

AI-Agent Application for Semantic Data Enrichment in Ventilation Systems Using National Nomenclature for IFC and GS1- Based Product Information

Otto Alhava, Tommi Arola, Osku Torro, Markus Järvenpää,
Tero Järvinen, Bettina Ruottinen

LDAC2025 - Linked Data in Architecture and Construction
July 9–11, 2025, Porto, Portugal



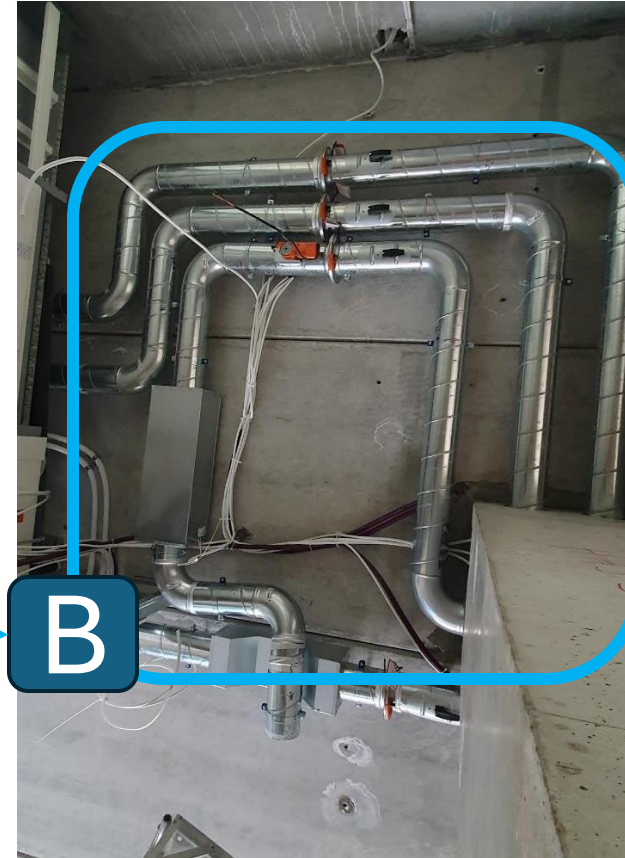
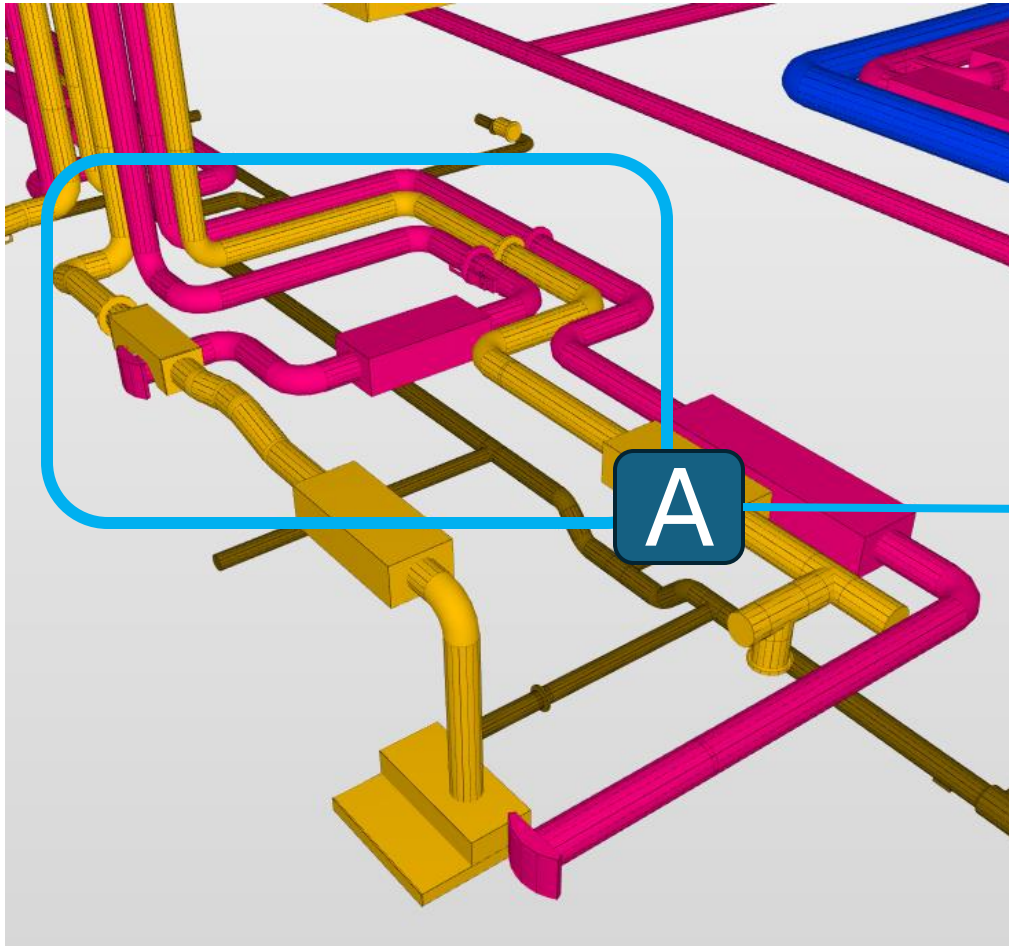
Introduction and research focus

- Construction still lacks industrial grade data-transfer standards and interfaces
- BIM tools are used primarily for design-phase coordination and have limited integration to other systems
- The research focuses on deploying an AI agent platform to bridge the identified digital discontinuities in ventilation system design and supply chain data
 - The aim is to (i) demonstrate how AI agents can enrich BIM-based E-BOM into M-BOM that support product-selection decisions, and (ii) to propose ways to simplify the underlying enrichment logic

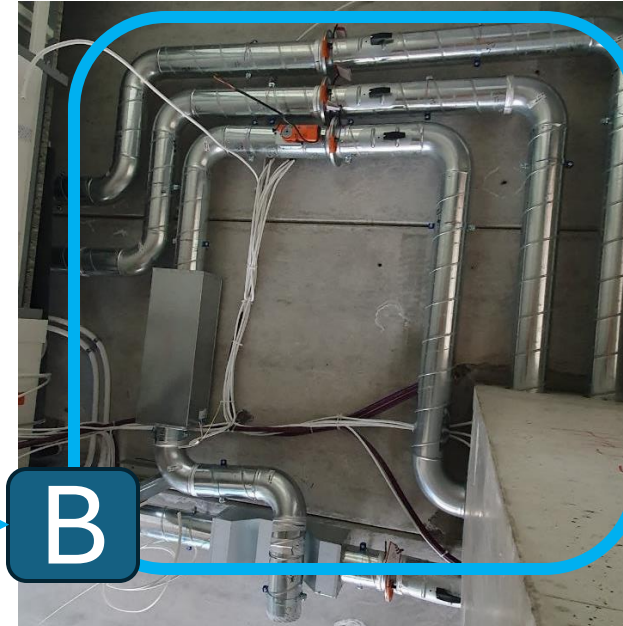
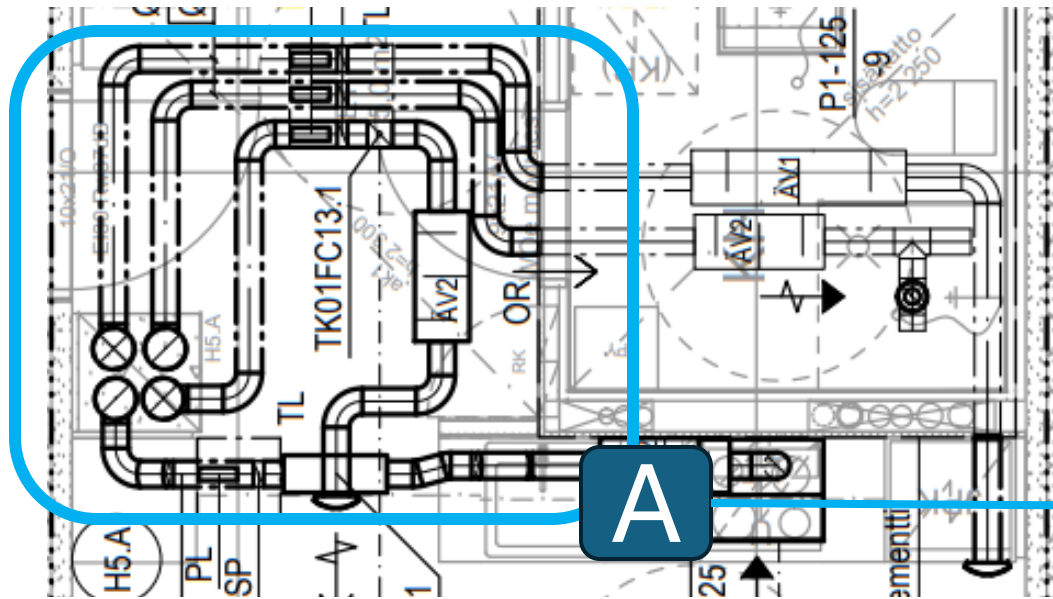
Background

- Importance of digitalisation and data flow in manufacturing and construction
 - Manufacturing has successfully **conceptualised its production model** and has **developed standard**
 - Such as the GS1 standard family along with PEPPOL standard
 - Climate change and the accelerating EU **regulations has driven other industries towards digitalisation**
 - Retailers utilise **machine data processing** to link consumer market and the **information is reported via mobile applications** to customers
 - However, **construction industry lacks essential digital management solutions** for inbound and on-site logistics, machine readable supply chain data, automated processing of design data and seamless data exchange between systems
- Digitalisation of the construction supply chain in Finland
 - The construction industry has aimed to solve the challenges of digitalisation with **collaboration between public and private sectors and research institutions**
 - Digital interoperability platform has been developed for the national standardisation of data content and the digitalisation of supply chains

Case MEP design and supply chain - How do we get from BIM to a completed HVAC installation?



Case MEP design and supply chain - How do we get from BIM to a completed HVAC installation?

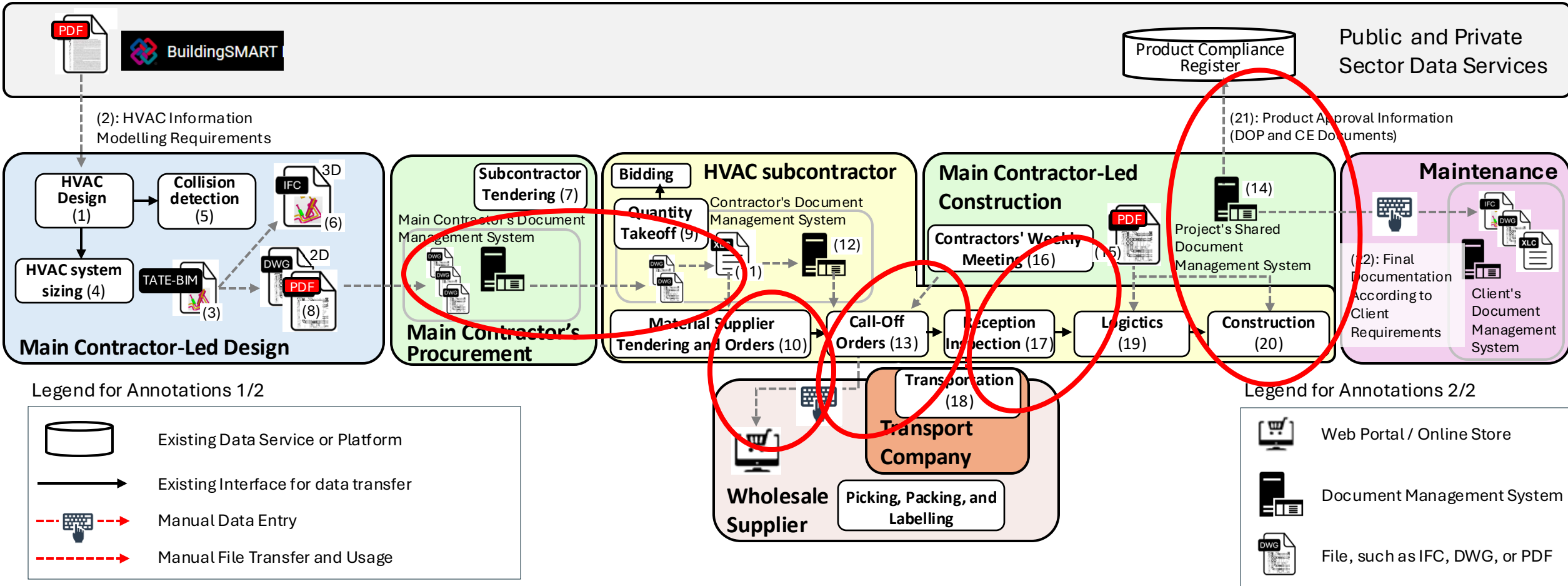


The route from A to B does not start with BIM but with 2D PDF files. Thus, BIM is not the starting point for data flow.

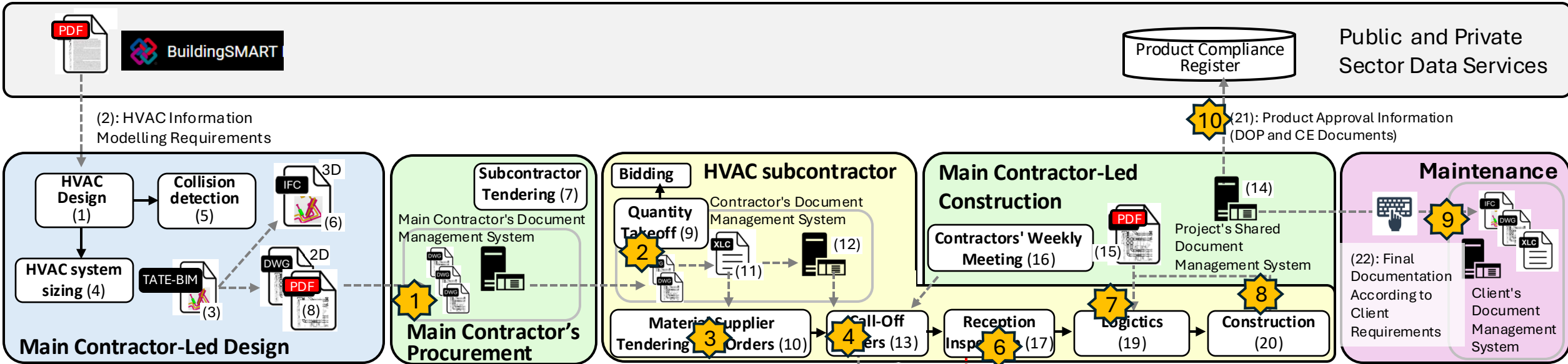
Construction information management is not a data factory but a PDF factory:

- Information is repeatedly collected from 2D PDF files for the same HVAC components
- Data management methods are based on repetitive manual work, files, email attachments and the use of Excel
- The same applies to other BIM models and structures of the building

Due to the use of 2D images and PDF files, information does not flow from the MEP design process to procurement, suppliers, logistics, or construction.



Due to the use of 2D images and PDF files, information does not flow from the MEP design process to procurement, suppliers, logistics, or construction.



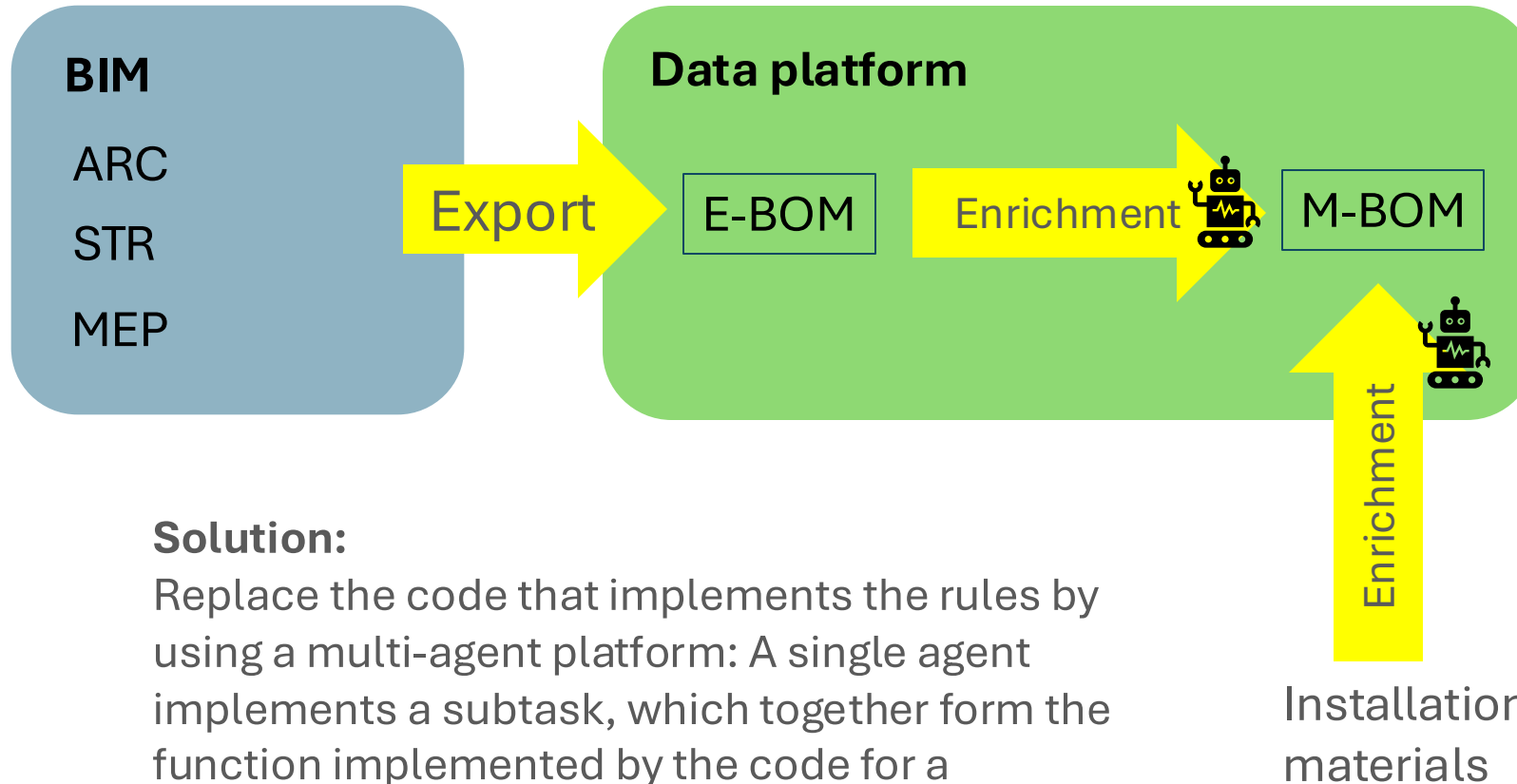
- The end-to-end process consists of digital discontinuities that require manual work
- Upstream inefficiency is one main reason why productivity in construction itself is not improving

....and the situation gets worse with CO₂ reduction and environmental reporting requirements.

Using a multi-agent platform for linking and enriching data

Problem:

As the number of products increases, the number of rules to be maintained increases linearly. Changes to products cause a cumulative maintenance burden.



Solution:

Replace the code that implements the rules by using a multi-agent platform: A single agent implements a subtask, which together form the function implemented by the code for a specific product.

Result:

Increasing the number of products does not require code maintenance work, as agents learn from previous products how to enrich the information for new products correctly.

Why AI Agents?

- Fast evolution of generative AI has made agent technology central to current innovation
- Development environments for multi-agent systems are emerging:
 - AutoGen Studio (Microsoft)
 - Agentspace (Google)
 - Amazon Bedrock Agents
 - Swarm (OpenAI)
- Agents are the most important strategic technology trend of 2025 (Gartner, 2024)
- Next-generation process automation
- Discontinuities in data flow are well-suited for agents, e.g., function/API calling, tool usage, etc
- Agents have the potential to mimic the existing human workflows (e.g., E-BOM -> M-BOM)

AI agent platform and does it fit for its purpose?

The image shows a screenshot of the Langflow AI agent platform interface. The interface is divided into several sections:

- Components:** A sidebar on the left lists various components like 'Extract Key', 'Filter Data', 'Filter Values', 'Merge Data', 'Parse Data', 'Parse JSON', 'Select Data', and 'Split Text'. It also includes a search bar and a 'Show Beta' toggle.
- Processing:** A central area where components are connected in a flow. The flow starts with 'BIM-model BOMExtractor' (receiving 'INPUT: JSON bill of material'), followed by 'Parse Data' (outputting 'Machine readable bill of material: all 3rd floor pipes'), then 'CO2 calculator' (outputting 'Calculator e.g CO2-calculations'), and finally 'Construction product ag...' (outputting 'AI agent with a specified task').
- Local AI language model (Ollama engine):** A component for generating text using Ollama Local LLMs, with a 'Base URL' set to 'http://localhost:11434'.
- Construction product information API endpoint:** A component for gathering real construction product information, with a 'Method' set to 'GET'.
- AI agent with a specified task:** A component where the agent's instructions are defined. The instructions are: 'You are a helpful construction product sourcing assistant. Your specified task is to find building bill of material and search the most suitable construction products based on the given CO2 limit. Your input is machine readable bill of material with suggested products'.
- User control loop:** A component for getting chat inputs from the Playground.

Annotations in blue boxes provide additional context:

- 'INPUT: JSON bill of material' points to the 'BIM-model BOMExtractor' component.
- 'Local AI language model (Ollama engine)' points to the 'Construction Ollama' component.
- 'Construction product information API endpoint' points to the 'Product information DB' component.
- 'AI agent with a specified task' points to the 'Construction product ag...' component.
- 'User control loop' points to the 'Agent instructor' component.

At the bottom left, a blue box contains the text: 'AI agent platform toolbox, example using Langflow <https://www.langflow.org/>'.

Limitations

- Focus on **MTS products of ventilation system's** design, procurement, and installation, limiting the generalizability of findings to the broader building services or construction industry
 - Expanding the research to other building services domains (such as plumbing), multiple system types and projects is necessary
- **A prototype with incomplete integrations**, and its ability to optimize product selection based on carbon footprint, cost, or installation efficiency is still limited
 - Further development requires structured databases, high-quality data, a trained language model, and a unified architecture for AI model selection and finetuning
- Ensuring semantic consistency in product data processing requires **linking structured data to databases while maintaining robust authentication and agent permissions**
- Orchestrating multiple agents warrants dedicated research as **a systemic solution** for construction data flow
 - How to coordinate many specialised agents, resolve conflicts and maintain global objectives in a live project environment remains an open question. Equally critical is establishing governance mechanisms to ensure that autonomous agent swarms act transparently and reliably across the project lifecycle

Conclusions

- Ventilation system design data is often transmitted as 2D drawings.
 - The entire downstream process relies on 2D drawings in PDF or paper format for data exchange
- Critical discontinuities were identified in the data flow, from which data enrichment from E-BOM to M-BOM was selected for the multi-agent analysis
- The study demonstrated that the AI agent platform's functionalities support the enrichment of IFC-based building services models using multi-agent system tools
 - **AI-agent platforms suits well** for modelling construction processes, supporting data flows, product selection, and real-time CO₂ calculations
 - **AI agent platforms effectively integrate** with external data sources when repositories are available
 - **AI-agent platform tools are directly feasible for AEC data management** purposes, such as web-based document searches, vector database queries, and mathematical computations
 - **AI-agent platforms efficiently transform** a machine-readable Engineering Bill of Materials (E-BOM) into a structured Manufacturing Bill of Materials (M-BOM) for construction

AI-Agent Application for Semantic Data Enrichment in Ventilation Systems Using National Nomenclature for IFC and GS1-Based Product Information

Thanks for your attention!

Questions?

Fira

RTS >

 Tampere University


Granlund

A!
Aalto University