

Automating Building Design & Consulting Workflow

From productivity AI to engineering AI: the properties that distinguish the two, and where current efforts are directed.

Davor Stjelja · AI Lead, Granlund · 17 April 2026

AI adoption in MEP: broad at the productivity layer, thin at the engineering layer.

PRODUCTIVITY LAYER — IN USE TODAY

Specification drafts, reports, literature review

Report generation, literature review, email, meeting summaries, and translation.

Measurable time savings across surveyed MEP firms (Bluebeam 2026; Medha 2026).

ENGINEERING LAYER — LARGELY UNREACHED

Routing, sizing, clash prediction, validation

Piping routing, equipment sizing for a specific building, cross-discipline clash prediction, HVAC design validation.

These require integration the productivity layer lacks.

General-purpose LLMs are a baseline, not a substitute for the integration and validation that engineering deliverables require.

Three failure modes, measured on real problems.

6×

Model-to-model variance on the same problem

The same MEP pump-station project was run through six leading AI models. Reported peak heating loads came back anywhere between 25 and 150 MBH — a sixfold spread on a deterministic load calculation.

Fruehan & Doyle, CSE Magazine, 2026

>400% vs ~1.5%

Stand-alone LLM vs. solver-integrated LLM

On structural-analysis benchmarks, stand-alone LLMs produced errors exceeding 400%. The same models, wired through Model Context Protocol (MCP) to a validated solver, held within ~1.5%.

Ávila, Ilbay & Rivera, Buildings 15(3190), 2025

87–95%

AI initiatives that never reach impact

An estimated 87–90% of AI initiatives stall before deployment. In AEC-specific surveys, up to 95% of GenAI projects fail to show measurable impact.

AWS MLOps 2025; DTO AEC Part 2, 2025

Integration is the binding constraint. What the model is plugged into, and what it is grounded on, is what decides whether engineering AI actually works.

PART 1

What engineering AI actually has to be

Four system-level properties that distinguish general-purpose LLMs from AI usable in engineering deliverables.



Granlund

Four properties of engineering-grade AI

The bar is set by the system, not the model.

01 INTEGRATION

Coupled to BIM and validated calculation engines

The LLM orchestrates computation performed by validated solvers (EnergyPlus, TRNSYS, FEM packages). Pattern demonstrated in MCP-style architectures (Ávila et al., 2025).

02 GROUNDING

Retrieval over the firm's data and the applicable rulebook

Generic models have no access to how a firm has sized chillers for Nordic hospital projects, nor to standards (SFS/EN), national regulations, company guidelines, and design templates that govern deliverables.

03 TRACEABILITY

Audit trail from input to output, by construction

Each AI-assisted output is linked to its sources, model, and validation step. The engineer of record remains liable; documentation must carry that signature under audit.

04 HUMAN OVERSIGHT

Validation as the primary quality mechanism

AI performs best on well-defined tasks with structured data and clear validation criteria. Outside that envelope, engineer judgement remains the mechanism for catching edge-case failures.

PART 2

What are we developing

Five directions, each summarised at the level of problem, approach, and designer-visible outcome.



Granlund

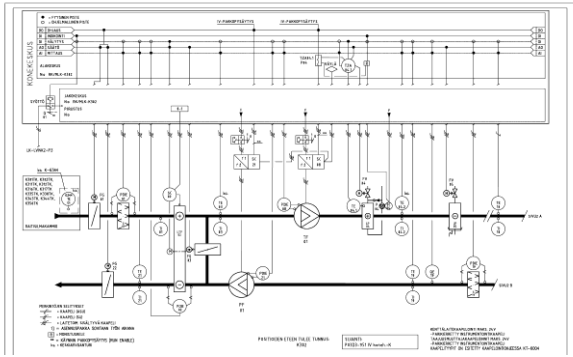
Reading & generating engineering schematics

Reading the documents of the trade

PROBLEM

Control and room-device schematics are redrawn by hand across hundreds of projects each year. They are among the most repetitive MEP artefacts, with large drafting-quality variance between engineers.

IN · RAW CONTROL SCHEMATIC



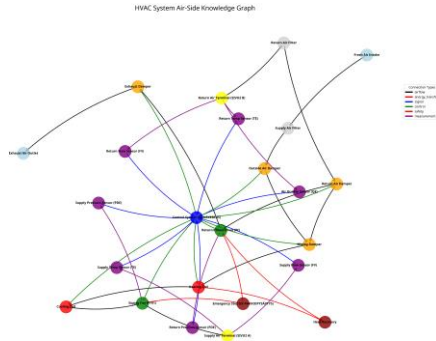
PDF / scan of a control diagram: hundreds of symbols, thousands of lines, project-specific conventions.

Designer-visible change: repetitive schematic production shifts toward review of machine-generated drafts.

TECHNICAL APPROACH

- Vision-language models extract components, signal chains, and control relationships from scanned or vector PDFs
- Structure written into a standardised engineering schema (DEXPI 2.0 / Brick-compatible)
- Operation-description text generated from the semantic graph, not the original document
- Phased delivery: quality-check first, then editing, then generation from specs

OUT · EXTRACTED SEMANTIC GRAPH



Machine-readable topology: sensors, actuators, and their connections, typed and traversable by downstream tools.

Spec-vs-supplier verification

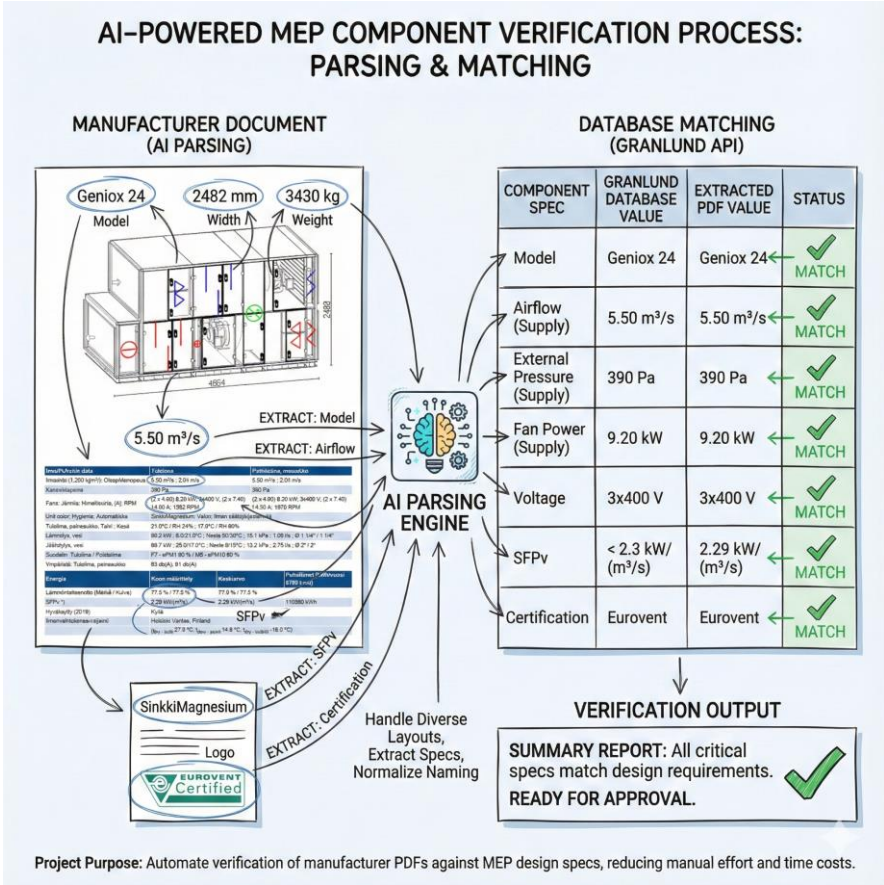
QA against the supply chain

PROBLEM

Manual cross-referencing of manufacturer documentation against project specification is time-intensive, inconsistent between engineers, and produces an error class that tends to surface only at commissioning.

TECHNICAL APPROACH

- Structured parameter extraction from heterogeneous supplier documentation
- Comparison against the project specification with explicit tolerances
- Discrepancy reports with per-parameter source linkage, structured for engineer review and sign-off



Classifying & tagging building-automation data

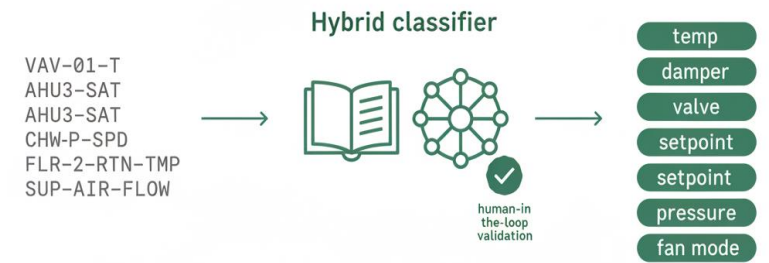
The design → operations bridge

PROBLEM

Each project generates thousands of BAS data points under project-specific naming conventions. Manual tagging for analytics, commissioning, and handover is a persistent, under-measured time cost across the industry.

TECHNICAL APPROACH

- Few-shot learning of project-specific naming conventions from engineer-validated examples
- Hybrid rule + LLM pipeline: deterministic rules handle high-confidence cases; LLM used only on ambiguous ones
- Human-in-the-loop feedback captured at project level and fed back into cross-project priors



Designer-visible change: design-phase tags propagate into operations without manual re-tagging at handover.

Structured extraction from drawings & models

Re-using information already in the model

PROBLEM

Many downstream workflows (compliance checks, quantity take-off, impact assessment, reporting) require re-reading the same drawings the designer produced, manually extracting materials and classifications. Typically, 30+ hours per mid-sized project on information already in the model.

TECHNICAL APPROACH

- Quantity and material extraction from vector PDFs, scanned drawings, and IFC models
- Automatic classification against project and regional taxonomies with confidence scoring
- Export into the existing downstream tool rather than creating a parallel data silo



Material	Thickness	Class
Concrete	300mm	Wall
Glass	10mm	Window
Wood	Window	Floor

95% Confidence

Designer-visible change: downstream workflows become incremental over existing design artefacts rather than parallel ones.

A connected knowledge layer under the tools

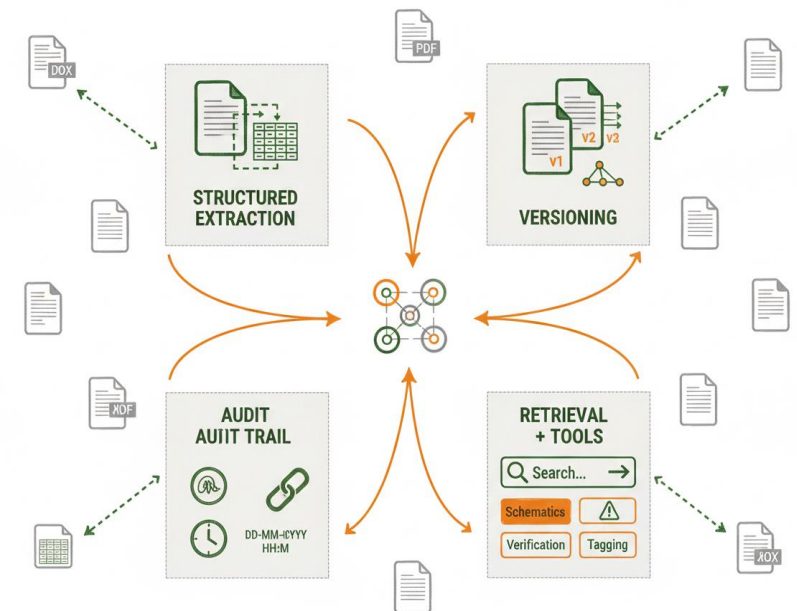
The compounding layer

PROBLEM

Each direction above is a specialist. They become more useful when they can read from and write back into one shared body of project data (documents, notes, contracts, drawings, cost estimates, IFC models) and the firm's knowledge base (standards, regulations, product data, design templates). Without that shared layer, every tool starts from scratch.

TECHNICAL APPROACH

- Captures project artefacts as they are produced — structured, versioned, searchable
- Joins project memory with the firm-level knowledge base behind one retrieval interface
- Lets specialist tools contribute back into the same connected graph rather than isolated silos
- Governance, audit trail, identity, and vendor-agnostic model routing handled at the platform layer



Designer-visible change: each project automatically contributes to a shared, versioned, auditable knowledge layer rather than starting over in every new tool.

Three cognitive loads, three different AI roles.

From Cognitive Load Theory: not all mental effort is equal.

01

Extraneous load — ELIMINATE



Mental effort wasted on mechanical overhead: reformatting, manual data entry, searching product data, unit conversion, syntax checking.

AI's legitimate job here is to make this disappear.

FirstQ Paper 2, 2025

02

Intrinsic load — MANAGE



The genuine difficulty of thermodynamic cycles, hydraulic networks, electrical load flows.

AI's role is to break these into manageable components and offer progressive disclosure, present complexity clearly, not remove it.

FirstQ Paper 2, 2025

03

Germane load — PROTECT



Productive engineering thinking: system optimisation, cross-discipline conflict resolution, design judgement.

Good AI deployment expands the cognitive budget available for it rather than replacing it.

FirstQ Paper 2, 2025

The test for any engineering AI tool: which load is it acting on, and is that the right one?

CLOSE

Engineering AI is bounded on two sides.

THE SYSTEM SIDE

The integration, data, and validation layer the model sits in.

THE HUMAN SIDE

Whether it frees the engineer's mind for judgement or erodes it.

OUR APPROACH

Build those foundations in parallel with practical tools rather than in sequence, eliminating extraneous load, managing intrinsic load, and protecting germane load, with the engineer of record in the loop at each validation step.



FirstQ whitepaper series on AI in MEP

Much of the framing in this talk is drawn from a joint whitepaper series we contribute to.



PAPER 1 • STRATEGIC

Generative AI in MEP Engineering: The Five-Year Transformation Ahead

A strategic overview for European MEP decision-makers.

PUBLISHED



PAPER 2 • ORGANISATIONAL

Transforming the MEP Workforce: Human-AI Collaboration in Engineering Practice

How roles, validation, and liability shift once AI enters engineering deliverables.

PUBLISHED

PAPER 3 • TECHNICAL

Foundations for MEP AI: data, governance, and AI-enabled design processes

The substrate beneath tools: data, governance, and engineering-grade integration.

UPCOMING

Thank you.

Questions, disagreements, or your own war stories from building AI into MEP workflows — all welcome.

Davor Stjelja · AI Lead, Granlund

davor.stjelja@granlund.fi



Granlund