SVAR – Systematic Verification and Acceptance of Requirements

Reference group meeting

April 29, 2025

Reference group

Pia Schönbeck – Sponsor. Project lead in systemic requirement management.

~~Oskar Permwall~~ – Specialist in systemic requirement management

~~Marit Jidemo~~ – Business developer in information management.

Erik Häggström – Area responsible (Background in BIM/GIS, information management in BIM

Rastkar Rauf – technical engineer, Digital project management

Susanne Van Raalte – BIM strategist

Karin Anderson – BIM specialist

Agenda

- Progress report
  - Objective 1: ACC Capability Maturity Model
  - Objective 2: TRVInfra requirements verifiability
  - Objective 3: Demonstration of verification methods
- Identifications of “Champions” for result hand-over

Project overview

Duration: October 1, 2023 – September 30, 2025

Three objectives, each with three work packages.

- **Objective 1:** Development of an Automated Compliance Checking Capability Maturity Model (ACC-CMM)
- **Objective 2:** Understand to what degree the compliance checking of requirements (TRVInfra, project-specific) is automatable
- **Objective 3:** Develop procedures for automated, reusable, verification of requirements

Objective 1: ACC Capability Model

ACC Capability Maturity Model

Done: Developed the model and discussed internally at BTH

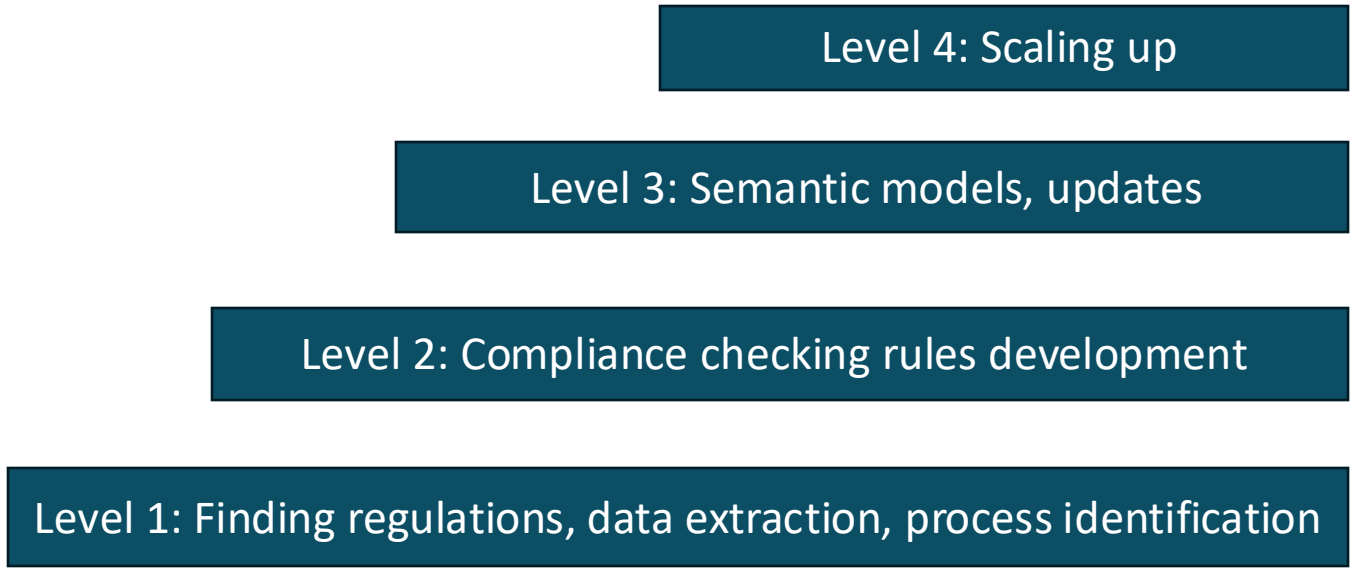
Done: Analyzed how the model fits into other Digitalmognad initatives

Done: Performed 9 interviews TRV (list provided by Susanne)

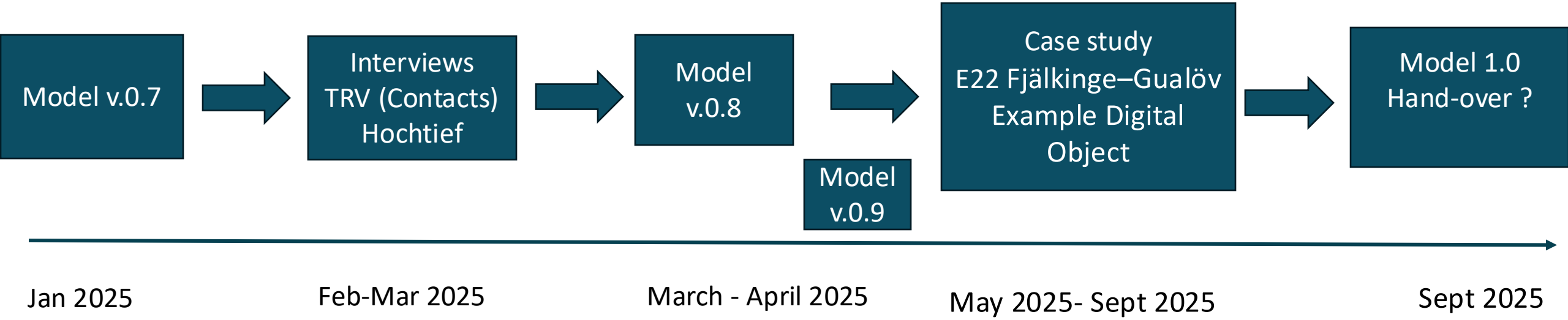
Done: Interviewed Tobias Odebjer, manager of the project E22 Fjälkinge–Gualöv

Next step: Writing the version 1.0 of the model and sharing the report with TRV.

Next step: Review of the assessment of E22 Fjälkinge-Gualöv



ACC CMM Model Evaluation



Objective 2: TRVInfra requirements verifiability

TRVInfra requirements verifiability

Purpose: We perform the classification to judge how verifiable the TRVInfra requirements are.

Approach: Classify requirements along 5 dimensions (target, nature, interpretability, reference, logic rule)

Goal: Create a ground truth to train a classifier (deep learning) to predict verifiability of 18.000 TRVInfra requirements

Next steps (from January 2025)

- Increase the data for training
- Continue the validation of the classification with Trafikverket
- Once we achieve 90%+ accuracy, classify whole TRVInfra dataset (18.000) requirements [WP04]
- Use chatGPT or IBM's PoC to reformulate non-verifiable requirements [WP04]
- Document software and usage instructions [WP05]

Classification status

- Validated with Martin/Oskar 72 requirements
- Classified and consistency checked: 250
- Implemented the classifier and published here: https://github.com/bth-dipt-research/SVAR

Classifier evaluation results (10-fold cross validation)

N=72	Accuracy	Variation	Confidence interval
Target	85.9%	6.4	82.7% - 89.1%
Nature	86.4%	11.6	80.6% - 92.2%
Interpretability	63.9%	9.0	59.4% - 68.4%
Reference	84.8%	7.2	81.2% - 88.4%

Interpretation:
- Excellent performance with very little data
- Interpretability is the most difficult dimension
- Larger ground truth likely to improve results

N=250	Accuracy	Variation	Confidence interval
Target	89.6%	2.3	87.3% - 91.9%
Nature	90.8%	4.7	86.1% - 95.5%
Interpretability	72.8%	2.7	70.1% - 75.5%
Reference	95.2%	3.0	92.2% - 98.2%

Interpretation:
- Even better performance with a bit more data
- Variation has gone down considerably
- Interpretability is still the most difficult dimension
- Larger ground truth has improved the results

TRVInfra statistics

target	Count - target
Documentation	871
Process	3130
Product	13371
Total Result	**17372**

nature	Count - nature
Mixed	580
Qualitative	14057
Quantitative	2735
Total Result	**17372**

interpretability	Count - interpretability
Ambiguous (artificial)	1292
Ambiguous (natural)	313
Non-ambiguous	15767
Total Result	**17372**

reference	Count - reference
External	969
Internal	860
Local	1729
No reference	13814
Total Result	**17372**

Using genAI to improve TRVInfra

Created the "Kravklar Expert" (Demo link):

- Analyzes a requirement w.r.t. ambiguity
- Explains why a requirement is ambiguous
- Suggests reformulation alternatives

Next steps

- Prepare report (scientific article)
- Validate classification results with Trafikverket
- Find recipient in Trafikverket: what to do with the results?

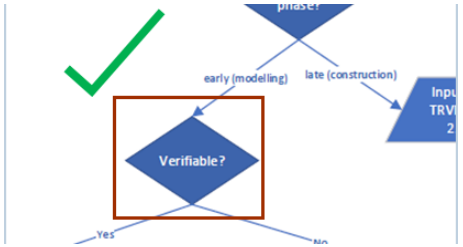
Objective 2: Machine-readable requirements

Objective 3: Demonstration of verification methods

Overview
Approach / Activities

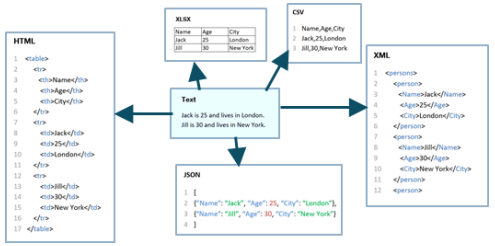
Work Packages 4/5 (Input)

List of verifiable Requirements



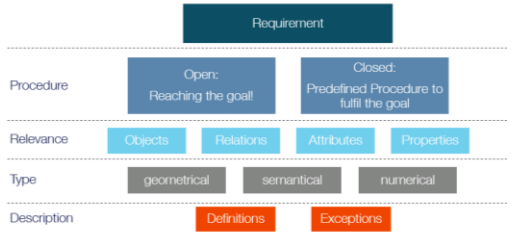
Work Package 6 (Preparation)

Methods to make Requirements machine readable



Work Package 7/8

Proof of Concept for Verification

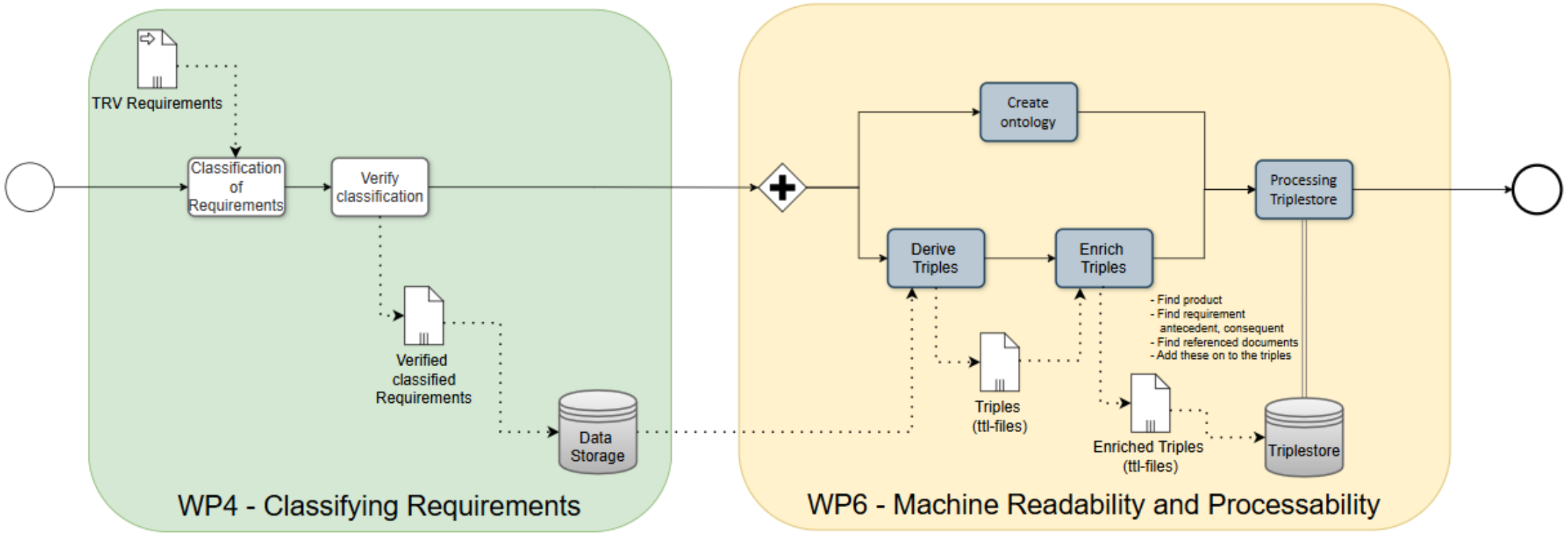


Objective 2

Objective 3

Objective 2 – Work Package 6

Transferring Requirements to Machine Readability



Objective 3

Develop procedures for automated, reusable, verification of requirements

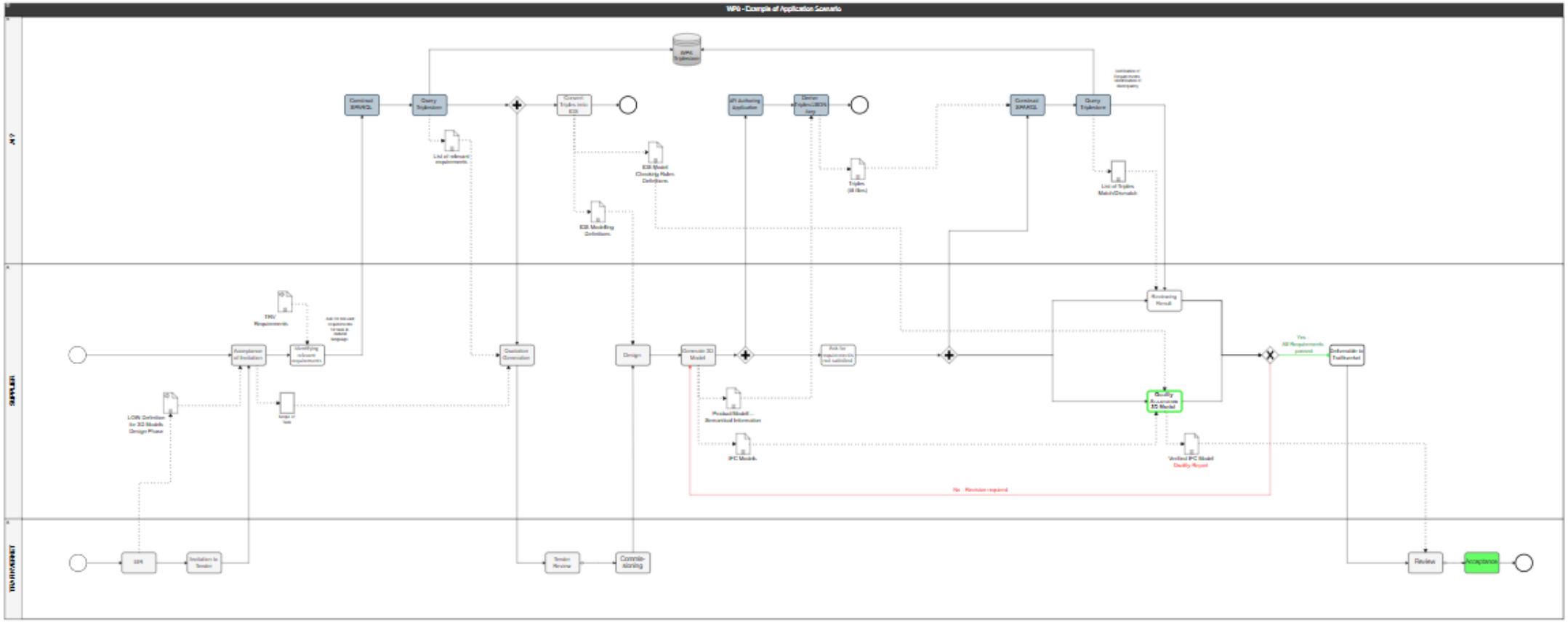
Work Package 07 – **Demonstration** of verification methods of models

Work Package 08 – **Evaluation** of verification methods

Work Package 09 – Roadmap and recommendations for implementation

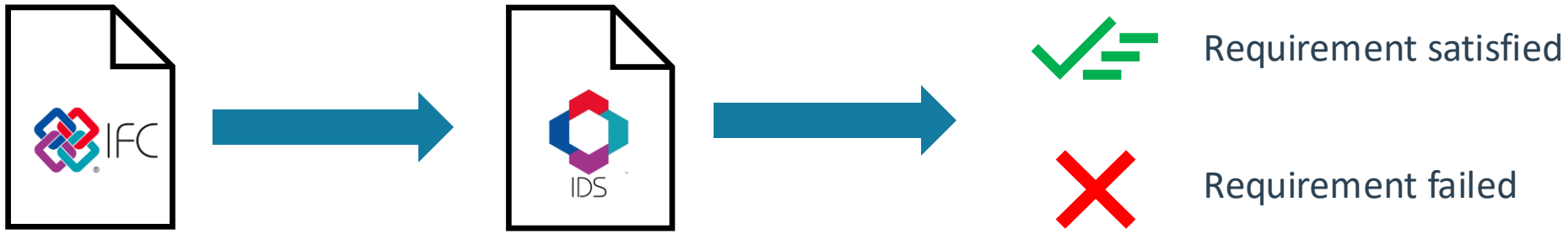
WP7/8 – Demonstration & Evaluation

Elaboration of an Application Scenario



WP7/8 – Demonstration & Evaluation

Elaboration of an Application Scenario



Input

IFC File

TRV Requirements

Given Definitions

TRV Property Sets

bsDD Property Sets

“As a designer I would like to validate my models in terms of compliance with TRV Infra regulations.”

BuildingSMART Data Dictionary

Relying on Standards

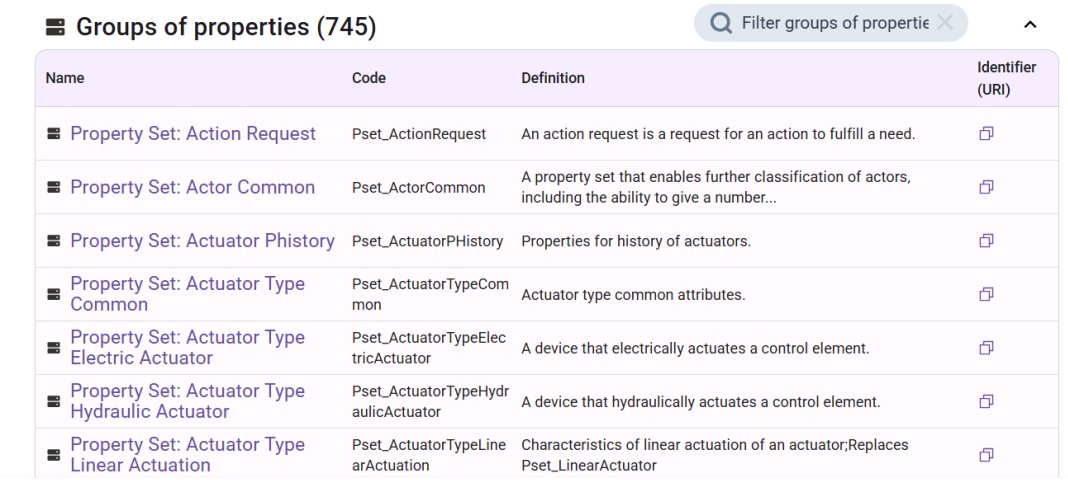
bSDD

Collection of **interconnected data dictionaries** with **definitions of terms** to describe the built environment

definitions of classes and properties - what is a Duct or how to capture a Diameter

relations between definitions, mapping between classifications, translations

no information about particular products or projects, only definitions of terms to describe them **(metadata)**



Groups of properties (745)

Filter groups of propertie

Name	Code	Definition	Identifier (URI)
Property Set: Action Request	Pset_ActionRequest	An action request is a request for an action to fulfill a need.	
Property Set: Actor Common	Pset_ActorCommon	A property set that enables further classification of actors, including the ability to give a number...	
Property Set: Actuator Phistory	Pset_ActuatorPHistory	Properties for history of actuators.	
Property Set: Actuator Type Common	Pset_ActuatorTypeCommon	Actuator type common attributes.	
Property Set: Actuator Type Electric Actuator	Pset_ActuatorTypeElectricActuator	A device that electrically actuates a control element.	
Property Set: Actuator Type Hydraulic Actuator	Pset_ActuatorTypeHydraulicActuator	A device that hydraulically actuates a control element.	
Property Set: Actuator Type Linear Actuation	Pset_ActuatorTypeLinearActuation	Characteristics of linear actuation of an actuator;Replaces Pset_LinearActuator	

WP7/8 – Demonstration & Evaluation

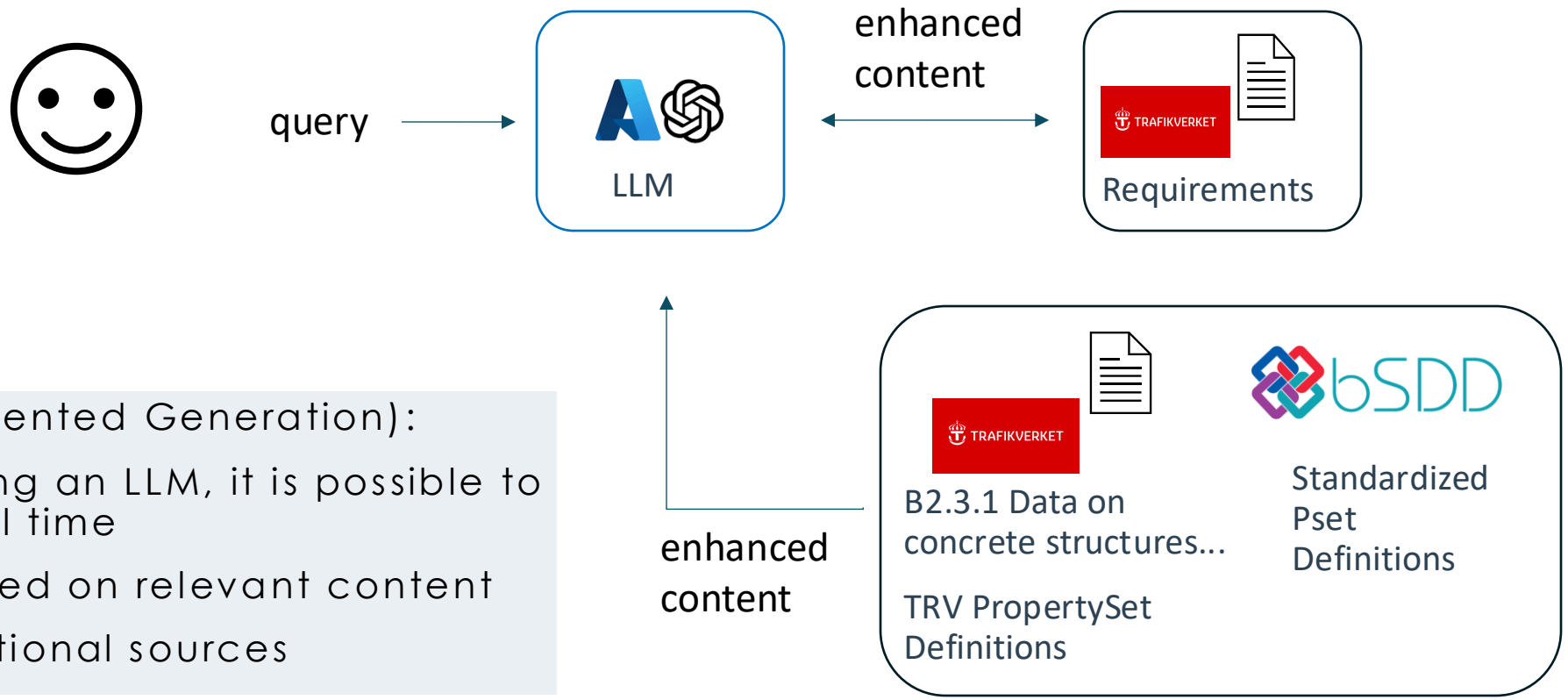
Current Approach: AI to assist

RAG (Retrieval Augmented Generation):

- Instead of fine-tuning an LLM, it is possible to **guide an LLM** in real time
- Limiting Output based on relevant content

WP7/8 – Demonstration & Evaluation

Current Approach: AI to assist



RAG (Retrieval Augmented Generation):

- Instead of fine-tuning an LLM, it is possible to guide an LLM in real time
- Limiting Output based on relevant content

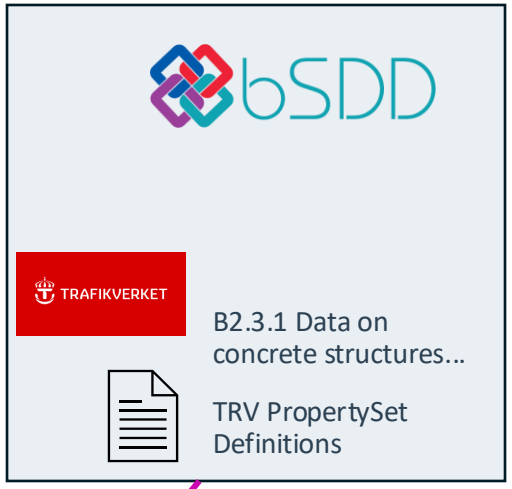
  ...AND adding additional sources

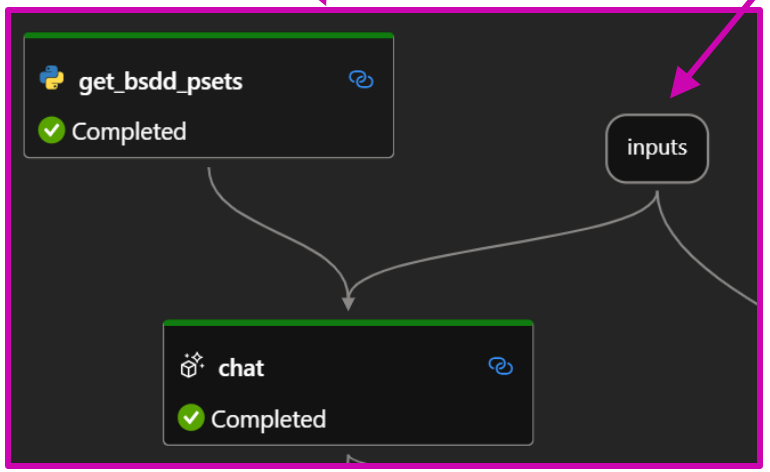
WP7/8 – Demonstration & Evaluation

Current Approach: AI to assist

System instruction: What are the relevant property sets in bsDD?

↳ Pset_ConcreteElementGeneral





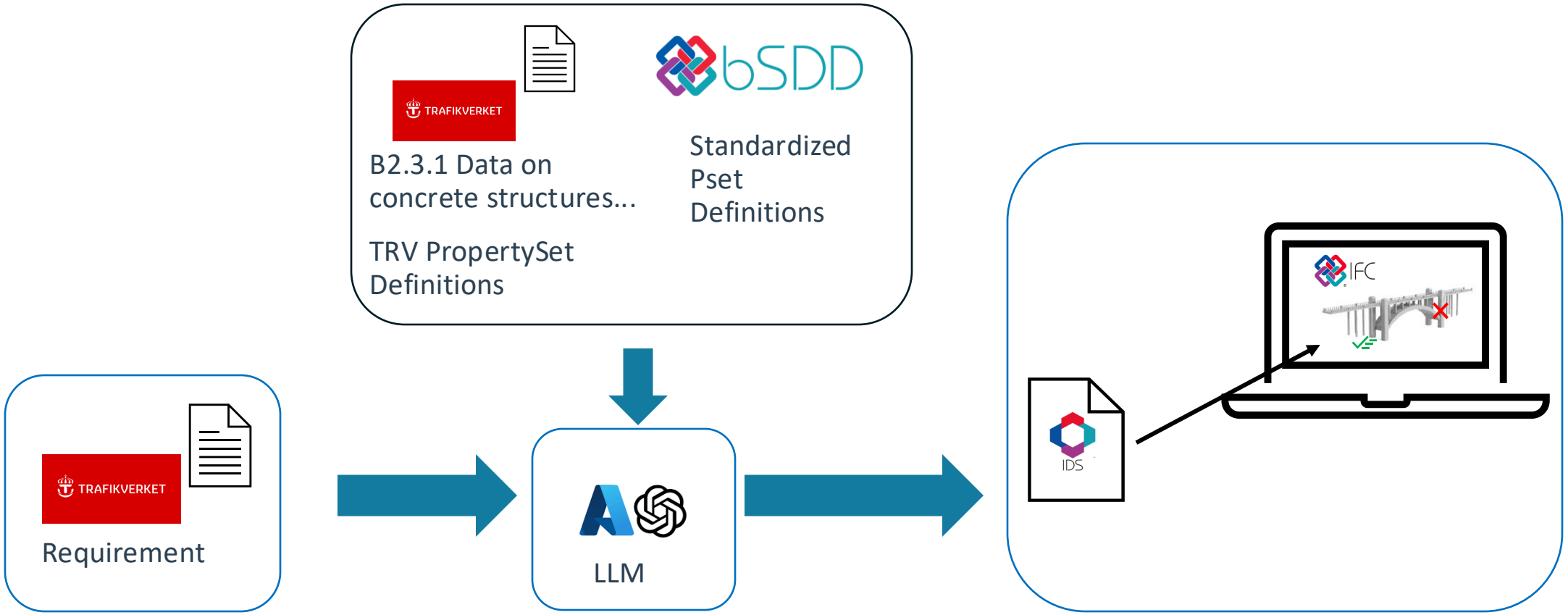
RAG (Retrieval Augmented Generation):

- Instead of fine-tuning an LLM, it is possible to guide an LLM in real time
- Limiting Output based on relevant content

  …AND adding additional sources
- A combination of chat nodes and code blocks in a flow

WP7/8 – Demonstration & Evaluation

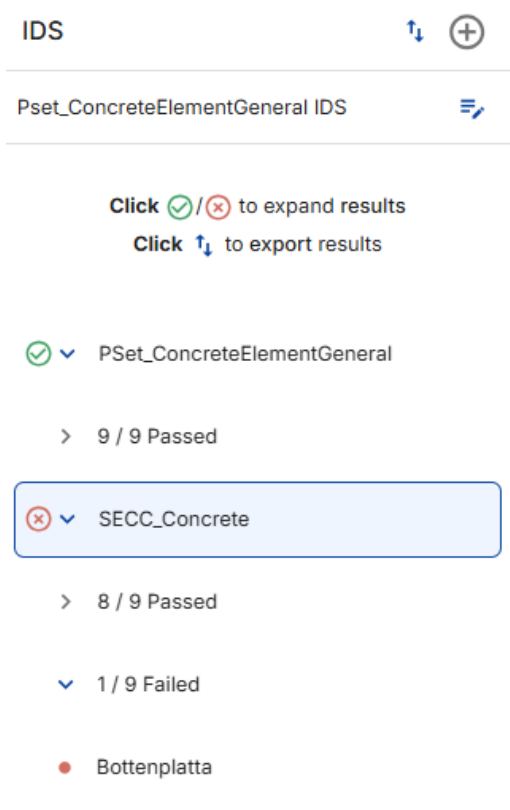
System Pipeline: How it works



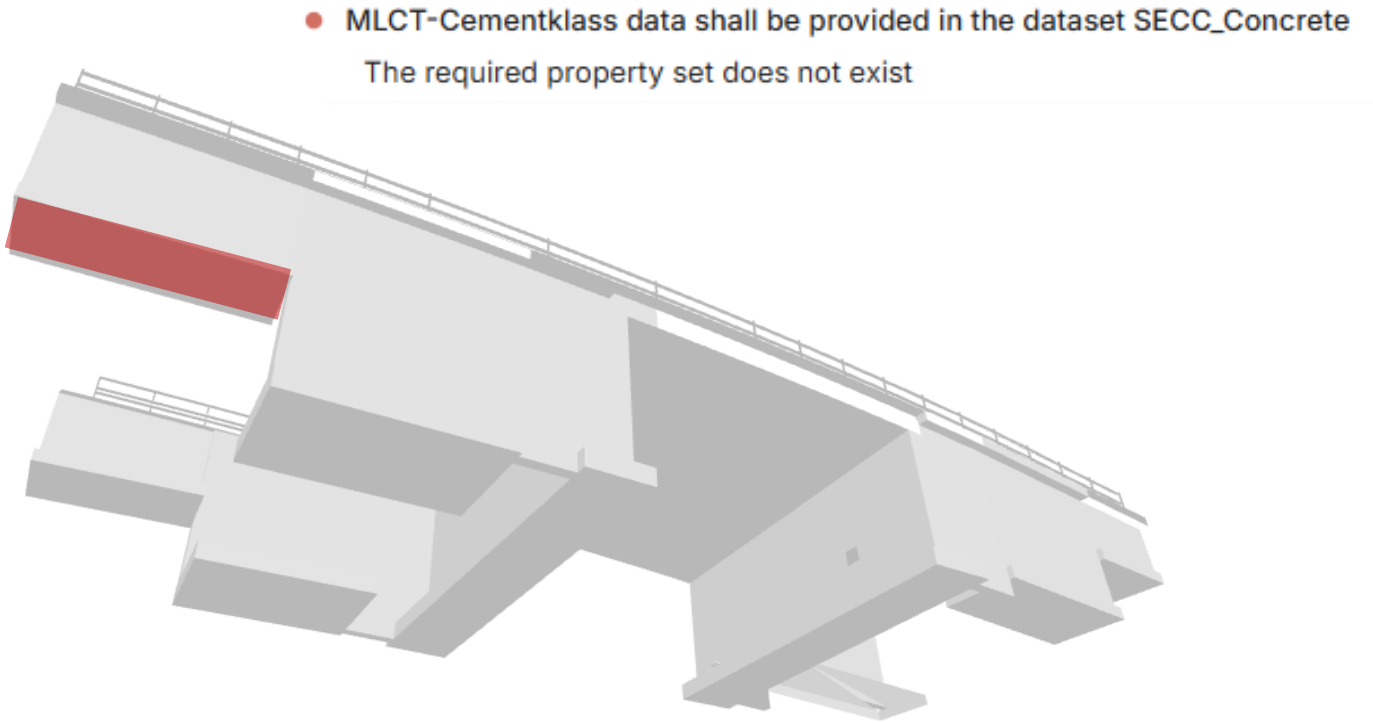
gpt-4o on Azure OpenAI

WP7/8 – Demonstration & Evaluation

Provide results in Viewer



IDS Specs created from the flow



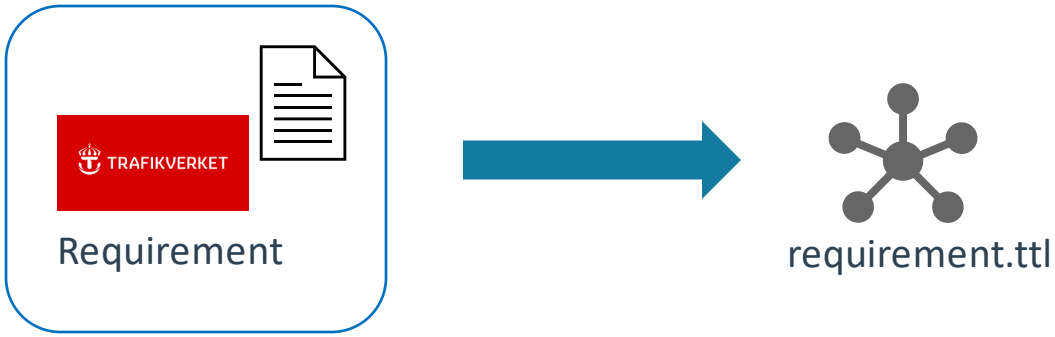
Free IFC Viewer with IDS checker from Sortdesk: Sortdesk | Free Online IFC Viewer

WP7/8 – Demonstration & Evaluation

Improvements – Ontology Approach

- While the approach works, sometimes the LLM fails to detect the targeted Pset
- Better approach:

→ Assisting the LLM with **structured data** to find the most relevant Psets



WP7/8 – Demonstration & Evaluation

Improvements – Generating RDF

Current: Natural Text

The drawing or description of the concrete structure shall contain at least:

1. Data on the concrete concerning
   - strength class
   - exposure class
   - execution class
   - cement type and cement class
   - consistency
   - water-cement ratio
   - air content
   - ballast property of importance.
   - Type of additive

Interim: SBVR Notation

1. Concrete structure has a drawing or description
2. Drawing or description contain concrete data
3. Concrete data contains strength class
4. Concrete data contains exposure class
5. Concrete data contains execution class
6. Concrete data contains cement type and cement class
7. Concrete data contains consistency
8. Concrete data contains water-cement ratio
9. Concrete data contains air content
10. Concrete data contains ballast property of importance
11. Concrete data contains type of additive

Final: requirements.ttl

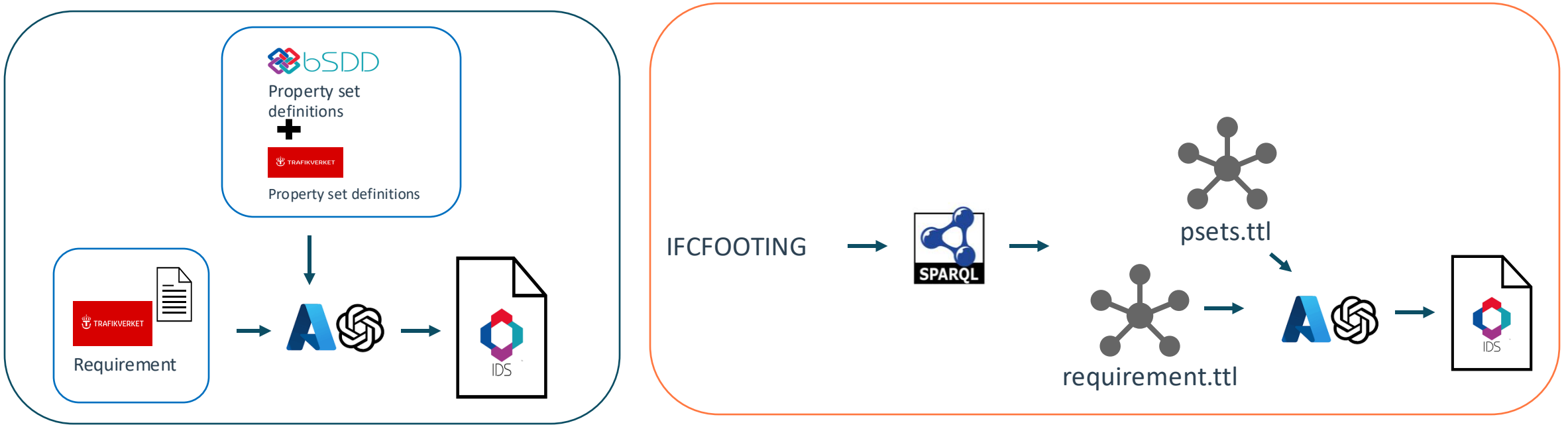
```
:data_1234    :contains    IFC-PSD:StrengthClass
.
:data_1234    :contains    IFC-
PSD:ExposureClass .
```

- Reconstruct the requirement using SBVR to generate the RDF
- SBVR - Semantics of Business Vocabulary and Business Rules
- RDF - Resource Description Framework

WP7/8 – Demonstration & Evaluation

In progress

- Overcome the challenge of finding the wrong Psets
- Constructing SPARQL query that returns the actual result in multiple steps using LLMs



WP8 – Verification Methods

SPARQL Queries

Enable

Application Scenarios

- AS 01: Quality Assurance 3D Model (IDS)
- AS 02: Construction Site Assessment
- AS 03: Daily Shift Reports/Daily Diaries
- AS 04: Site Logistics

AI Search: Establishing different workspaces for stakeholders, Indexing of relevant documents

Manual Checks

WP7/8 – Demonstration & Evaluation

Other application scenarios

Other Resources/Standards:
- Site reports
- Design documents
- Product sheets
- Etc.

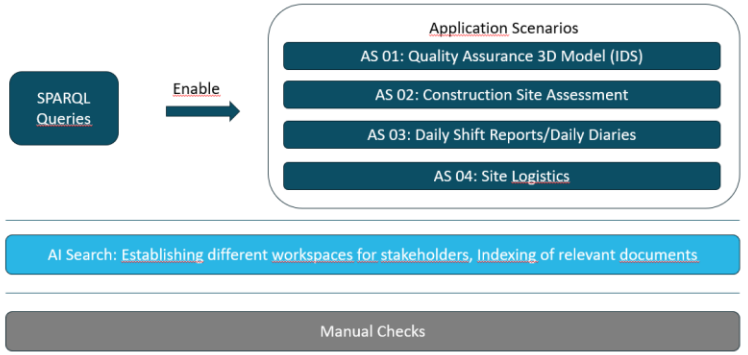
Structured Data:

RDF
XML
JSON
CSV

TRAFIKVERKET

Requirement

LLM



Queries on Chaos to check compliance

Queries on other databases (SQL)/triplestores(SPARQL)

Objective 3
Upcoming Activities

Work Packages 7
(Demonstration)

- Choose five requirements ✓
  - Different classified requirements
  - Referring to different file types
- Example ✓
  - Create suitable Ontology
  - Identify information in document
  - Identify information in model
- Verify Requirement ✓
  - Via Ontology approach

Work Package 8
(Evaluation)

- Use this verification methods ✓
  - for IFC, Excel, PDF, etc.
  - PoC for Ontologies
- Create several Checks ✓ ~
  - based on Ontologies
- Reporting of Evaluation ✓ ~
  - Findings
  - Limitations
  - Challenges

Work Package 9
(Recommendation)

- Summary of all achievements ~
  - Processes
  - Examples
  - Demonstrators
- Possibilities and Opportunities
  - Proof of Concepts
  - Identified Benefits
- Requirements for implementation
  - What is missing?
  - What needs optimization?
  - Where to (possibly) integrate AI?

Champions for project outcomes

Motivation: critique from previous research projects that results are not transferred to TRV

Idea: have one person from TRV "champion" the results and drive dissemination/adoption in TRV *after* the project

Goal: find in 2024 champion(s), based on the results we achieve.

Ambition: start in 2025 with dissemination/promotion, before the project ends in September

Champions

1. ACC-CMM: Susanne

2. TRVInfra verifiability procedure: Susanne + Erik

3. ACC demonstrator: Karin

Next steps

- Prepare handover:
  - Contact champions
  - Prepare deliverables
- Date for next (last?) reference group meeting


BLEKINGE TEKNISKA HÖGSKOLA
BTH