Data-driven Construction Planning with Knowledge Graphs and Reasoning

Photo by Sven Mieke on Unsplash

Accurately planning a build is one of the biggest challenges facing the construction industry, particularly considering the increasing complexity of buildings and the range of disciplines required in the process (structural, electrical, landscaping, etc.). Project data is often kept in different silos and it is difficult to create an integrated and interoperable information flow across all stages of the process.

NIRAS is an international engineering company currently running thousands of commercial projects around the world, spanning sectors from construction and infrastructure to energy and climate. They are experts in the space and all too familiar with the challenges facing the industry.

By implementing knowledge graph technology and exploring the potential of the powerful reasoning of RDFox, the team at NIRAS has laid the groundwork to overcome these hurdles.

...

A construction project generally starts with a client needing a new building. With the client's wishes in mind, the architect and engineers will design the building in its entirety. This then guides the procuring of contractors, outlining the contractual requirements and allowing them to form a construction plan.

The work required for completion has to be estimated according to the design. It's not unusual for this to rely on conjecture when—depending on the phase of planning—the required information is not yet available. This frequently leads to increasing costs beyond budget and delays.

While it is common to encounter these issues, there are a few things already in place that attempt the reduce their effect.
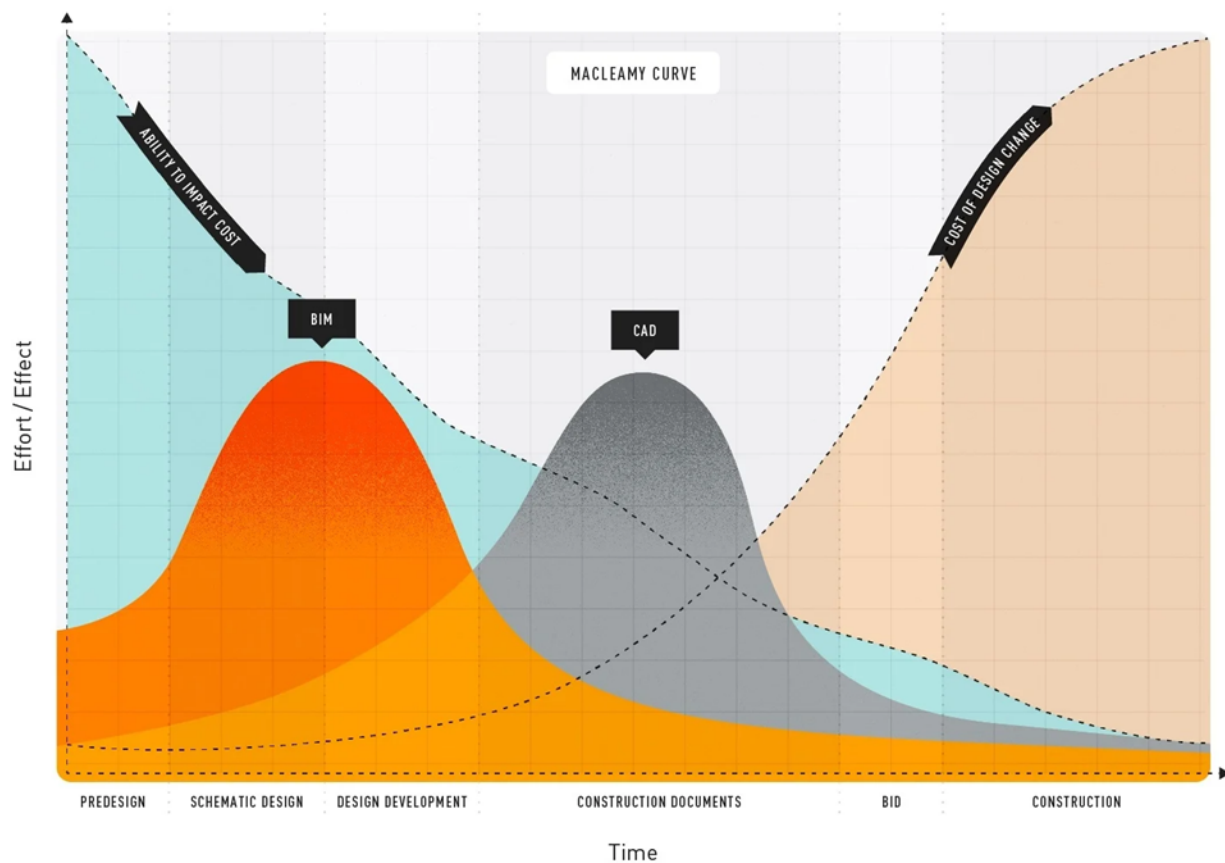
...

The MacLeamy Curve

The MacLeamy curve schematically describes the benefits of applying data-driven planning to construction projects over the different phases. This process is known as Building Information Modelling (BIM), and it allows us to model the entire building with several 3D representations, grouping the elements by their relevant profession. Every element within these models can be assigned information to provide additional details: its material, location, specific properties, etc. This whole process serves the main goal of gathering as much valid information about the project as early as possible.

As described by the MacLeamy curve, the ability to impact cost and functional capabilities of the building decreases as the project progresses, and in parallel, the cost of changes increases. The approach of BIM aims to shift pivotal decision-making moments to earlier stages of the project by applying better information management, increasing the likelihood of a smooth success and keeping costs down.

That's the theory but, in reality, it is a tremendous task to share information efficiently and openly between all project participants. Even with BIM, many projects breach time and budgetary targets—clearly, an improvement is needed.

. . .

Work Estimations

Accurate information flow is not the only issue plaguing the industry; the act of estimating construction time is a constant cause of problems itself. Estimating the time it takes to complete a complex building is (unsurprisingly) incredibly difficult, and is a skill possessed by just a handful of experts. It takes years, if not decades of experience before an expert is able to make estimates accurately, thus this skill is not a widely available resource and leaves companies reliant upon individuals.

Ultimately, this leaves individualsresponsible for the planning of multi-billion dollar projects rather than the organisation which can prove problematic. If an expert suddenly leaves a project, all their expertise leaves with them. Without someone who understands the project to the same extent, this often leads to disastrous results.

This human-centred expertise together with a lack of continuous and adequate information flow presents a huge problem to the industry. In order to address the issue, we raised the question: how can we formalize the expert knowledge into a computer-readable format and integrate all relevant project data to formulate a plan based on transparent and reliable information?

---

[1] https://blog.vectorworks.net/4-keys-to-successful-data-modeling-implementation

In the rest of this article we will answer this question with a linked-data solution characterized by transparency, reliability, interoperability, and reusability of the created project data.

…

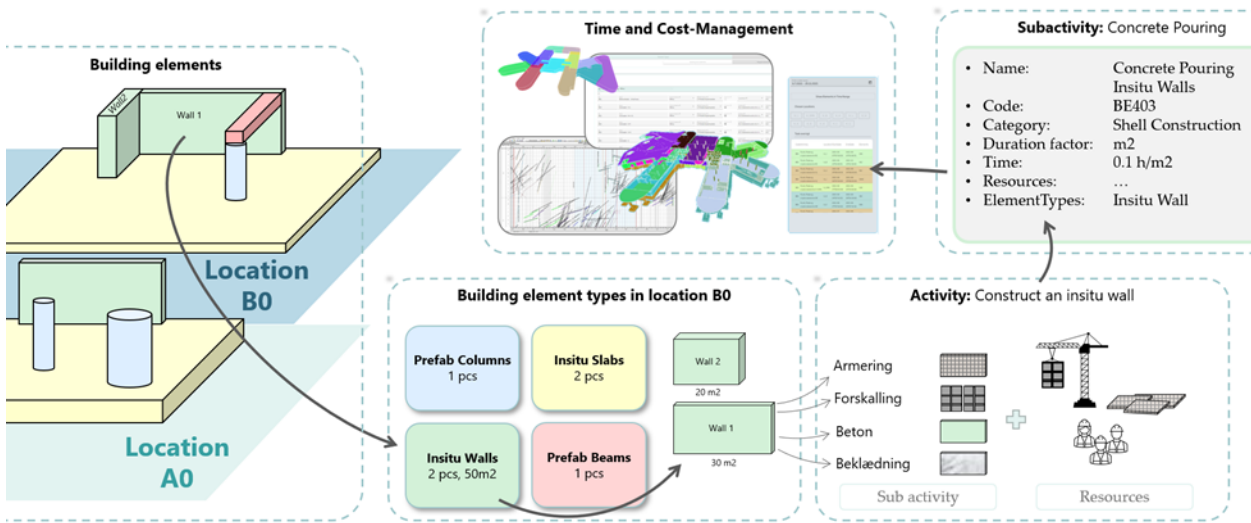# Using Linked Data to Design Workflow

Fig. 2: Data-based planning workflow

Here we have a high-level visualization of the concept behind our data-driven planning workflow–a workflow designed around BIM in the MacLeamy curve. In this integrated process of location- and quantity-based planning, the model is used to split the project into distinct building elements. Associated with each is information about the skills and resources required for its construction. This structured view allows us to combine what we know about each element (the weight, dimensions, material, etc.) in a knowledge graph.

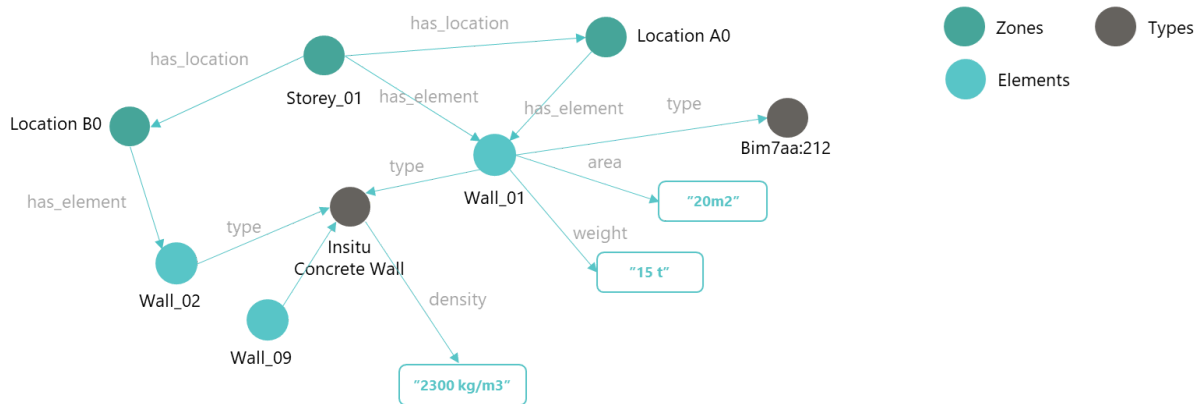Below you can see what this would look like for an insitu concrete wall.

By doing this with every component we can plan the construction process. The information has been extracted entirely from the building design and now allows us to accurately allocate resources where they are needed.

Enriching the data with experience—expert advice

For each type of building element, we can list the procedures required to construct or install it on-site. Based on experience we can define standard properties for each process that allow us to calculate the time required to complete it.
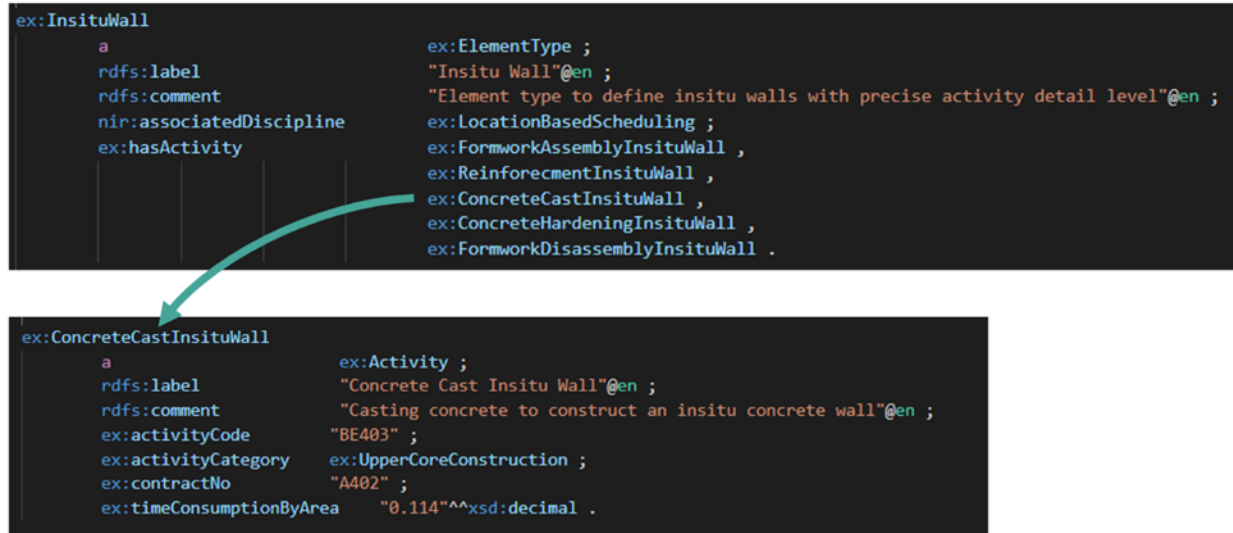

Fig. 4: Insitu wall standard activity example

In the example shown above, the element type 'insitu wall' has an associated set of activities that are necessary for its construction. One such activity is concrete casting, represented above as 'ex:ConcreteCastInsituWall'. It too has some associated properties; a code, a category, the responsible contractor, expected time consumption and so on. We know the quantity of this type of element from the building model (previously translated into a graph data model), so we can use this combined data set to create a more accurate project time estimation based solely on reliable information.

We can add this information directly to our graph, combining knowledge of the building elements, processes they are involved in, and the duration of those processes.
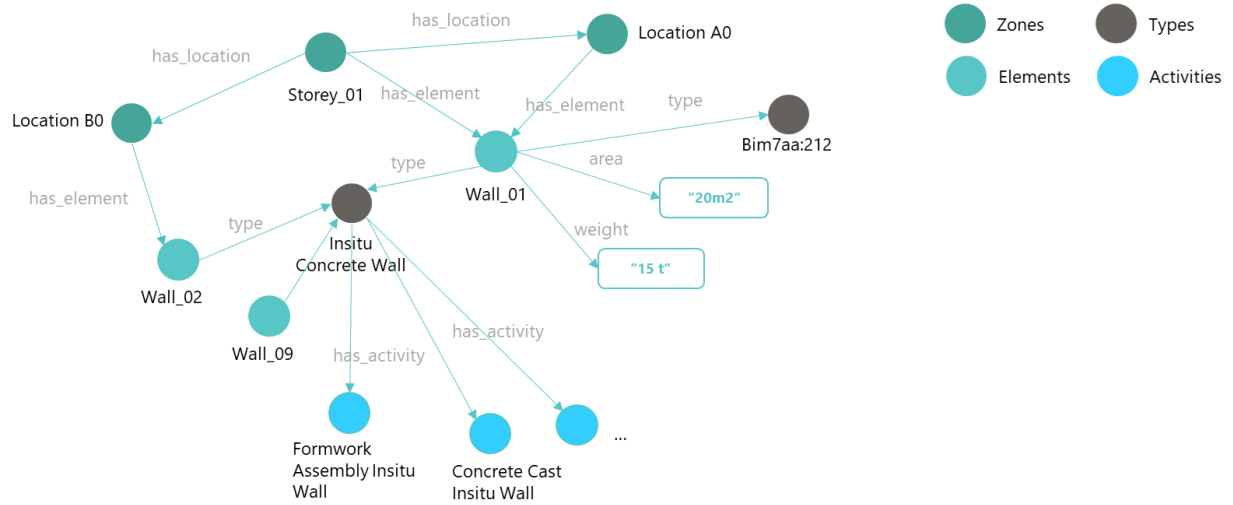
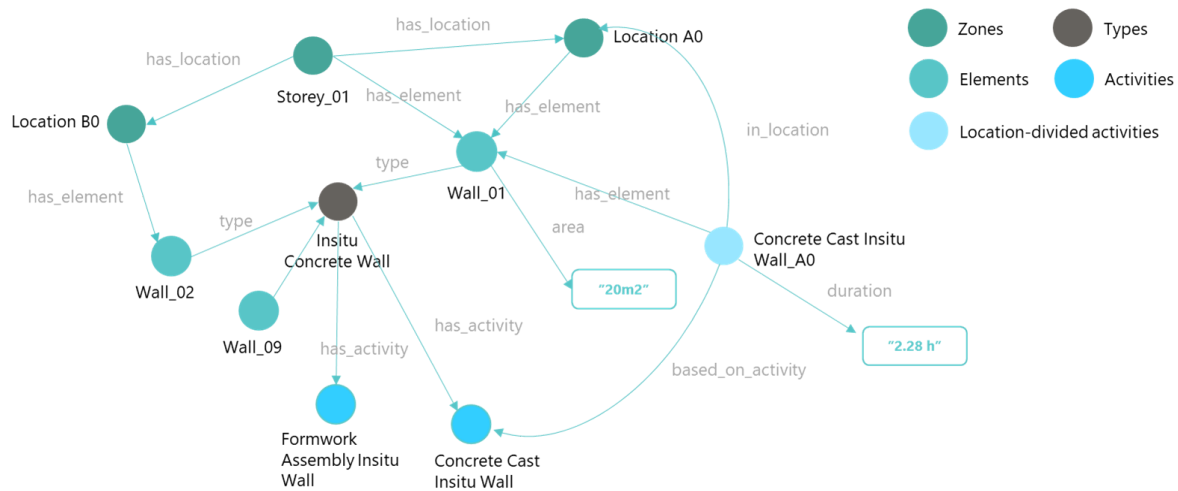**Fig.** 5: Adding activity information to element types

Fig. 6: Deriving location-divided activities and calculating their duration

. . .

Enhancing Expertise with Rules and Reasoning

As mentioned, RDFox is the knowledge graph and reasoner behind this solution. Its reasoning capabilities allow us to take the project data to another level by defining specific rules that infer further knowledge from the existing information. Rules in RDFox are written in Datalog, an expressive declarative rules language. With these rules, we can incorporate the knowledge and

experience regarding construction planning, and broadly apply it across all projects to upgrade the data.

The following are some examples of rules and the functionality they added.
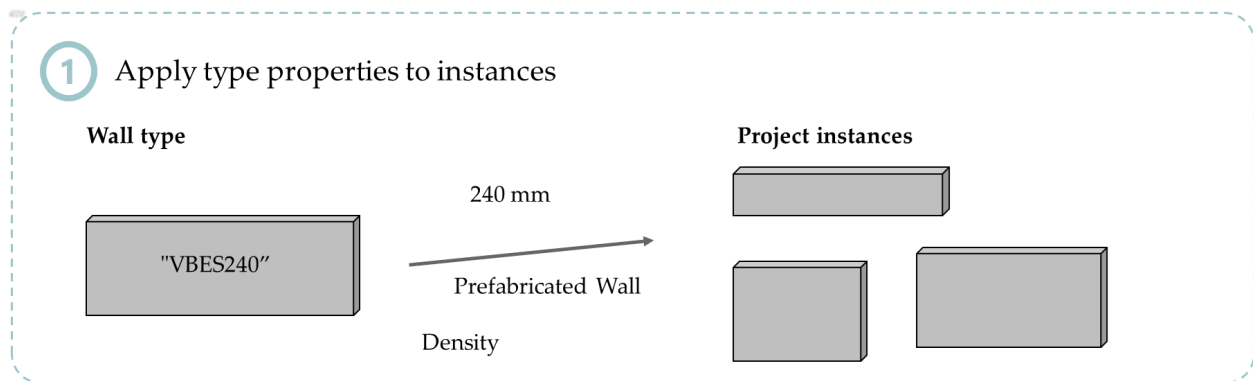
1. Apply type properties to instances

Fig. 7: Datalog rule example 1

The first is a very general rule that helps us to work with type and instance properties. Most BIM applications differentiate between individual properties and group properties for any instance of a component. Take a prefabricated concrete wall as an example: we can define its length, height or location as an individual property because the information is only valid for that specific instance of the wall. We can also categorize the wall as a VBES240 to assign it properties that it shares with other individuals in that group such as the material, density, fabrication process and width.

```
[ ?type , nir:density, ?val] :-
      [ ?type , lbs:revitStructuralMaterial, ?material ] ,
      [?material, nir:density, ?val] .
```

When planning construction or conducting analysis, we almost always have to include properties at both the individual and group level, which leaves the corresponding queries bloated, confusing and slower than they need to be. Therefore, with the application of a simple rule that applies group properties directly to each instance of that group we make the process of managing and querying data much more efficient. All related properties of an instance can be retrieved by just querying the instance.

```
[?instance nir:length, ?length],
[?instance nir:volume, ?volume],
[?instance nir:height, ?height],
[?instance nir:density, ?density],
[?instance nir:thickness, ?thickness],
[?instance nir:columnDiameter, ?columnDiameter],
[?instance nir:columnWidth, ?columnWidth],
[?instance nir:columnDepth, ?columnDepth],
```

```
[?instance nir:structuralFramingHeight, ?structuralFramingHeight],
[?instance nir:structuralFramingWidth, ?structuralFramingWidth]
:-
        [?instance bot:elementType, ?type],
        [?type, nir:height, ?height],
        [?type, nir:density, ?density],
        [?type, nir:thickness, ?thickness],
        [?type, nir:columnDiameter, ?columnDiameter],
        [?type, nir:columnWidth, ?columnWidth],
        [?type, nir:columnDepth, ?columnDepth],
        [?type, nir:structuralFramingHeight, ?structuralFramingHeight],
        [?type, nir:structuralFramingWidth, ?structuralFramingWidth],
        [?type, nir:length, ?length],
        [?type, nir:volume, ?volume].
```

2. Splitting of prefabricated elements based on max. weight
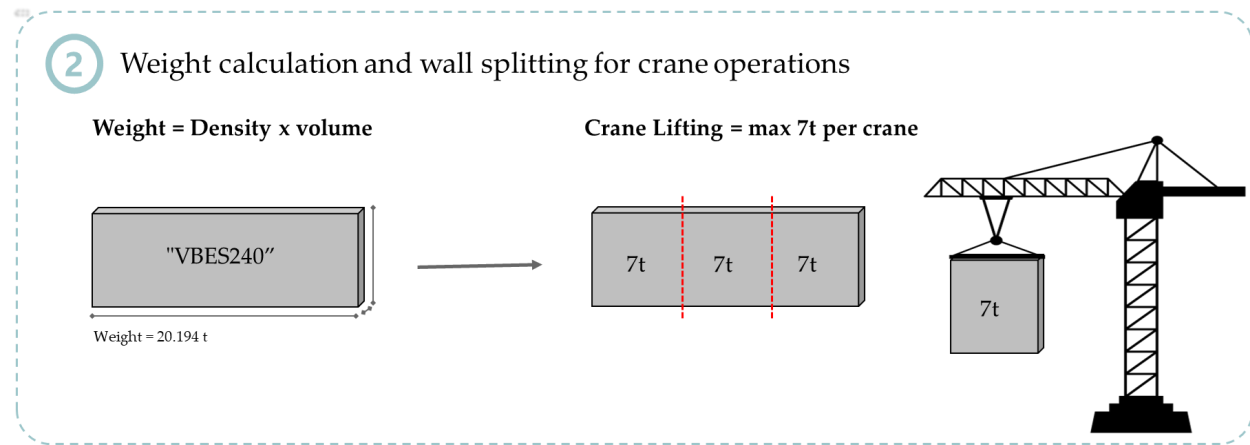

Fig. 8: Datalog rule example 2

In the process of installing prefabricated elements, it is crucial that the time estimation is based on the right assumptions. To make sure these assumptions are acted upon correctly, we need to equip our data with expert experience. In the case of lifting elements with a crane, there is a maximum weight limit–around 7 tonnes on average. With this knowledge, we can define a rule that will split all of the elements heavier than 7t into chunks that our equipment can manage. In the example of the prefabricated wall, we can calculate its weight by multiplying the group density by the instance volume. This particular element weighs in at just over 20t, and thus the wall must be split into three manageable parts. By applying the rule, this information is saved in the data model so we can use it to plan the work for installing prefabricated wall elements.

```
[ ?element , nir:weight, ?weight ] :-
        [ ?element , bot:elementType, ?type ],
        [ ?type , nir:density, ?dens ],
        [ ?element , nir:volume, ?vol ],
        BIND( (?dens * ?vol) AS ?weight ).
```

```
[ ?element , lbs:minLiftingDivision, ?splitSections ] :-
      [ ?element , a, lbs:Element ],
      [ ?element , nir:weight, ?weight ],
      BIND( IF(ROUND(?weight/7000)=0,1.0,ROUND(?weight/7000)) AS ?splitSections ).
```

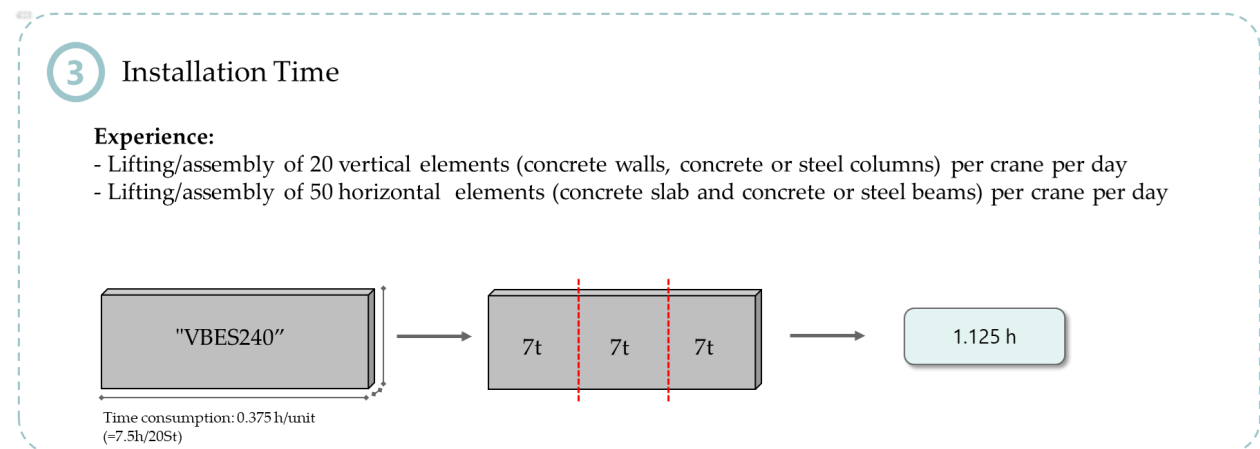## 3. Calculating the installation time for prefabricated elements

Fig. 9: Datalog rule example 3

The next step in planning the construction activities is to estimate how long it will all take, and to do so we need to estimate the time consumption of each activity from experience. In our scenario, we know that we can lift and assemble around 20 vertical elements with a crane every day. If we take into account a working day of 7.5h, this results in a time consumption of 0.375h per lifted object. With the aforementioned wall that we just split into 3 elements for lifting, we can calculate a total installation time of 1.125h.

```
[ ?element, lbs:installTime, ?installTime ] :-
      [ ?element, bot:elementType, ?elementType ],
      [ ?elementType, lbs:hasActivity, ?activity ] ,
      [ ?activity, lbs:timeConsumptionByPcs, ?timeConsumption ] ,
      [ ?element, lbs:minLiftingDivision, ?splitSections ],
      BIND( ROUND(?splitSections*?timeConsumption*100)/100 AS ?installTime ).
```

This rule can be applied to all vertical elements that are made out of either concrete or steel. For horizontal elements or other materials, other values or calculation methods must be applied.

By applying these three rules we are able to automatically build an estimate for the total installation time from known qualities of the building. The Datalog rules combine experience with context buried in linked data. The same concept can of course be applied to any construction

activity that can be related to a quantifiable resource in either a building model or any other data source.

…

Improving Legacy Planning Methods

Why is this so much more reliable and transparent than the legacy methods in place today?

The main difference is that all the data is integrated and that, due to the incremental reasoning of RDFox, the system maintains internal consistency and is always kept up to date without the need for a restart. If the design changes, the data for construction planning is updated, and all the consequences of the rules (the time calculations, logistical conclusions, etc.) are updated automatically too. If say the area of the wall is changed with a design alteration, the area property of Wall_01 will update accordingly along with the time estimate for the activity Concrete Cast Insitu Wall_A0.

Suddenly the planning process becomes parametric because the construction schedule is directly linked to the building models. As we change data in the models, the schedule changes accordingly. This enables us to explore alternatives and to see the impact of doing things differently.

How much time would we save if we only install prefabricated concrete elements? If we know the activities for installation, a simple update of the material in the building model will give us adequate information on the alternative approach. By evaluating and comparing the results, multiple designs can be tested for their viability.

Other factors to test include the number of cranes, machines or human resources, operating on-site. All of these scenarios become simple to simulate, helping planners to find the optimal schedule.

…

Where will this go in the future?

In the future, this concept solution can be further developed with more rules and experience to cover the entirety of a construction project. With an application that is able to visualize data and integrate it into the planning process, the concept can be tested and further refined. Niras is currently working on such a solution and is constantly engaged in finding new cases for utilizing Linked Data technologies to improve the way we plan construction projects.

If you would like to find out more about NIRAS and their solutions, please visit their website niras.com.

Likewise, if you would like to learn more about RDFox, graph technology, or semantic reasoning, you can read about other applications on the Oxford Semantic Technologies blog. You can even try out RDFox for yourself for free and start solving your challenges with graph technology today.

About the Author, Alex Schlachter

Alex Schlachter is a construction professional and consultant at NIRAS A/S working in the field of digitalization with focus on location-based planning and automatization. His team aims to develop valuable and digital solutions for common challenges in both design and construction, based on Web and Linked Data technologies.

NIRAS

NIRAS A/S is a Danish engineering company with 51 offices and around 2,400 employees worldwide. With over 7,000 current projects within the sectors of construction, infrastructure, utilities, environment, climate and energy, the company has a wide range of projects and gathers experts in all the fields of engineering. Niras' mission is closely linked to sustainability and digitalization and runs numerous big and small initiatives to fulfil this mission.

RDFox

RDFox is a high-performance in-memory knowledge graph and semantic reasoner. Optimised for speed and advanced reasoning, it offers query and load times that are orders of magnitudes faster than alternative triplestores, while also achieving greater insights into the data. RDFox is developed by Oxford Semantic Technologies—an Oxford University spin-out founded by leading academics who led decades of cutting-edge research in semantic web technologies.