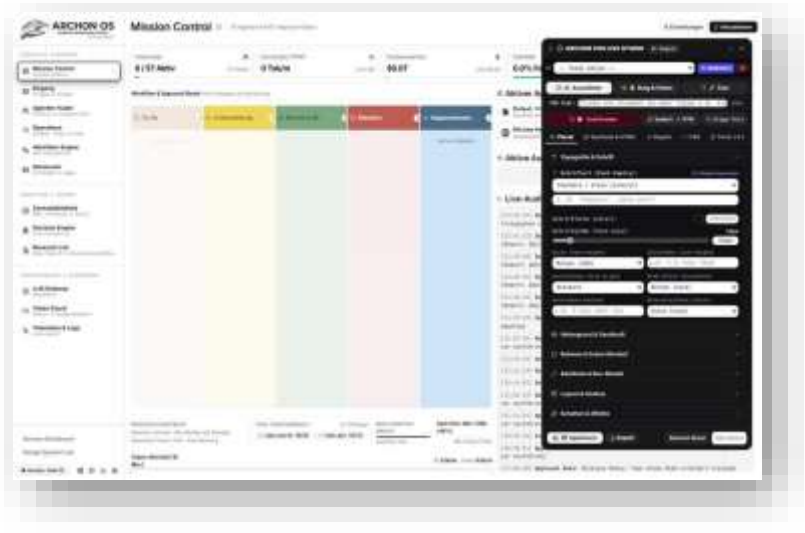
[Tip] Practical Tip & Guardrail: Apply the 5-field structure systematically across all complex delegations. Agent outputs become substantially more precise and error-free, requiring minimal manual rework.

[Prompt] Example Command / Prompt for the Assistant in This Area:
"Open the Prompt Studio and show me the 5 best templates for contract due diligence reviews."

In-depth guide: For prompt versioning and canary releases, see the Markdown system documentation (docs/system/PROMPT_REGISTRY_CANARY_GUIDE.md).

19. Special Feature: Visual CSS Live Studio & In-Place Inspector

Real-Time Design Studio, In-Place DOM Inspector & Design Token Management



Core Concept & Architecture:

The Visual CSS Live Studio (keyboard shortcut: Alt + S) is the design and customization hub of Archon OS. It allows you to inspect, measure, and adjust the entire visual interface in real time during live operation—without restarts, code compilation, or page reloads. Changes to colors, typography, spacing, or borders become visible instantly and are saved persistently in the browser.

Key Controls & Sections at a Glance:

In-Place DOM Inspector (crosshair mode): Click any UI element on screen: The studio jumps directly to the matching node in the CSS tree, displaying padding, margin, font families, and color values.
Color & typography sliders: Sliders and color pickers for primary colors, accents, border widths, corner radiuses, and font sizes.
Theme presets & dark/light synchronization: Enables seamless switching between color moods while maintaining WCAG-compliant contrast ratios.
CSS export & reset: Exports custom adjustments as a clean `custom_studio.css` or resets to the factory profile.

Practical Step-by-Step Workflow:

1. Press 'Alt + S' on your keyboard at any time (or click the palette icon in the bottom left).
2. The CSS Live Studio expands as a panel on the right.
3. Activate the inspector and click the card or button you want to customize.
4. Change color hues or border radiuses—the interface updates instantaneously.
5. Close the studio with 'Alt + S' again—your settings remain saved permanently.

[Tip] Practical Tip & Guardrail: Use the CSS Studio to tune contrasts to your workspace lighting conditions. With the export function, you can transfer your corporate color scheme cleanly into source code repositories.

[Prompt] Example Command / Prompt for the Assistant in This Area:
"Open the Visual CSS Live Studio and adjust the navigation accent color."

Appendix I: 5 Practical Use Cases

Step-by-step deployment scenarios for daily business operations

The following five use cases demonstrate how system modules interact in daily practice. They are structured by increasing complexity: from second-fast ad-hoc inquiries via Telegram call-signs, to semi-autonomous workflows requiring human approval, up to fully autonomous 24/7 background monitoring.

I.1 Use Case 1: Direct Telegram Interaction via Personal Call-Signs (/anwalt, /robert)

[Fast] Complexity Level 1: Low (Ad-hoc direct dialogue)

Context & objectives:

You are on the move and need an immediate legal initial assessment, a definition, or the current system status. Instead of creating a formal Kanban task, you want a well-grounded answer directly via mobile chat.

System workflow (step by step):

1. Open the Telegram app on your smartphone and start a chat with your bot (@Beatman_bot).
2. Send the command using the designated call-sign of the target agent: '/anwalt Do overtime hours strictly need to be recorded under current labor law?'.
3. The Archon OS Telegram gateway recognizes the '/anwalt' prefix, looks it up in the Specialist Fleet, and routes the message immediately to Dr. Maximilian von W. (Commercial Lawyer).
4. The legal specialist agent accesses its role profile and the assigned model (OmniRoute auto/best-free), drafts the response, and delivers it to your Telegram chat within 3 seconds.
5. The conversation is mirrored synchronously on the dashboard under 'Inbox & Event Triage' (Column 3: Telegram Stream) for complete audit documentation.

Involved system modules: Telegram Bot (@Beatman_bot) → Call-sign router → Specialist agent (Hermes) → OmniRoute AI Gateway → Inbox (Triage Stream)

Sample command / Telegram command:
"/anwalt Do overtime hours strictly need to be recorded under current labor law?"

[Tip] Practical tip: Ideal for rapid inquiries without administrative overhead. The call-sign bypasses extra steps and eliminates manually creating task cards.

I.2 Use Case 2: Incoming Partner Email: Triage, Response Draft & Human-in-the-Loop Approval

[Fast] Complexity Level 2: Medium (Email triage & HITL approval)

Context & objectives:

A business email containing a partnership inquiry arrives in the company inbox. A specialist agent prepares a response draft, which for security reasons must not be dispatched automatically and requires your approval.

System workflow (step by step):

1. The email inbox synchronizes every 180 seconds via IMAP; the new email appears in 'Inbox & Event Triage' (Column 2: Email Inbox).
2. The 'Email Guard' plugin automatically scans the email text for malicious prompt injections and phishing patterns.
3. You review the email in the inbox and delegate it with one click to Katharina Völker (Lead Go-to-Market).
4. Drawing upon corporate guidelines from the Central Library, the agent drafts a professional response.
5. Because sending an email constitutes an external write action, the Human-in-the-Loop (HITL) guard blocks dispatch: a yellow approval card appears in the inbox.
6. You review the draft, make minor text adjustments if necessary, and click 'Approve'—only then is the email physically dispatched to the sender via SMTP.

Involved system modules: Inbox & Event Triage (IMAP/SMTP) → Email-Guard Plugin → Hermes Agent → Central Library (RAG) → Human-in-the-Loop (HITL) Approval Gate

Sample command / Telegram command:
"Analyze the latest partnership email in the inbox and create a response draft for my approval."

[Tip] Practical tip: The HITL gate prevents agents from dispatching emails to real clients or partners prematurely or without oversight.

I.3 Use Case 3: Knowledge Building: 80-Page PDF Expert Report via Auto-OCR into Central Library & RAG

[Fast] Complexity Level 3: Advanced (Knowledge ingestion & OCR)

Context & objectives:

You receive a new technical report or zoning plan as an 80-page PDF. You want all 56 agents to master this specialized knowledge immediately without pasting facts into prompts manually every time.

System workflow (step by step):

1. Open the 'Central Library' view (sub-tab: Document Ingestion) and drag and drop the PDF into the upload area.
2. The integrated auto-OCR pipeline (Docling) automatically extracts continuous prose, tables, and headings, converting them into clean Markdown.
3. The formatted document is stored directly in the '02_Wissensbibliothek' directory of your local Obsidian vault on disk.
4. In parallel, the ingestion service segments the file into logical chunks and computes vector embeddings for the Central Library RAG index.
5. Switch to the 'Obsidian Vault' sub-tab—the integrated Markdown viewer displays the formatted note immediately.
6. Whenever a user asks a question about this construction project, the responsible agent (e.g., Dipl.-Ing. Bastian Braun, Construction Management) retrieves the facts accurately via vector search.

Involved system modules: Central Library (Document Ingestion) → Auto-OCR Pipeline (Docling) → Markdown Converter → AnythingLLM (RAG) → Obsidian Vault

Sample command / Telegram command:
"Extract the core findings of the newly ingested report in 02_Wissensbibliothek and list all relevant setback boundaries."

[Tip] Practical tip: Storing files in the local Obsidian vault ensures all data remains backed up permanently as open Markdown on your own drive.

I.4 Use Case 4: Strategic Project Development: Delegation by Robert (CEO), Research Lab & Kanban

[Fast] Complexity Level 4: High (Autonomous project delegation)

Context & objectives:

A complex major assignment needs execution: 'Münster site and zoning law analysis for modular construction'. The task requires web research, legal review, and civil engineering calculations.

System workflow (step by step):

1. Send Robert the command via Telegram: '/robert Create a comprehensive site and zoning analysis for modular construction in Münster'.

2. Robert (Executive Decision Architect) accepts the overall assignment and decomposes it into 3 sub-tasks on the Operations Board:
• Sub-task A to Research Lab: Research market data, rental levels, and land prices in Münster.
• Sub-task B to Dr. Maximilian von W.: Legal review of zoning plan requirements and permit pathways.
• Sub-task C to Dipl.-Ing. Bastian Braun (Construction Manager): Analysis of modular construction methods and interface calculations.

3. On the Operations Board, all three sub-cards move from 'To-Do' to 'In Progress'. The WIP limit prevents system overload.

4. The Research Lab carries out evidence triangulation and stores interim analyses in the Showroom.

5. Robert consolidates all deliverables, compiles a multi-page dossier in the Showroom, and advances the task to 'Review & QA'.

6. You receive a notification on Telegram, open the task card, and inspect the finished deliverable via the 'Open in Obsidian' button.

Involved system modules: Telegram Bot → Robert (CEO) → Operations Kanban (WIP Limits) → Research Lab → Domain Leads → Showroom → Obsidian Vault

Sample command / Telegram command:
"/robert Create a structured site and zoning analysis for modular construction in Münster including legal review."

[Tip] Practical tip: Robert acts as your digital CEO: he relieves you entirely from manually coordinating individual domain specialists.

I.5 Use Case 5: Fully Autonomous Background Monitoring, OmniRoute Failover & Emergency Escalation

[Fast] Complexity Level 5: Very High (24/7 automation & failover)

Context & objectives:

Archon OS operates continuously in the background, scanning industry tenders, recovering autonomously from external API failures, and escalating controlled alerts to you when infinite loops or incidents occur.

System workflow (step by step):

1. Every business day at 06:30 AM, the n8n workflow engine triggers an autonomous scan workflow via a cron node.
2. A research agent scans defined industry portals for new modular construction tenders.
3. During the search run, the primary LLM provider fails (rate limit / HTTP 429).
4. The OmniRoute gateway detects the failure within 180 ms, switches seamlessly to the configured fallback model, and resumes the run.
5. If a malfunctioning tool call enters an infinite loop, the Guardian loop killer steps in after exactly 15 steps: it hard-stops the process, marks the task as 'Blocked / Escalation', and alerts Robert.
6. Robert summarizes the incident and sends an emergency Telegram push notification. In the telemetry log stream, the incident is documented with second-by-second precision using cryptographic hash chains.

Involved system modules: n8n Workflow Engine (Cron) → OmniRoute AI Gateway (Failover) → Guardian (Loop Protection) → Operations Kanban → Telegram Push → Log Audit

Sample command / Telegram command:
"Set up a daily workflow in n8n that scans tenders at 06:30 AM and escalates errors to Robert."

[Tip] Practical tip: Thanks to OmniRoute failover and Guardian loop protection, your system continues operating stably and within budget even when external AI servers fail.

Appendix II: IT Security, Encryption & GDPR Overview

This appendix summarizes the key security, encryption, and GDPR capabilities of Archon OS for daily corporate operations and data protection audits. Detailed configuration examples for administrators are available as Markdown documents in docs/security/.

II.1 GDPR Compliance & Zero-Egress Privacy Routing

Data subject rights (Art. 15, 17 & 30 GDPR): Data access exports (JSON) and complete erasure requests are executed atomically across all SQLite databases ([ANONYMISIERT_DSGVO_ART17]) and recorded with tamper-proof SHA-256 deletion receipts (docs/system/GDPR_COMPLIANCE_GUIDE.md).
PII scrubber & privacy routing (Art. 28 & 32 GDPR): Sensitive inputs (names, addresses, IBANs, contract terms) are automatically masked before reaching external cloud models (PII scrubber + local rehydration) or strictly routed to local Ollama models or certified EU endpoints (docs/security/LOCAL_LLM_PRIVACY_ROUTING_GUIDE.md).

II.2 Encryption (In Transit & At Rest)

Protection Layer	Implemented Method	Protected Data & Practical Benefit
Network & Web UI	TLS 1.3 / HTTPS with HSTS & security headers	All traffic routed through the Caddy reverse proxy including Authelia SSO and brute-force protection (docs/security/REVERSE_PROXY_SSO_GUIDE.md).
User accounts (RBAC)	PBKDF2-HMAC-SHA256 (100,000 rounds + salt) / Argon2id	Passwords and role permissions (admin, operator, viewer) in backend/auth_service.py (Chapter 17.10) or Authelia SSO.
Emergency Vault (Backup)	PBKDF2-HMAC-SHA256 (200,000 rounds) + CTR-HMAC-SHA256 + Auth-Tag	Encrypted offline archive break_glass_vault.enc (.env, config/, SQLite DBs) with 256-bit master key (docs/security/BREAK_GLASS_RECOVERY_GUIDE.md).
API Keys & Secrets	Salted SHA-256 hash (one-time display) & .env isolation	Tokens are never stored in plaintext; pre-commit scanners (audit_secrets.py) prevent accidental Git leaks (docs/security/SECRETS_MANAGEMENT_UND_ROTATION.md).

II.3 Audit Compliance (GoBD), Prompt Protection & Approvals

GoBD & ISO 27001 audit log: All business-critical actions, AI approvals, and key changes are immutably logged via cryptographically linked SHA-256 hash chains (Chapter 16.3 & docs/security/GOBD_ISO27001_AUDIT_LOG_GUIDE.md).
Prompt injection protection & HITL gates: Inbound emails from unknown senders pass through an isolated dual-LLM guard; critical actions (finances, dispatch, deletion) strictly require human approval (Chapters 2.6 & 14), including an optional vacation proxy.

II.4 Further Security & Operations Documentation

All detailed guides for IT administrators and security officers reside in the repository under docs/security/ (including API Keys/Rate Limits, Break-Glass Recovery, Dual-LLM Guard, GoBD/ISO 27001, Privacy Routing, Reverse Proxy/TLS, Secrets Rotation) and under docs/system/ (GDPR Compliance, Approval Policies, Backup/Restore, ERP Specifications). Complete catalog of all modules: ARCHON_OS_FEATURE_LIST_2026_EN.pdf · Boardroom methodologies: ARCHON_OS_BOARDROOM_FRAMEWORKS_2026_EN.pdf.

Appendix III: Glossary & Alphabetical Index

Alphabetical reference of key operational, security, and domain terms including exact page numbers in the manual.

Term	Brief Definition	Page(s)
360° Customer File (Contact Drawer)	Full view of a CRM contact with master data, customer/vendor account numbers, timeline, and receipts.	p. 54, 55, 56, 58
4-Eyes Quality Gate	Automated cross-check in Kanban by a reviewer agent before completing a task.	p. 13
Accounting	Finance area for incoming/outgoing vouchers, analytics, contracts, and dunning.	p. 6, 62, 64, 65
AnythingLLM	Integrated RAG and document platform in the Central Library for PDF indexing and knowledge queries.	p. 37, 47, 48, 49
API-Keys & Access	Management of access keys and service tokens in the System Control Center with hashed storage.	p. 83
Approval Policies (Human-in-the-Loop (HITL) Approval Gate)	Binding approval rules (Human-in-the-Loop) for finances, emails, and critical system operations.	p. 6, 23, 67, 68
Archon Prompt Studio	Dedicated studio with 832 prompt templates and Fabric SOPs for recurring tasks.	p. 87
Argon2id	Memory-hard password hashing in the optional Authelia SSO gateway to protect against brute-force attacks.	p. 94
Articles & Services (ART)	ERP catalog for services, flat rates, SaaS subscriptions, and hardware including inventory.	p. 59, 60, 61
Audit Log (GoBD / ISO 27001)	Tamper-proof logging via chained SHA-256 hash values in Telemetry & Logs.	p. 6, 40, 74, 94
Boardroom (Decision Engine)	Structured AI panel session incorporating multiple specialist perspectives and 55 solution frameworks.	p. 38, 50, 52, 53
Bot Screen (Web Navigator)	Visual live screen for browser and desktop actions of autonomous agents.	p. 29
Break-Glass Vault	Encrypted emergency archive (break_glass_vault.enc) for disaster recovery.	p. 81, 94
Capture (Document OCR)	Central Library intake tool for text recognition and tagging of PDFs and receipts.	p. 27, 48, 78, 91
Central Library (Pillar 2)	Knowledge base consisting of AnythingLLM RAG index, OCR capture tool, and Obsidian vault.	p. 37, 47, 48, 49
Context Guide (?)	Integrated contextual help in every module view with explanations and operation tips.	p. 8
Corporate & Multi-Tenant	Management of corporate entities, including separate SQLite DBs, letterheads, and DATEV number ranges.	p. 66
CRM (Addresses & Contacts)	Customer and vendor master data in Pillar 3, including debtors (10000+) and creditors (70000+).	p. 54, 55, 56, 58
DATEV Export (EXTF / CSV)	Transfer interface for invoices, receipt images, and charts of accounts to financial accounting.	p. 54, 55, 64, 65

Term	Brief Definition	Page(s)
Deal Pipeline & Forecast	Weighted revenue forecast across active sales opportunities in Operations & Contracts.	p. 62
Deep Research v2	Multi-stage research loop in the Research Lab with source verification and dossier generation.	p. 44, 45, 78
Docker & 1-Click Installer	Container runtime and guided installation scripts (setup.bat, setup.ps1, install.sh).	p. 9, 16, 73, 76
Dual-LLM Prompt-Injection Guard	Isolated two-stage inspection of external emails and web content against hidden injection instructions.	p. 21, 75, 76, 94
Dunning Management	Automated monitoring of overdue invoices, including multi-stage dunning runs.	p. 65
Email Approval Queue	Review and approval view for agent-drafted emails prior to actual dispatch.	p. 21, 23
Encryption	Multi-layer security using TLS 1.3, PBKDF2-HMAC-SHA256, Argon2id, and AES/CTR vault.	p. 6, 68, 86, 94
GDPR (Art. 15, 17, 28, 30, 32)	Data privacy functions: data disclosure export, deletion anonymization with receipt, PII-Scrubber, and Zero-Egress Privacy Routing.	p. 6, 9, 58, 94
GiroCode (EPC-QR)	Standardized QR code for wire transfers directly on outgoing invoices and letterheads.	p. 64, 66
HITL (Human-in-the-Loop)	Mandatory human approval as a safeguard for critical actions.	p. 23, 67, 68, 90
Inbox & Event Triage	Unified inbox for emails, webhooks, messengers, and approval requests.	p. 18, 19, 20, 21
Internet Watchdog	Autonomous web scanner with topic buckets for continuous monitoring of news and sources.	p. 6, 37, 42, 43
Kanban Board	Visual task and status management from Pending through Review to Done.	p. 7, 11, 13, 17
Messenger & Telegram Bridge	Control via smartphone app, including direct handle invocations (/robert, /anwalt).	p. 22, 75, 89, 92
Mission Control	Central dashboard for container status, token usage, and active tasks.	p. 6, 11, 12, 56
n8n Automation	Visual workflow engine for scheduled routines and webhook integrations.	p. 6, 11, 15
Obsidian (Knowledge Vault)	Interlinked local Markdown knowledge base in the Central Library.	p. 9, 47, 49, 91
OmniRoute & Providers	Central AI provider management and automatic routing between local and cloud models.	p. 78, 81, 82, 93
Operations (Pillar 1)	Workspace for projects, Kanban, org chart, automations, and team radio.	p. 11, 13, 14, 15
PBKDF2-HMAC-SHA256	Cryptographic key and password hardening (100,000 or 200,000 rounds) for logins and backups.	p. 86, 94
PII-Scrubber	Automated redaction of personal data prior to external AI calls with local restoration.	p. 94
Quotes & Orders (ANG / AB)	End-to-end sales transactions from quote to order confirmation to invoice.	p. 62, 63

Term	Brief Definition	Page(s)
RAG Index (Central Library)	Semantic knowledge base for targeted querying of proprietary PDF and document records.	p. 40, 46, 47, 48
Research Lab	Specialized module for in-depth research, source verification, and dossiers.	p. 37, 44, 45, 46
Showroom	Storage and preview view for generated reports, media, code artifacts, and final deliverables.	p. 36, 39, 40, 41
SOUL.md & Personnel File	Role profile, work style, permissions, and performance history of each specialist in the Specialist Fleet.	p. 32, 33
Specialist Fleet	Control center for all 70 specialized AI agents, including Chat, Operations, Bot Screen, Soul, and personnel file.	p. 24, 25, 27, 28
System Control Center	Central console for all 11 system sections from Hermes Core to API keys to system updates.	p. 75, 76, 77, 78
System Update	Integrated version and update management in the Control Center including safety snapshots.	p. 9, 86
Telemetry & Logs	Monitoring of Docker containers, agent costs, and audit logs.	p. 6, 72, 73, 74
Token Guard	Cost and quota control to prevent excessive token consumption.	p. 6, 69, 70, 71
Transactions & Contracts	Management of quotes, orders, sales pipeline, and contract durations.	p. 6, 59, 62, 63
Users & Roles (RBAC)	Role-based user management (admin, operator, viewer) with PBKDF2-secured passwords.	p. 75, 86, 94
Visual CSS Live Studio	Tool for visual customization of colors, spacing, and UI components.	p. 6, 8, 86, 88
Webhooks	HTTP endpoints in the inbox to trigger automated tasks via third-party systems.	p. 6, 15, 19, 75
Zero-Egress Privacy Routing	Routing sensitive requests to local Ollama models or EU servers instead of US clouds.	p. 9, 75, 81, 94