AALBORG UNIVERSITY
STUDENT REPORT

2021

Advanced visualization and contextualization of energy performance and indoor environmental quality operational data of BIM models

Master Thesis
MSc in Construction Management and Building Informatics

Abhishek Relekar & Piotr Smolira

Aalborg University

Department of the Built Environment

2020-2021

Title: Advanced visualization and contextualization of energy performance and indoor environmental quality operational data of BIM models

Affiliation: Department of the Built Environment, Aalborg University

Program: MSc in Construction Management & Building Informatics

Semester: 4th, Master Thesis

Project hand-In: January 8th, 2021

Supervisors: Associate Professor Kjeld Svidt
Assistant Professor Ekaterina Aleksandrova Petrova

Authors: Abhishek Relekar
Piotr Marek Smolira

Acknowledgments

The authors of this study would like to express their gratitude to the thesis supervisors Ekaterina Aleksandrova Petrova and Kjeld Svidt for their guidance, advice and helping us connect to the industry professionals while preparing the report. It was great to learn from their experience. Furthermore, we would like to thank Torben Østergård and Steffen Maagaard for describing us the issues from the industry and providing us the data for developing the use case. We would also like to thank the developers at Autodesk for providing us the information about different resources used during the prototyping. The received inputs were of the essence for this research. Finally, we are very grateful to our families and friends for the support and encouragement received during the process.

AALBORG UNIVERSITY
STUDENT REPORT

Abstract

The AEC industry (Architecture, Construction, and Engineering) is going through a technological paradigm shift due to the growing user needs, and the demand for making sustainable built environment to improve people's quality of life. Due to the establishment of BIM in the AEC industry, design data is stored and can be utilized for connection to other data sets in order to develop innovative user-oriented applications. Although this is generating challenges for the users from different knowledge areas of the construction sector in collaborative decision making while interacting with and understanding such data. Through a technological perspective, there is a demand to integrate different fields extending from Building Information Models, IoT devices and ultimately the end user services to build intelligent, user-oriented applications. A broader theme is presently trending in the construction industry about cyber physical systems also known as Construction Digital Twins. This project report is looking into integrating BIM and IoT data to make user-oriented tools for visualizations while increasing the understanding and aiding the communication process between project stakeholders as well as corelating with the developing digital twin paradigm in the construction industry.

AALBORG UNIVERSITY
STUDENT REPORT

Table of contents

Abbreviation list

Abbreviation	Full name
AEC	Architecture, Engineering, and Construction
AI	Artificial Intelligence
API	Application Programming Interface
BIM	Building Information Modelling
CDT	Construction Digital Twin
CPS	Cyber Physical Systems
CSS	Cascading Style Sheets
CSV	Comma Separated Values
DOM	Document Object Model
DGNB	Deutsche Gesellschaft für Nachhaltiges Bauen
DT	Digital Twins
FM	Facility Management
HTML	Hypertext Markup Language
IC-Meter	Indoor Climate Meter
IDE	Integrated Development Environment
IDE	Integrated development environment
IoT	Internet of Things
MQL	Mongo query language
ODBC	Open Database Connectivity
PLM	Product life cycle management
REST	Representational State Transfer
SME	Subject Matter Expert
SVG	Scalable Vector Graphics
VS	Visual studio code

1. Introduction

Evolving Building Information Modelling (BIM) tools and technologies are progressively changing the way information in the AEC (Architecture, Engineering, and Construction) sector is produced, stored, and exchanged between involved stakeholders. Therefore, the advancement of BIM should be meticulously planned while considering people, processes, and these evolving technologies in a developing inter-connected world (Batty, 2018).

Simultaneously to the developments within the construction sector, there is a growing challenge for a smarter built environment by more determined energy and carbon emissions programs across the world. The addition of IoT (Internet of Things) and AI (Artificial Intelligence) is demanded to produce improved energy efficiency and reduce operation costs (Howell et al., 2017).

BIM in this context represents a minor building-level view within the broader environmental context. Although, BIM has been utilized to incorporate lifecycle management of built assets, the current level of BIM is not compatible with IoT integration (Howell et al., 2017).

BIM is deficit for its semantic completeness in subjects such as control systems, sensor networks integration, social systems, and urban artifacts beyond the scope of buildings, hence it requires a comprehensive and a scalable semantic approach that can handle dynamic data at multiple levels (Boje et al., 2020).

The use of the IoT technology in the construction of intelligent buildings can widen the practicability of intelligent systems, help develop resource allocation strategies and improve the management and service capabilities of intelligent buildings, thereby enhancing people's quality of life (Kong & Ma, 2020). Whilst this efforts are aimed for increasing the use of virtual models, the problem of integrating the virtual models and the physical construction world such as to enable adaptive interaction or bi-directional coordination has not been satisfactorily addressed (Akanmu et al., 2013). To achieve this bi-directional coordination, computational resources are required to securely integrate the virtual and physical domains such as the changes in one environment are autonomously mirrored to the other (Akanmu et al., 2013).

The CPS (Cyber-Physical system) approach is especially critical in the construction industry as it presents an opportunity for real-time seamless data flow between the design and construction teams, therefore giving opportunities for swift and informed decision making (Golparvar-Fard et al., 2011).

To develop the CPS approach in the construction industry, there is a requirement to address the demand for the integration of sensor data and digital representations of the built environment for fostering stakeholder collaboration management within the area of Real Estate 4.0 and Facility Management (FM), particularly in a spatial representation context (Stojanovic et al., 2019).

A major challenge in FM is visualization of historic and current sensor data within as-is built environment representations such as virtual models and plans of buildings, produced from a building's DT (Digital Twin). The DT is an alternative notion correlated with cyber–physical integration (Tao et al., 2019). Sensors calculating building-related operational data and man-made events can provide information and insights about the current operational status of a building or a site (Stojanovic et al., 2019).The process of visualizing BIMs continues to be an issue according to the veracity of construction project typologies. The visualization requires to be accurate, enhance understanding of the as real world structure and specific to different user's needs, hence aid collaborative and coordination consultations, from a functional and ergonomic outlook (Kubicki et al., 2019).

Data visualization is a subject of ongoing development. Almost every software or service provides some kind of data representation, even limited one. There are tools which helps contextualize and present data in accessible for a user way, such as D3, Lyra, Tableau (Fisher et al., 2011; Veglis, 2017), currently researched Project Dasher, plugins, add-ins for programs like Autodesk Revit. Even though they provide great utilities in itself, they do not contribute to the implementation of visualization as frameworks (Kamran Sedig & Parsons, 2016). They can provide ideas which can be an inspiration for creating new ways of interaction with information, but not in a systematic manner. In other words, it is possible to look at those tools as examples. However, it is tough to synthesize them and use as one coherent, sensemaking, framework (Kamran Sedig & Parsons, 2016).

Foremost, data visualization can be divided into two main categories– static and interactive (Kamran Sedig & Parsons, 2016; Veglis, 2017). First one uses predefined by designer, developer, or implementation schemes, and views. While it can be useful in some areas, it has its shortcomings, such as fixed data selection, data types, range or types, 2D-only perspective (Ward et al., 2015). Relationship of different data types, data structures are getting more complex and they grow in size in a built environment, and sufficient solutions used with static visualization are being replaced with interactive approach (Kim et al., 2017; Natephra & Motamedi, 2019a, 2019b). The reason for this is that a user can contextualize given information, dynamically adjust it to different perspective and sub-visualizations, making it faster to process computationally and mentally (Kamran Sedig & Parsons, 2016). To have complete picture, state of certain situation, outcome or prediction, simulation, the end-user have to use various sets of data, or even the same sets, but presented in particular ways (K.-M. Chang et al., 2018). Interactive data visualization, unlike static, enhances the analysis, process, and complex activities, especially in the age of big-data, which has to be properly aggregated and visualized (Dou et al., 2020; Khan & Hornbæk, 2011; Po et al., 2020).

Even great tools are meaningless without a user. Together they create one cognitive body which is meant to accomplish given assignment (Liu et al., 2008). There are five spaces in such cognitive system (K. Sedig et al., 2012)– information, computing, representation, interaction and mental. It covers whole human-information interaction. Starting from the beginning, some information exists and captured, then it is stored and processed. Thereafter, the data is represented by means of media. The user access data from representation space through interaction space (Parsons & Sedig, 2014). Here activities– what activities, and decision are done, and made are defined in the mental space (what user wants to do with such information, how the user wants to perform)– influence, form, sculpt representation space to user needs (Kamran Sedig & Parsons, 2016).

Data visualization in the built environment industry does not enclose in what the end-user sees. The concept is broadened to the whole system architecture, and variety of stakeholders, ultimately creating workflows for such processes. It covers data acquiring system (information space), database and backend solution (computing space), data visualization and frontend solution (representation space), end-user exposure on data representation, its manipulation and modification (interaction space), and cognitive activity of the end-user regarding tasks that end-user must perform (mental space).

2. Problem formulation

In Collaborative and Multi-Stakeholder organizations such as construction industry, communication is crucial to project success. Therefore, project management must consider suitable methods and practices to include the stakeholders in a collective decision-making. Especially context is considered as a key concept in the design of new innovative technologies (Kubicki et al., 2019).

Models and visualization of the context of an AEC project is generally considered to be necessity in the decision-making procedures (Kubicki et al., 2005). Relying on multi-dimensional information datasets, hence it should meticulously handle context visualization to aid the decision-makers(Kubicki et al., 2005).

Color patterns and dynamic animations are the main features of 4D models, helping the viewer to understand the data better (Chen et al., 2012). As BIM is adopted rapidly, "*the management of information "included in" and "related to" digital building models is challenging"* (Kubicki et al., 2019) , particularly as the number of linked datasets grows exponentially. The access to the necessary and relevant information at the appropriate time requires formalization and effective management of all collaborative information exchange processes (Kubicki et al., 2019) (Po et al., 2020).

The problem formulation of this thesis project is a synthesis of what have been found in the literature review– area of data visualization and contextualization, BIM and IoT approach– and challenges of communication and data visualization described by the building energy engineers of a construction engineering consulting company in Denmark. The problem the energy engineers face sparked the need to look closer into the literature to find the research gap corresponding to the problem. Here we quote the visualization and communication challenge engineers encounter in their stakeholder meetings:

"*Our problem (as a consultant and energy/ indoor climate engineer) could be to disseminate results from measurements and simulations to building owners, architects and end-users who may have a more "visual" need to understand such technical data (Sensor data of buildings). As engineers, we love time-series plots, heat maps, tables, bar graphs, etc. And for this, we like to use programs like Excel and PowerBI - with the limitations they have. In addition, such tools will probably help to understand and visualize the dynamics of the building, which both our customers and we may have a hard time seeing from the graphs and the project participants would like to develop better* *methods for communicating and understanding data and what it means for the indoor climate, health and possibly energy consumption"*.

Furthermore, adhering to the research gap of deficiency in sematic completeness in subjects such as sensor networks integration, social systems beyond the scope of buildings with the need of handling dynamic data at multiple levels and the demand of a web based integration of IoT with the eventual merge between virtual models and sensing converging on a common semantic web platform (Boje et al., 2020) to develop a cyber physical systems approach in buildings the following problem formulation is established:

"How to contextualize building sensor data with BIM (Building information models) to create advanced visualization techniques according to user requirements?"

The above-mentioned problems will be addressed in this project. The foremost step will be to establish the method how the literature review is going to be proceeded as well as what type of methodology will be adopted to the use case (chapter "*3 Methodology"*). This will help to focus on the case as a proof of concept and treat it as a rigorous tool to evaluate the undertaken work process and findings. Thereafter, a literature review (chapter "4 *Literature Review"*) of the subject and its core components will aid to understand the latest research, technical aspects, and the gaps in the literature to develop a holistic, scalable, and innovative application framework.

There would be two use cases based on two different BIM models. The structure of the BIM models is different thus would require various approaches to prepare, integrate and scale for the framework application, this would be discussed in detail in the use case chapters.

The start of working on the solution begins with implementation of methodology stated before (chapter "*3 Methodology"*). Here, each point presented in *Methodology* Sub-Chapters *3.2* and *3.3* are addressed, and utilized as a part of solution development process to show practical application of the chosen systematic approach. Another reason, as much important as the first one is to create foundation for the technical part of the project – developing system architecture for solution (chapter *"6 Development of System Architecture")*. In this section will be described actions to accomplish tasks, and milestones necessary to form the solution for the project's problem.

Our expectations/purpose of this project are to develop data visualization and contextualization tool using systematic approach of the field of user-centered design by following each step and evaluating its feasibility.

3. Methodology

Well established methodology is a key for systematic written report. It helps to categorize and show reasoning of the actions taken in developing the paper. The aim of this chapter is to establish research approach applicable for a study area of this project. One method will suit for an overall outline of the paper, another will be for literature review, data collection, and analysis. However, they should create holistic framework how the research is conducted.

3.1. Methodological approach and research design

In the case of this report, it is important to find out what the AEC industry is settled on nowadays in terms of data visualization – advancement on BIM in general and what are the current processes for working with data (reading, refining, analyzing). This would not only give a good overview of the subject, but also start discussion why certain aspects are currently on the particular stage of development, e.g., process and framework, data flow (speed, accuracy, interoperability), software robustness and human interaction with information. The last is particularly interesting since it concerns the end-user regarding the given solutions– it can be software, web-application, or physical tools and devices. Therefore, the methodology and processes will be adopted which put a user in a center of the design.

The interconnected term– a user-centered design was introduced by Norman & Draper (1986) in their work about human-computer interaction. Various processes go along with the user-centered design. One of them is Contextual Design, for which the foundation was developed by Karen Holtzblatt and Hugh Beyer in the early 1980's. The concept has been improved and developed since then to be fully presented in *Contextual Design: Defining Customer-Centered Systems* (Beyer & Holtzblatt, 1997). Holtzblatt & Beyer (2014) also contributed to the book *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* by Interaction Design Foundation (Donald A. Norman, 2014) with the chapter *Contextual Design* highlighting key principles of that concept in a comprehensive manner. This process emphasizes on using technology and solutions suitable for an environment of the subject of IT implementation. The Contextual Design process will be used together with the framework specifically dedicated to data visualization developed by Kamran Sedig & Parsons (2016) in *Design of Visualizations for Human-Information Interaction*. The plan is to work with those two approaches simultaneously to use their advantages, to fill some shortcomings they may have, and to use their synergy and complementariness, which at the end create cohesive framework body.

3.2. Contextual design

Contextual Design consists of two major sequences – I. Requirements & Solutions, II. Define & Validate Concepts. Figure 1 shows the steps included in the whole concept. The similarities of some of them will be visible while presenting Five Spaces of Cognitive System.

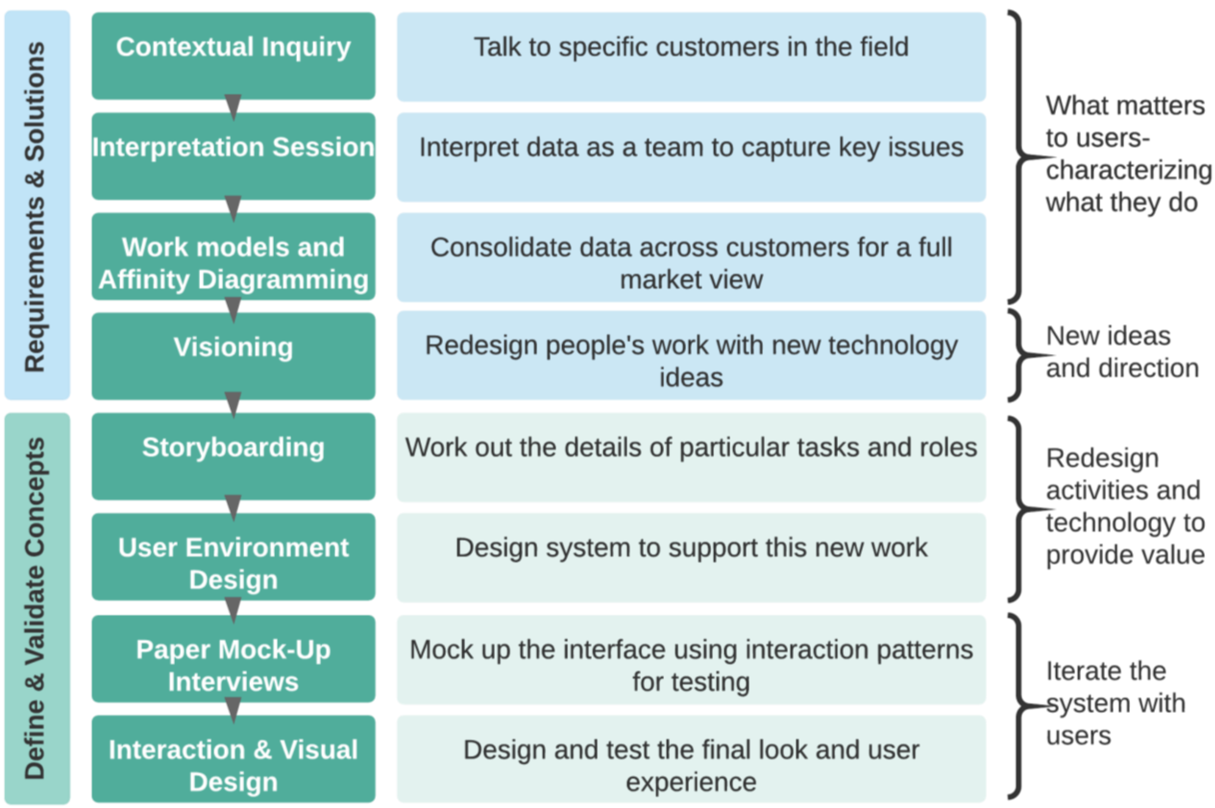


Figure 1: Contextual Design process. Adopted from Holtzblatt & Beyer (2014)

3.2.1. Contextual Inquiry

The point of this step is to understand the end-users, who have direct contact with the subject of implementation – indirect-users whose information provided helps end-user finishing the task and managers who supervise the workflow for given action and are responsible for implementation. Usually the end-user is a focal point of solution design process since his action will depend on the actual speed, accuracy, and overall performance of the tasks performed (Holtzblatt & Beyer, 2014).

In order to settle the status of current solution it is necessary to perform contextual inquiry. It often means conducting interview along with inquiring how the user works and observing his daily activities. This is also the plan for this case. However, the inquiry of daily actions will be limited due to the restriction on visiting workplace and general remote-work policies. During the interview it will be possible to ask for quick demo how the current tools work, and how the user operates with them and on what kind of information.

The first step is to identify the participant (the end or the indirect user) for contextual inquiry who would give insight into his work routine. It is obligatory to note that the term 'end-user' greatly depends on the implementation phase of the project. It means that, based on the current strategic goal, the end-user might not necessarily be the customer, building owner, public facility or building user, or even the person who browses the Internet or employs web application in one's work. The focus leans towards a specific user as the implantation advances.

There are multiple actors in the project or enterprise, and they create one network, where an actor receives information from the predecessor but also creates one and passes it on further down the workflow to another actor. During the full development and implementation of a solution, all participants and their roles are examined and given appropriate attention to their activities. It would be possible to go ahead and to study someone before or past the engineer, closer to absolute receiving end of information of the whole process, such as building end-user. However, the project revolves around visualization of sensor data where the main beneficiary of well represented information in the context of building model is the energy engineer consultant – it increases work efficiency by acquiring more complete picture of the data and generating information faster and more understandable for other stakeholders. He manages information – both received and created – with his tools to generate artifacts which are utilized for his own work (i.e., main task actions regarding data from sensors, simulations, energy calculations and assessments), or communication with other participants (e.g., the presentation for the meeting with stakeholders, discussion with engineers, architects, or departments). For the purpose of this project, it is decided to declare the energy engineer consultant as the primary user of focus.

Another point is an interview. The questions for it are prepared by doing the first step of design process, i.e. information space and task space (Kamran Sedig & Parsons, 2016). In their book included examples of questions which can be asked on the stage of inquiring information. For purpose of having as clear and simple interview as possible, questions are encoded (transformed) and provided with examples when it is necessary, maintaining the original questions and their structure. After the interview, answers and questions are decoded to see how the responses are related to the original inquiries. This is done in a such way to maintain the structure of the systematic approach and to make navigation between steps more comfortable with a possibility of recollecting thoughts, ideas, or set back to previous stages or iterations.

3.2.2. Interpretation Session

During this session, people from different fields and profession discuss various aspects of the information received during the contextual inquiry. It results having different perspective and considering them simultaneously (Holtzblatt & Beyer, 2014). It reduces bias and does not allow to focus on one specific aspect of the solution. After all, the implementation is not only about the tool itself, but it must fit with the overall company's business model.

For this interpretation session the plan is to discuss what information is gathered during the interview and demo presented by the interviewee. Also, in this chapter it is more suitable to focus more on the answers, and show summary of the contextual inquiry, highlighting the main points and main area of the interest, such as information type, data visualization techniques; from whom-to-whom information flows, etc. Although, this is going to be a session from one user point of view, he will provide insight what other stakeholders might want based on his and company's workflow– what he receives and provides, what others need, and how the whole body operates. Some suggestion will be received from the end user– either direct, e.g., when he mentions something specific what he desired or indirect when he points out weak links in current work process or application. This will help to recognize the issues and the user's needs which would be the subject of change making and will be brought back in the visioning stage.

3.2.3. Work Models and Affinity Diagramming

Visual representation how current work processes are done is a great tool to create a summary of actions performed by different stakeholders not only on individual level but also on interdisciplinary level (Holtzblatt & Beyer, 2014). Isolated actions by actors can be helpful, and it is possible to work on improvement in microscale by, for example, reducing navigation in a software, or automatizing small sub-tasks. However, without any other insight or depth of tasks performed by the actors, their intentions, and motivation might be never discovered or compared between each other to see if they are aligned together; to check if some misconception lies in roots of understanding the purpose of a process. Moreover, without such knowledge, it would be even tougher to evaluate how the process is well adjusted to the strategy of the company (Holtzblatt & Beyer, 2014). It helps to have different points of view on a single process.

There are different ways to show various perspective. (Holtzblatt & Beyer, 2014) highlight five contextual work models. Flow model (I) emphasizes relation between different actors, which can be formal or informal. Additionally, division of roles and interaction between actors or group of actors are captured in the workflow. It reveals type of information exchanged between participants, its frequency, iterations, amount, etc. In general, it helps to understand what actions done by one actor affects the others and how they influence the rest of stakeholders. In the case of this report, it can be communication outline between engineers, architects, a project manager, an owner, and their way of executing the tasks as well as exchanging information. Therefore, process map is outlined to know at what state of the workflow which an energy engineer has to deal with.

The cultural model (II) tackles the aspect of constraints and policies which people affected by them have to deal with, either by total consent or working around. Cultural model might include external factors, like market driven demands, or passing legal bill regarding built environment industry which compels companies to adjust their internal ways of performing job. It can be as well company's policy itself that will affect current working processes of employees. According to Sørensen et al. (2009), this model is especially important when it comes to implementation of the new solution, change of the present workflow of activities since it affects directly employers, employees, and even people seemingly being outside the scope – citizens who are concerned when the new building is being built, or a new technology starts surrounding them, such as RFID or sensors. As for this model, it is expected to draft relation between actors in terms of expectations, responsibilities and highlight possible conflicts.

The sequence model is used to illustrate detail steps of which takes the user to accomplish a task from the beginning to the end. It focuses more on the end user, his individual performance, and motivation or thought behind his action. In the use case of the report, it might be well established sequence of what an engineer does with information data from sensors and what kind of additional, side steps he takes during that process. It helps to have better in-depth understanding of what the engineer does, what actions he is forced to do, what actions are his initiation and how flexible he can be with them.

The physical model is another form of how to better understand the user. Such mapping unfolds the user's navigation process through his environment (Holtzblatt & Beyer, 2014). It can be in the physical setting such as building site or office, or software. User's path would give insight if there are some unnecessary actions in his process which can be improved (e.g., going back and forth to some areas whilst it could be avoided), or if there are obstacles which prevent the user perform more straightforward action. In the example of building site, it might be the question if the routes are convenient enough and properly marked which dissolves ambiguity or if workers themselves can organize their work in such a way to reduce time for carrying equipment, etc. In the case of this project, physical model will be set in software as this is primary tool and environment which the end-user works in. We could acquire information how often the user changes applications' windows, views, tabs; how often he has to compare or confirm sensor data information with 2D/3D digital representation of the building.

The artifact model gives information what artifacts have been involved in a user's process either as an input, means, or an output. This signifies the thought process of an individual user and his organization (Holtzblatt & Beyer, 2014) – this mostly refers to means and output. On the macro scale, it is worth to know if the type input the end-user of the interest gets is the one, he needs; and why he receives files, or information in such way. This would uncover how strongly his actions of work are affected by an external determinant and how strict his output is preordained in company's process workflow. In built environment the type of files might be an artifact and it can be, by any means, valid question why certain type was used rather than another. It would help to identify issues, shortcomings of a particular way of handling an information, but also see some benefits which might be carried on in the future implementation. During this study it is of the essence to learn more about freedom of choice when it comes to tools, the ways of performing a task, and user's motivation.

The last part of this step is consolidation. The purpose of it is to find common factors and patterns among users, but at the same time to maintain diversity by presenting individual issues presented in affinity diagram (Holtzblatt & Beyer, 2014). The representation is read from utmost important, most flexible elements to the least, which are derivative of the former. This diagram together with consolidated models mentioned in this chapter– where singular user's models of each type (e.g., artifact model) are combined– create one piece of workflow in an organization, a cohesive system.

There is no diversity in great extent in terms of users in the use case. Although, they exist, there is only one type which stays as a focal point. However, the paper will highlight the main influencing aspects for other actors after implementation of the solution for the end-user of this project.

3.2.4. Visioning

This is the last of the part "Requirements & Solutions". This is also the first time the design team will approach designing of the new environment for users. The challenge here is to analyze how the technology, and overall design solution will improve the actions performed by the actors (Holtzblatt & Beyer, 2014). Therefore, the new working process should be outlined. At this stage it does not have to be detailed to the smallest scale possible, but rather it – "high-level design" (Holtzblatt & Beyer, 2014)– must capture general purpose, functionality, interconnection between actors, individual action flow in the new system, its architecture and database layout.

The project will mainly focus on drafting scheme for the one end-user– energy engineer – and how his work is influenced by others and how his actions are influencing others. The paper will cover as widest scope of the system as possible to see how the solution would fit in the total work process and organization structure.

3.2.5. Storyboarding

The first step of rewriting tasks, activities, and solutions architecture begins with storyboarding. The objective here is to narrow the spectrum of level of detail of the whole working process – from general indication of relations and interactions between actors to more concrete description of functions and actions performed by users. This is an intermediate step between visioning ("high-level") and more concrete and structured – User Environment Design (Holtzblatt & Beyer, 2017). Each storyboard should reflect the step of one task that user should do to accomplish certain activity (Holtzblatt & Beyer, 2017, 2014). Although, a description of individuals is important, it must be assured that the placement of users' actions in the work process structure are in right order and correlation. In addition, the coherence of actors' activities and tasks which can be exchanged between each other – with a intended sequence– must be maintained, which the purpose of this step (Holtzblatt & Beyer, 2014).

In the case of the report, the most crucial storyboards will be drafted. It will emphasize navigation of energy engineer around the system workflow, e.g., what steps he takes to ensure that he has everything ready to perform a task– what information he has to receive; what activities he must perform to accomplish his tasks; what measures he takes to provide everything the next actor will need to perform his task– what information he has to exchange. Due to limited amount of information how exactly the process is done by predecessor and successor of the main end-user, the project will rely on the answer of the interviewee. Connection of the three main conditions– prior, during, and after end-user's task– will assure that most of considerations surrounding the energy engineer– the stakeholder in focus– will be addressed.

3.2.6. User Environment Design

The functions of users and their activities described in visioning and storyboard must be considered while designing a new workspace. The User Environment Design ensures that space for a task outlined in a storyboard has sufficient area to be performed. It means that user environment should be treated as a floor plan for a house where it is not explicitly said what tools and equipment are located in the kitchen, but general structure and "rooms" with specific functions (Holtzblatt & Beyer, 2017, 2014). Authors' intention of comparison to the architectural outline is to help visualize what should be done in user environment stage – defining size, function, room-to-room flow, constraints, all in par with user requirements, lifestyle, i.e., with client's brief. It is all relevant to the software design where size of areas in the interface defines the freedom the user has; where every space has designated function and purpose, called also "Focus Area" (Holtzblatt & Beyer, 2017, p. 339); where connection from one view to another influence user experience.

The challenge for this project is to capture user's requirements and design structure (like a floor plan) and focus areas (rooms) which will include their purpose and function. The design of storyboards and user environment is a subject of continuous improvement and being in a feedback loop. Therefore, the project will have at least one iteration with at least one storyboard and corresponding user environment focus area. As mentioned in previous sub-chapters, the focal end-user will be energy engineer, and this is most of the requirements and insights will come from.

3.2.7. Paper Mock-up Interviews

Presenting a system in the early stage of development to the user does not correspond with the necessity of having a working software, web solution. Having that application already implies to involve software engineers for the backend and frontend to write a code and sure it is working. That option is not only time-consuming but also cost-consuming (Holtzblatt & Beyer, 2014). The resources are allocated into something which will not necessarily see the daylight. The process of making a paper prototype is able to involve more people from different disciplines at once with a possibility of instant change of the mockup. Moreover, iterations between design team and the users happen more frequent. The paper prototype should look clean and should consist of all parts which will be a coherent solution, but also it should give the user an impression that views, functions, interactions can be changed (Holtzblatt & Beyer, 2017).

Although, the primary focus at this point is good structural layout of the user environment, interaction patterns are something which has to be taken into account on the later iterations. It means that the good basis for system architecture has to be laid down before an attempt to design the patterns. Holtzblatt & Beyer (2017) mentions in Chapter 15 that they do not provide techniques, principles for interaction patterns. Sedig & Parsons (2013) present list of patterns together and Sedig & Parsons (2016) index micro-level objects and macro-level aspects of interactivity. First refers to a fundamental characteristic of a singular interaction (e.g., immediate or delayed activation), while the second look at the relations between interaction which ultimately will reflect users' performance of their activities. The concept of will be elaborated in a Chapter *Methodology, Five spaces of cognitive system*.

The report will present the prototyping process with paper, or other media tools. The prototyping with interaction patterns, interview will be conducted and interpreted with guidelines from Chapter 15 and 17 of Holtzblatt & Beyer (2017). The interaction patterns techniques from Sedig & Parsons (2016) is a part of five spaces of cognitive system method and its own internal steps. They will be adopted in the use case and used as a supplement for contextual design. Based on the interview, the project will go the last phase of iteration of a paper mock-up before continuing with the last stage of the contextual design.

3.2.8. Interaction & visual design

After one or two paper mockup sessions with the interview– when the structure is more refined and stable, and interactions patterns and views are conceptually established– the development is ready for an implementation of storyboards and user environment in the virtual setting (Holtzblatt & Beyer, 2017). There are online tools which supports prototyping. It could be any mean which imitate interaction (e.g., dimmish of various views and layers, tiles, on-click actions, etc.) or even already working code which will be from that point of time the subject of improvement. Authors mention that this is a time to focus on "low-level" software elements such as drag-and-drop in interaction patterns, or specific visual design aspects, e.g., colors of windows, tabs, dialogue boxes, charts.

The plan for this part in project is to transition from paper mockup interview and to proceed with at least one iteration of an interview with working interface. The key aspects to focus are interaction and design. In the first is important to pay attention whether it is clear for the user to navigate to understand. In the design, the overall structure, design, colors, sizes of views, charts, 3D models are examined if they are displayed correctly, and if they convey the exact message based on the provided content, i.e., information, data. Iterations of this step should clear ambiguity, so it is crucial to listen to the user's feedback and see if he is confused by some actions, layout, colors; if something important from the dataset is missing which creates operational gap.

3.3. Five spaces of cognitive system

In order to discover what has to be done in regards to data visualization, it was decided to follow aforementioned systematic approach for data interaction and visualization using five spaces of cognitive system introduced by K. Sedig et al. (2012), and later developed and proposed as human-interaction design system by Kamran Sedig & Parsons (2016), which is represented in Figure 2. It depicts the system presented by the authors and shows how aforementioned spaces are connected. It starts from the source of information which can be data from sensors, other measuring tools and instruments. Such data would go through computational operations which its function is to store, encode, and have refined data ready for analytics. In the representation space VR means visual representation and here the information is displayed through various means. Interaction space work connects representation and mental space by adding functionalities to work with data such as its transformation in the visual layer. The last space regards to users' activities on the cognitive layer such as information reasoning, memorizing such data.

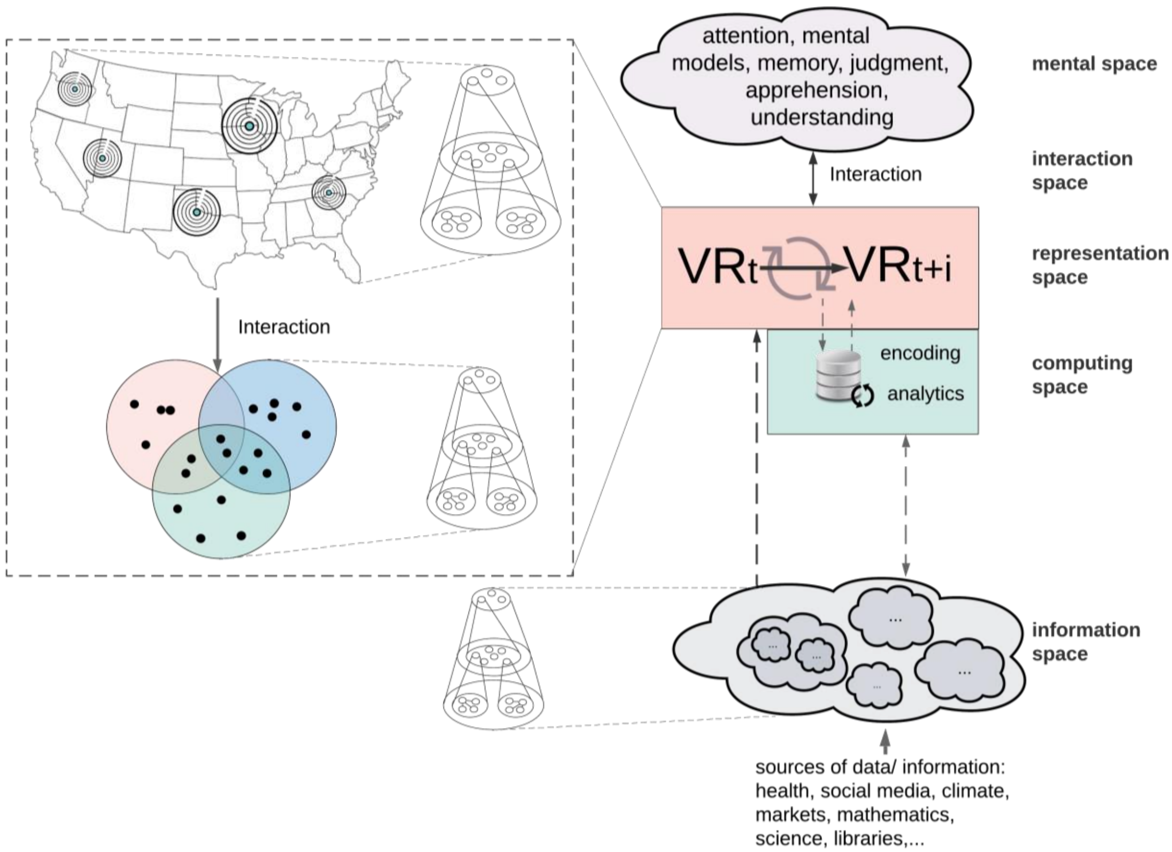


Figure 2: Overview of cognitive system. Adopted from (Kamran Sedig & Parsons, 2013)

As mentioned before, there are five areas, which are going to be explored in this case. Each of them has its own characteristics. In addition to that Kamran Sedig & Parsons (2016) includes four stages of design process which are tied with the cognitive system, and allocated to corresponding spaces. They iteratively so, even though there is a preferable sequence of steps to follow, it can move between steps, and even skipping.

1. Information space and task space.
2. Patterns, blending, and abstract structures.
3. Visualization techniques.
4. Concrete encoding and interaction.

3.3.1. Information space

This is the spatial system where data from sensors – metered, historical, from simulation – is acquired. This is an also place for any kind of predefined information/data, expressed with database, spreadsheets, specific file extensions, logs.

Information space is a part of the first step of design process. To acquire sufficient information about current data, data sources and its different properties it can be advisable to review current system by conducting interview and perform first part of contextual design – finding requirements and solutions. By going through assessment of information space system characteristics it is possible to gather information, for example, about data relationship, type of system (open/closed), etc.

3.3.2. Computing space

In this area data is processed, stored, and prepared for later representation – data refining, and filtering. It is important that end user is presented with clear, easy-to-read, and unambiguous information. It must be decided what kind of database system should be used, as well as backend and frontend solution.

This space overlaps with information space which means that data stream is continuously streamed, updated, refreshed, refined. Therefore, the first stage of design also encloses the computing space, since information, for example, about data relationship and cardinality is strongly tied with computing performance.

3.3.3. Representation space

The space is the start– the window– for human-information interaction. Here, information is encoded in visually accessible form (e.g., computer screen). Aspects, such as representation types (e.g., graphs, plots, histograms, 2D and 3D plans), navigation items (e.g., buttons, navigation wheel, floor view navigation thumbnail), and the whole UI.

To support representation space development, we investigate blending patters, and abstract structures. We should consider what kind of information will be displayed, and what pattern will be efficient to map and presented that data. For example, blending areal, and graph structures can an outcome with 3D building model representation (areal structure), with temperature chart (graph structure).

3.3.4. Interaction space

That part is a bridge between representation and mental space. Action performed by user, and reaction to the changes in the representation (alternatively also in computing space) are taken into account while designing such interaction (e.g., zooming in and out, changing floor plan view to cross section, etc.).

3.3.5. Mental space

This is probably the most abstract space of all five as it is about mental activity, events, and operations. User during the interaction with the information memories data to some extent, encode it format, and saves it in storage in one way or another (locally or in cloud). The point of that space is to assess how much a normal user will remember and has to remember while going from one view, or chart to another.

The brief overview of those five spaces, and four stages of design will be reviewed and used in designing solution for the case.

3.4. Methods for data collection

Data collection is of the essence in order to propose a viable solution. Therefore, in this chapter the main data collection methods namely the literature review, interview with the industry and observations are presented and clarified in reflection to this study.

3.4.1. Literature review

The aim of literature review is to systematically gather information about the topic of the research as well as background information, such as what has been done in the area of data visualization in building industry.

Snowballing was the main technique conducting the literature review. By the use of this approach, papers which were referred in the work (backward), or papers which cites the examined work (forward) were scrutinized (Wohlin, 2014). Primarily, keywords related this study were defined. Then suitable titles were identified. Thereafter, as mentioned before, forward, and backward iterations are continued for related research.

3.4.2. Interviews

Interviews is another broadly used method for data collection. In reflection to this study interview with an industry professional is as well a part of the contextual inquiry in the contextual design explained in chapter "*3.2.1Contextual Inquiry*".

Semi-formal interview is conducted during this study in order to retrieve the latest information and practical inputs from the industry. As well it is a prerequisite for prototype development as the solution must be compliant with the user needs. Semi-formal interview approach was chosen in order to create an open environment leaving space for the interviewee input and considerations. Furthermore, to make sure the interview results would not be influenced by the interviewers, suggestive tone is avoided, and questions are prepared with an open end.

3.4.3. Observations

As this study aims at proposing a practical solution for industry professionals, a thorough understanding of the internal processes is crucial. Observations is a well-known method used to lead researchers to that direction.

However, as was mentioned before, there were some restrictions regarding the physical office observations. Therefore, on-line demonstrations of the user workflows where carried out, where the researchers were able to observe the common workflows. The reflections during the process were documented and analyzed as inputs for the contextual design process leading to creation of the first prototype.

3.5. Chapter summary

In this chapter two key methodological approaches namely Contextual Design and Five Spaces of Cognitive System were introduced and clarified in the context of this thesis. The main aim of using both methodologies is to gain better understanding of the current practices regarding valuable data utilization for energy engineering processes in AEC industry as well as to identify present user workflows and possible improvements. It is clear that the user here plays the main role and therefore, user-centered approaches are employed.

Contextual design was presented as an iterative process combined of eight main steps namely contextual inquiry, interpretation, work modelling, visioning, storyboarding, UI design, paper-mock up designing and finally interaction and visual design. It is important to understand, that all these processes are user driven and interconnected. With the help of contextual design, the collected user inputs can be analyzed and represented in a structured and unambiguous way. Regarding this study, the energy engineering consultant was identified as the main user and methods to collect inputs from the user were chosen to be in a form of interview and observation. Subsequently, five work models to visualize collected data were introduced, namely flow, cultural, sequence, and physical models. These models are deemed to give better understanding of internal processes, actor communication, environment, etc. leading to a more precise storyboards in the following step. The final three steps of contextual design were explored to be UI design, paper mock-ups and digital prototype suggestion. The value of paper mock-ups during the process were highly emphasized as it helps to prevent errors as well saving money and time.

In addition, Five Spaces of Cognitive Design were introduced as complimentary methodology for contextual design. The five identified spaces are namely information space, computing pace, representation space, interaction space and mental space. Similar, to contextual design steps these spaces as well are interconnected, and user centered. Furthermore, they closely relate to the contextual design processes. It was explained that for example, information space makes use of the first two steps of the contextual design being the contextual inquiry and interpretation.

Both of these methodologies will be further used in the study aiming to fully utilize the concept of use case with the cooperation with end-user. Together– contextual design and five spaces of cognitive system– processes will be considered as a part of proof of concept which aims is to successfully adopt both process for designing advanced visualization techniques for energy engineers.

4. Literature Review

Evolving Building Information Modelling (BIM) tools and technologies are progressively changing the way information in the AEC (Architecture, Engineering and Construction) sector is produced, stored, and exchanged between involved stakeholders (Howell & Rezgui, Beyond BIM: Knowledge management for a smarter built environment, 2018). The advancement of BIM should be meticulously planned within a paradigm while considering in people, processes, and these evolving technologies in a developing inter-connected world (Batty, 2018).

BIM is both a technology and a process evolving in the AEC sector (Eastman, Teicholz, Sacks, & Liston, 2011). It has been developed for embedding the buildings 3D computer aided design (CAD) model with additional data related to building specification, time schedule, cost estimation and maintenance management also defined as 4D, 5D and 6D BIM respectively to decrease costs by avoiding errors (Lee et al., 2015). BIM is also utilized as a platform for maintaining a precise and interoperable record of building information to improve planning construction and maintenance over the lifecycle of a facility (Volk et al., 2014).

BIM is being applied in AEC industry along with FM (Facility Management) functions for design visualization and consistency, clash detection, lean construction, and augmenting stakeholder interoperability (Volk et al., 2014). It is extensively documented that beyond the technological challenge, BIM introduces new ways of collaborating among AEC as well as FM value chains (Kubicki et al., 2019). The context of BIM, the 4D modelling process provides a virtual representation of an additional dimension (time), which means that all characteristics of the BIM process (3D models, cost, safety issues, scheduling) can now be represented, visualized and evaluated with a temporal perspective (Boje et al., 2020). Studies on BIM (Multi-dimensional BIM) modelling asserts that every one of these aspects is an additional dimension of BIM in their own methods (Ding et al., 2014). These dimensions generate new levels of complexity, with additional input data expected by each domain. This data is often derived from heterogeneous sources (simulations, sensors, building management systems, etc.), which should be correlated to existing BIM models on object levels, be consistent across project models and documentation, as well as with the temporal dimension (Boje et al., 2020). The pursuit for 3D BIM real-time visualization occurs due to the communication requirements between several actors in engineering, construction, and architectural companies(Dave et al., 2018). Integrating IoT deployment in the built environment and developing user interfaces for it continues to be a major challenge in the construction industry (Dave et al., 2018).

Therefore, in order to experience the full power of BIM and satisfy the ever-increasing user demands other technologies should be implemented supplementing BIM technologies. Further literature review is focused on identifying those technologies with reflection to the main problem question.

4.1. BIM & IoT

Simultaneous to developments within the construction sector, there is a growing challenge for a smarter built environment by more determined energy and carbon emissions programs across the world. The addition of IoT (Internet of Things) and AI (Artificial intelligence) is demanded to produce improved energy efficiency and reduce operation costs (Howell et al., 2017). The Internet of Things (IoTs) is merged within the communication network to turn into a specific extension and expansion application of the Internet. The physical environment is entirely perceived by installation of smart devices such as sensors, and the perceived information is pinpointed; then, the utilization of the transmission function and interactive function of the Internet is to aid the interconnection and dock the information through the characters and objects, to precisely actuate and manage the real physical environment (Mousavi et al., 2017). The addition of BIM in this context represents a minor building-level view within the broader environmental context. Although BIM has been utilized to incorporate lifecycle management of built assets, the current level of BIM is not compatible with IoT integration (Howell & Rezgui, Beyond BIM: Knowledge management for a smarter built environment, 2018).

BIM is deficit for its semantic completeness in subjects such as control systems, sensor networks integration, social systems, and urban artefacts beyond the scope of buildings, hence it requires a comprehensive and scalable semantic approach that can handle dynamic data at multiple levels (Boje et al., 2020). Sensor networks in buildings are growing in number and heterogeneity, inhabitants can be empowered to enhance the control of their environment for comfort maximization and energy minimization (Khan & Hornbæk, 2011). Since buildings are the prime consumers of energy and are the leading cause of greenhouse gases, applications that assist occupants to understand and control the interactions with a building could be tremendously useful to society. Although, the immense raw data sets that are collected must be aggregated and visualized to utilized which offers significant data handling, information visualization, and interaction challenges (Khan & Hornbæk, 2011). The use of the IoTs technology in the construction of intelligent buildings can widen the practicability of intelligent systems, help develop resource allocation strategies and improve the management and service capabilities of intelligent buildings, thereby enhancing people's quality of life (Kong & Ma, 2020).

4.2. Cyber physical systems approach

Whilst this efforts for increasing the use of virtual models, the problem of integrating the virtual models and the physical construction world such as to enable adaptive interaction or bi-directional coordination has not been satisfactorily addressed (Akanmu & Anumba, 2015). To achieve this bi-directional coordination, computational resources are required to securely integrate the virtual and physical domains such as the changes in one environment are autonomously mirrored to the other. This is defined as a cyber-physical systems (CPS) approach (Akanmu & Anumba, 2015). The CPS approach facilitates the integration of physical devices such as sensors and cyber that is digital or virtual information components to produce situation-integrated analytical systems that react intelligently to dynamic alterations of real-world scenarios (Tang, et al., 2010). Figure 3 depicts a visual diagram of bidirectional co-ordination between virtual models and physical construction.

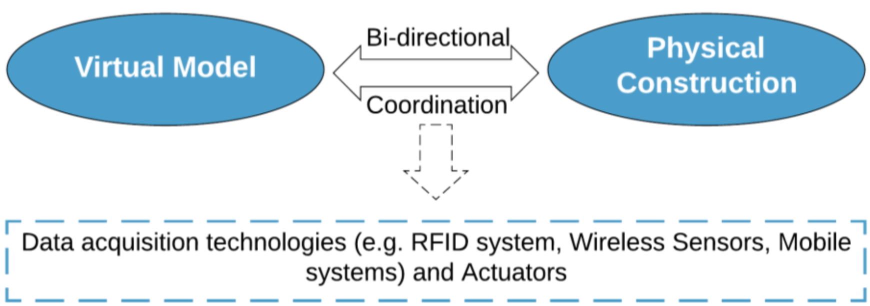


Figure 3: Bi- directional coordination between virtual modes and Physical construction. Adopted from (Akanmu et al., 2013)

The CPS approach is especially critical in the construction industry as it presents an opportunity for real-time seamless data flow between the design and construction teams, therefore giving opportunities for swift and informed decision making (Golparvar-Fard et al., 2011). To develop the CPS approach in the construction industry, there is a need to address the demand for the integration of sensor data and digital representations of the built environment for fostering stakeholder collaboration management within the fields of Real Estate 4.0 and Facility Management (FM), particularly in a spatial representation context (Stojanovic, et al., 2019).

4.3. BIM visualization & scalability

Another fundamental challenge in FM is visualization of historic and current sensor data within *as-is* built environment representations such as virtual models and plans of buildings, derived from a building's DT (Tao et al., 2019). Sensors calculating building-related operational data and man-made phenomena can provide information and insights about the current operational status of a building or a site (Stojanovic, et al., 2019). The process of visualizing BIMs continues to be an issue according to the veracity of construction project typologies. The visualization requires to be accurate, enhance understanding of the as real-world structure and specific to different user's needs, hence aid collaborative and coordination consultations, from a functional and ergonomic outlook (Kubicki et al., 2019).

One of the key factor in developing a systems approach is scalability (Duboc et al., 2013). Scalability can be defined as – *"The ability of a system to maintain the satisfaction of its quality goals to levels that are acceptable to its stakeholders when characteristics of the application domain and the system design vary over expected operational ranges."* (Duboc et al., 2007).

Present simulation processes are not completely incorporated in the design life cycle. Several scalability problems have occurred when integrating standard BIM tools and simulations (Lipman, n.d.). Hence an analysis of user needs, and perception and technical aspect is required to develop an ergonomic framework for enhancing the decisions of stakeholders through accurate data contextualization. In the subsequent chapters 4.5 & 4.6 we discuss the latest trends from selected scientific literature and projects. We divide it in two topics: 1. Data visualization and human perception which describe papers which are about the basics of visualization, i.e., graphical perception, design for human-information interaction. It would generate a basic understanding on how human/user perceive data, what is important in the presentation of data (color, pattern, data ranges, additional information). It would also include how human/user interact with data or – more precisely – the tools/media (application, desktop, tablet, etc.) and 2. Data visualization in terms of technical aspect here we describe the technical components and how they work together and their functions. Furthermore, (1) Gathering the data from sensors/BMS/IoT, (2) refining data, presenting data (3). In gathering, the aspect of historical, simulation, real-time data.

4.4. Data visualization and human perception

Decision making is connected how people obtain information, compare it and make a final selection (Jin et al., 2019). This phenomenon is present in any profession and every human faces it on daily basis – choice of daily products, choice between construction materials for better energy efficiency over the other while taking into account also the cost of solution. Before making any decision, people are faced with information which affects the choice to be made (Kleinmuntz & Schkade, 1993). Humans are not able to make absolutely rational decision (Simon, 1956). It is due to the time to make such decision, or limited information presented or obtained within given timeframe (Rachlin, 2003; Simon, 1956). The heuristics, so called “mental shortcuts” (Elstein, 1999), or “rule of thumb” helps to simplify the judgement. They aim to reduce the effort in decision making, information integration (Gigerenzer & Gaissmaier, 2010). It can also lead to reduction of effort around the task, activity (Shah & Oppenheimer, 2008).

There are a few influencing factors which leads to an outcome. One ties-up with cost-benefit concept, which derives from the concept of heuristics. (Kleinmuntz & Schkade, 1993) take two dimensions into consideration. First being cognitive effort to analyze information before making a choice. It is either measured by the time spent on finalizing the operation or a number of cognitive activities which has to be taken in order to perform the task. Those operations in the decision process are lined up in a sequence (Svenson, 1979). Iteration might appear where a decision maker evaluates information, due to either specific design of information display or extra quality assurance steps.

The way it is presented, displayed influences the final result made by decision maker and accuracy of such choice. Some decisions or analyzing processes can be prolonged due to unsuitable data representation (Svenson, 1979). As a result, it increases sequence of, for example, comparison of information (like presenting historical data of sensor record using table or pie chart rather than time series graph). Not only pure data itself influences the decision, but a particular way of representation (i.e. various kinds of tables, or graphs) can resonate different verdict on the same set of information (Jarvenpaa & Dickson, 1988; Kleinmuntz & Schkade, 1993). Moreover, Jarvenpaa & Dickson (1988) advocate for having framework regarding creating tables, and charts depending on the style and present their own guidelines to convey data findings as close to the reality as possible without distortion. Those deformities in presenting the information are an effect of manipulating vividness (showcasing wanted set of data), evaluability (ability to comprehend, evaluate and compare information), or framing (changing reference points, data ranges on arguments and values)

(Lurie & Mason, 2007). In the result they create biases, either by purposely manipulation or by badly designed visualization system which amplifies cognitive biases (C. J. Chang & Luo, 2019). Such biases are not necessarily derived from only the design, but also user's competences, and relying on intuition. Therefore, visualization type should not be only user's preferences, but rather predetermined thorough selection suited the given task as inexperienced user tend to pick realistic, more pleasant to eye display rather than efficient one (Hegarty, 2011).

Hegarty (2011) covers the subject of cognition in the design of visual representations. There are three types of visual display which encapsulate all possible categories. The first one is iconic which represents entity placed in real world (called, referent by (Hegarty, 2011; J. Zhang & Norman, 1994)) – its properties, shape, colors. Those kinds of displays try to show reality, e.g., 3d model of building, cross-section (which holds information unable to be obtained in real world), schematic connection of room in facility (where distance, true measurements are disregarded in favor of providing only necessary information – distorted iconic display). Relational display shows relationship between two factors, properties, which are not present or visible in the represented object, e.g., room temperature – surely, they exist and can be measured, but they are not obtainable on the first sight. However, the temperature data can be gathered and plotted onto the suitable charts. Finally, there can be a mixture between first two displays, which allows to allocate non-visible properties on the visual-spatial object (Hegarty, 2011). It can be temperature heatmap overlaying the rooms of the building, enabling more context and depth of cognition.

Visual representations can augment perception of data related to its referent. Foremost, they enable to allocate more resources in external representation rather than internal– the first is about knowledge about the environment, its constraints, properties and can be retrieved with perceptual processes, whereas internal representation is about schemas and meanings of objects and uses cognition to retrieve necessary information (J. Zhang, 1997). Together they create distributed representations (J. Zhang & Norman, 1994). Both, however, might stand independently, where they can change nature of activities, and they do not necessarily reflect one another (Luo, 2019). Freeing mental activities and bringing represented models closer to the mental one is also important for planning mental space for interaction (Kamran Sedig & Parsons, 2016; Tominski, 2015; J. Zhang, 1997).

Data and information are widely used in daily life and businesses – from the weather forecast to complex engineering calculations. Data sets grow everywhere in the world (Po et al., 2020)– new emails, accounts, sensor data, etc. No matter the type of information and the source of it, disciplines– there is always the need for visual represent the data (Kamran Sedig & Parsons, 2016).

In the cognitive field of research, there has been increased focus on visualization subject (Kamran Sedig & Parsons, 2016). One of the goals which interest anyone is to develop framework for the design process. However, most of the frameworks are made to solve particular problem with some visualization techniques already known to designer's mind – to modify existing ones, create derivatives (Kamran Sedig & Parsons, 2013), not to design visual representations systematically anew (Ainsworth, 2006).

Burkhard (2004) proposes the framework for knowledge visualization. He used "The Knowledge Visualization Cube". It is divided into knowledge, recipient, and visualization type. This three-dimensional framework allows to define the information type, match it with the end-user, and use certain visualization. Due to definite and predefined visualization types, this might be a suitable proposal for creating quick visual representations. However, the limitation might appear during the design of completely tailored cognitive system.

Javed & Elmqvist (2012) introduce "composite visualization views" framework. Their focus is to create a structure and environment for working with multiple views of information for different visual structures. There are four visual composition– juxtaposition, superimposition, overloading, and nesting– where each of them has its own characteristics. The paper proposes examples of visualizations which can be used in their framework, stating that it can be also any form of visual representation. This is a method which can be utilized in a later stage of design process when the types of visualizations are already established.

4.5. Data visualization in technical terms

The next four subchapters of this section of literature review is dedicated to exploring the latest scholar contributions exploring the concept of digital twin (DT), web design and object-oriented modeling, relevant APIs and their importance to the study as well as available knowledge bases.

4.5.1. The Digital Twin Paradigm

'The concept of a Digital Twin conveys a more holistic socio-technical and process-oriented characterization of the complex artefacts involved by leveraging the synchronicity of the cyber-physical bi-directional data flows' (Boje et al., 2020).

According to (Schluse et al., 2018) Digital twins are physical entities along with their data, functions, and communication capacities in the virtual world. As nodes within the internet of things, it enables networking and thus the automation of intricate value-added chains hence achieving complex control processes, advanced user interfaces, or mental models for intelligent systems.

The key concept of DTs is to produce a digital copy or virtual model for physical objects to simulate and mirror their state and actions by modeling and simulation analysis, additionally to predict and manage their future states and behaviors through feedback (Hochhalter et al., n.d.). The DT incorporates all elements, the complete business, and the process data for ensuring consistency (H. Zhang et al., 2019).

The control in CPS and DTs contains two components the physical objects or processes affecting digital depiction, digital representation and the control of physical objects or processes. For the controlling part, the physical system is dynamic, the same object would show unique properties at various times (Hu et al., 2012).

For it to stay consistent, real-time data from the physical objects is gathered with the help of deployed sensors and is transmitted to the digital system to propel the digital elements in synchronization with the physical objects. Particularly for a DT, physical real-time data propel the virtual models to simulate the physical processes and its development (Tao et al., 2018). Adhering towards the controlling part, the digital world utilizes this data to compute the control of the output and transmit it to the actuators for the physical implementation (Fawzi et al., 2014).

Within a construction sector perspective, the Digital Twin paradigm intends to develop the current construction processes and models (nD BIMs), along with their underlying semantics within the context of a cyber-physical synchronicity, through which the digital models are a mirrored reflection of the construction physical assets including their temporal dimension (Boje et al., 2020). According to (Tao et al., 2019) CPS and DT appear to be conceptually similar although DTs concentrate more on the virtual models and representation of physical assets for simulation analysis, so DT technology can be considered an essential foundation for constructing CPS. CPS consider sensors and actuators as the key modules, while DTs adhere to a model-based systems-engineering approach that accentuates data and models (Vachalek et al., 2017). Additionally, a DTs purpose is to create high-fidelity visual simulation, since a DT focuses majorly on creation of models, containing geometric shape, rule, behavior, and new constraint models (Tao et al., 2019).

The models provide an interaction and recording process to assist and understand the performances of machines or systems and to predict the future state based on real-time data, historical data, experience, and knowledge, as well as on data from models. Thus, models and data can be deemed as the main components of a DT (Tao et al., 2019).

The main DT components are:

1. The physical components
- Sense
- Monitor
- Actuate
2. The virtual models
- Simulate
- Optimize
- Predict
3. The data
- BIM integration and scalability
- IoT
- Knowledge storing and querying.

Contained by the interdependent relationship amongst the physical and virtual parts of the DT, BIM is often considered as a DT sub-component (Boje et al., 2020). BIM is now part of the initial procurement and the design-construction-demolition stages, the prominence of the CDT ought to be at the pre-construction and construction stages while the "Physical Twin" gets built. BIM processes and model should be able support and facilitate improved collaboration (Boje et al., 2020). Furthermore, users from different social and educational settings should be able to interface with the CDT, which varies corresponding to application domains over the lifecycle of an asset (Schluse et al., 2018).
(Boje et al., 2020) proposes a progressive evolution approach and describes the multiple domains, tiers, and uses of a CDT. They consider the execution efforts for a CDT to be gradual, although constant over the building's lifecycle, whilst considering the supply chain

integration and the complexity of technologies implemented. A web-based integration of IoT is demanded across the board. The eventual merge between virtual models and sensing would converge on a common semantic web platform (Boje et al., 2020). Figure 4 describes the 3-tier generation evolution and functions of the Construction Digital Twin as analyzed by (Boje et al., 2020).

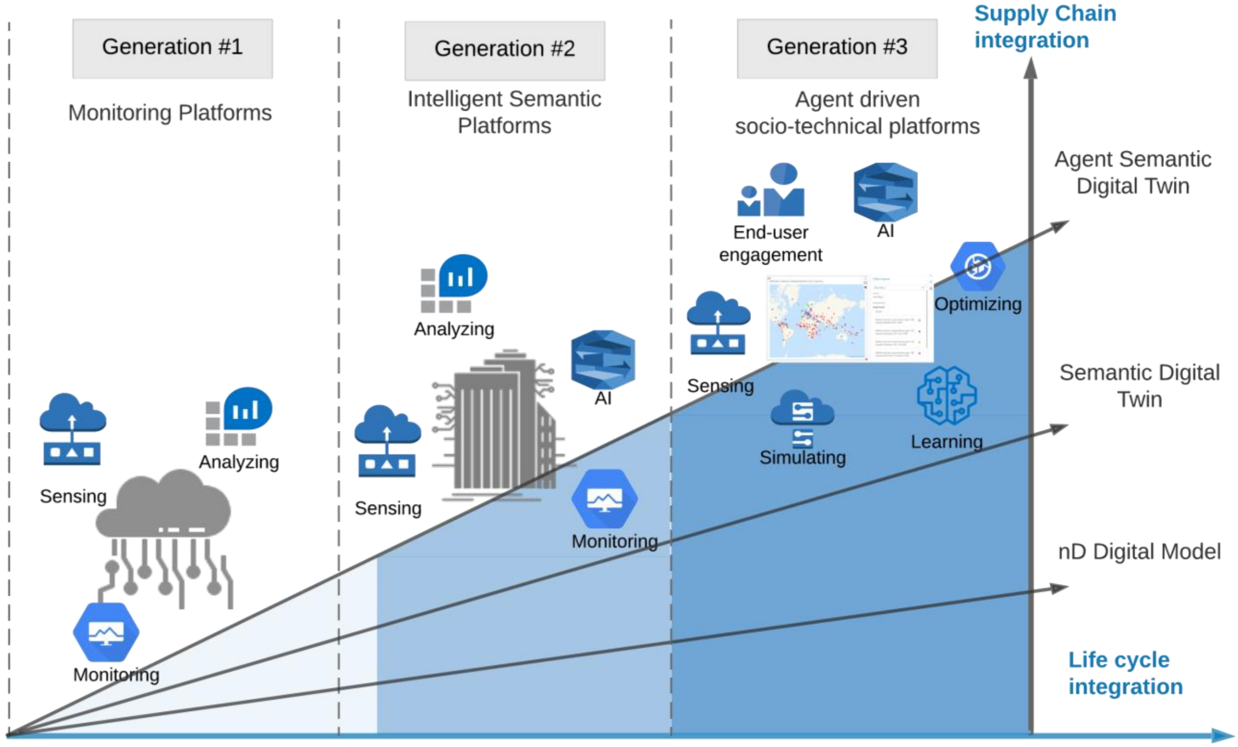


Figure 4: Tier generation evolution and functions of the Construction digital twin. Adopted from (Boje et al., 2020)

This envisions the DT in a convoluted socio-technical dilemma, the DT needs to acclimate and respond in real time to its users and dynamic changes which occur regularly. Additionally, to the semantic layer, the social aspects of the building needs to be addressed. Providing a user-driven experience is necessary, the DT should be able to deliver to social requirements and engage with end-users to assist holistic decision-making (Boje et al., 2020).

According to (Moyne et al., 2020) DT technology has been executed in numerous ways in the industry. Often the technology is not described to as "digital twin", and perhaps, it has occurred prior to the phrase "digital twin" was devised. A common factor amongst the most applied DT solutions is the utilization of a broad variety of models and analytics. These are chosen based on the purpose that drives the creation of the DTs, The modeling methods for DTs in and out of manufacturing, even though varying, can be believed as delivering intelligence so that each DT can perform its planned functions in a satisfactory manner in a specific environment over an adequate time period (Moyne et al., 2020).

A second perception is that the domain of applicability should be defined for a DT to be efficient. This can be done with detailing out the contexts and relations as ‘‘equipment ‘X’, process ‘Y’, for product ‘Z’’’ so the applicability of the result is constrained to the domain specified in the context definition. This influence both the human and cyber intelligence in the context-defined area (Moyne et al., 2020).

A third perception is that the efficient DT implementation should have a mechanism to retain a degree of intelligence in the solutions so the required value can persist. This typically implies that a method of DT maintenance should be a component of the solution, and this process must assist an revise of the DT intelligence to comprehend and adjust to alterations in the application environment (Iskandar & Moyne, 2016).

The life cycle of a DT has two phases to categorize it – The off-line development and the on-line deployment and maintenance. At the off-line development phase, data, analytics, and SME (Subject Matter Expert) are iteratively brought collectively to envisage, design, improve, authenticate, and validate DT solutions before deploying on-line (Moyne et al., 2020).

The off-line data also called as “Data at rest” is usually historical data, which is sourced alongside with SME and analytics to better understand the application environment for e.g., data quality, ability to merge data, and level of supervision or the degree to which ‘‘output’’ data is available to relate to input data of the data set ,to determine the feasibility of building a DT that can offer benefit, develop initial conceptual solutions, and quantitatively verify and validate these solutions (Moyne et al., 2020).

At the on-line deployment and maintenance phase, the evaluated off-line solution is implemented, utilized, continuously evaluated, and maintained. The deployment includes the integration of the DT capability into the existing system, including DT data, interfaces, services, and behavior, so that the DT capability is effectively used by DT clients. Once deployed, the DT uses select run-time data from its operation environment also known as data in motion to assess the state or condition of aspects of the environment and make recommendations, thereby fulfilling its intended purpose (Moyne et al., 2020). Figure 5 describes a high-level view of a DT lifecycle.

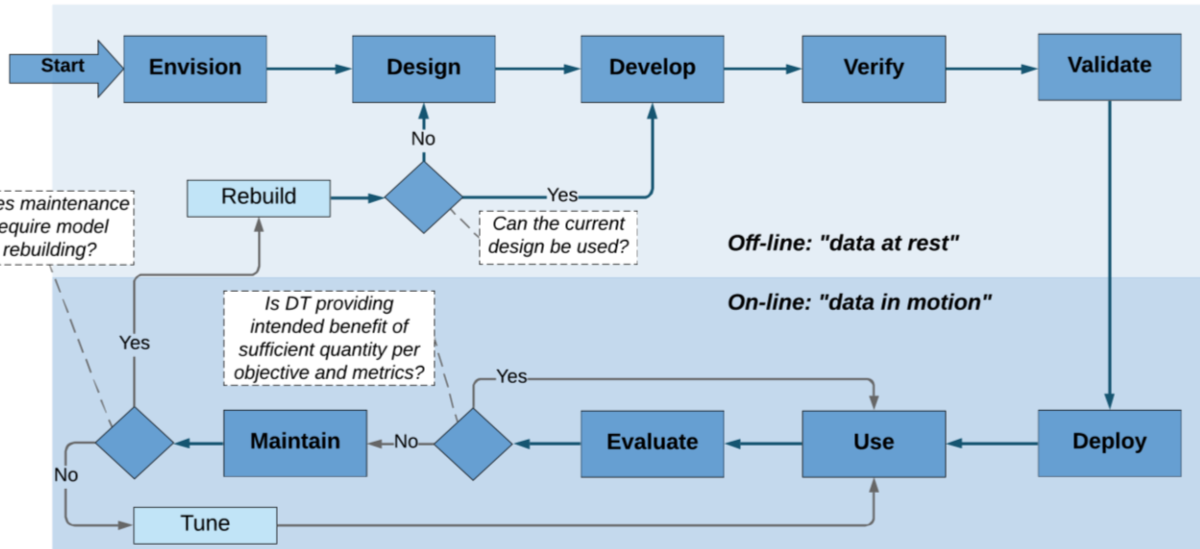


Figure 5: Process of a DT lifecycle. Adopted from (Moyne et al., 2020)

According to (El Jazzar, 2020) to utilize the Digital Twin to its full ability, the conceptualization and initialization should be at the beginning of the building design phase. The data collection must begin at the design phase utilizing a BIM model. Data must then be constantly updated and collected throughout the project lifecycle to achieve an operation and maintenance phase; the model combines data from different sensors. The data is collected and analyzed through cloud-computing. The virtual representation is then updated in real-time with the necessary data and predicts the performance of the physical facility. *This functionality provides the owner, the facility manager, and the operator of the facility the capability to make informed decisions. The bidirectional communication between the physical and virtual facility also enables proactive maintenance* (El Jazzar, 2020). Additionally, benefit of employing this concept is to enhance the next generations of construction projects utilizing the knowledge encapsulated in the Digital Twin. Figure 6 shows the required framework for construction digital twin implementation (El Jazzar, 2020).

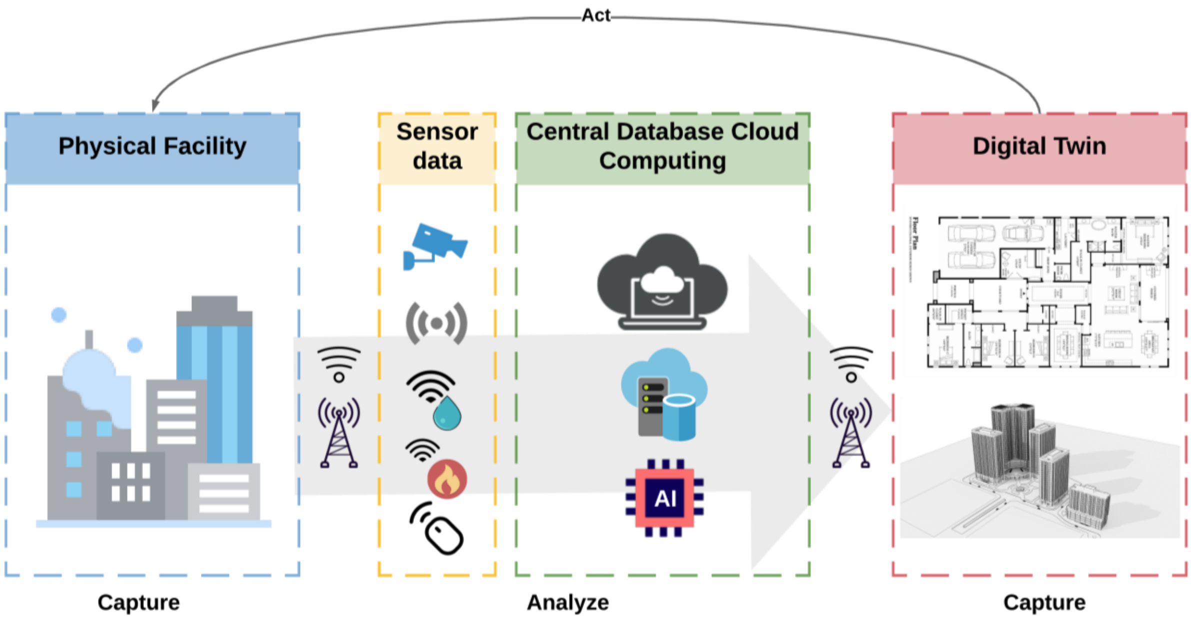


Figure 6: Framework for implementation of Digital twin in construction. Adopted from (El Jazzar, 2020)

(Zheng et al., 2019) proposes an application framework of DT for product lifecycle management which comprises of three spaces, physical space, virtual space, and information-processing layer as shown in Figure 7. In the application process, the DT technology can achieve complete-physical system mapping, life-cycle dynamic modeling and the entire process real-time optimization. The bidirectional mapping and interoperability of physical space and virtual space are attained due to data interaction. Intelligent decisions are based upon iterative optimization and regulatory interaction in between the two spaces.

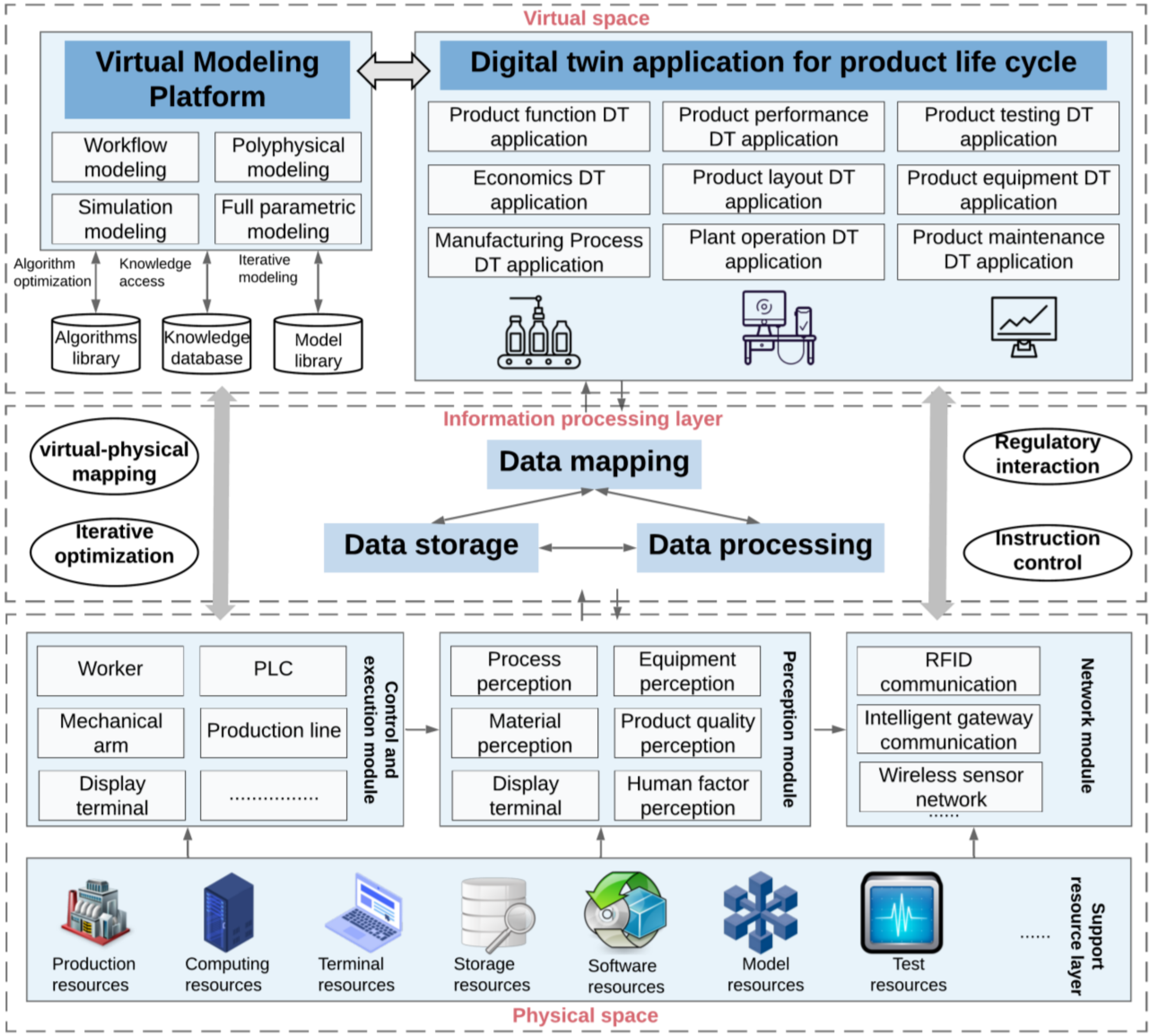


Figure 7: Application Framework of Digital twin technology. Adopted from (Zheng et al., 2019)

The bidirectional mapping of the physical part and virtual part are connected through the information processing layer and is accumulation of the data interaction in this layer. This layer has three core function modules, data storage, data processing and data mapping (Zheng et al., 2019). The data storage consists of further two divisions for data captured from the physical space and the virtual space. Data processing of this layer has further four levels data acquisition, data preprocessing, data analysis and mining, plus data fusion. Data source includes the different databases of different manufacturing systems. The raw data is gathered by an ODBC (Open Database Connectivity) interface and additional technologies. Afterwards the set of raw data are preprocessed, and the procedure mostly comprises of the rule-based data cleaning, data structuring, and primary clustering (Zhang J, 2016).

In the virtual part, the modeling of physical objects is developed by acquiring the attributes of the virtual model from the database, and the response of 3D models would be stored in the database with utilizing subsequent interfaces. In the Meantime, the DT, can not only enhance comprehension to the visualization of products, but likewise will realize the enhancement of simulation for complex systems (H. Zhang et al., 2019).

(Tao et al., 2018) proposes a new product design process for creating digital twins, not particularly for the construction sector but from recognizing the issues of implementing it from the PLM (Product life cycle management) and manufacturing enterprises. The virtual product would mirror the entire lifecycle process of the related physical product. Based on digital twin, the product design process can be split in three phases conceptual design, detailed design, and virtual verification as shown in Figure 8.

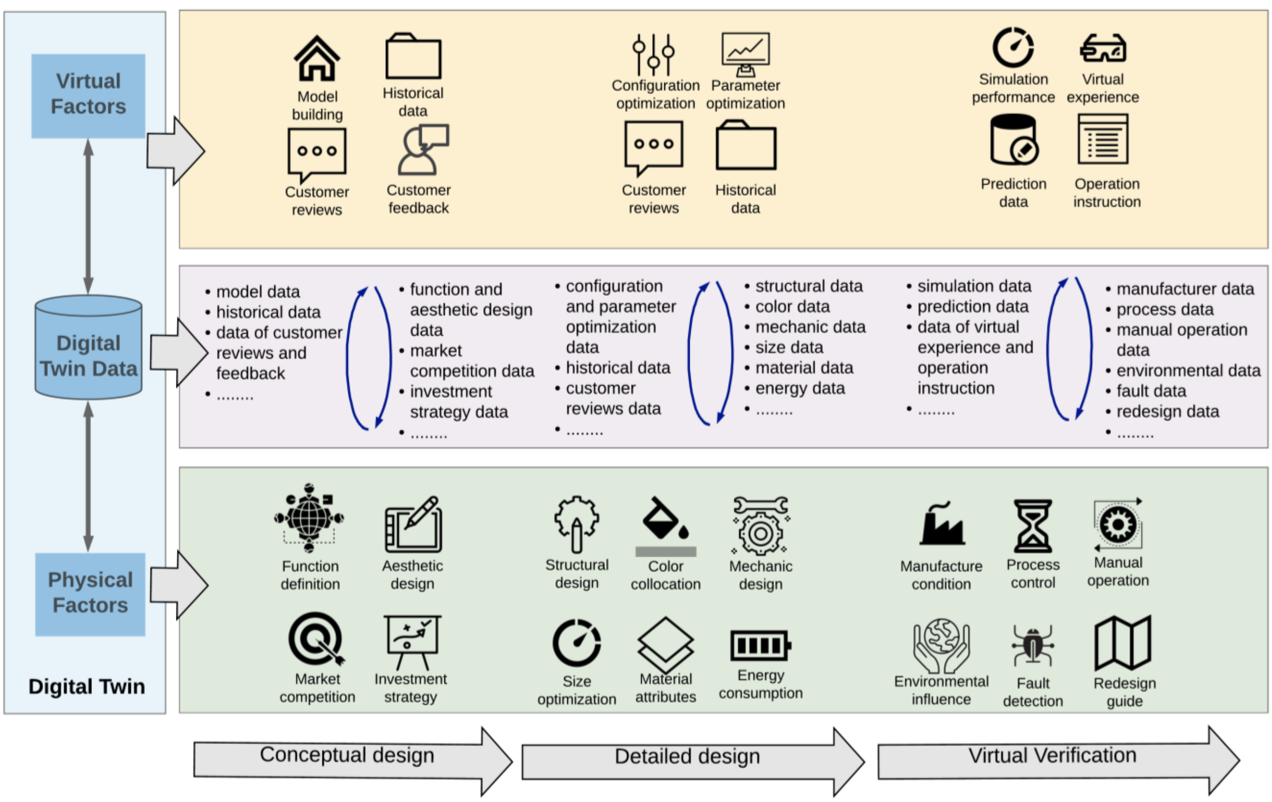


Figure 8: Digital twin-based product design. Adopted from (Tao et al., 2018)

Each phase describes the factors of components and the data processed for a fully integrated digital twin product.

Present literature on the topic of Digital Twin in construction is not easily found because the term “Digital Twin” is not explicitly mentioned in most papers and, is occasionally referred to as BIM or BIM-FM (facility management).

High-level-performance processing of time-series data is the crucial for the effective implementation of Digital Twins. Due to the following two reasons that data collected from the physical world is largely discrete time data, and there is precise time requirement for a Digital Twin (Lu et al., 2020).Data-focused computation, constant streams and timestamped data are the leading problems that have occurred while dealing with streaming data, such as through sensor data acquisition in monitoring requirements (Mehmood et al., 2017).

4.5.2. Web design and object-oriented modelling

Web browsers utilize JavaScript as the leading language to develop web applications and programming complex applications for web (Jensen et al., 2009).The world wide web utilizes HTML to describe web pages. It uses markup tags to define structural semantics of a website by denoting its elements such as headings ,tables, paragraphs and others.HTML also enables the inclusion of external resources into the web pages like videos, pictures and other objects, while the final presentation and style lies in the domain of CSS (cascading style sheets) (Jakus et al., 2010). Moreover, apart from CSS, a JavaScript scripting language is utilized together with HTML. JavaScript is interpreted by a web browser and delivers web pages with interconnections and dynamics. The JavaScript code interacts with the DOM (Document Object Model) through the different API (Application Programming Interface) libraries based on various methods of user-triggered events(Jakus et al., 2010).

“JavaScript is an object-based language it utilizes prototype objects to model inheritance. As virtually all predefined operations are accessed via prototype objects. Objects are mappings from strings (property names) to values. In general, properties can be added and removed during execution and property names may be dynamically computed.”(Jensen et al., 2009).

Graphic rendering on the web pages has only been possible with the aid of plug-ins, such as Flash or Silverlight (Jakus et al., 2010). With the HTML5, Graphical rendering of web pages was earlier made through help of plug ins like Silverlight or flash, with the new version of HTML i.e. HTML5 the functionality required for graphics rendering is executed in browsers in the type of SVG (Scalable Vector Graphics)(Jakus et al., 2010). The graphical elements are entirely incorporated into HTML and a component of the document object model (DOM). Their style can be marked through CSS and can be controlled through the JavaScript(Jakus et al., 2010).The SVG (Scalable Vector Graphics) standard makes it possible to represent complex graphics by a compilation of graphic vector-built primitives, extending the opportunities of modern web designing also including the traditional raster images such as: scalability, resolution independence, etc.(Battiato et al., 2005).

(Garett et al., 2016) discusses the major elements that play a major role in website designing oriented towards the users specifically 1. graphical representation, 2. organization ,3. content utility 4. purpose 5. simplicity ,and 6. Readability,7. Navigation (Garett et al., 2016) advises to consider these 7 elements along with the unique user needs when developing user engagement strategies for designing web applications.

4.5.3. APIs

The Web was originally developed as a client- side server approach for human-oriented hypertext documents(The Original Proposal of the WWW, HTMLized, n.d.) Previously, Web application interconnections was limited to following links and translating documents. Afterwards, forms were integrated into HTML and https was broadened to incorporate the capability to send out data and request it. Successively, the W3C developed XML and the ability of using Web technologies as a program for distributed systems, became popular (Kopecký et al., 2014). The initial Web just had one kind of client-side program - the browser. Few systems tried to upgrade to develop a applications based client which were planned for desktop tools, and for the integration within additional existing systems (Kopecký et al., 2014).

Gradually the browser support for JavaScript and XML Http Request developed into Ajax(Garrett, 2005). This shift is stated to as web 2.0, (Kopecký et al., 2014). As shown in Figure 9.

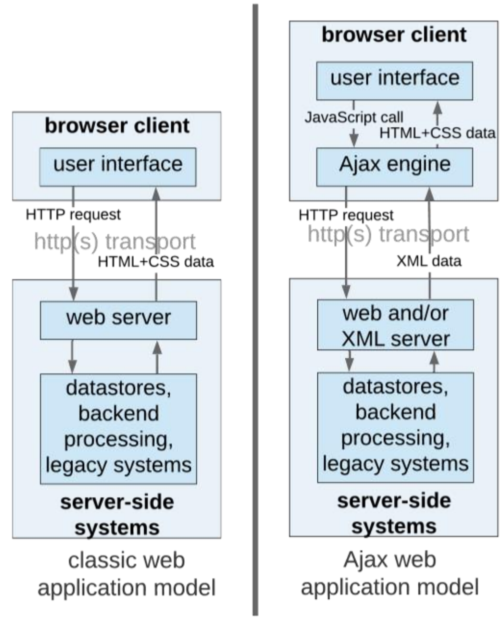


Figure 9: Earlier model of web application with the new (left) Ajax model referred as web 2.0. Adopted from (Garrett, 2005)

A Web 2.0 websites execute interconnectivity in their web pages, the JavaScript code in the web pages turn into a program that requires an API to access the server (Kopecký et al., 2014). The introduction of managed web-API s has gained popularity. The foremost feature of the Web APIs is that they require an API key from the client that should be received at every request (Kopecký et al., 2014). API keys are usually utilized to authenticate the user and recognize the client system by enabling several managerial functions (Kopecký et al., 2014).

The amount of services offering openly available web application programming interfaces has been increasing precipitously (Maleshkova et al., 2010). Numerous studies provide evidence that developers are increasingly keen towards deploying Representational State Transfer (REST) web services, as the method for consumers to utilize their services(Maleshkova et al., 2010), (Kopecký et al., 2014). REST interfaces entirely depend on Uniform Resource Identifiers (URI) for resource recognition and interaction, and generally on the Hypertext Transfer Protocol (HTTP) for message transferal(Fielding, 2000), (Costa & Pires, 2020).

The web can be called as the biggest information system made in the world. Moreover, it is the greatest storage of open data which is available for exploring innovative works and for machine processing, automation and generating visualizations through combination of multiple data sources(Kopecký et al., 2014).

4.5.4. Knowledge bases

The effectiveness of the DT is based on the capability and viability to retrieve data and the semantics accurately, and make the accurate data sets available for processing (Boje et al., 2020). Also (Qi & Tao, 2018) consider the databases, data and data processing storage crucial to retrieve knowledge as a part of further AI capabilities, permitting them to learn and take decisions.

There are multiple types of data storage options designed for various categories of data formats and application requirements, broadly, Relational databases and Non-relational databases. Relational databases are an appropriate option for programs that include managing of compound database transactions and substantial data analysis, because of the integrity to referencing or relations (Lu et al., 2020).

Non-relational databases are applied where managing larger sets of diverse and regularly updating data, generated from distributed systems. They do not have to be structured according to the rigid schemas associated with relational databases (Lu et al., 2020). In relation to smart manufacturing, key-value databases are being used to store software log files from the connected DTs. Likewise, documents are collected in document databases to enable high flexibility and stability. Knowledge can be structure and stored in graph databases to quicken the querying and reasoning processes. Practically, these databases will require to complement to the whole framework to ensure minimal latency (Lu et al., 2020).

Digital Twin appliances including real-time monitoring, prediction, require precise time data processing to be precise and reduce latency. Therefore, the data processing architecture should prioritize design for latency (Weber et al., 2017). Parallel computing technologies ensure less latency for processing data. The benefit of parallel computing is to segregate the tasks into batches of similar sub processes and combine them later when it is completed (Lin, 2013). Tools such as Apache spark, Apache Flink, Kafka, Google data flow, Apache storm are some distributed real time processing platforms (Lu et al., 2020).

4.6. Summary of the Literature Review

The literature review presented in this chapter assisted in identifying the latest contributions and revealing the current research gaps related to this study. First of all, the growing potential of BIM technology was presented and the prospects of integrating it with facility management explored. However, it was found that with expanding possibilities the amount of information increases, bringing higher levels of complexity. In order to successfully integrate the data to FM processes it is required to have all relevant data consistent in one BIM model, thought at the moment the data is scattered across multiple sources. Furthermore, the need for real-time 3D visualization was expressed in order to support the communication requirements between the actors in different industry domains.

Internet of Things (IoT) was found to be one of the biggest opportunities and challenges for the AEC industry. Research showed that the integration of IoT and BIM could assist in determining energy consumption and carbon emissions in buildings. Furthermore, sensor data can be used to get a better understanding about energy consumption by aggregating, visualizing, and interpreting the data. However, BIM is currently not compatible with IoT integration and has limitations capturing information related to control systems, sensor network integration, social systems, etc., therefore cannot handle the dynamic data.

Further, the cyber physical systems approach was explored. It was revealed that these systems can bring important benefits for faster decision making between site and design teams. Nevertheless, there is still a big issue while integrating the physical world with the virtual model more specifically the bi-directional connection. It is argued that computational resources could assist to solve this problem.

BIM visualization and scalability are another exceedingly referred challenge. It was determined that BIM visualization still is highly limited due to the veracity of the construction projects typologies. Simulations was found as a useful approach while predicting building's energy needs. However, the integration of simulations with BIM raise scalability issues. Therefore, the need for an ergonomic framework for enhancing the decisions of the stakeholders through accurate data contextualization was expressed.

Human perception was identified as an important factor for data visualization as the information provided to the people affects their decisions. Therefore, it came to the attention that visualizations can enhance the decision-making process. Even more the way how the data is presented and visualized was found of the essence. However, some pitfalls were identified as well related to the time and scarcity of the data provided as well as the prospect of bias due to poor visualizations or data interpretation.

The Digital Twin Paradigm was introduced and explored. It was found that using CPS the real-time data from the physical objects can be transmitted to the digital system to propel the digital elements in synchronization with the physical objects or processes. DT technology could be considered as a foundation for constructing CPS. During the literature review, DT were introduced in the light of different perceptions and reflected in the life cycle of the concept. However, there were found some live-data limitations like Data-focused computation, constant streams, and timestamped data.

Finally, the web design and object-oriented modeling, APIs, and knowledge bases were reviewed. These concepts are vital to develop a digital platform for visualizations on the web. Knowledge bases were revealed to be the key element for the effectiveness of DT as it relies on data retrieval. Subsequently, different types of data storage were reviewed, however it was found that for real-time monitoring and prediction reflective platform is vital.

5. Pathway to prototype development

In this chapter previously suggested methods are applied in the context of this study in order to collect the necessary data, analyze it and propose a user-centered solution. Energy engineer is identified as the main user and is the key process driver. Therefore, a dedicated interview questionnaire is created combined from different sets of questions retrieved from the relevant publications and insights gained from the literature review. The question mapping process is introduced and supplemented by a snippet from the interview summary fully represented in Appendix I. The collected data is interpreted and analyzed using the thematic analysis method. Identified results are further visualized in multiple work models leading to the visioning stage where connections between actors and processes are created. Storyboards are drawn in order to reveal the concrete functions of the developed application which further are reflected in the UI design. Finally, layout spaces for the first prototype are drawn.

5.1. Collecting the data

The contextual inquiry consists of three main steps– user identification, interview, and interpretation session. The foremost step of contextual design in this project is to identify the end-user. It is followed by conducting interview with the user with forecreated questions. Finally, in interpretation session the team evaluates answers, group and categorize them for the utilization in later processes.

5.1.1. User Identification

As mentioned in *Methodology, Contextual Design* chapter, there are different end-users depending on the role in implementation phase. In this situation, this is an energy engineer. The user, his challenges, objectives, and desires were briefly introduced in the preliminary e-mail exchange for setting-up the cooperation. That is how the user was identified.

5.1.2. Contextual Interview

Later, the interview with the user was developed. This part was supported by the first stage of the design process from *Methodology, Five Spaces of Cognitive System*. The intention was to base the systematic approach on the high-level concepts developed by Sedig & Parsons (2016). They include definition-specific questions which are helpful to stay on the rigorous track of design. Moreover, some of them were found to be too abstract which would require knowledge of the meaning of some concepts, known mostly to the people who are interested in human-information interaction field. Therefore, the aim was to maintain original questions and unfold each to understandable, more specific to user's profession and current type of work. Some of the questions did not come from the initial ones, but they were necessary to include in the contextual inquiry. As a result, there is a mixture of standalone and derived inquiries.

The overview of the process of preparation to the interview is shown in the Figure 10. It captures the transition of questions from high-level concepts to low-level. It was done in a way to make the entire interview understandable and friendly. The interviewee should concentrate solely on answers rather than trying to decipher the meaning behind the query and to learn new definitions. Initially, 20 questions were created, from which 10 were own invention. The second half were adopted from Sedig & Parsons (2016) with numbers of supporting question to each topic. By doing so it was easier to transition from general, definition-specific queries to more suitable, and accessible for the use case. Some of the supplementary explanations were still a little bit too complex. Therefore, they required additional simplification, or they were omitted from the interview, but they were still constructive in the interview preparation and interpretation. Creating map and connection between aforementioned inquiry sets made it easier to navigate through interpretation session, and remaining stages of the contextual design such as developing own vision and solution for the use case.

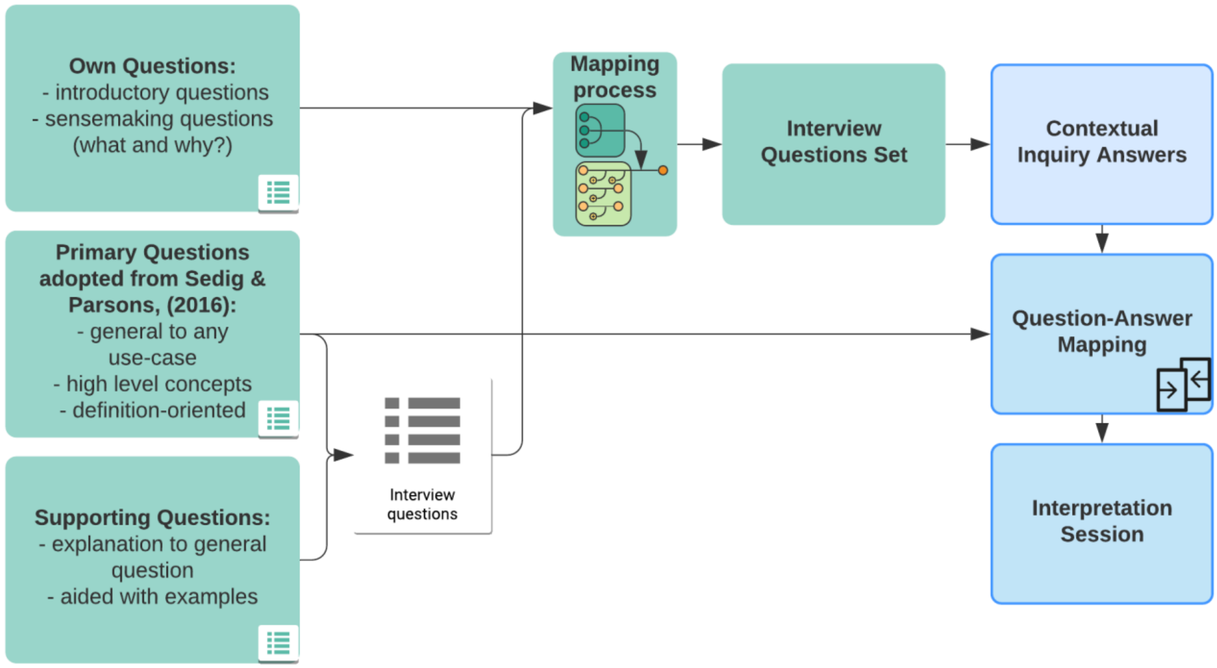


Figure 10: Contextual Inquiry Process.

The list of "Primary questions" was adopted from Kamran Sedig & Parsons (2016). It goes as follows:

1. What is the information source?
2. Does the information have open or closed space?
3. Does the information space acquire new information?
4. Are there any sub-spaces in information space?
5. Are there many levels in information space?
6. If so, how deep is information space, how many levels?
7. Is it heterogeneous or homogeneous space?
8. Do sub-spaces consist of any relations between each other?
9. What characteristics do sub-spaces have?
10. What is the system architecture and how elements are arranged in the network?
11. To what extent granularity should be applied for encoding the components of the system?

Another set of question is the final one which was presented to the interviewee and to which answers were inquired. This is set is called “Own questions” and goes as follows:

1. What is your professional background and what is your role at MOE? Can you tell us about the main goals of MOE and the department you are a part of?
2. Can you briefly tell us about the current work model and the participants working in your group and their competencies/roles?
3. Who are the interested stakeholders or participants who need to see BIM, sensor, and simulation data visualizations? What type of meetings would you define these as? Who are the participants?
4. How are you currently representing data from the readings for the meetings you mention in the brief apart from Graphs and Bar charts? How do you show the 3d model (E.g., In Revit,3d images etc.)? How many different parameters do you use in the visualization (e.g., not only value like temperature but room id, sensor id, etc.)?
5. Why are you using this type of visualization? What are the pros of using that kind of visual representations? What are the shortcomings with showing the data in form of tables and graphs?
6. Do you have fixed properties of visualization representation you just mentioned (certain way of expressing with bar, pie charts; with different sizes, ranges)? If not, how do you decide which way of representation is most suitable for you? Do you have some schema?
7. What exactly is the purpose of the visualization? What needs to be communicated through these visualizations? What is usually the later stage after data representation in the meetings that you mention in the brief?
8. What instruments or physical objects do you use to acquire data?
9. Which type of Simulations is the data generated from and from what software tools?
10. Is there constant flow of information from sensors that affects simulations, predictions, assumptions, analysis or are there almost fixed values (e.g., temperature for room type) used for analysis, calculations, simulations? Does your system readjust values of factors, or some constants needed for calculations, based on acquired data from sensors? Do you make that kind of calibrations?
11. How is the different type of sensor/simulation data (Co2, Humidity etc.) currently been compared? Can you dynamically compare/blend sensor/simulation data (e.g.,

Information about temperature blended with humidity etc.) with adjusting range, areas, visual representation (e.g., from 2D to 3D)? What is something you want to have to include in your current visualization process?

12. Is live data (Real time data) required? Or just visualization from historical data and simulation data? What do you think would be the benefit of contextualizing this data with the virtual building model rather than just representing it through bar charts?
13. Along with specific unit of measurement from sensors, what more information do you receive (e.g., date, file format, type of document)? How do you catalogue and store information? Do you archive them?
14. On what level the building is the data recorded by the sensors – are sensors only placed in main building area or in all floors, then rooms, and further down to each workspace, cubicle, unit? Do you have any tree-structure of information or is it more linear, so everything has sort of the same priority?
15. If so, how deep can you "zoom in" to information – going from overall building, main parts, particular floors, rooms, workspaces – that you will get no more new information?
16. What is the information you wish you would blend/mix, but it is not possible now (e.g., lack of settings, lack of data) and what is the reason you are not doing so?
17. Do you have established information structure that you use in every project or does this vary from case to case? If so, what is the structure and how you document that?
18. Are there any established connections (relationship) between the type of data you mentioned before? For example, once you gathered data do you use them in combination – e.g., putting temperature in the context of CO_2, amount of people? If so, do you make this that kind of linking for every project or is vary from case to case?
19. How detailed (e.g., data from every room, every area, and how precise, like 2 digits after comma, etc.) or frequent (e.g., once/two times a day) data should be acquired to make your work efficient? How is it now in your current system?
20. Are you familiar with overall/partial technical architecture of the current system – from data gathering to visual representation of data and interaction with it? If so, can you briefly introduce us to it? What changes would you want to see in current architecture?

Combination of the “Primary questions” with “Own questions” is listed below. Furthermore, Table 1 represents the results of mapping the questions and some of the interviewee provided answers. All four columns (“Primary questions”, “Supporting questions”, “Interview questions”, “Answers”) are linked by rows – information in each row correspond to each other. Numbering next to “Interview questions” and “Answers” are related to the numbering of list of questions called “Own questions” above. Full Table 1 is presented in Appendix I.

Table 1: Primary and Own questions mapping supplemented by some answers

Primary questions	Supporting questions	Interview questions	Answers
1. What is the information source?	• What are the information types are acquired (temperature, CO_2 distribution, etc.)? • What instruments do you use to acquire data?	(8.) What instruments or physical objects do you use to acquire data? *Confirmation question: What are the information types acquired?* - information received in CSV file before the interview.	(8) "For this case of healthy homes, we use IC meter and Netatmo – It measures, Co2 humidity, noise, temperature. It has a ppm showing the values if the co2 levels are acceptable and it notifies the user. This is for the building occupants to monitor their room air quality. Facility manager has the data and we have request them to make analysis. They are updating the data set every night and providing it in a csv format. We need to send a HTML request for accessing the data from the sensor API."
2. Does the information have open or closed space?	• Is there constant flow of information from sensors that affects simulations, predictions, assumptions, analysis or are there almost fixed values (e.g., temperature for room type) used for analysis, calculations, simulations? • Does your system readjust values of factors, or some constants needed for calculations, based on acquired	(10.) Is there constant flow of information from sensors that affects simulations, predictions, assumptions, analysis or are there almost fixed values (e.g., temperature for room type) used for analysis, calculations, simulations? Does your system readjust values of factors, or some constants needed for calculations, based on acquired	(10) "We do not have a link between measured data to simulation data. When performing simulations, we do it from the industry guidelines. We do not know how the end users are going to use the building, what their preferred set points are going to be how their occupancy patterns are going to be. So, with guidelines we try to standardize that. But usually, it's very different from reality and there is a performance gap in simulations and reality, and it has been proven many times that simulations don't predict the real scenarios. We must be aware about comparing different data because occupants

	data from sensors? Do you make that kind of calibrations?	data from sensors? Do you make that kind of calibrations?	can change over time and we would have different patterns. We see that in research, people trying real time optimization for fault detection, but we haven't seen building owners demanding that or willing to pay for that. But we are considering using it because we think it will be beneficial for us as building energy engineers."
3. Does the information space acquire new information?	• Is there a constant batch of information revived from data sources (e.g., sensors)? If so, how often? • Do sensors are a subject of periodic calibration/check? If so, why and how often? • Are the sensors being relocated from time to time? If so, why and how do you make assessment on new location? • How do you check quality of data reads? How do you check the quality, accuracy before it's settled/approved to run?	(12.) Is live data (Real time data) required? Or just visualization from historical data and simulation data? What do you think would be the benefit of contextualizing this data with the virtual building model rather than just representing it through bar charts? *Confirmation question: Is there a constant batch of information revived from data sources (e.g., sensors)? If so, how often?* – information received in CSV file before the interview.	(12) "So, if we are making an analysis, we don't need the real time. But if we are trying something like the Netatmo application for the building end users to monitor then we need the real time data."

Table 1 represents first set of adopted questions with supporting ones. The left column represents starting ones, the right– modified. Although, some of the questions could have been answered with what was presented in the overview, it was decided that for the sake of art of systematic design that the questions would be added to the contextual inquiry as they not only would maintain the coherence of the inquiry, but also could reveal some aspects not presented in the early description. Some other questions were indirect answered with the BIM model and sensor data shared by the engineer. In order not to prolong the interview, it was decided to shorten them by only confirming some initial questions with the interviewee, but not omitting them.

5.2. Interpretation

The method for interpretation of the interview is inspired from(Gale et al., 2013) which aided us to understand the knowledge attained and use it for further stages of the contextual design processes.

5.2.1. Thematic analysis

STAGE 1: TRANSRIPTION

The detailed transcription of the interview can be found at Appendix I.

STAGE:2 FAMILIARISATION WITH THE INTERVIEW

A video of the whole interview was recorded with the help of screen recording feature. The video was reviewed to help familiarize with the interview and extract codes to further categorize them into themes by the writers of this project.

STAGE 3: CODING

Labels were added and highlighted, following we present the codes identified from the interview:

1. Sustainability, Indoor environment, healthy buildings.
2. Guiding decision makers.
3. Efficient design and code compliance with DGNB.
4. Exploring new ways to analyze sensor data.
5. Evolve in the Facility management area.
6. Spatial recognition of a building for explaining to FM managers, Building owners and HVAC engineers.
7. Finding reasons for discrepancies.

8. Hard time to measure and analyze sensor data.
9. Trying to visualize a lot of sensor data but it generates lot of problems.
10. Different users need different visualizations to understand sensor data.
11. Occupancies change of building units, so patterns change.
12. Simulations cannot predict, very different from reality.
13. Analysis of multiple factors to diagnose the causes in design or systems.
14. No geometric context of the building creates problems.
15. Facility managers require to find spatial elements of buildings faster.
16. HVAC consultants need to find reasons for discrepancies.
17. Mapping occupancy patterns is important.
18. We use Excel, Power BI tools. But there are limitations. We sometimes use dynamo to automate data transfer from excel to Revit
19. We represent it with Bar charts, heatmaps, graphs, Pie charts.
20. Values mostly required and recorded are Temperature, Co2, Humidity, noise.
21. We use color codes for depicting different values and darker colors for showing exceeded thresholds.
22. We use CFD simulations for HVAC sometimes.
23. We are using a data explorer tool to compare different kinds of data to each other quicker we want something similar as well.
24. We want to show the clients the discrepancies in the values.
25. After the meetings we make a report and perform a sensitivity analysis on what needs immediate attention.
26. Contextualization of sensor data with building geometry would benefit us as well as our clients.
27. The data from the homes is collected through IC meter other projects have a app called Netatmo for the building occupants.

The interview is segregated into five themes to categorize and extract relevant knowledge attained from the interview explained as follows:

1. Purpose – The labels under this category would congregate the motives of the building energy consultant to accomplish tasks in a project.
2. Issues – The labels under this theme category describe the issues described by the building energy consultant during the interview which he faces to accomplish the purpose,
3. User needs – The labels under this theme category lists the different needs for understanding the data required from the different stakeholders to the building energy consultant.
4. Values and standards - The labels under this theme category list the different standards and value used by the Building energy consultant to standardize values in a construction project.
5. Methods - The labels under this theme category list the tools. Process and ways the building energy consultant uses to accomplish the tasks for different data visualization.

STAGE 4: CHARTING DATA INTO FRAMEWORK MATRIX

In the following framework matrix (Table 2) five themes were identified and the labels were categories according to the relevant theme category –

Table 2: Framework matrix

Purpose	Issues	User needs	Values and standards	Methods
1. Sustainability, Indoor environment, healthy buildings.	9.Hard time to measure and analyze sensor data.	15.Facility managers require to find spatial elements of buildings faster.	20.Values mostly required and recorded are Temperature, Co2, Humidity, noise.	18.We use Excel, Power BI tools. But there are limitations.
2.Guiding decision makers	10.Trying to visualize a lot of sensor data but it generates lot of problems.	16.HVAC consultants need to find reasons for discrepancies.	3.DGNB standards and Bsim simulations.	19.We represent it with Bar charts, heatmaps, graphs, Pie charts
3.Efficient design and code compliance with DGNB.	11.Different users need different visualizations to understand sensor data	17.Mapping occupancy patterns is important.	22.CFD simulation for HVAC	21.We use color codes for depicting different values and darker colors for showing exceeded thresholds.
4.Exploring new ways to analyze sensor data.	9.Communicating sensor data to other stakeholders is a problem.	7.Detect malfunctions.		23.Multi objective optimization with Data explorer tool

Purpose	Issues	User needs	Values and standards	Methods
6.Spatial recognition and orientation of a building for explaining to FM managers, Building owners and HVAC engineers.	12.Simulations cannot predict, very different from reality.	6.Hard for Building owners to understand sensor data.		27.IC meter sensors to collect data and Netatmo
7.Finding reasons for discrepancies.	11.Occupancies change of building units, so patterns change. Mapping becomes complex.	26.Contextualization of sensor data with building geometry would benefit us as well as our clients.		18.Dynamo scripts for automating BIM data to excel
25.Sensitivity analysis	13. Analysis of multiple factors to diagnose the causes in design or systems.	13.Relating different sensor data to itself for comparisons.		18.Slicer tool to filter data to showcase
23.Optimizing for future designs	14.No geometric context of the building creates problems.	27.Building end users need to monitor their indoor environment.		27.HTML request to the API every day to collect sensor data
25.Reporting and summarizing data	14.Tables and graphs are difficult to perceive for architects, building owners.			

5.3. Work Modelling

The following information was captured through the interviews and information was summarized and diagrammed through the work modelling process of the contextual design methodology. Following we present the different work models to understand the user intents, needs and information exchanges required.

5.3.1. Flow model

The flow model in Figure 11 describes the roles of each participant and methods through which they collaborate to exchange information of the project to complete their individual tasks. The arrows depict which information is requested from which project participant. The yellow lightning symbol highlights the doubts and questions raised due to poor understanding of the exchanged data.

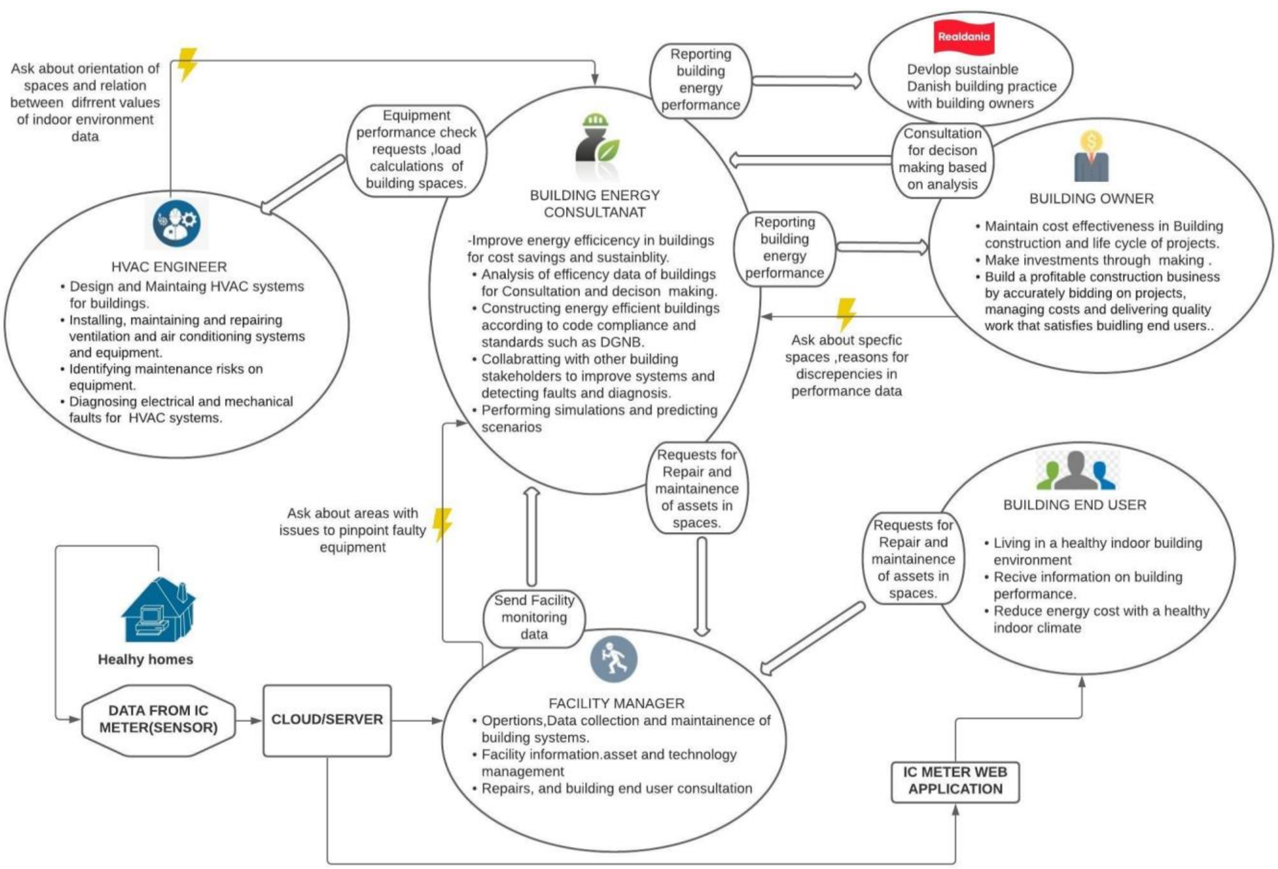


Figure 11: Flow model representing information exchange between actors.

5.3.2. Sequence model

The sequence depicted in Figure 12 describes the detailed steps performed and tools utilized by the Building energy engineer to gather relevant data for presenting it to the stakeholder for the collaboration meetings. The yellow lightning symbol in between the tasks show the unautomated processes and processes which sometimes generate incorrect data/representation issues as described by him through the interview.

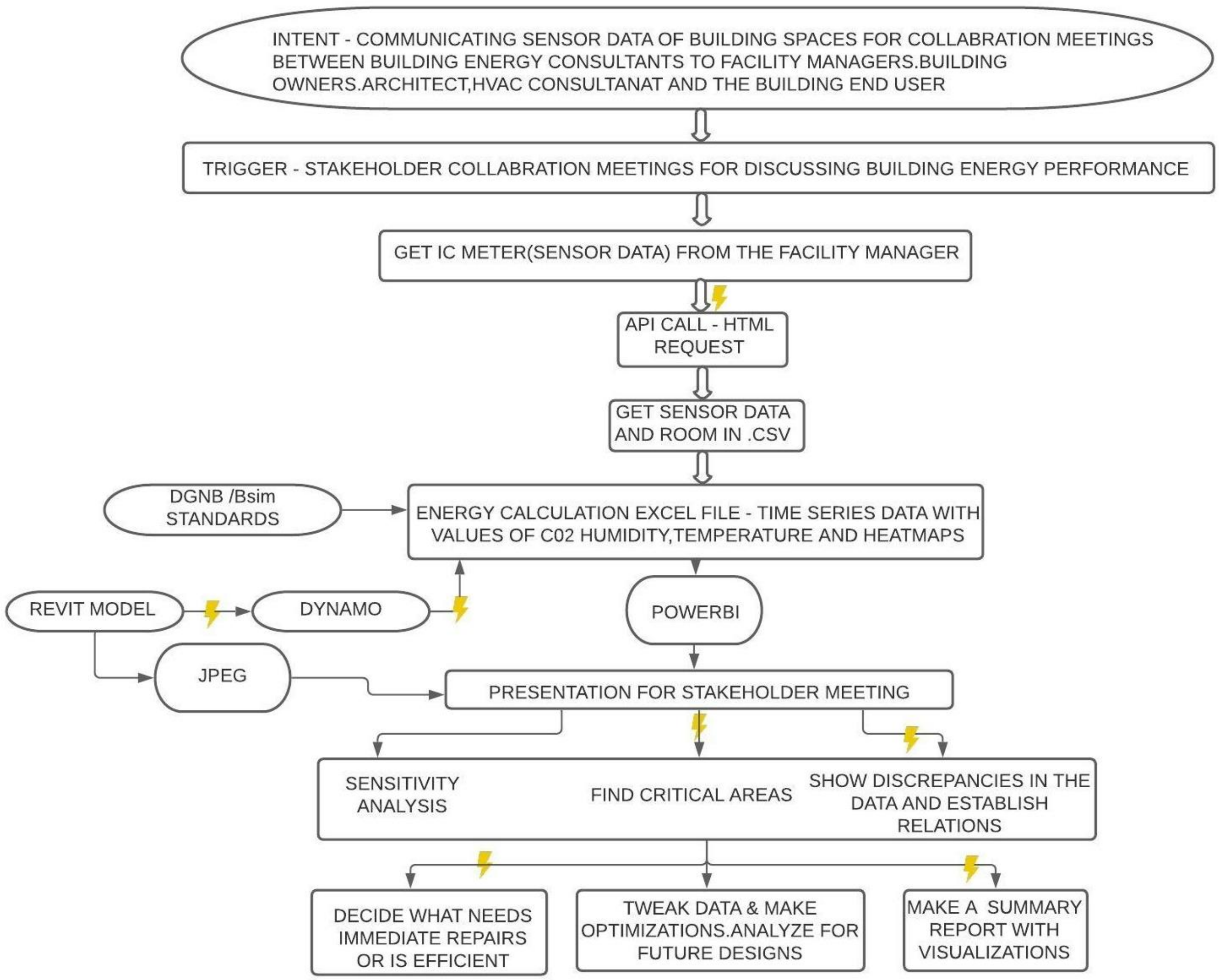


Figure 12: Sequence model.

5.3.3. Physical models

The physical models show the steps undertaken by the building energy consultant to collect the data, refine and make a presentation for the stakeholder meetings. In Figure 13 the steps taken by the energy engineer to tweak and display specific data according to color codes to generate a heatmap graph is shown.

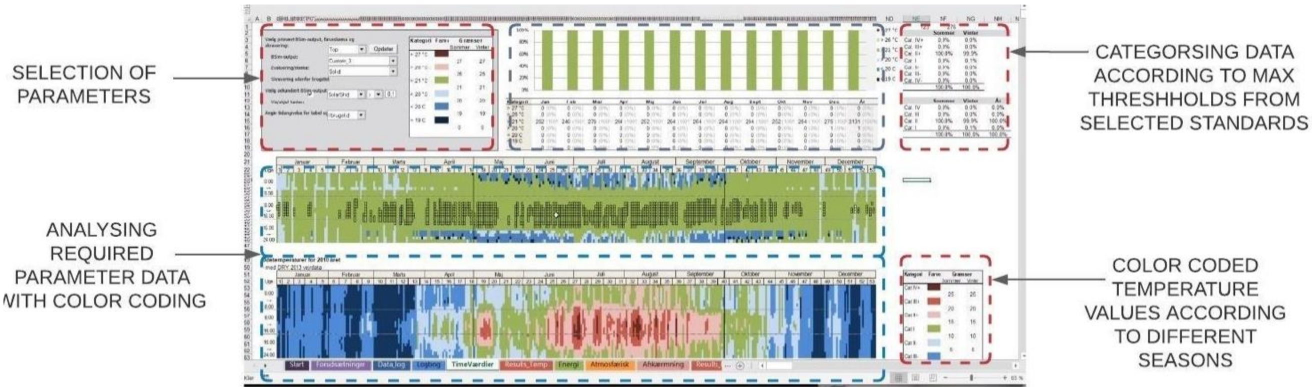


Figure 13: Representation of indoor environment data sheet managed by the building energy consultant.

Figure 14 further describes the functionalities of in the excel data sheet, the selection of time durations and standards such as DGNB to visualize the data in form of a graph. Furthermore, the building energy consultant uses multi objective parameter optimizations for different energy design parameters. Currently it is not used to compare data from the sensors, but the energy consultant plants to incorporate it for future projects.

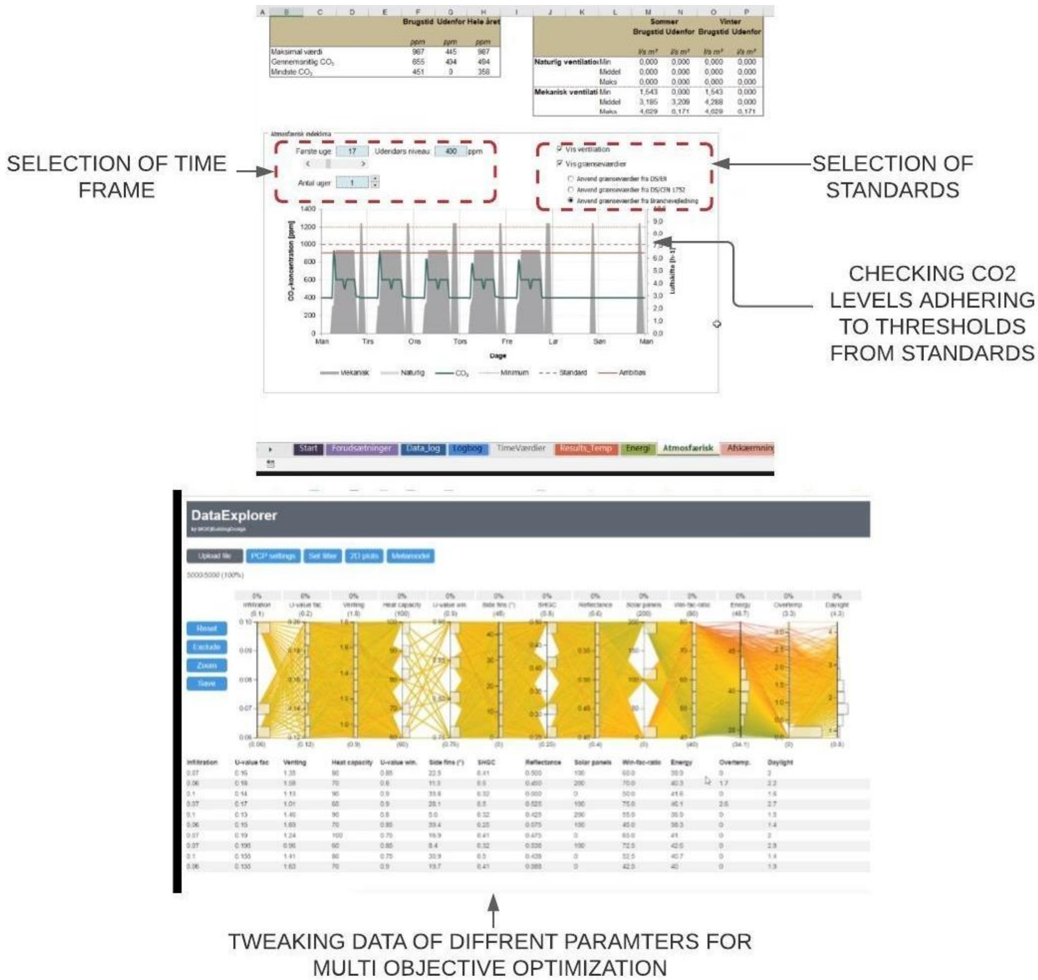


Figure 14: Components of the Building energy consultant's data sheet.

Figure 15 shows the Power BI tool interface which is connected to the excel sheet. The energy consultant uses it specially to visualize data for the stakeholder meetings. The Power BI interface has a slicer tool to dynamically select the time duration of the required sensor data and the slicer tool to select the corresponding rooms. The tools also contain color gradients related to range of the values of the data from the sensors.

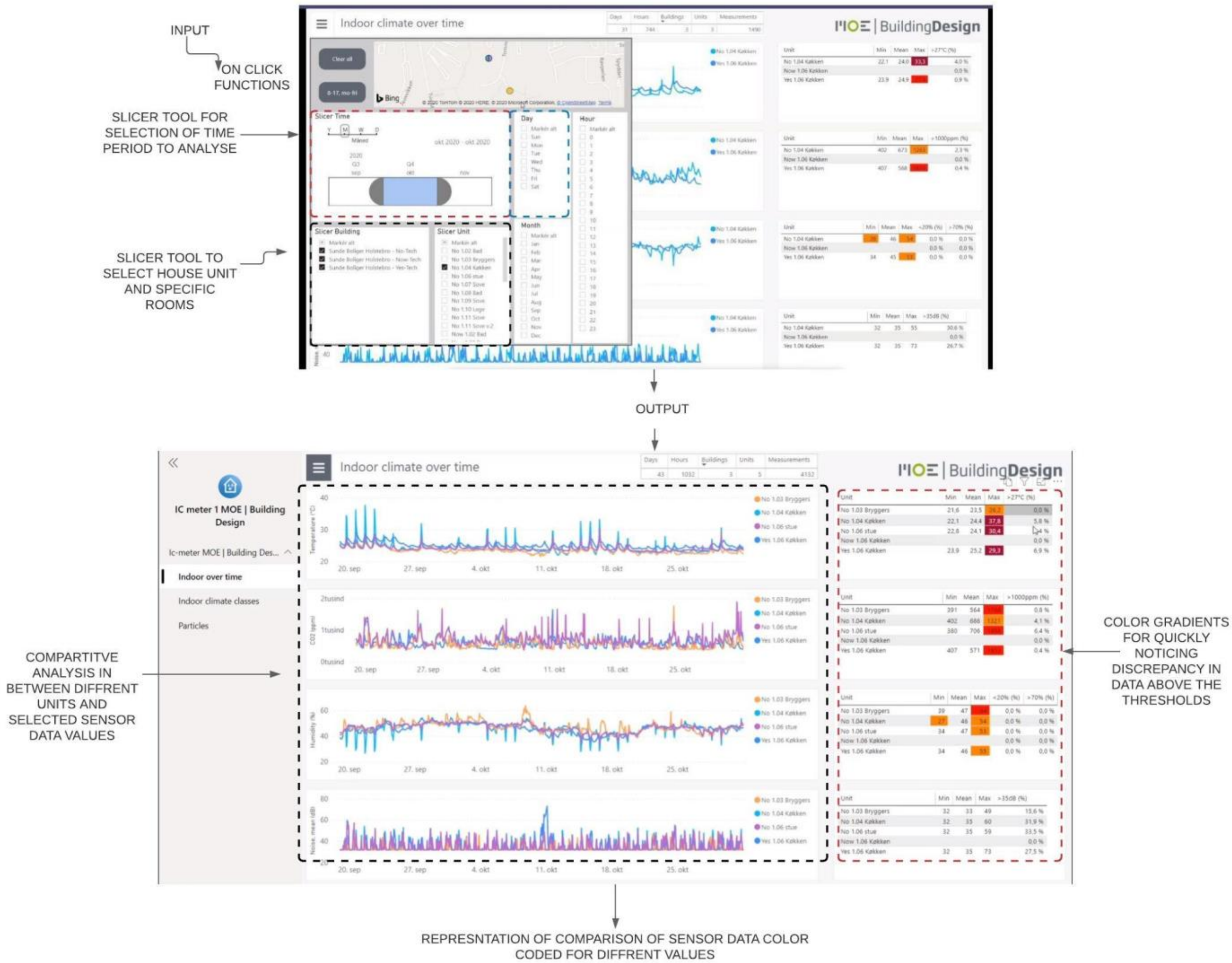


Figure 15: Visualization thru Power BI for the stakeholder meetings.

Figure 16 describes the final visualization and reporting structure visualized through using the Power BI tool to the different stakeholders.

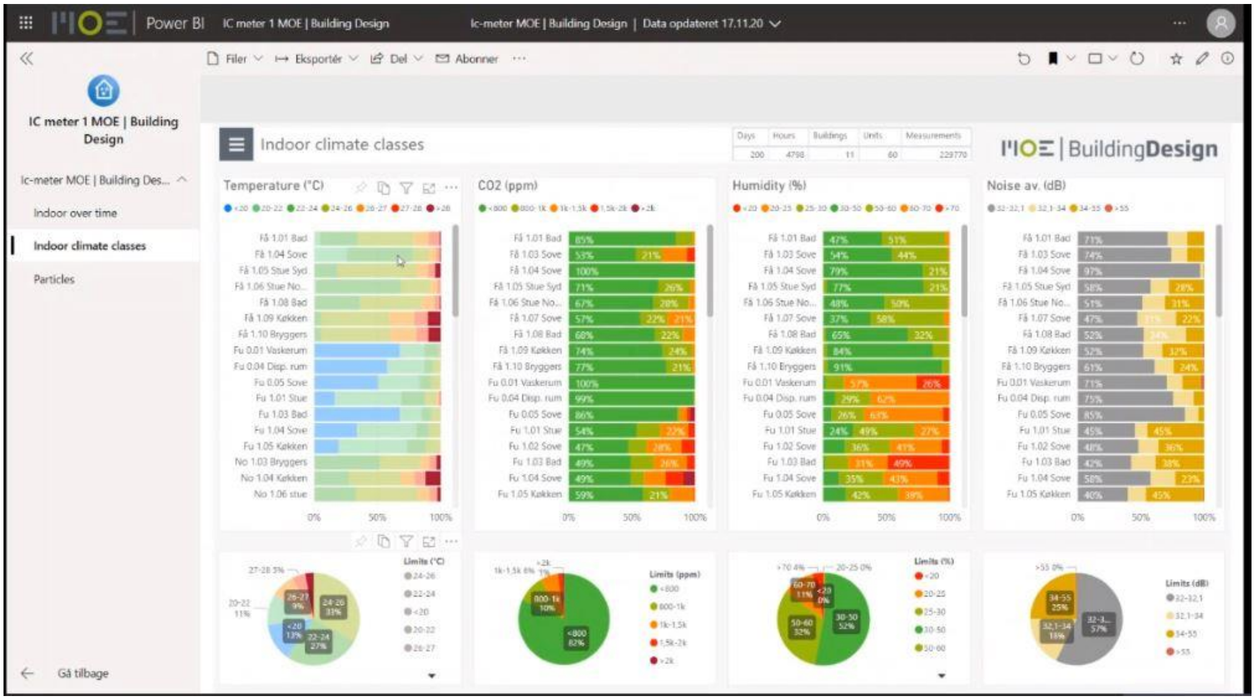


Figure 16: Final presentation template for the stakeholders meeting and reporting.

5.4. Visioning

Information gathered and interpreted in the contextual inquiry; mapped and consolidated in work models helps to successfully start the next step. In the previous one, it was important to create the relations between actors, overall processes, and singular actor's actions. Different mapping settings allowed to capture various aspects of the organization processes with specific theme and subject of focus– such as connection of actors, sequences of tasks, workspace setup and navigation through the environment.

In larger projects there are many stakeholders. Hence, the affinity diagram is predecessor of the visioning. Conceptually, it helps to have more structured information on experience of multiple users. In this case, even though some of the stakeholders were included in the flow diagram, there is one user who was focused on. Moreover, the interpretation session was aided by "Thematic analysis of the interview" by Gale et al. (2013) . It turned out that the session went on par with organizing work models. Therefore, the visioning was based on both work models and thematic analysis.

Figure 17 shows the visioning diagram developed for the use case. The idea for the starting point derived solely from the user– energy consultant engineer. By taking into account user's needs, and expectations hinted in the contextual inquiry, it is possible to map ideas as visions – what should be included in the future application; is it feasible and doable from technical point of view (although, not a thorough evaluation); does it fit the model and aspirations of the company (Holtzblatt & Beyer, 2017).

The visioning consists of two parts. First, more detailed, displays design ideas for the energy consultant which is the end user for this project and, therefore, more thought process was put into that regard. Second part address users who are brought to attention in work models to see how they potentially might be able to utilize new system solution. The main idea is to have different types of users participating in the project. Participants would have various access level privileges depending on the role they have. Additionally, the set of dashboards, tool, viewing panels, ability for accessing data would differ between them. For example, building owner could be permitted to see overall energy consumption, yearly/monthly cost of energy, reports, maintenance alerts, model, and overview of household, and other assessments with just his credentials without the necessity of contacting the company personally, calling for meeting every time he wants have an insight to the data. The level of detail of provided information should also be decided in the project documentation. The visioning also includes rough ideas for a building end-user, HVAC engineer or facility manager and how they can interact with project information.

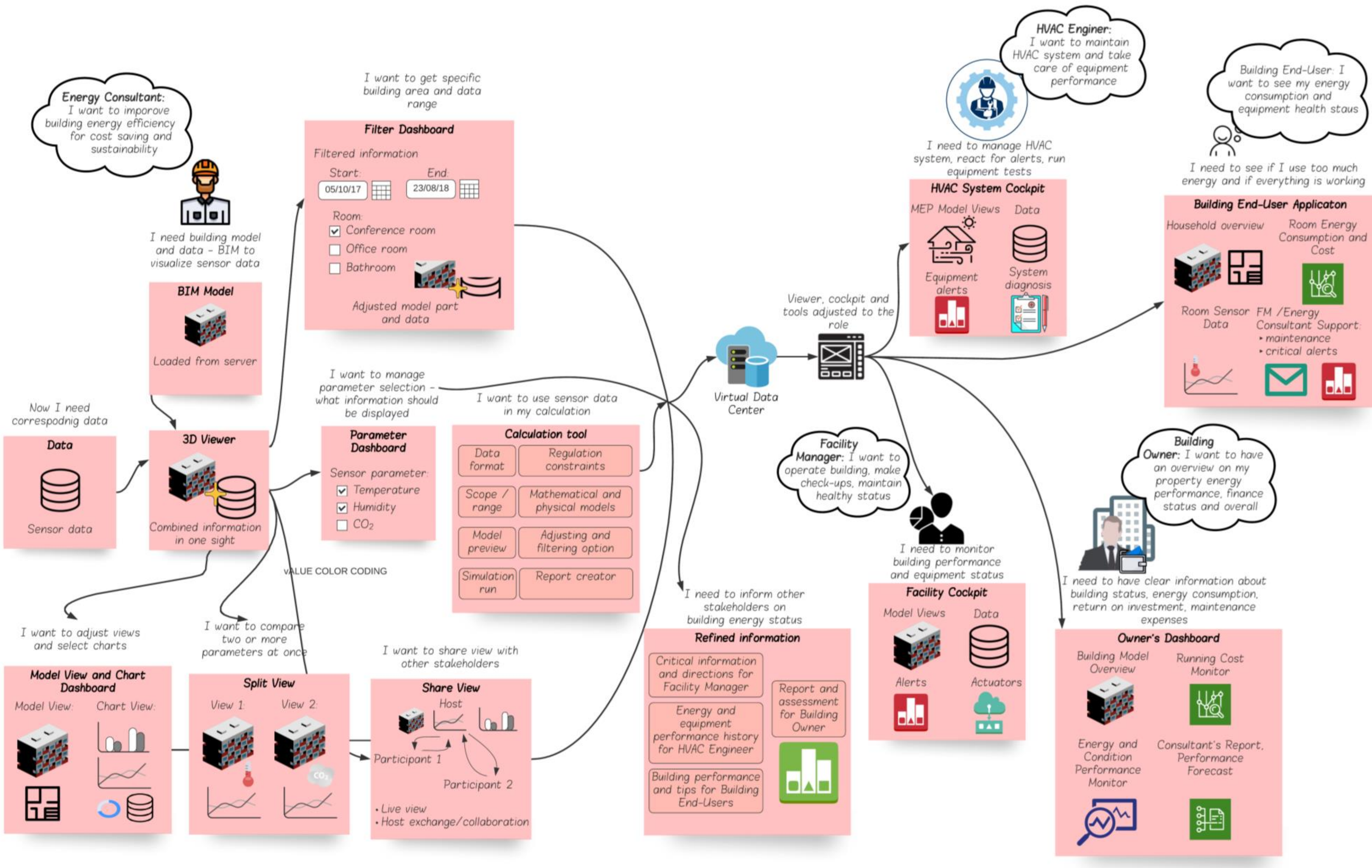


Figure 17: Visioning diagram for the use case.

Configuration of users, and list of their privileges are flexible, expandable and could be tailored to the individual needs with private cloud, virtual data center or other solution fitting the organization to manage the application and fit into existing company's infrastructure.

The focus user is the energy engineering consultant. Hence, more thoughtful ideas shown in Figure 17 are regarding the consultant. The development of these design boards (visions) starts from the initial actions and the initial user need. From that point, the creation of other elements should unfold as description of user's actions– a story. Other it is free to add other forms of cognitive activities.

The intermediate stage between visioning and storyboarding is creating concepts. Product concepts emerge from the from vision. Each product concept presents the group of actions, tasks, solution features, and functions based on their coherence and relation– how far they are each other in terms of cognition, how strongly they are connected.

The visioning part of energy consultant was categorized into four concepts. The example how it looks were created for one of them, and it hints how it might look for others.

5.4.1. Model Viewer

Figure 18 depicts the first concept. The intention was to gather solution elements and actions regarding the elementary viewing experience of model and data. Ideas from Figure 17 such as: 3D Viewer, Model View and Chart Dashboard, Filter Dashboard, Parameter Dashboard were categorized as one cohesive structure which enables performing a cognitive action of viewing the model, analyzing data, contextualizing it with displaying indoor climate values directly on the model.

The user should be able to get an overview of the situation in the project without going into specifics of data, detailed descriptions of the projects and visual cues, surrounded by all possible tools. This product concept serves as main hall/corridor where the occupant decides to what room (focus area) wants to go to perform another action and task.

Said arrangement of the concept helps to visualize possible layout of wireframe for that particular set. There is, for example, a possibility of loading different models, changing modes (i.e., concept views), arranging views, navigating inside the model, filtering data to be displayed, arranging charts, inspecting project and objects properties. Ideally, sole location of focus areas should be enough to make the user understanding how to navigate through the application, what he can expect, or if it is relatively easy to find necessary features.

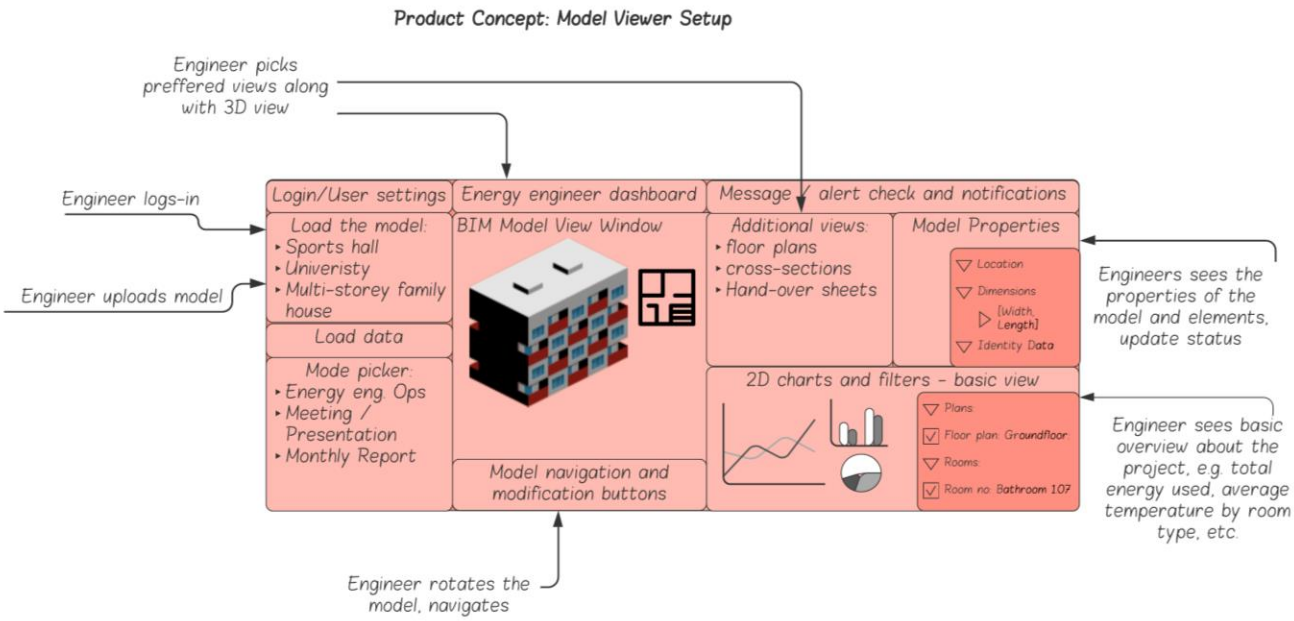


Figure 18: One of the Product Concepts - Model Viewer.

5.4.2. Energy Engineering Operations

This product concept revolves around main actions and tasks strictly engineering. The user does not lose the functionality of previewing project and building data. Conversely, the level of detail of the information features, and utilities is expanded.

This product concept is inspired by acquired information and data analysis, energy performance simulations, and overall calculations performed by energy consultant engineer which were presented in Physical Models. In the current situation, the engineer uses web tools, i.e., custom tool for energy simulation (Figure 13 and Figure 14), PowerBI (Figure 15 and Figure 16), and withal 3D model representation in another display media like Revit or PDF viewer for the preview of plans or sections.

The idea was to extract main activities and of the user's current model and combine it into the new solution. For example, in the spreadsheet about representing indoor environment data Figure 13 and in the overview of components of energy engineer consultant's data spreadsheet from Figure 14 is the common factor of constraints and regulations appearing. This lead to include "Regulation constraints" as one of the features of Energy Engineer Operations. It can be seen in Figure 17 under "Calculation tool" box.

5.4.3. Stakeholder Meeting – Presentation Mode

Another aspect extracted from the Contextual Inquiry and derived from work models is stakeholder meeting. There might be a meeting between engineers, architects, and departments, but also with the building owner or other actors.

The energy engineer consultant described the process of such meetings. Report with most important factors, values, and project status which ought to be presented by the engineer. Additionally, 3D model preview, or floor plans (any necessary view) are attached. Usually, architect puts everything together in one presentation and this is the final product ready for the meeting with stakeholders.

It is the most common way to present the results and project status. The biggest shortcoming is its static nature. Certainly, the presentation can include GIF or short video showing the model from different perspectives as the camera moves around the model, or data visualization for a particular sensor data. However, it lacks the flexibility and interactivity for showing other information than initially intended, which the need emerged as the meeting advances.

Therefore, it was decided to include product concept called “Stakeholder Meeting – Presentation Mode”. There are a few possible functionalities which would enhance the experience of the meetings. One of them is that the speaker is able to dynamically share data (e.g. indoor climate sensor values), show the data visualisation not only in the chart, but also corresponding values visible in the BIM model. Moreover, stakeholder would independently discover shared view and data. It would be especially functional in the time of increased number of remote meetings. Sharing the information in such way, might be also beneficial for general project coordination. This concept would utilize Share View, Split View, Model View and Chart Dashboard (Figure 17) and specific draft for it should be prepared, which ultimately would evolve into storyboard with its own User Environment Design.

5.4.4. *Report Creator*

The final, as for now, concept concerns creating the report of work done, project status, energy usage over a month or equipment usage. As in previous examples, the idea emerged from the consulting engineer's needs. There are guidelines which the engineer must follow, like uniformed way of plotting the chart or special type of chart for a particular data.

The intention is to have all necessary information in the line of sight or under a few mouse clicks. The box "Refined information" in Figure 17 shows the main idea of this concept. There would be structured information which can be provided to other participant, e.g., facility manager would like to have more detailed information about critical information of a building performance, whereas the building owner would be more interested in overall building efficiency, energy usage, and economy values. The engineer should be able to create a report with desired information – model views, data with arguments and value in an efficient way. It means that the separate canvas the user would pick the view to be included. Then, he would choose element filters (if particular set of elements should be highlighted). Additionally, he would decide on ranges of sensor data, or other data for that matter, and types of data covered in the report.

As much as technical skills and software capabilities would allow, this concept would leverage of the fixed structure of report (each kind of report would have different layout), where there are placeholders designated for specific BIM model elements (2D/3D representation, charts, tables, etc.).

5.5. Storyboarding

The storyboard is a visual representation of how users would perform action within a singular product concept. It means that for each concept there should be separated set of storyboards. Number of steps in a sequence might vary depending on the granularity of thereof concepts and the intended level of detail for the stories. Some might require to be narrowed down after an initial attempt. From that point until the prototyping a few iterations should be considered. Sometimes it is necessary to take a step back, change the action process, clarify layout arrangement and interactions until functional prototype. The iteration loop involves:

1. Storyboarding – to alternate actions, tasks, processes, views, arrangements.
2. User Environment Design – to clarify and rearrange focus areas with functions and objects. Interaction patterns and visualization model are also involved in this step.
3. Prototyping and interaction with the user – to account user's feedback into the design and to reactively adjust the storyboards or User Environment Design.

The loop is interchangeable which means that from any point it possible to go back and change any other the cause of identified issue. After performed change, it is important to validate other planes design and adjust them accordingly. Then, each concept can be broken down in individual storyboarding.

In this example is Product Concept: Model Viewer Setup (Figure 19). It consists of four steps where each encapsulates activities, tasks thematically, and cognitively related. Storyboards are not linear, despite numbers. The sequence gives clue what is the general path, but nothing stops to go step back if necessary, or at least, it should not stop it. Breaking down the structure might help to understand the reason of this arrangement:

1. Work setup for the project

Here, the user logs in with his credentials, picks the project and model which he has access to. Tab about loading data regards mostly new project/model setup or updated version. With the data received with CSV or JSON file, it might be tiresome as it requires refreshing and updating information periodically, but the end goal is to have connection. This point of concern, however, should not move how the activities are solved from the technical side as it will be also to the time during Development of System Architecture, but rather if connection between them activities makes logical sense. There is also called "Mode picker" (workspace) – it serves as a hall in building analogy - which enables user to enter different "room", where he is faced with cognitively different assets and tasks in each mode.

2. Workspace arrangement

After choosing a particular mode to work in, it is a time to arrange the space, so it fits the user work style, pace, and project characteristics. On the top of the board there is a notification and message tab. It gives easy access to communication channel regarding project which currently open. Notifications are sent when project changes happen, critical information or events occurred. The engineer is able to arrange windows in terms of sizes and, to some extent, of content. He is free to choose additional views, aside from the 3D representation, and set of properties. He sets up navigation bar of 3D model. Moreover, he arranges the files, documents, reports, other data for a quick look and for reference.

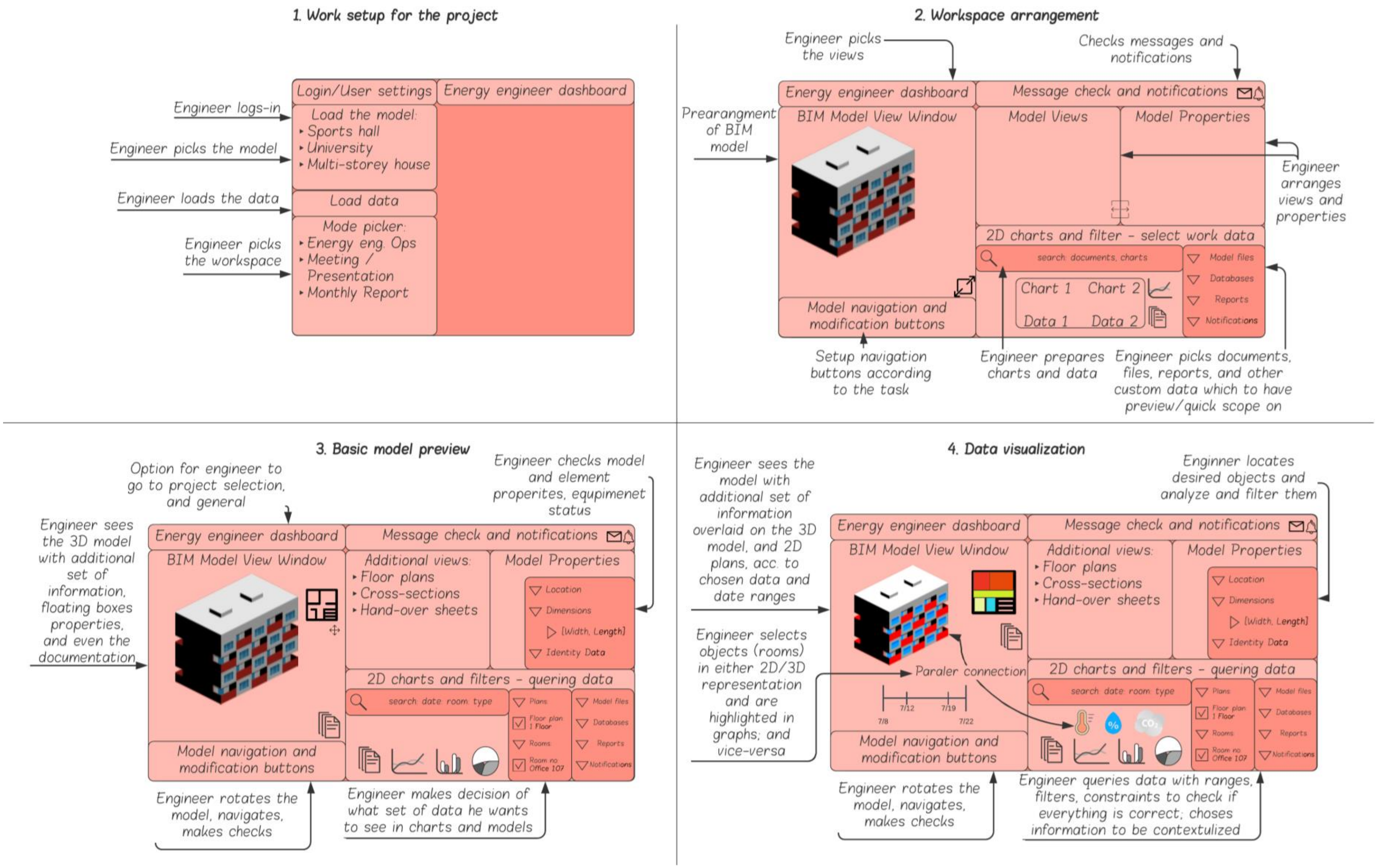


Figure 19: Storyboard set of Model Viewer product concept.

Finally, the engineer preliminary loads charts, and make sure that the correct data is ready to be used.

3. Basic model preview

Preparation of working areas is followed by the main model previews. Here are utilized main functions which can be found in viewer. The most vivid is the ability of moving, rotating model, zooming in/out, arranging perspective. The engineer can also open various floating boxes which enhance the viewer experience, such as project and model properties, project browser, 2D mini map of floor plan, etc. The user might also want to view, and inspect elements, their current efficiency, maintenance status and similar. The last, as for now, function is model and data filtering. This is inspired by the current workflow of the engineer. He has the opportunity to choose multiple areas (room) of the building, selected by function or instance, and adjust date ranges of concerned data. It allows to view specific areas, which otherwise the scope and amount of information could be too large to comprehend and hence impractical.

4. Data visualization

The point of converge of this product concept storyboard is the visualization of sensor data in the context of spatial representation of the building. The first performed task in this part is to make sure that the correct data is prepared. Objects to be contextualized and examined ought to be filtered not to overcomplicate and not to overload visual perception capabilities of the user. It regards both charts and data overlaid in the 2D plan 3D model.

Once elements filtering is done, another option is to reduce or extend date range of which data will be displayed. It should also address dynamically change granularity of data while extending (zooming out – it would equal to larger intervals, making data display more coarse, more general.) or reducing (zooming in – it would be on par with having more detailed information, shorter intervals, etc.). Manipulating data range would be accessible in charts, but also in model viewer, but they would react interchangeably – one changes as another is altered. It is also linked to relation between what is displayed on chart and the model, and how selection on one entity triggers selection on another.

The last piece is to involve variety of information types, such as sensor data of temperature, humidity, CO_2. The type of data might be filtered in the same manner as elements, plans, objects in the model. It is mostly up to user preferences how many and how should be displayed. However, some constraints might occur when developing visualization interaction models. Data would be queried according to specified query operations.

Storyboard puts slightly tunnel vision into the designer mind as naturally he wants to solve the problem for a particular case (Holtzblatt & Beyer, 2017). However, preparing universal solution in the first iteration might not be worth the time spent, but rather support yourself with the next steps, because as mentioned earlier – it is viable to back and make changes according to the newly gathered information. The next section helps to generalize the steps, functions, and areas to make it suitable for greater range of situation.

5.6. User Environment Design

Successively as the storyboard design is completed as of now, it is time for proceeding with the User Environment Design. It is a transition from high-level vision to more structured, but not yet low-level workspace (Holtzblatt & Beyer, 2017). Some specific, concrete ideas are hard to be omitted during the storyboards, but freehand sketch of the areas, steps helps to avoid detailed thinking. The User Environment Design holds similar function. It assists with transition from high-level, more general ideas initiated in the product concepts, and finalized with storyboards to more concrete abstraction of Focus Areas, which are characterized by purpose, functions, and objects included. The User Environment Design serves as a buffer and prolong the time of initiation low-level workspaces.

The produced User Environment Design for this project is shown in Figure 20. The key aspect of the diagram is Focus Area and its relation between multiple areas. Such spaces should be structured in a way that each has assigned one purpose – one cognitive activity. It is not a singular small, task, but rather set of tasks which are related to each other. Each Focus Area in the use case consists of:

1. Purpose – generalized intention of an activity. It should consist of one distinctive cognitive activity, i.e., a goal, supporting the performance.
2. Functions –multiple tasks (sub-category of an activity) helping to achieve purpose.
3. Objects – list of elements which would allow to make use of function (fulfil a task) and navigate within the focus area or transition to another.
4. Links – relationship with other Focus Areas.
5. Restrictions and risks – limitations, concerns of area which affect functionalities.

Storyboards presented in the previous chapter were decapitated and organized into eight Focus Areas. One of them is (1) Model viewer. It is a place to view and navigate through the model using functions, supported with objects. It has multiple links (Figure 20). Among them are (2) Data visualization charts, (3) Model data visualization. They are not connected directly but converged at Model viewer. Their sole purposes coexist without direct effect on one another, until they meet at one point (1). It essentially allows to contextualize data powered by (2) charts and (3) model with their purposes and functions.

The User Environment Design can be now re-evaluated within itself or iterated and refined between storyboards, product concepts and future stages as it was described earlier. Emerged steps are interaction pattern and prototype.

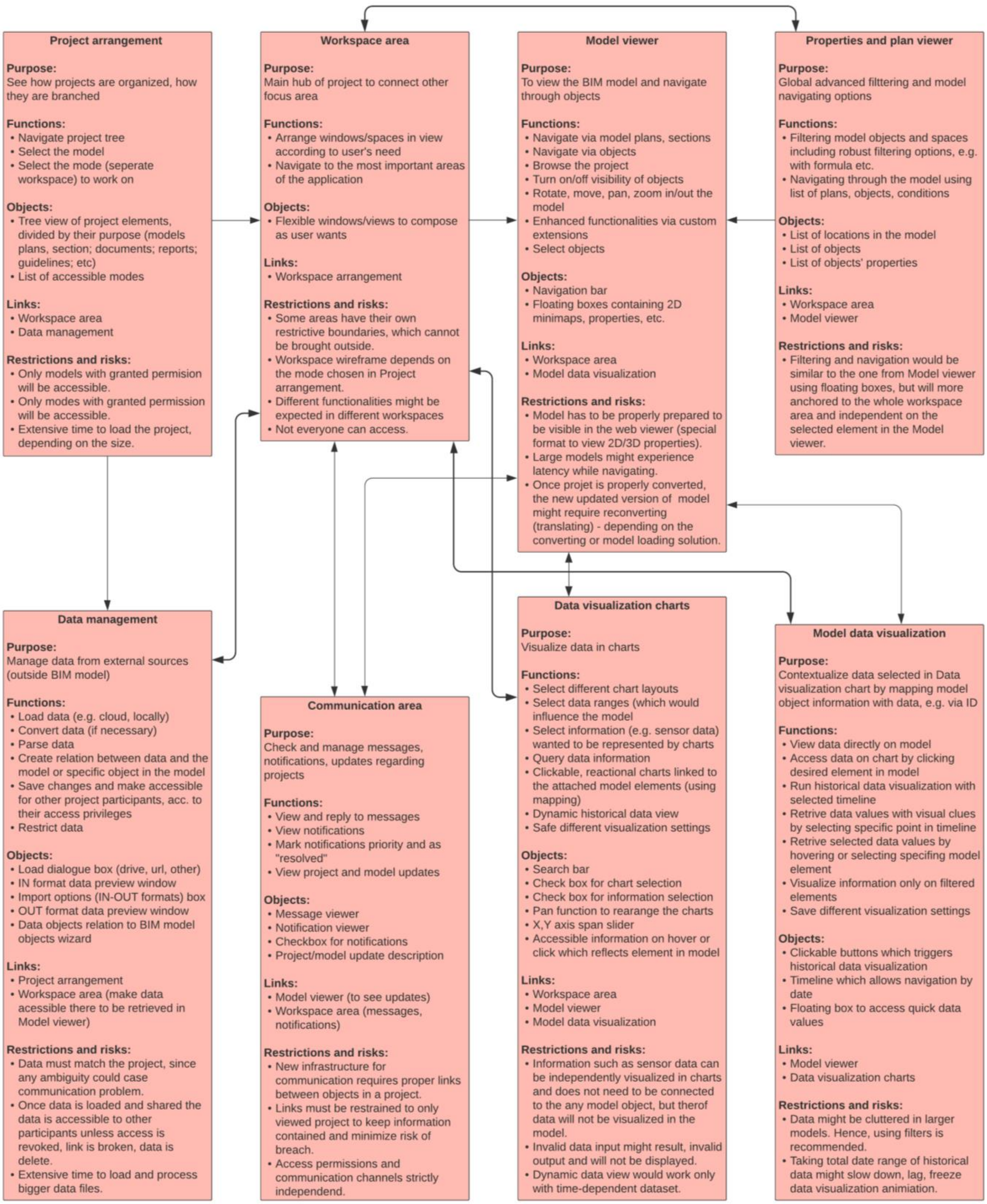


Figure 20: User Environment Design in consideration Model Viewer product concept.

5.7. Prototyping and Interaction

Prototyping is the final stage where technical challenges, aspects and solutions are not fully involved. It is, however, the time when the consideration leans towards interaction models and patterns. It means that outline of the User Environment Design receives closer inspection, and it evolves into even more concrete entity. Those interaction patterns describe how the user navigates through Focus Area. It informs how functions are affect each other and what are the dynamics between objects. Finally, it describes how interactions affect content of the application – the main user focus, object of work, e.g., building, data with graphs.

In the use case, interaction patterns and prototype were created at once (Figure 21). It was supported by previously completed stages. However, ideally, those two steps could have been split not to be influenced by instant feedbacks interaction which are accessible in the prototype-creating software/tool, and they are hard to ignore.

For the use case basic interaction patterns were adopted, such as moving object (2D/3D representation), zooming (model, data set in charts), selecting (elements), dynamic querying (data values, types, related entities).

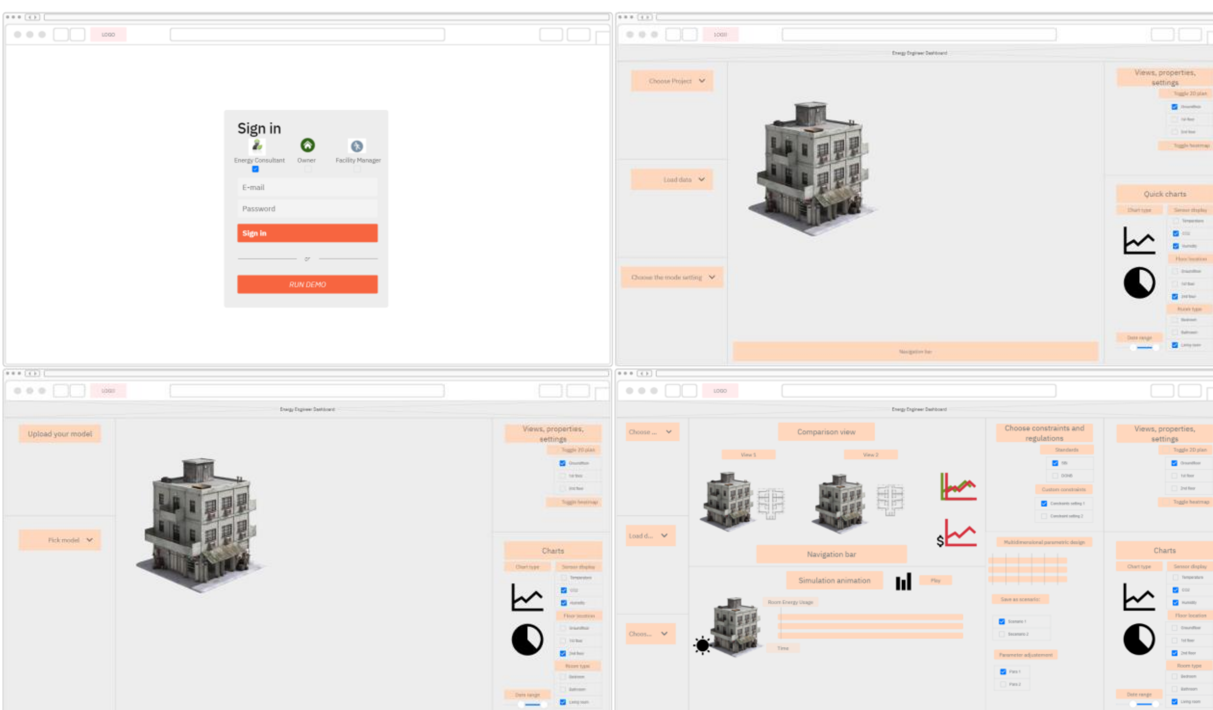


Figure 21: Layout of spaces for 1st prototyping session.

The prototype was sent to the user and the meeting was arranged to be able to listen to his feedbacks. The intention was to check user's behavior around different views. Not all Focus Areas and their functions were brought to the prototype. The prototype consisted of four scenes.

1. Login scene – user picks the role assigned to him; or opportunity to have demo preview of the viewer.
2. Demo scene - showcase of the application with limited functionalities and preloaded model and data.
3. Model viewer scene – basic visualization of the building representation, showing how panels (such as project and model property, charts with filter options) are arranged. There is an option going to project assigned to the user, loading data, change mode (workspace)
4. Engineering operation mode scene – this is the scene for which product concept and storyboards were not designed. It was decided to put the Split view from Visioning chapter, Simulation animation, panels regarding workspace configuration (constraints picker; view properties filter; chart viewer with filtering; multidimensional parametric design – which was not explicitly mentioned in the Visioning chapter).

User's was positive, but the main drawback was how cluttered was the 4th scene. The study on scene-to-scene transition and interaction patterns were not fully determined. Moreover, the certitude was that the engineer should have access to all major functionalities within one click. Luckily, it was the first iteration of the prototyping session. Lesson learned that the prototype should go along what was decided during the previous stages and not to try show everything at once, even though singular functions and objects could be attractive.

Conceptually after completing the first prototyping session, it is time to list the feedback, go as much to back as it necessary, for example, to product concept, storyboarding or the User Environment Design.

6. Development of System Architecture

Based on the developed contextual design models, literature review and the problem formulation a holistic process is proposed for the visualization and contextualization of indoor environment data collected from the sensors. Adhering to the (Boje et al., 2020) 3 tier generation evolution of the construction digital twin, and models analyzed from different literature aiming to develop the digital twin concept and the visualization issues consolidated in work models from the building energy consultant directed towards the communicating data to the different users. Moreover, the framework is based on the identified user needs, information flow sequencing from the interviews, validated concepts from the contextual design processes. A basic web-based user interface is built through applying the processes contained in the framework, which would require further development.

The proposed system architecture serves the visualization and data contextualization needs and adheres to the other components of the Construction digital twin concept. The use of a web-based solution substantiates that the proposed framework would further be aligned to other components such as actuation, monitoring etc., to become a full-fledged construction Digital twin in the future. Although the framework focuses on developing visualization and interaction techniques for the end users. It is one of the elements to the larger idea of a construction digital twin.

The framework is segregated into the spaces derived from (Kamran Sedig & Parsons, 2016) to create a holistic understanding and division of technical components, particularly which would aid the developer to specifically rectify or enhance the needs of the user through relating to these spaces as shown in Figure 22. Two different BIM models and sensor data was utilized for prototyping which is divided in use case 1 and 2.

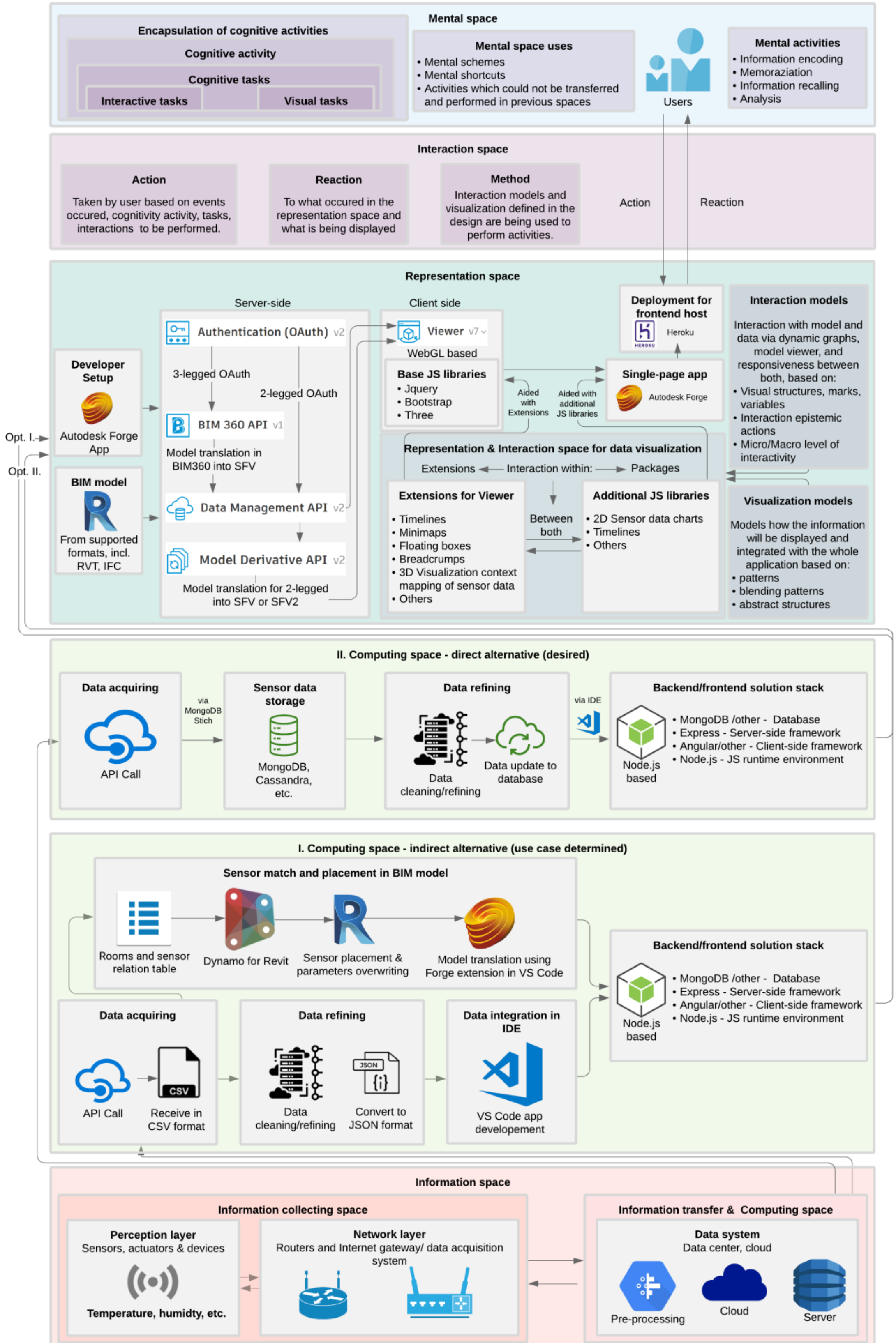


Figure 22: Proposed System Architecture.

6.1. Information collection space

The information space contains two layers as shown in Figure 23. The first is the *Perception layer*, which consists of physical objects such as sensors and devices which are capturing different values of the physical environment such as temperature, Co2, humidity, noise etc. The second layer is the *Network layer* which is connected to the internet and transmits the collected data of the values to the cloud in the information and computing space. For example, in the *Use case 2* an IC-Meter was used to sense the environment in the buildings *"Indoor Climate Meter (IC-Meter) measures visualizes and analyses indoor climate in a room or home. The concept consists of a meter, a server solution, and an APP/ Website. The measuring unit (IC-Meter) provides accurate measurements of temperature, humidity and - as something new in the consumer market - CO2. Measurements are uploaded every 5 minutes to a server via the customer's Internet/ Wi-Fi. The server also retrieves local weather forecasts, energy measurements, etc., from the servers of public and commercial actors."* (ic-meter, n.d.).

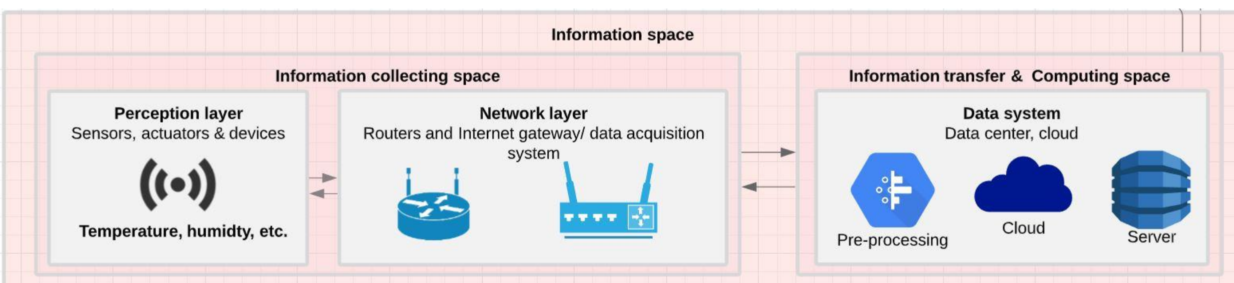


Figure 23: Information collection space and sub spaces.

6.2. Information transfer and computing sub space

In the information and computing space the data is stored on the cloud which can process the large amount of continuous data on the distributed server. For example, in our use case 2 IC meter (Indoor climate meter) was used in the buildings which has its own dedicated cloud server to store and fetch data.

6.3. Computing Space – Direct alternative

The computing space has four layers the *Data acquiring layer*, the *Storage layer*, the *Refining layer,* and the *backend/front end layer containing MEAN stack* for web development.

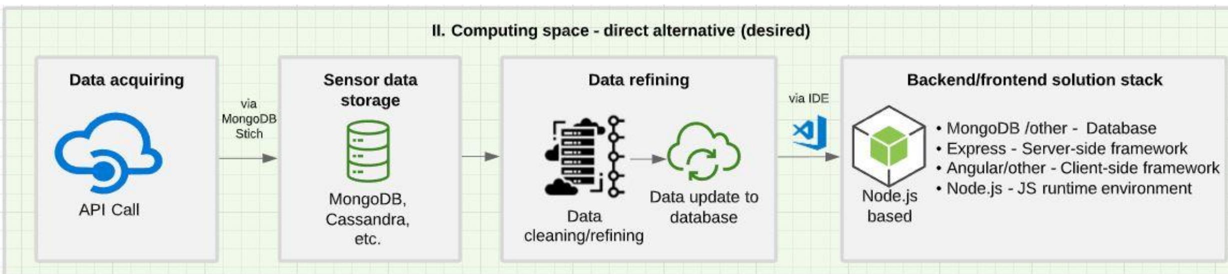


Figure 24: Computing Space – Direct alternative.

In the Data acquiring layer the data is collected from the information transfer and computing space through an API call which is specific to the product used to sense or actuate. The acquired time stamped data can be stored in a NoSQL database such as MongoDB with the help of MongoDB stich libraries which are JSON based and data can be fetched through an API service. Refining, structuring, and querying data if required can be done in the MongoDB compass which is an application interface for the database service, querying is based on MQL (Mongo query language). In the *back end/front end layer* for web development, *MEAN* solution stack is utilized. It is an open source and free of charge software stack for developing dynamic websites and applications on the web.

MEAN derives its name from the four components that together offer mutually client & server-side mechanisms for interactive web applications: Mongo DB offers the object-database; Express.js which delivers a framework for web routing; Angular.js intended for web applications; and Node.js — the JavaScript engine built on chrome, and web server component(Linnovate/Mean, 2013/2021). Node.js is the runtime and *npm* is the Package Manager for Node.js modules.

AALBORG UNIVERSITY
STUDENT REPORT

6.4. Computing space – Indirect alternative

The alternative of the computing space has to be drawn since the case advanced it has been realized that with the material given, and resource is not feasible to solve problems with fully integrated solution presented chapter below. Therefore, this computing space consists of additional steps (Figure 25).

Data is acquired in CSV (Comma separated value) format. It sometimes requires cleaning and refining. Purely cleaning and rearranging data in desired manner can be done using programming language Python and designed for data science library – Pandas.

If the model does not contain all information, for example relation between sensors, and rooms, or not sensors at all, it needs to be adjusted. It can be done using Dynamo for Revit, which is a visual programming specifically written for Revit and using Revit API. Developing the script allow to place objects in the model under set conditions, change parameters using just table with room-sensor relation. Model, thereafter, is translated via Forge extension in Visual Studio Code, using Model Derivative API. Refined data is placed in the code repository in IDE.

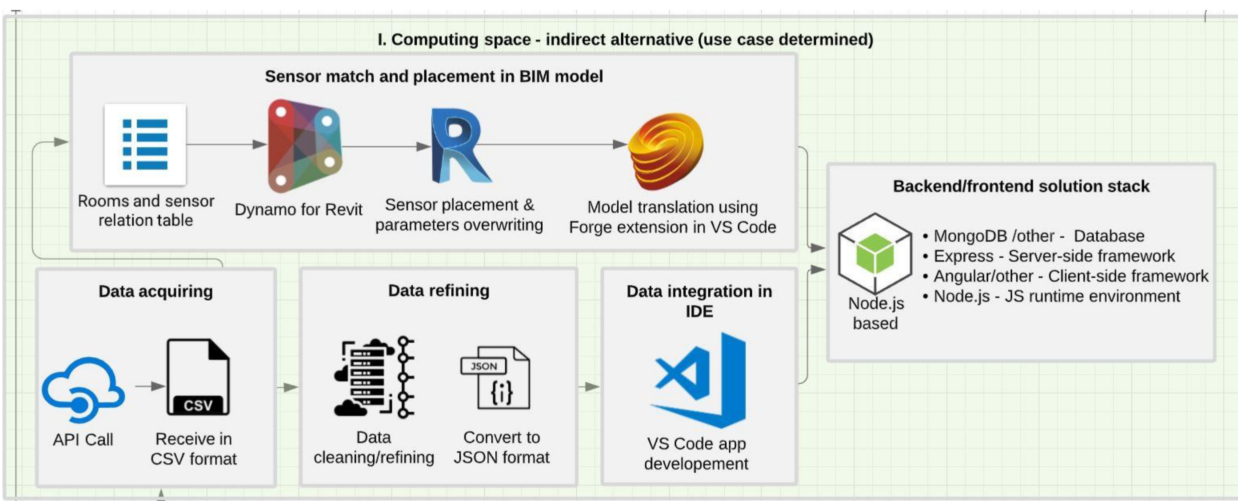


Figure 25: Computing space-indirect alternative.

6.5. Representation Space

The *Representation space* contains two layers the *server side* and the *client-side* the *server side* has components of the front-end web development tools that can project the user interface and viewer through which the end user is going to interact with the system from the *client side*. The *Representation space* further provides the developer to leverage a vast collection of open-source JavaScript based libraries which aid in data representation and web designing.

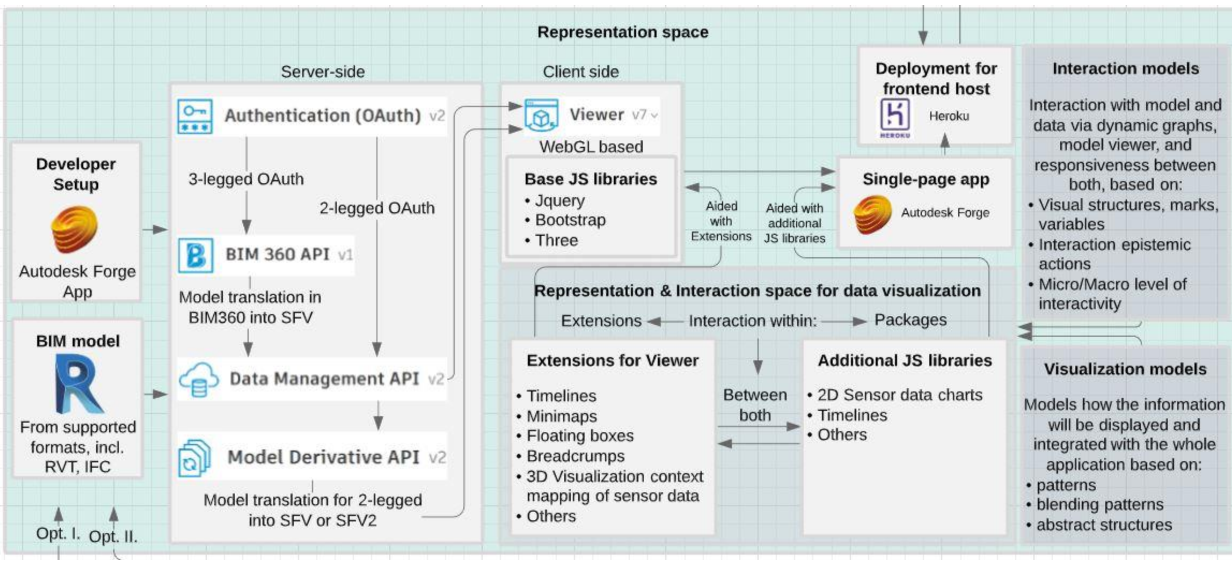


Figure 26: Representation space.

The *Representation space* is developed using Autodesk Forge API services. The Forge Platform provides APIs and services that facilitate to access and use design, engineering data by the use of the cloud (*Autodesk Forge*, n.d.-a).Forge tools enables to utilize design and engineering data to make custom applications for the AEC sector(*Home - Autodesk Forge*, n.d.). The forge services have the following features:

- Rendering 3D models in the web browser: The Viewer is WebGL based which allows to embed, control, and retrieve meta data from design files.
- Data management: The Data Management API allows to access data across A360, Fusion, and the Object Storage Service from Autodesk.
- Preparing files and data for the web: Model Derivative API aids to prepare files for the Viewer, extract geometries, retrieve metadata from multiple formats.

(*Home - Autodesk Forge*, n.d.).

6.6. Interaction Mental Space

The section consists of two spaces since they are closely connected. One influences another as seen in Figure 27, and they also flow through *Representation space* – there is no solid boundary. Interaction regards *Reaction* – user reacts to events shown in the *Representation space*; and *Action* – user leads the initiative.

The main medium in *Interaction space* are Interaction models and Visualization models powering *Representation space* (Figure 26). In the scope of whole system architecture, it acts as a bridge between *Mental* space and *Representation* space, but it is neither not lesser nor less important than the others. There are three levels of interaction – low (detailed, specific, e.g. button clicks), intermediate (broadening the abstract level of exploration, e.g. selecting, connecting, navigating), and high(complete abstract concept of reasoning and relating to mental models) (Tominski, 2015).

Mental space is more abstract is directly referred to perception, cognition and decision making. The crucial aspect is that it starts with Cognitive activity, through cognitive task, down to interactive and visual task. This space also uses mental shortcuts (heuristics). The more activities are delegated to *Representation* and *Computational* space, the less intensive can be mental activity.

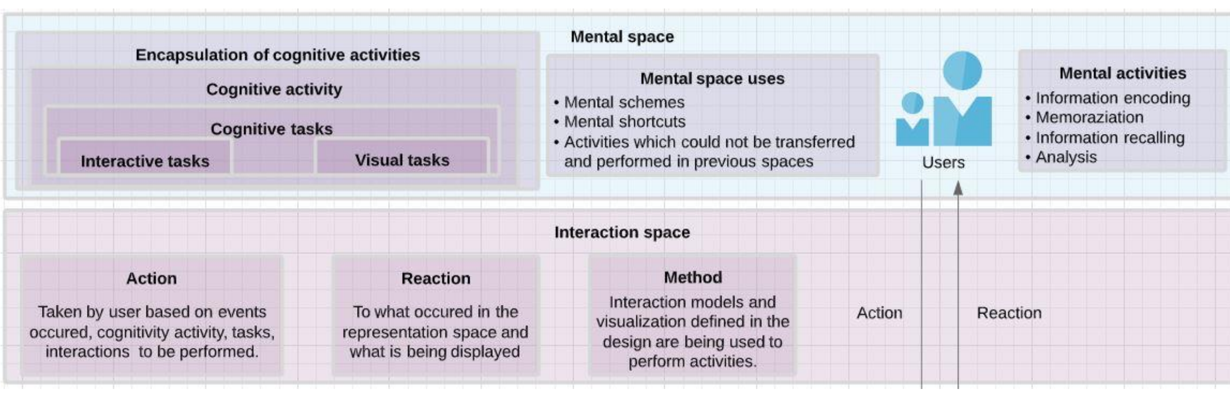


Figure 27: Interaction and Mental spaces.

7. Use cases

Two use cases were conducted for providing a test and development of the system architecture. It is a place where theory meets practice not only to see if the idea of the system workflow is feasible as the end solution, but also to describe, evaluate the steps taken to achieve desired results and discuss possible improvements.

In both use cases, the intention is to visualize sensor data of the building in real world (referent). Foremostly, it requires to have its own 3D representation in representation medium (it must be a software with compatible 3D kernels, or web-oriented library for 2D/3D rendering), and sensor data acquired from corresponding instrument in the building which can be also location in the 3D model.

In both cases, 2-legged authentication of Autodesk Forge OAuth API was utilized instead of 3-legged. The primary reason for this choice was the ability of preloading the model for debugging, troubleshooting, and running the code of application without the necessity of navigating through repository tree ("buckets" in Autodesk Forge documentation) and picking the model every time the changes in occur and the local server for the app is refreshed. Accessing developer options in BIM360 service as well as learning the new environment was another reason to choose the 2-legged solution over 3-legged. Each way has their own advantages and disadvantages, but the differences have not been thoroughly examined and compared. Some of the steps described below – regarding model, data, or code – are predetermined by choice of using the 2-legged authentication, and the focus is to illustrate actions taken, report findings based on that factor.

7.1. Use case 1

In the first use case, the BIM model of one floor area of a building called BloxHub located in Copenhagen was utilized to prototype, each space maintains a wall mounted wireless sensor capturing different values of the indoor environment– temperature, humidity, Co2, motion etc. The data contained historical sensor data values in CSV format to begin the prototype with the visualization. Both accommodate real-word data– 3D representation of the building and time series of indoor climate values. In the BIM model the sensor objects were already accurately included in the properties with embedded Room IDs so creating a relation to the values and room names from the CSV was not required which would be essential later in the IDE as shown in Figure 28.

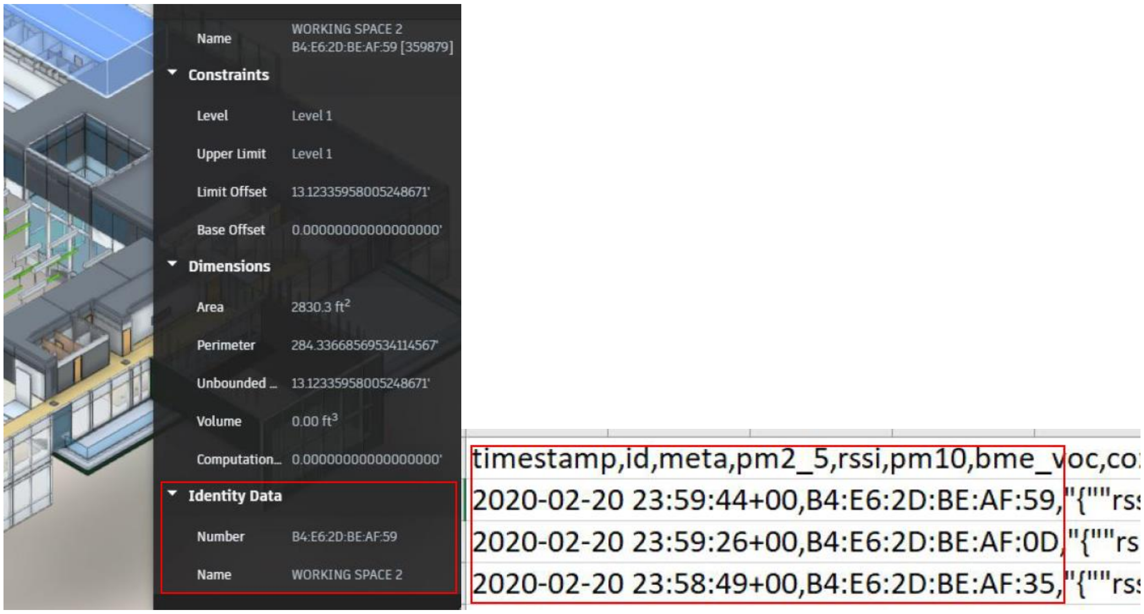


Figure 28: Room object metadata in the BIM model and relation in CSV file.

As the CSV data was accurately related to the room objects of the BIM model, it was possible to directly transfer the data to the IDE environment in JSON format from the CSV. The JSON data file containing sensor values and time stamps was imported into the IDE (Integrated Development Environment), VS code (Visual studio code) was used as the ideal IDE. The next step was to translate the BIM model to be represented in the Forge Viewer. BIM models in RVT format or other 3D formats can be translated into SVF format for viewing in the *Chrome* web browser using the forge extension for VS code. One important aspect during translation from RVT to SVF files is to generate it through the SVF2 format and use the *"Generate Master views"* options, this makes sure that the designated rooms and properties are included as well (Autodesk Forge, n.d.-b). Otherwise only room boundaries are generated. A snip shot of Forge extension in VS code for the translation options is shown in Figure 29. Furthermore, a custom URN is generated of the model. It would be used later to embed and identify the model from the forge viewer API.

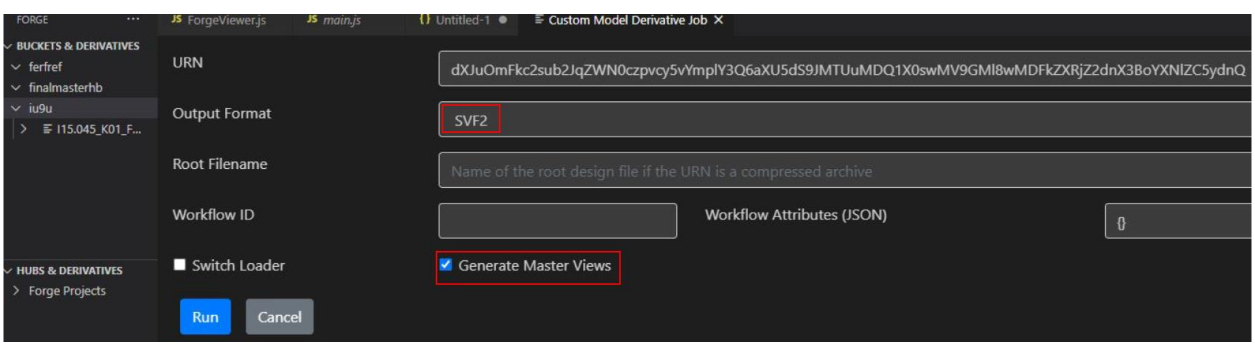


Figure 29: Options during translating BIM models to SVF format for the web browser.

For the development of the MEAN stack for Backend/front end (Computing space) a file structure is necessary and different packages of the components are required to be installed through the IDE to develop a real time connection to the web and running, debugging, and testing the code. Following the methods are described.

7.1.1. Accessing the APIs and user authentication

Foremost, to access the Forge APIs an application needs to be created through the Autodesk Forge portal; this provides the developer with the environment variables required for authentication. While authenticating the application internal tokens and public tokens are created. Internal tokens are to be maintained at the server side and public tokens is required by the forge viewer API for viewing access at the front-end (*OAuth - Autodesk Forge*, n.d.).

7.1.2. Viewing the model on the browser

For viewing the model on the browser and setting up the environment on the web a designated web server is required to be created. This is achieved directly in the IDE as MEAN stack is being utilized, Node.js was used for the connection. Further, *npm* packages are needed to be installed specifically *npm* packages for express, *multer* for file uploads and forge API packages. Through these steps a package JSON file is created with references in *start.js, launch.js* and *config.js* which contain components to run and debug the models accurately in the browser through the viewer API. For accessing the model derivative API buckets need to be created extract the geometry and metadata of the BIM model. For the client-side viewing and custom web design a HTML component and CSS file containing the script is required.

7.1.3. Creating custom UI extensions and querying values

For each values of the sensor data a button was created for the user to select to visualize the needed data with designated color schemes and accurately relating the values to the specific spaces. Extensions deliver a method to create custom code that interacts with the viewer. Each extension requires a skeleton which is coded in a separate JavaScript file. Loading the extension to the UI and adding custom symbols or design can done using CSS or referencing the JPEG/PNG in the HTML script (Viewer Extension - Autodesk Forge, n.d.). Each extension encompasses a basic skeleton file which contains CSS styling and *main.js* where methods. Functions, variables are scripted for the custom extension. The files structure in VS code is depicted in Figure 30.

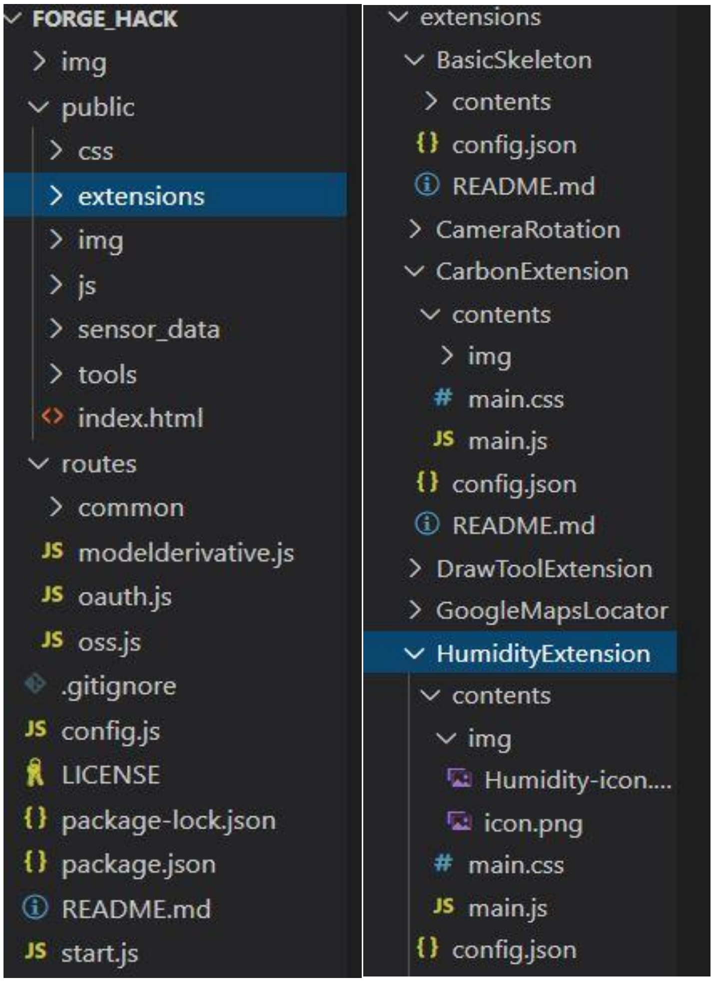


Figure 30: File structure in VS code.

For this use case the sensor data of just 24 hours was used to tweak the data and test the prototype. The sensor data values are stored in the IDE as a JSON file under the extension's loader in key value pairs as shown in Figure 31.

```
oneday = [{
        "timestamp": "2020-02-20 23:59:44+00",
        "id": "B4:E6:2D:BE:AF:59",
        "CO2": 442,
        "temperature": 23.1,
        "pressure": 996.56,
        "humidity": 31.047
    },{
        "timestamp": "2020-02-20 23:59:26+00",
        "id": "B4:E6:2D:BE:AF:0D",
        "CO2": 413,
        "temperature": 22,
        "pressure": 996.52,
        "humidity": 33.602
    },{
```

Figure 31: Sensor data stored in JSON in VS code.

To tweak the sensor data and assigning color schemes three variables were created assigned to the color vectors to three different color densities according to the value ranges as shown in Figure 32.

```
var red = new THREE.Vector4(1, 0.8, 0.8, 1);
var red2 = new THREE.Vector4(1, 0.6, 0.6, 1);
var red3 = new THREE.Vector4(1, 0, 0, 1);
```

Figure 32: Variables created of different color densities.

Let variables were generated to query the room id arrays from the sensor data json file Figure 33.

```
let viewState = viewer.getState();

let lookUp = () => {
    let rooms = [];// Array to store Room Objects

    let room;// be an Room Object
    oneday.forEach((row)=>{
        room=new Room();// New Room Object
        Object.assign(room,row);// Assign json to the new Room
        rooms.push(room);// Add the Room to the Array
    });

    let roomOne = []; // "B4:E6:2D:BE:AF:0D"
    let roomTwo = []; // "B4:E6:2D:BF:0D:35"
    let roomThree = []; // "B4:E6:2D:BE:AF:59"

    rooms.forEach((rr) => {
        if(rr.id == "B4:E6:2D:BE:AF:0D")
            roomOne.push(checkTemperature(rr.temperature));
        else if(rr.id == "B4:E6:2D:BF:0D:35")
            roomTwo.push(checkTemperature(rr.temperature));
        else
            roomThree.push(checkTemperature(rr.temperature));
    });
```

Figure 33: Code snippet of querying room ids and assigning color gradients to sensor data value.

If and *Else* functions were assigned to check the values for different value ranges if temperature values are below 20 degrees, it returns lighter gradient of red if higher then red 2 and temperature values higher than 23 the room displays red 3 which is the darkest gradient as shown in Figure 34.

```
function checkTemperature(temp) {
    if(temp > 20 && temp < 22) {
        return red
    }
    else if(temp >= 22 && temp < 23) {
        return red2
    }
    else {
        return red3
    }
}

Autodesk.Viewing.theExtensionManager.registerExtension('TemperatureExtension',
    TemperatureExtension);
```

Figure 34: Utilizing If and Else functions.

To set the speed of the data stream for the color variables a timer function was created as seen in the code snippet in Figure 35, and it was needed to refer it to the color coding and the room ids. The speed can be edited as per the requirement.

```
let timer = setInterval(function () {
    let tempColor;

    if(roomOne.length == 0 && roomTwo.length == 0 && roomThree.length == 0 ){
        clearInterval(timer);
    }

    if (started != false && roomOne.length > 0) {
        tempColor = roomOne.pop();
        viewer.model.setThemingColor(7818, tempColor);
        viewer.model.setThemingColor(7807, tempColor);
        viewer.impl.invalidate(true);
    }
    if (started != false && roomTwo.length > 0) {
        tempColor = roomOne.pop();
        viewer.model.setThemingColor(7819, tempColor);
        viewer.model.setThemingColor(7792, tempColor);
        viewer.model.setThemingColor(7805, tempColor);
        viewer.model.setThemingColor(3748, tempColor);
        viewer.impl.invalidate(true);
    }
    if (started != false && roomThree.length > 0) {
        tempColor = roomThree.pop();
        viewer.model.setThemingColor(7830, tempColor);
        viewer.model.setThemingColor(7799, tempColor);
        viewer.model.setThemingColor(7793, tempColor);
        viewer.impl.invalidate(true);
    }
}, 2000);
```

Figure 35: Setting speed for the color variable with the room ID.

A room class was created referring to the earlier created room arrays and methods to set and return time stamps ids from the sensor data values as shown in Figure 36.

```
class Room {
    set timestamp(timestamp) {
        this._timestamp=timestamp;
    }
    set id(id) {
        this._id=id;
    }
    set CO2(CO2) {
        this._CO2=CO2;
    }
    set temperature(temperature) {
        this._temperature=temperature;
    }
    set pressure(pressure) {
        this._pressure=pressure;
    }
    set humidity(humidity) {
        this._humidity=humidity;
    }
    get timestamp() {
        return this._timestamp;
    }
    get id() {
        return this._id;
    }
    get CO2() {
        return this._CO2;
    }
    get temperature() {
        return this._temperature;
    }
    get pressure() {
        return this._pressure;
    }
    get humidity() {
        return this._humidity;
    }
    constructor() {
    }
}
```

Figure 36: Created Room class to get and return sensor data values.

7.1.4. User interface of the web application

The first prototype of user interface of the web application is shown in Figure 37. The forge viewer API makes it possible to display the translated model and the reference to the URN selects the precise file derived while the translation process to the SVF format. The viewer API has basic toggle options to rotate the model and select elements. The extension tool panels contain two options to view custom and built-in extensions. On ticking the tabs, it adds the extension selected in extension bar at the bottom in the viewer which further creates a window of the extension on click. This proved important as it reduces confusion during viewing models and keeps the UI clean and precise. On the right a data stream dashboard is available which displays the sensor data values and the time from the *sensordata.json* file in the form of a bar chart. This is currently under development as it shows static values currently but would need more development to show dynamic data and onclick options from selecting the rooms/spaces directly in the viewer.

Figure 37: User interface of the web application.

There are three extensions to run the visualization of the selected values from the sensor data i.e., temperature, Co2 and humidity. The icons were added through the HTML script for the front end of the UI as seen at the bottom right in the extension bar panel.

Selecting the temperature extension icon activates it to show different color gradients according to the different temperature values of the rooms as shown in Figure 38. Here it can be observed by the users that the two areas with darker gradient of red has temperature higher than 23 degrees. This allows the user to understand the discrepancies in the data and the specific areas relating to it and make comparisons. Hence, contextualizing the indoor environment data from sensors to the BIM model.

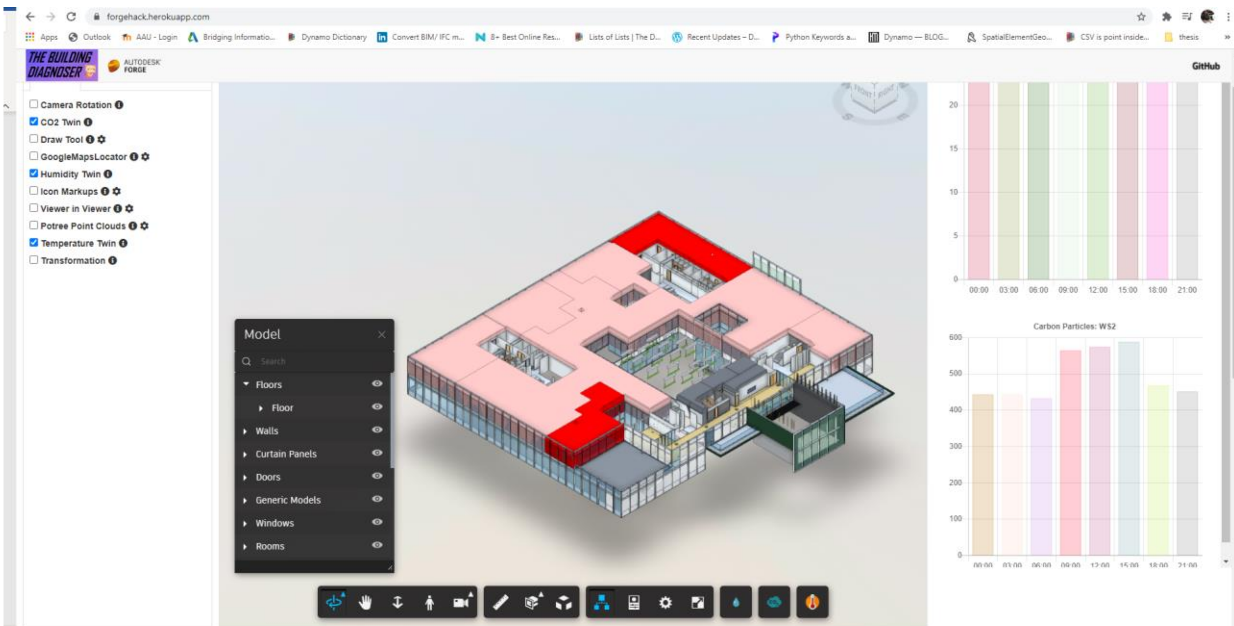


Figure 38: Color schemes for different values based on values of temperature when the temperature extension is activated.

Furthermore, specific rooms can be selected to view the meta data in model browser extension and isolating the room to view data of just one space as shown in Figure 39.

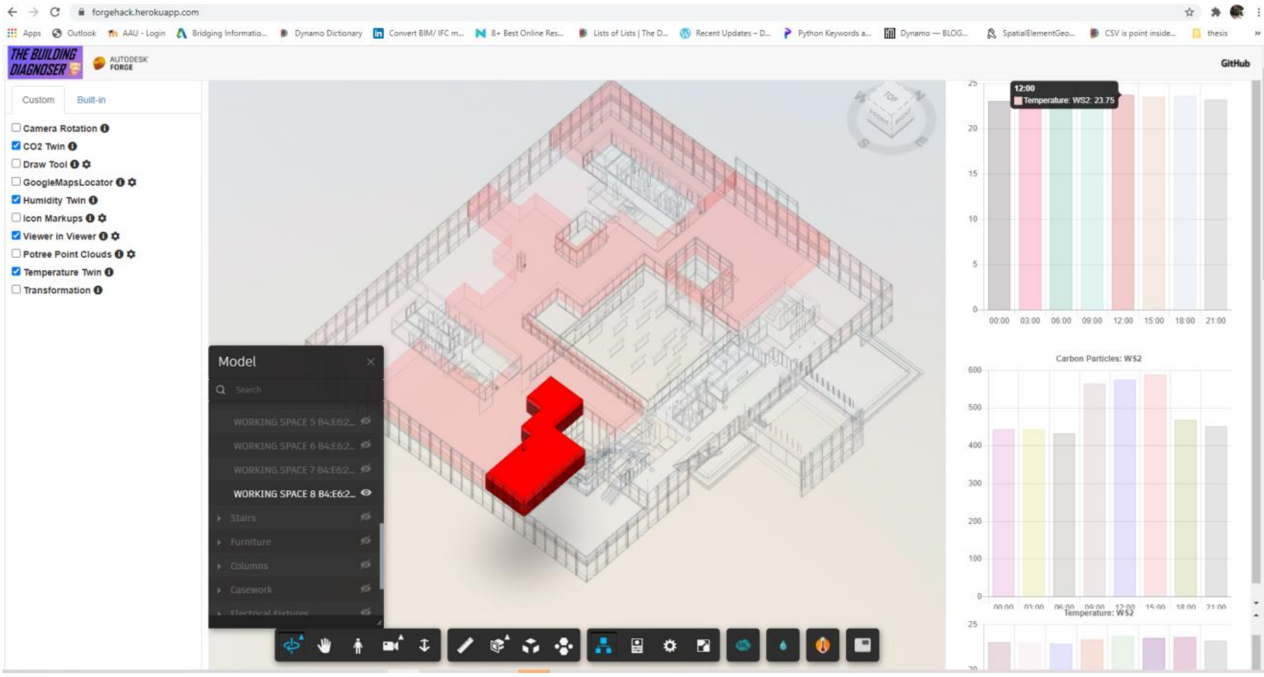


Figure 39: On click Temperature extension and specific room area.

Similarly, the other two extensions Co2 and humidity are activated on click and the visualization animation runs the specific data for one day. Each value is depicted with a different color gradient of green for humidity and blue for Co2 as shown in Figure 40

Figure 40 Web application displaying humidity values related to spaces.

7.2. Use case 2

The first use case gave an idea how to visualize sensor data linked to the sensor in BIM model. For the second case, the intention was to replicate the process with another model. In this way it was possible to see what steps could be replicated in any kind of scenario, what were useful only for the singular case, and what improvements could be made to address modularity, scalability, and universality of the solution.

7.2.1. Acquired artifacts for work

Complete set of artifacts were given by MOE – energy consultant company – to be able to proceed with the second use case, i.e., BIM model and sensor data. The first thing done was to examine if they are set-up sufficiently and consist of everything necessary to execute data visualization app with a presented model. Adjustments were done after issues emerged during the evaluation of the artifacts.

For visualizing sensor data in the context of the 3D representation of the building are necessary mostly two artifacts– BIM model and corresponding sensor data. Both are going to be evaluated if they consist of everything it is needed for the whole procedure. Therefore, there will be a listing of preconditions to be met, actual conditions of the given model and data, and which must be resolved to satisfy the set-up ready for the operation.

7.2.2. BIM model

The BIM model provided by MOE was in RVT file extension– format of proprietary software Revit, governed by Autodesk. Revit helps to model 3D representation of the future or existing referent with incorporated meta-information and properties. It is not the only option for BIM modelling, but it is the format which can be hardly ignored while developing tools for design and construction operation as of the time being.

The model itself is from project of "Healthy Homes" where MOE in collaboration with Real Dania have set-up sensors (IC-meters) in 3 existing apartments for measurement of indoor climate to help understanding and evaluating the indoor climate change overtime and how to share gathered data in accessible way with other stakeholders, not only with graphs but also with a support of 3D model as a context background and reference.

Learned from the experience on the previous case, there were a few aspects which had to be checked before uploading the model into the viewer and some new emerged only in the second, which are listed as follows:

- Construction phase set-up

In some situations, e.g., for refurbishment, there is a need for distinguishing between existing construction and a one to be built. This model contained both *Existing* and *New Construction*. To display the model on the web, it is necessary to translate it as SVF or SVF2 (Autodesk Forge, n.d.-b). Moreover, to display rooms in the list of objects and in the 3D representation it is obligatory generate *Master Views* as they will make rooms visible, which is a default option in BIM360; whereas the default translation of the basic Forge skeleton does not have the option of creating the views enabled (Vandecar, 2019). Therefore, it must be done with Forge extension for Visual Studio Code (Figure 29). Master views are generated for each phase of the project after translation.

During the initial approaches of translating the model as mentioned in the previous example and displayed on Figure 29, the results were unsuccessful, and the model could not be displayed in the Forge Viewer. After some investigation, the issue was cleared with reducing the model to only one phase– *New Construction*. The reason why it had to be this way has not been yet fully resolved in those experiments. However, it is noted that some example models among the Forge developers' community were successful with having multiple viewable (different phases) exported and the only matter was to switch between different phases. However, it was not accomplished, and it was necessary to rely on changing all rooms phases to the *New Construction*.

- Central file aspect

The provided model was given as central file, which enables the collaboration between project participants among different professions. Hence, the nature of central file and its influence of performing changing of phases was another point of consideration. To prevent the situation of inability of alternating the phase due to unrelinquished items in work sets (i.e., when somebody else is the "owner" of the item and has not relinquish its ownership before closing the file), it was decided to detach the file from the central model and to discard all work sets.

- Room and sensor identification

The model was provided with room names and numbers (embedded in Revit Room properties), sensor data (CSV files), and corresponding list of relation between rooms (Excel (XLSX) table with room name, number, and related sensor ID). There was no information about the evaluation of the room information consisted of: 1) which rooms in Revit are listed in the Excel table, 2) which rooms are omitted, 3) is there any anomalies/errors with the listing.

First two steps were done relying on visual feedback using crosscheck method– to go through the Excel list of all rooms and to check if some rooms listed are unattended (XLSX =>RVT); and to do reverse process, i.e., check rooms in Revit and to check if some rooms are not listed in Excel file (RVT => XLSX). There will be another way to verify relations of rooms during the process of placing sensors in corresponding rooms using Dynamo for Revit and issuing additional parameters which will be useful for matching sensor data with the room according to room number and sensor ID in future Forge app. Last, but not least is to check if there were any records of information which does not seem right and might be even invalid using also visual feedback. It was found out that in the Excel list were two sensors linked to one room. It is not particularly invalid, since there can be allocated multiple sensors in a single space (many-to-one relationship), especially if the room, area is big enough to have multiple sensor or user requirements or scientific, research purposes demand it, but the provided information about project said it was a single sensor per room. In order to make sure that no information is wrongfully deleted– because another instrument can be installed for some reason (such as to double check data, to uniform values, to check sensor accuracy, etc.), data from those two sensors was checked. One had recorded information in the same manner as spotted in other sensors, but the another was empty. Only after this procedure it was safe to assume that no crucial information was held. More information about cleaning and preparing the data for visualization will be described in the next section.

7.2.3. Sensor data - Revit

In sensor data there are two aspects to be taken into account– data itself and also the relationship between room and sensor. The list of Rooms and corresponding sensors were given in XLSX format. The challenge was to place sensor in the mentioned rooms and attach correct sensor ID to each room. It was done using Revit for Dynamo (whole script in Appendix II).

Breaking down the steps the procedure goes as follows:

1. Import data from Excel.
2. Clean items and get only Room Name, Room ID, Sensor ID.
3. Prepare dictionary – to preserve list order and match it with Revit room order.
4. Get all rooms from Revit with parameter Name and Number.
5. Filter rooms according to list from Excel.
6. Sort the rooms by number parameter.
7. Get room boundaries.
8. Filter walls to only internal – where the sensor is going to be placed.
9. Filter out walls with null surfaces.
10. Get only 1 wall and 1 surface.
11. Generate sensor.
12. Assign parameters with the list from Excel from step 3.

The main reasons for Dynamo scrip development are:

- Importing room list from Excel (Figure 41) – it has to be ensured that there is same data to work with (Step 1.).
- Arranging data in dictionaries (Figure 42) – in order to match the list of rooms generated by Dynamo (Step 3.).
- Placing sensor and setting parameters into room (Figure 43) – to create relationship via sensor object (Step 11. and 12.).

Figure 41: Importing room and sensor list to Dynamo script.

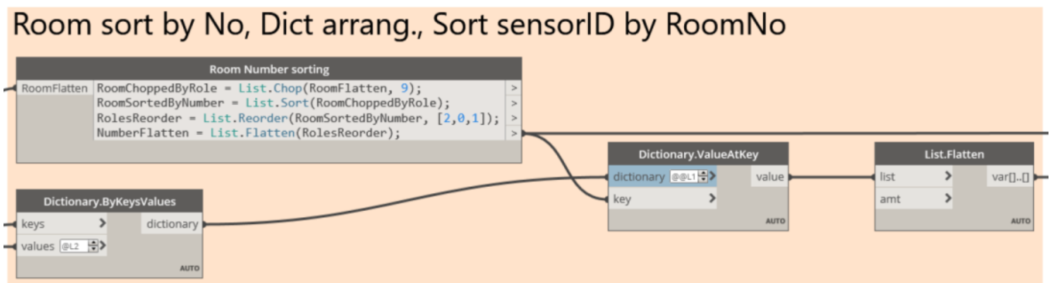


Figure 42: Creating dictionary to preserve list order.

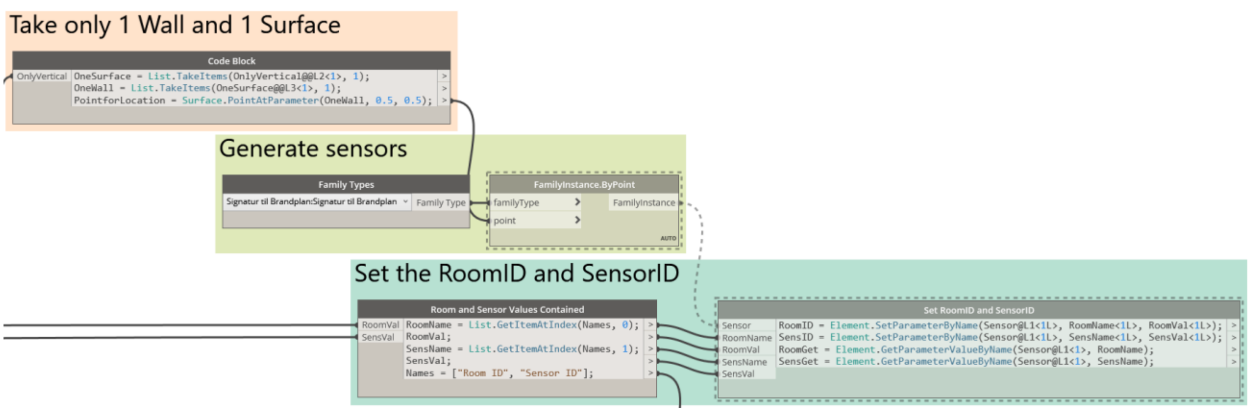


Figure 43. Generating sensor by Family with automatically selected wall and setting parameters.

7.2.4. Sensor data – Python

Sensor data provided had to be cleaned, most headings of columns were empty for the reasons unknown. With one CSV file it was sufficient to use Power Query Editor in Excel– it manages the structure of CSV and helps to reorganized. However, with bulk number of CSV files it is getting tougher, which was around 30.

It was decided to use Python and Pandas library to clean, refine data and ultimately import to JSON as this format turned out in the *Use case 1* to be flexible when it comes to connection of data with model in Forge Viewer.

The order of processes which had to occur is following:

1. Import libraries.
2. Getting the path of all CSV files in the directory.
3. Creating numbers of arrays which will be filled later.
4. Running *for* loop for all CSV files.
5. Creating data frames from each file.
6. Removing files which have fewer than 3 lines (number of lines is arbitrary).
7. Getting room names and splitting to only have room number.
8. Append room numbers and sensor id to the list which will be a new column list.
9. Running another *for* loop for remaining file.
10. Adding new columns (room number, sensor id) at the beginning of table.
11. Renaming column names.
12. Reversing rows order
13. Changing data type for each column.
14. Merging all data frames into one.
15. Exporting to JSON.

The outcome is arranged in "records" but can be view as a standard object tree.


```
{
    "id": "Yes 1.09",
    "sensor_id": 5324,
    "timestamp": "2019-12-31T19:19:54.000Z",
    "temperature": 24.90625,
    "humidity": 37.53125,
    "CO2": 442.0,
    "noise average": 32.0,
    "noise peak": 32.0,
    "particle sensor": 145.0
},
{
    "id": "Yes 1.09",
    "sensor_id": 5324,
    "timestamp": "2019-12-31T19:24:53.000Z",
    "temperature": 24.90625,
    "humidity": 37.5,
    "CO2": 453.0,
    "noise average": 32.1875,
    "noise peak": 32.1875,
    "particle sensor": 147.0
},
```

```
1 {"id":"Yes 1.09","sensor_id":5324,"timestamp":"2019-01-
2 {"id":"Yes 1.09","sensor_id":5324,"timestamp":"2019-01-
3 {"id":"Yes 1.09","sensor_id":5324,"timestamp":"2019-01-
4 {"id":"Yes 1.09","sensor_id":5324,"timestamp":"2019-01-
5 {"id":"Yes 1.09","sensor_id":5324,"timestamp":"2019-01-
6 {"id":"Yes 1.09","sensor_id":5324,"timestamp":"2019-01-
```


Figure 44: Data in final JSON format – tree and records.

The key steps were 4-8 where room numbers were obtained (Figure 45).

```python
for file in all_files:
    temp = pd.read_csv(file, sep=';')
    if temp.shape[0] < 3:
        invalid_csv.append(file)
        all_files.remove(file)
        #dfs.remove(files)
        print(' not valid csv', file)
   # info = temp.info(verbose=True)
    temp_column_name = list(temp.columns)

    room_name = temp_column_name[2]
    room_split = room_name.split()
    room_connect = room_split[0] + ' ' + room_split[1]
    room_number.append(room_connect)
```

Figure 45: Loop through all CSV files and obtaining only room number without room names (Steps 4-8)

Final procedure of setting explicitly type of data was important since using in built converter number were not always recognized and put as string Figure 46. The export itself had to be specified (e.g., orient as "records", date format as "iso", and *lines=True* assures that each record is in separated line).

```python
    dfs = dfs.astype({
                    'id' : 'string',
                    'sensor_id' : 'int16',
                    'timestamp' : 'datetime64[ns]',
                    'temperature' : 'float16',
                    'humidity': 'float16',
                    'CO2': 'float16',
                    'noise average': 'float16',
                    'noise peak': 'float16',
                    'particle sensor': 'float16'
                    })

    dfs = dfs[::-1]

    frames.append(dfs)

result = pd.concat(frames)

tojson = result.to_json(orient="records", date_format="iso")
parsed = json.loads(tojson)
tojson_file = result.to_json(path_or_buf="to_print.json", orient="records",
                             date_format="iso", lines=True)
print(json.dumps(parsed, indent=4))
```

Figure 46: Data type column change and final export to JSON.

7.2.5. Data in charts

Importing data and visualizing into was not fully successful, but the achievement managed was to create short JavaScript (Figure 47) file which parses and arranges CSV data to be displayed by Chart.js library (Figure 48).

The code fetches only two columns [2,3], which are temperature and date. Later, global variable from the code is imported to the chart script.

```
export const xAxis = []
export const yAxis = []
    getData();
    async function getData() {
        const response = await fetch('sensor2.csv');
        const data = await response.text();

        const table = data.split('\n').slice(1);
        table.forEach(row =>{
            const columns = row.split(',');
            const date = columns[2];
            xAxis.push(date);
            const temperature = columns[3];
            yAxis.push(temperature);
            console.log(date, temperature);
        });
    }
```

Figure 47: JavaScript code for parsing CSV data series.

Data is displayed, however, in reverse order. This JavaScript file was written before work with CSV through Python. In the Python code it was ensured to reverse order date.

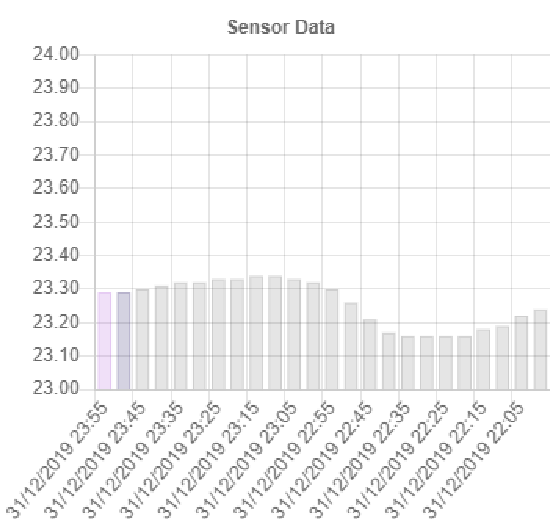


Figure 48: Visualized data series via Chart.js

8. Discussion

The visualization needs of BIM and IoT data integration with BIM led to explore the latest efforts put forward by researchers as well as the needs of the industry professionals in real world projects. The continuous prototyping sessions through a methodological approach helped to recognize and breakdown the needs of the user and processes. A technical approach was established through testing and prototyping the system architecture. The authors were successful in developing a basic prototype and structure for making a web application incorporating a UI and sensor data streams from IoT devices with contextualization of BIM. However, moving from the one case to other authors found other difficulties regarding incorporation data sets with model. Therefore, robust and more structured way from technical side is needed to ensure that the approach of data visualization process is replicable, reusable and scalable without extensive project curating. Furthermore, the needs of preparing and structuring the meta data of BIM models plus refining, relating the data from the sensors for such deployments in an automated way due to the large data sets was examined through the use cases. Developing the UI for every user need is an ongoing process, although the system architecture provides a scope for extending it to add such utilities through the proposed web-based framework. The fact that it is web based allows the usage of numerous APIs and resources like JS libraries readily available to be added coherently. Furthermore, visualization of sensor data can be done in several ways and possibly easier and better for example through embedded tools like dynamo available in Revit, but the reason for a web-based solution was adopted was to keep it in sync with the broader perspective of cyber physical systems approach and digital twin paradigms technical infrastructure so it can be extended to serve the needs in the future.

The user needs, and processes documented will help create a user-oriented application. Through the visioning process it was possible to identify the different needs of multiple users participating in construction project. The primary user for this project report was the building energy consultant but the solution was envisioned to extend to cater the needs of other participants as well as described in the visioning diagram.

9. Conclusion

The purpose of the project report was to develop a data visualization and user interaction framework aligned to the broader prospect of developing a construction digital twin. From the literature review and interviews conducted with the industry professional it was validated that there is a need to cater to the visualization demands for different end users as well as to the digital twin paradigm, respectively. The interpretation sessions and contextual design methods assisted us to identify the user processes through work modelling and the need to contextualize BIM and indoor environment data providing them with a visual interface to foster collaboration and understanding of information between project stakeholders. There are multiple aspects and specific knowledge areas which were taken into consideration for developing the solution, from refining and structuring of the data from the sensors to development of a web-based user interface. Although the proposed framework would serve as a concept to develop a holistic solution further, and the interface would require much more work modelling sessions to attain feedback and development to achieve a complete solution catering to all the user needs and visualization needs of project stakeholders. The developed use case is validated based on the practical working of the basic web application based on the proposed system architecture. Furthermore, the fact that it is developed aligning to the digital twin approach is beneficial for the future development of a holistic solution. Therefore, after the efforts put towards answering the main question:

"How to contextualize building sensor data with BIM (Building information models) to create advanced visualization techniques according to user requirements?"

The authors of this project report were able to establish the needed technical tools and consolidate it in a system architecture adhering to multiple needs and aspects required for it. The conducted research and methods used to approach the problem statement into specific spaces which eventually lead to the roots of it, also enabled us to develop possible solutions for it. Furthermore, it aided to identify the technical needs, preparation of BIM models to catering to such needs in construction projects.

10. Future work

Developing an application is a continuous process and cannot be achieved in the time span that was allotted for this project report. Therefore, it was realized there were much more needs from the user that were recognized during work modelling and prototyping sessions. Furthermore, adhering to those needs observed through the work modelling and prototyping sessions and further extending the application to include multiple aspects from the digital twin paradigm, the following list mentions the topics for future works:

1. Connecting a real time data stream with Mongo stich to the IoT device.
2. Actuating components such as HVAC systems.
3. Running multiple data values together with custom shaders.
4. Comparison/twin viewer.
5. User Sign in options and direct model uploads on the web application.
6. Report generator.
7. 2D mini maps on click object selection in 3D view.
8. Dynamic graphs on room object selection.
9. Better visualized graphs, heatmaps and data streaming tools.
10. Improvement to proposed framework.

Furthermore, more input and ideas would be generated with more prototyping sessions, testing with different project participants and feedback. Following is the repository link to view the entire code and run the web application developed until now –
git clone - https://github.com/qwe0254/Forge_Hack.git

11. References

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183–198. https://doi.org/10.1016/j.learninstruc.2006.03.001

Akanmu, A., & Anumba, C. J. (2015). Cyber-physical systems integration of building information models and the physical construction. *Engineering, Construction and Architectural Management*, *22*(5), 516–535. https://doi.org/10.1108/ECAM-07-2014-0097

Akanmu, A., Anumba, C., & Messner, J. (2013). Scenarios for cyber-physical systems integration in construction. *Journal of Information Technology in Construction (ITcon)*, *18*(12), 240–260.

Autodesk Forge. (n.d.-a). Autodesk Forge. Retrieved 3 January 2021, from https://forge.autodesk.com/

Autodesk Forge. (n.d.-b). *Model Derivative—Translate a Revit File, Generating Room and Space Information*. Forge Docs. Retrieved 4 January 2021, from https://forge.autodesk.com/en/docs/model-derivative/v2/tutorials/prep-roominfo4viewer/about-this-tutorial/

Battiato, S., Blasi, G. D., Gallo, G., Nicotra, S., & Messina, G. (2005). *SVG Rendering of Digital Images: An Overview*. 3.

Beyer, H., & Holtzblatt, K. (1997). *Contextual Design: Defining Customer-Centered Systems*. Morgan Kaufmann Publishers Inc.

Boje, C., Guerriero, A., Kubicki, S., & Rezgui, Y. (2020). Towards a semantic Construction Digital Twin: Directions for future research. *Automation in Construction*, *114*, 103179. https://doi.org/10.1016/j.autcon.2020.103179

Burkhard, R. A. (2004). Learning from architects: The difference between knowledge visualization and information visualization. *Proceedings. Eighth International Conference on Information Visualisation, 2004. IV 2004.*, 519–524. https://doi.org/10.1109/IV.2004.1320194

Chang, C. J., & Luo, Y. (2019). Data visualization and cognitive biases in audits. *Managerial Auditing Journal*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/MAJ-08-2017-1637

Chang, K.-M., Dzeng, R.-J., & Wu, Y.-J. (2018). An Automated IoT Visualization BIM Platform for Decision Support in Facilities Management. *Applied Sciences*, *8*(7), 1086. https://doi.org/10.3390/app8071086

Chen, Y.-H., Tsai, M.-H., Kang, S.-C., & Liu, C.-W. (2012). *SELECTION AND EVALUATION OF COLOR SCHEME FOR 4D CONSTRUCTION MODELS*. 19.

Costa, B., & Pires, P. (2020). Evaluating a Representational State Transfer (REST) Architecture. *Anais Do Concurso de Teses e Dissertações Da SBC (CTD-SBC)*, 55–60. https://doi.org/10.5753/ctd.2015.10002

Dave, B., Buda, A., Nurminen, A., & Främling, K. (2018). A framework for integrating BIM and IoT through open standards. *Automation in Construction*, *95*, 35–45. https://doi.org/10.1016/j.autcon.2018.07.022

Ding, L., Zhou, Y., & Akinci, B. (2014). Building Information Modeling (BIM) application framework: The process of expanding from 3D to computable nD. *Automation in Construction*, *46*, 82–93. https://doi.org/10.1016/j.autcon.2014.04.009

Dou, S. Q., Zhang, H. H., Zhao, Y. Q., Wang, A. M., Xiong, Y. T., & Zuo, J. M. (2020). RESEARCH ON CONSTRUCTION OF SPATIO-TEMPORAL DATA VISUALIZATION PLATFORM FOR GIS AND BIM FUSION. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-3/W10*, 555–563. https://doi.org/10.5194/isprs-archives-XLII-3-W10-555-2020

Duboc, L., Letier, E., & Rosenblum, D. S. (2013). Systematic Elaboration of Scalability Requirements through Goal-Obstacle Analysis. *IEEE Transactions on Software Engineering*, *39*(1), 119–140. https://doi.org/10.1109/TSE.2012.12

Duboc, L., Rosenblum, D., & Wicks, T. (2007). A framework for characterization and analysis of software system scalability. *Proceedings of the the 6th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on The Foundations of Software Engineering - ESEC-FSE '07*, 375. https://doi.org/10.1145/1287624.1287679

Elstein, A. S. (1999). Heuristics and biases: Selected errors in clinical reasoning. *Academic Medicine: Journal of the Association of American Medical Colleges*, *74*(7), 791–794. https://doi.org/10.1097/00001888-199907000-00012

Fawzi, H., Tabuada, P., & Diggavi, S. (2014). Secure Estimation and Control for Cyber-Physical Systems Under Adversarial Attacks. *IEEE Transactions on Automatic Control*, *59*(6), 1454–1467. https://doi.org/10.1109/TAC.2014.2303233

Fielding, R. T. (2000). *In Information and Computer Science*. 180.

Fisher, B., Green, T. M., & Arias-Hernández, R. (2011). Visual Analytics as a Translational Cognitive Science. *Topics in Cognitive Science*, *3*(3), 609–625. https://doi.org/10.1111/j.1756-8765.2011.01148.x

Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology*, *13*(1), 117. https://doi.org/10.1186/1471-2288-13-117

Garett, R., Chiu, J., Zhang, L., & Young, S. D. (2016). A Literature Review: Website Design and User Engagement. *Online Journal of Communication and Media Technologies*, *6*(3). https://doi.org/10.29333/ojcmt/2556

Garrett, J. J. (2005). *Ajax: A New Approach to Web Applications*. 5.

Gigerenzer, G., & Gaissmaier, W. (2010). Heuristic Decision Making. *Annual Review of Psychology*, *62*(1), 451–482. https://doi.org/10.1146/annurev-psych-120709-145346

Golparvar-Fard, M., Peña-Mora, F., & Savarese, S. (2011). Integrated Sequential As-Built and As-Planned Representation with D4AR Tools in Support of Decision-Making Tasks in the AEC/FM Industry. *Journal of Construction Engineering and Management*, *137*(12), 1099–1116. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000371

Hegarty, M. (2011). The Cognitive Science of Visual-Spatial Displays: Implications for Design. *Topics in Cognitive Science*, *3*(3), 446–474. https://doi.org/10.1111/j.1756-8765.2011.01150.x

Hochhalter, J. D., Leser, W. P., Newman, J. A., Glaessgen, E. H., Gupta, V. K., & Yamakov, V. (n.d.). *Coupling Damage-Sensing Particles to the Digitial Twin Concept*. 15.

Holtzblatt, K., & Beyer, H. (2017). *Contextual Design: Design for Life.* (Second). Morgan Kaufmann Publishers Inc. https://www.sciencedirect.com/book/9780128008942/contextual-design#book-info

Holtzblatt, K., & Beyer, H. R. (2014). Contextual Design. In *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* (2nd ed., Vol. 1). Interaction Design Foundation. https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/contextual-design

Home—Autodesk Forge. (n.d.). Retrieved 3 January 2021, from https://learnforge.autodesk.io/#/?id=learn-autodesk-forge

Howell, S., Rezgui, Y., Hippolyte, J.-L., Jayan, B., & Li, H. (2017). Towards the next generation of smart grids: Semantic and holonic multi-agent management of distributed energy resources. *Renewable and Sustainable Energy Reviews*, *77*, 193–214. https://doi.org/10.1016/j.rser.2017.03.107

Hu, L., Xie, N., Kuang, Z., & Zhao, K. (2012). Review of Cyber-Physical System Architecture. *2012 IEEE 15th International Symposium on Object/Component/Service-Oriented Real-*

Time Distributed Computing Workshops, 25–30. https://doi.org/10.1109/ISORCW.2012.15

ic-meter. (n.d.). Hvad er IC-Meter? *IC-Meter*. Retrieved 3 January 2021, from https://www.ic-meter.com/dk/hvad-er-ic-meter/

Iskandar, J., & Moyne, J. (2016). Maintenance of virtual metrology models. *2016 27th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*, 393–398. https://doi.org/10.1109/ASMC.2016.7491083

Jakus, G., Jekovec, M., & Tomaži, S. (2010). *New technologies for web development*. 9.

Jarvenpaa, S. L., & Dickson, G. W. (1988). Graphics and managerial decision making: Research-based guidelines. *Communications of the ACM*, *31*(6), 764–774. https://doi.org/10.1145/62959.62971

Javed, W., & Elmqvist, N. (2012). Exploring the design space of composite visualization. *2012 IEEE Pacific Visualization Symposium*, 1–8. https://doi.org/10.1109/PacificVis.2012.6183556

Jensen, S. H., Møller, A., & Thiemann, P. (2009). Type Analysis for JavaScript. In J. Palsberg & Z. Su (Eds.), *Static Analysis* (Vol. 5673, pp. 238–255). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-03237-0_17

Jin, M., Ji, L., & Peng, H. (2019). The Relationship Between Cognitive Abilities and the Decision-Making Process: The Moderating Role of Self-Relevance. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01892

Khan, A., & Hornbæk, K. (2011). Big data from the built environment. *Proceedings of the 2nd International Workshop on Research in the Large - LARGE '11*, 29. https://doi.org/10.1145/2025528.2025537

Kim, T., Saket, B., Endert, A., & MacIntyre, B. (2017). VisAR: Bringing Interactivity to Static Data Visualizations through Augmented Reality. *ArXiv:1708.01377 [Cs]*. http://arxiv.org/abs/1708.01377

Kleinmuntz, D. N., & Schkade, D. A. (1993). Information Displays and Decision Processes. *Psychological Science*, *4*(4), 221–227. https://doi.org/10.1111/j.1467-9280.1993.tb00265.x

Kong, L., & Ma, B. (2020). Intelligent manufacturing model of construction industry based on Internet of Things technology. *The International Journal of Advanced Manufacturing Technology*, *107*(3–4), 1025–1037. https://doi.org/10.1007/s00170-019-04369-8

Kopecký, J., Fremantle, P., & Boakes, R. (2014). A history and future of Web APIs. *It - Information Technology*, *56*(3). https://doi.org/10.1515/itit-2013-1035

Kubicki, S., Guerriero, A., Schwartz, L., Daher, E., & Idris, B. (2019). Assessment of synchronous interactive devices for BIM project coordination: Prospective ergonomics approach. *Automation in Construction*, *101*, 160–178. https://doi.org/10.1016/j.autcon.2018.12.009

Kubicki, S., Laboratory, M.-C., & Halin, G. (2005). *ASSISTANCE TO BUILDING CONSTRUCTION COORDINATION – TOWARDS A MULTI-VIEW COOPERATIVE PLATFORM*. 22.

Lee, H. W., Oh, H., Kim, Y., & Choi, K. (2015). Quantitative analysis of warnings in building information modeling (BIM). *Automation in Construction*, *51*, 23–31. https://doi.org/10.1016/j.autcon.2014.12.007

Lin, J. (2013). Mapreduce is Good Enough?If All You Have is a Hammer, Throw Away Everything That's Not a Nail! *Big Data*, *1*(1), 28–37. https://doi.org/10.1089/big.2012.1501

Linnovate/mean. (2021). [TypeScript]. Linnovate. https://github.com/linnovate/mean (Original work published 2013)

Lipman, R. (n.d.). *Developing Coverage Analysis for IFC Files*. 10.

Liu, Z., Nersessian, N., & Stasko, J. (2008). Distributed Cognition as a Theoretical Framework for Information Visualization. *IEEE Transactions on Visualization and Computer Graphics*, *14*(6), 1173–1180. https://doi.org/10.1109/TVCG.2008.121

Lu, Y., Liu, C., Wang, K. I.-K., Huang, H., & Xu, X. (2020). Digital Twin-driven smart manufacturing: Connotation, reference model, applications and research issues. *Robotics and Computer-Integrated Manufacturing*, *61*, 101837. https://doi.org/10.1016/j.rcim.2019.101837

Luo, W. (2019). User choice of interactive data visualization format: The effects of cognitive style and spatial ability. *Decision Support Systems*, *122*, 113061. https://doi.org/10.1016/j.dss.2019.05.001

Lurie, N. H., & Mason, C. H. (2007). Visual Representation: Implications for Decision Making. *Journal of Marketing*, *71*(1), 160–177. https://doi.org/10.1509/jmkg.71.1.160

Maleshkova, M., Pedrinaci, C., & Domingue, J. (2010). Investigating Web APIs on the World Wide Web. *2010 Eighth IEEE European Conference on Web Services*, 107–114. https://doi.org/10.1109/ECOWS.2010.9

Mehmood, N. Q., Culmone, R., & Mostarda, L. (2017). Modeling temporal aspects of sensor data for MongoDB NoSQL database. *Journal of Big Data*, *4*(1), 8. https://doi.org/10.1186/s40537-017-0068-5

Mousavi, S. M., Shams, H., & Ahmadi, S. (2017). Simultaneous optimization of repair and control-limit policy in condition-based maintenance. *Journal of Intelligent Manufacturing*, *28*(1), 245–254. https://doi.org/10.1007/s10845-014-0974-8

Moyne, J., Qamsane, Y., Balta, E. C., Kovalenko, I., Faris, J., Barton, K., & Tilbury, D. M. (2020). A Requirements Driven Digital Twin Framework: Specification and Opportunities. *IEEE Access*, *8*, 107781–107801. https://doi.org/10.1109/ACCESS.2020.3000437

Natephra, W., & Motamedi, A. (2019a, May 24). *Live Data Visualization of IoT Sensors Using Augmented Reality (AR) and BIM*. 36th International Symposium on Automation and

Robotics in Construction, Banff, AB, Canada. https://doi.org/10.22260/ISARC2019/0084

Natephra, W., & Motamedi, A. (2019b). *BIM-BASED LIVE SENSOR DATA VISUALIZATION USING VIRTUAL REALITY FOR MONITORING INDOOR CONDITIONS*. 10.

Norman, Donal A. (Ed.). (2014). *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* (2nd ed., Vol. 1). Interaction Design Foundation. https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/contextual-design

Norman, Donald A., & Draper, S. W. (1986). *User Centered System Design; New Perspectives on Human-Computer Interaction*. L. Erlbaum Associates Inc.

OAuth—Autodesk Forge. (n.d.). Retrieved 5 January 2021, from https://learnforge.autodesk.io/#/oauth/

Parsons, P., & Sedig, K. (2014). Adjustable properties of visual representations: Improving the quality of human-information interaction. *Journal of the Association for Information Science and Technology*, *65*(3), 455–482. https://doi.org/10.1002/asi.23002

Po, L., Bikakis, N., Desimoni, F., & Papastefanatos, G. (2020). Linked Data Visualization: Techniques, Tools, and Big Data. *Synthesis Lectures on the Semantic Web: Theory and Technology*, *10*(1), 1–157. https://doi.org/10.2200/S00967ED1V01Y201911WBE019

Qi, Q., & Tao, F. (2018). Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. *IEEE Access*, *6*, 3585–3593. https://doi.org/10.1109/ACCESS.2018.2793265

Rachlin, H. (2003). Rational Thought and Rational Behavior: A Review of Bounded Rationality: The Adaptive Toolbox. *Journal of the Experimental Analysis of Behavior*, *79*(3), 409–412. https://doi.org/10.1901/jeab.2003.79-409

Schluse, M., Priggemeyer, M., Atorf, L., & Rossmann, J. (2018). Experimentable Digital Twins—Streamlining Simulation-Based Systems Engineering for Industry 4.0. *IEEE*

Transactions on Industrial Informatics, *14*(4), 1722–1731. https://doi.org/10.1109/TII.2018.2804917

Sedig, K., Parsons, P., & Babanski, A. (2012). Towards a Characterization of Interactivity in Visual Analytics. *J. Multim. Process. Technol.*

Sedig, Kamran, & Parsons, P. (2013). Interaction Design for Complex Cognitive Activities with Visual Representations: A Pattern-Based Approach. *AIS Transactions on Human-Computer Interaction*, *5*(2), 84–133.

Sedig, Kamran, & Parsons, P. (2016). Design of Visualizations for Human-Information Interaction: A Pattern-Based Framework. *Synthesis Lectures on Visualization*, *4*(1), 1–185. https://doi.org/10.2200/S00685ED1V01Y201512VIS005

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin*, *134*(2), 207–222. https://doi.org/10.1037/0033-2909.134.2.207

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138. https://doi.org/10.1037/h0042769

Sørensen, K. B., Christiansson, P., & Svidt, K. (2009). Prototype development of an ICT system to support construction management based on virtual models and RFID. *Journal of Information Technology in Construction (ITcon)*, *14*(19), 263–288.

Stojanovic, V., Trapp, M., Hagedorn, B., Klimke, J., Richter, R., & Döllner, J. (2019). Sensor Data Visualization for Indoor Point Clouds. *Advances in Cartography and GIScience of the ICA*, *2*, 1–8. https://doi.org/10.5194/ica-adv-2-13-2019

Svenson, O. (1979). Process descriptions of decision making. *Organizational Behavior and Human Performance*, *23*(1), 86–112. https://doi.org/10.1016/0030-5073(79)90048-5

Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. (2018). Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced*

Manufacturing Technology, *94*(9–12), 3563–3576. https://doi.org/10.1007/s00170-017-0233-1

Tao, F., Qi, Q., Wang, L., & Nee, A. Y. C. (2019). Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison. *Engineering*, *5*(4), 653–661. https://doi.org/10.1016/j.eng.2019.01.014

The original proposal of the WWW, HTMLized. (n.d.). Retrieved 7 January 2021, from https://www.w3.org/History/1989/proposal.html

Tominski, C. (2015). Interaction for Visualization. *Synthesis Lectures on Visualization*, *3*(1), 1–107. https://doi.org/10.2200/S00651ED1V01Y201506VIS003

Vachalek, J., Bartalsky, L., Rovny, O., Sismisova, D., Morhac, M., & Loksik, M. (2017). The digital twin of an industrial production line within the industry 4.0 concept. *2017 21st International Conference on Process Control (PC)*, 258–262. https://doi.org/10.1109/PC.2017.7976223

Vandecar, K. (2019, November 10). *New RVT->SVF Model Derivative parameter generates additional content, including rooms and spaces*. Autodesk Forge Community Blog. https://forge.autodesk.com/blog/new-rvt-svf-model-derivative-parameter-generates-additional-content-including-rooms-and-spaces

Veglis, A. (2017). Interactive Data Visualization. In L. A. Schintler & C. L. McNeely (Eds.), *Encyclopedia of Big Data* (pp. 1–4). Springer International Publishing. https://doi.org/10.1007/978-3-319-32001-4_116-1

Viewer extension—Autodesk Forge. (n.d.). Retrieved 5 January 2021, from https://learnforge.autodesk.io/#/tutorials/extensions

Volk, R., Stengel, J., & Schultmann, F. (2014). Building Information Modeling (BIM) for existing buildings—Literature review and future needs. *Automation in Construction*, *38*, 109–127. https://doi.org/10.1016/j.autcon.2013.10.023

Ward, M. O., Grinstein, G., Keim, D., Grinstein, G., & Keim, D. (2015). *Interactive Data Visualization: Foundations, Techniques, and Applications, Second Edition*. A K Peters/CRC Press. https://doi.org/10.1201/b18379

Weber, C., Königsberger, J., Kassner, L., & Mitschang, B. (2017). M2DDM – A Maturity Model for Data-Driven Manufacturing. *Procedia CIRP*, *63*, 173–178. https://doi.org/10.1016/j.procir.2017.03.309

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, 1–10. https://doi.org/10.1145/2601248.2601268

Zhang, H., Zhang, G., & Yan, Q. (2019). Digital twin-driven cyber-physical production system towards smart shop-floor. *Journal of Ambient Intelligence and Humanized Computing*, *10*(11), 4439–4453. https://doi.org/10.1007/s12652-018-1125-4

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, *21*(2), 179–217. https://doi.org/10.1016/S0364-0213(99)80022-6

Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, *18*(1), 87–122. https://doi.org/10.1016/0364-0213(94)90021-3

Zheng, Y., Yang, S., & Cheng, H. (2019). An application framework of digital twin and its case study. *Journal of Ambient Intelligence and Humanized Computing*, *10*(3), 1141–1153. https://doi.org/10.1007/s12652-018-0911-3

Appendix I – Contextual Inquiry process

Primary questions	Supporting questions	Interview questions	Answers
1. What is the information source?	• What are the information types are acquired (temperature, CO2 distribution, etc.)? • What instruments do you use to acquire data?	(8.) What instruments or physical objects do you use to acquire data? *Confirmation question: What are the information types acquired?* - information received in CSV file before the interview.	(8) "For this case of healthy homes, we use IC meter and Netatmo – It measures, Co2 humidity, noise, temperature. It has a ppm showing the values if the co2 levels are acceptable and it notifies the user. This is for the building occupants to monitor their room air quality. Facility manager has the data and we have request them to make analysis. They are updating the data set every night and providing it in a csv format. We need to send a HTML request for accessing the data from the sensor API."
2. Does the information have open or closed space?	• Is there constant flow of information from sensors that affects simulations, predictions, assumptions, analysis or are there almost fixed values (e.g., temperature for room type) used for analysis, calculations, simulations? • Does your system readjust values of factors, or some constants needed	(10.) Is there constant flow of information from sensors that affects simulations, predictions, assumptions, analysis or are there almost fixed values (e.g., temperature for room type) used for analysis, calculations, simulations? Does your system readjust values of factors, or some constants needed for	(10) "We do not have a link between measured data to simulation data. When performing simulations, we do it from the industry guidelines. We do not know how the end users are going to use the building, what their preferred set points are going to be how their occupancy patterns are going to be. So, with guidelines we try to standardize that. But usually, it's very different from reality and there is a performance gap in simulations and reality, and it has been proven many times that simulations don't predict the real scenarios. We must be

Primary questions	Supporting questions	Interview questions	Answers
	for calculations, based on acquired data from sensors? Do you make that kind of calibrations?	calculations, based on acquired data from sensors? Do you make that kind of calibrations?	aware about comparing different data, because occupants can change over time and we would have different patterns. We see that in research, people trying real time optimization for fault detection, but we haven't seen building owners demanding that or willing to pay for that. But we are considering using it because we think it will be beneficial for us as building energy engineers."
3. Does the information space acquire new information?	• Is there a constant batch of information revived from data sources (e.g., sensors)? If so, how often? • Do sensors are a subject of periodic calibration/check? If so, why and how often? • Are the sensors being relocated from time to time? If so, why and how do you make assessment on new location? • How do you check quality of data reads? How do you check the quality, accuracy before it's settled/approved to run?	(12.) Is live data (Real time data) required? Or just visualization from historical data and simulation data? What do you think would be the benefit of contextualizing this data with the virtual building model rather than just representing it through bar charts? *Confirmation question: Is there a constant batch of information revived from data sources (e.g., sensors)? If so, how often? –* information	(12) "So, if we are making an analysis, we don't need the real time. But if we are trying something like the Netatmo application for the building end users to monitor then we need the real time data."

Primary questions	Supporting questions	Interview questions	Answers
		received in CSV file before the interview.	
4. Are there any sub-spaces in information space?	• Does the information you receive have sub-spaces; more than one type of information received? What kind of information are there (e.g., divided by units, file format, type of document)? • How do you catalogue them? E.g., Do you acquire all temperature info in JSON file?	(13.) Along with specific unit of measurement from sensors, what more information do you receive (e.g., date, file format, type of document)? How do you catalogue and store information? Do you archive them?	(13) "We get a time stamp we can get as a JSON or CSV format. In this case we archive in a secure location for the data set and we have the power bi that updates every night. But we keep it in CSV files for backup to retrieve the data later. Because its sensitive data Powerbi has limitations and the users need a subscription to Powerbi and that proves very useless."
5. Are there many levels in information space?	• Do you have broken-down structure of acquired information, e.g. Some global/general information can have more detail information as you decompose them to the atomic level (single point or entity which is the base source of information)? • On what detail level of the building is the data recorded by the sensors?	(14.) On what level the building is the data recorded by the sensors – are sensors only placed in main building area or in all floors, then rooms, and further down to each workspace, cubicle, unit? Do you have any tree-structure of information or is it more linear, so everything has sort of the same priority?	(14) "They record every 5 minutes. We also have hourly resolution and a daily resolution level, so it gets the average accordingly. One sensor in every rooms in these healthy homes project. There is no hierarchy in which rooms are more important. We can select all of them, filter them according to specific the room category. We can also do it with this manual checkbox. We have a slicer tool and so we can select the time and data period for the selected rooms."

Primary questions	Supporting questions	Interview questions	Answers
	E.g., from corridors to all rooms further until each cubicle or cabin?		
6. If so, how deep is information space, how many levels?	• If so, what is the depth of that information space? How deep can you "zoom in" to information that will no longer have extra information within?	(15.) If so, how deep can you "zoom in" to information – going from overall building, main parts, floors, rooms, workspaces – that you will get no more new information?	
7. Is the information space homogeneous or heterogeneous? (need to rephrase it)	• What are the types of information you acquire (among all you receive)? • Are there homogeneous (e.g., Just temperature)? Or heterogeneous (e.g., CO2 condensation)? What do you think of advantages and disadvantages having one over the other? • What kind of information do you blend together? What is the information you wish you would blend? What is the reason you're not doing so?	(16.) What is the information you wish you would blend/mix, but it is not possible now (e.g., lack of settings, lack of data) and what is the reason you are not doing so?	(16) “Especially looking at the spatial information how it is related to each other, how do people move in the rooms to find building usage patterns. Make visualizations of the heatmaps for quickly communicating it to other stakeholders. If we had the similar data for bigger buildings like educational institutions, we could see the occupancy data relates to the humidity and Co2.How the data is dependent and related to make a diagnosis. We can learn something from that show it the architects”

AALBORG UNIVERSITY
STUDENT REPORT

Primary questions	Supporting questions	Interview questions	Answers
	• Is information space structured? If so, why? If no, why?		
8. Do sub-spaces consist of any relations between each other?	• Are there any established connections between sub-spaces? E.g., once you gathered data do you use them in combination? For example, putting temperature in the context of CO2, amount of people?	(17.) Do you have established information structure that you use in every project or does this vary from case to case? If so, what is the structure and how you document that?	(17) "We have a lot of BIM coordinators they know how to structure a specific project. For us the file structure does not represent our work, I just need a different setup for development. Of course, there is way at MOE the way we structure data. We need to keep a track of who is updating the data. We work with our linked model and the BIM coordinator takes care of the management work of the model."
9. What characteristics do sub-spaces have?	• How are sub-spaces organized? • What is the cardinality of sub-spaces (like room id, sensor if, etc.)? • What are the properties of information represented (like preferred color of charts; kind of representation – charts, bars, etc.; size – e.g., is it fixed year range or you can crop and adjust, is its fixed range of values you're interested in)? Are they presented in 1D (e.g.,	(4) How are you currently representing data from the readings for the meetings you mention in the brief apart from Graphs and Bar charts? How do you show the 3d model (E.g., In Revit,3d images etc.)? How many different parameters do you use in the visualization (e.g., not only value like	(4) "We are currently representing sensor data for Temperature, Co2, Humidity for this healthy home project. We are using Bar charts, Line graphs, Heat maps and Pie charts with colors to represent the data with Excel and Power Bi. With that we show scree captures of the BIM models or zoomed in images of the spaces or 2D plans of the rooms. Pie charts have been proven to create confusions and are least understood for senor data. Apart from the traditional graphs we are using the Data explorer tool from three js library. It's a tool to make multi objective optimization. We have used input parameters with

Primary questions	Supporting questions	Interview questions	Answers
	Dots where larger dot represent larger quantity which represents), 2D (any sort of charts which have two variables, 3D (3 variables, or 3D representation of the building where readings from sensors are mapped into building model)? • Can you dynamically compare/blend those sub-spaces (e.g., Information about temperature blended with humidity etc.)? Can you freely change them? If so, how is your current process? • Do you think the comparison you are doing right now can be improved? • Is it possible to create motion picture of batch of data so it gives more depth how specific information (e.g., Temperature, humidity) is changing over time? Is it something you also want to have	temperature but room id, sensor id, etc. (5) Why are you using this type of visualization? What are the pros of using that kind of visual representations? What are the shortcomings with showing the data in form of tables and graphs? (6.) Do you have fixed properties of visualization representation you just mentioned (certain way of expressing with bar, pie charts; with different sizes, ranges)? If not, how do you decide which way of representation is most suitable for you? Do you have some schema?	individual coordinates which can be toggled with different parameters like daylight analysis factors or solar panel placement to generate simulations and compare designs. This can be used to compare sensor data as well. But this something we would use ourselves as a consultant. Different users would require different kind of visualization to understand and compare the data. We have workflow for daylight analysis that we would also for extracting geometric data from a Revit model using Dynamo to excel and calculate. But there can be lot of problems due to inaccuracies in the BIM model." (5) "Tables and graphs are difficult to perceive for architects, building owners. Making comparisons is very difficult to explain without context or interactions between the data. We engineer will do it individually but to communicate to others is very difficult in these ways. We have an excel sheet with symbols and colors with their relevant values to compare for example -when the solar shading was on and when it was off, we can compare it with other parameters like temperature. Also, different colors

Primary questions	Supporting questions	Interview questions	Answers
	and compare findings with simulation results?	(11.) How is the different type of sensor/simulation data (Co2, Humidity etc.) currently been compared? Can you dynamically compare/blend sensor/simulation data (e.g., Information about temperature blended with humidity etc.) with adjusting range, areas, visual representation (e.g., from 2D to 3D)? What is something you want to have to include in your current visualization process?	with DGNB standards or Bsim. That can be hourly data but its limited by the size and only tables with no visualization with the building, rooms or facades. Also, to monitor the efficiency for the overall year with comparison of different seasons. But tools like this can help make standardized graphs to understand it faster, rather than each engineer making his own style of analysis and representation for reports very often we have code compliance and KPIs that we need to evaluate. Also, during explaining the building owner and the facility manager spatial recognition of the building and orientation for other factors like daylight ca prove important to find out the reasons for the discrepancies." (6) "We have standard tools to make the same visualizations, every time. We have the data explorer to see the KPIs and make beneficial decision on based on that, graphics are less only bar chart and pie charts, sometimes we do use CFD simulations. We use excel to make our calculations and represent data amongst the engineers. We have preset color coding

Primary questions	Supporting questions	Interview questions	Answers
			according the range of values. We have the heatmap plotting style that is very standard for us. We use them for the building performance reports. The architect has his own graphics with the BIM model and diagrams, but we only have graphs and plots, and it takes time to explain the data to others and point out the spaces and relating the data to the BIM model. After the building is handed over to the facility manager the architect gives the responsibility so that is the problem in the industry is with the commissioning process. So, when we hand it over, we don't know how the building is performing, so that data is very important for us to know how correct the assumptions during the design process were. We must ask the facility manager to provide the data to us." (11) "Most often it's the temperature that dominates the design and then Co2 and then humidity. In the data explorer tool, we have different parameters we compare. Indoor quality, thermal comfort, lighting…. We are comparing different data with bar charts and line charts. We have some guidelines limit according

Primary questions	Supporting questions	Interview questions	Answers
			to DGNB for maximum discrepancies allowed. So, we can see thru the graphs if it is above it. Sometimes we do compare the building performance data from the first year to the second and we get requests about increasing the efficiency or finding the problems for the inefficiency, so we have to meet with other stakeholders to point out. But especially while showing this data to the building owner or facility manager there are labels for the spaces of the buildings in the graphs. But it doesn't give a spatial recognition that two rooms are next to each other, so the values are dependent on each other. Also, if we want to see if it's the north facing façade or south, we cannot see how the daylight factor may have been affecting the temperature. The spatial elements if connected to the data would be very beneficial, the tools we have are only showing the sensor data as heat maps or comparison with a pie chart. We think it very insufficient and for others to comprehend and use this tool. We lack a lot of graphic visual of the building and we want to demonstrate with the spatial characteristics and different types of data that we acquire."

Primary questions	Supporting questions	Interview questions	Answers
10. What is the system architecture and how elements are arranged in the network?	• What is network topology of your current architecture? • What changes would you want to see in changes of architecture/network topology?	(20.) Are you familiar with overall/partial technical architecture of the current system – from data gathering to visual representation of data and interaction with it? If so, can you briefly introduce us to it? What changes would you want to see in current architecture?	(20) "So, for this case it is like a collaboration between us and the software developers for APIs and data engineers to structure the data and Power BI specialist who could help us setup the app. But with this thing it is very innovative, it's our first attempt. We have explored PowerBI a lot. For the future this is exactly what we plan to work and there is lot of iteration and how can we more efficiently use this data for fault detection. Much of it very novel ground. We have been also trying to store data in database and make queries according to the time stamp to the room name. Although there is no geometric context from the building yet. Also, we haven't done any outlier readings".
11. To what extent granularity should be applied for encoding the components of the system?	• How detailed data should be mapped and then represented? • At what level of granularity should the items or data should be mapped from information space to representation space?	(19.) How detailed (e.g., data from every room, every area, and how precise, like 2 digits after comma, etc.) or frequent (e.g., once/two times a day) data should be acquired to make your work efficient? How is it now in your current system?	(19) "We are getting the data in an interval of five minutes. If we are looking at specific conditions, then maybe we need a deeper insight if there is too much variable data in a particular area. If the building relies on venting for thermal comfort, then we need high resolution to see the functioning of the system in relation to opening of windows and flow. But we don't have a specific need right now"

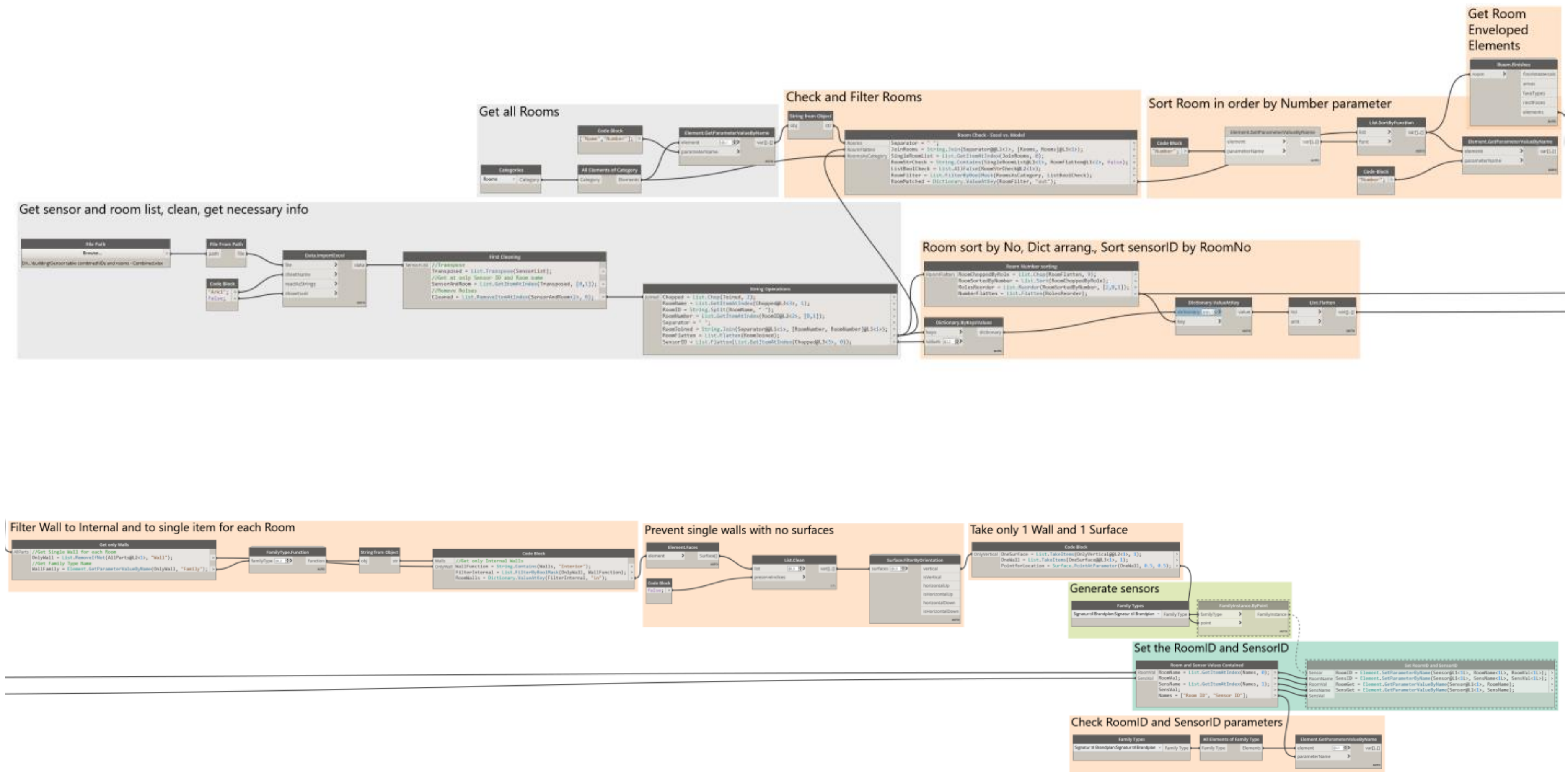
Primary questions	Supporting questions	Interview questions	Answers
		(1) What is your professional background and what is your role at MOE.? Can you tell us about the main goals of MOE and the department you are a part of?	(1) "I am working at MOE-Artelia group as a Specialist in Building energy for indoor climate and as an architectural engineer. I am also a Post -doc researcher at Aalborg university. MOE is one of Denmark's largest consulting engineering companies. My department is around of 12-14 people and we focus on sustainability, indoor environment and health of buildings, working with innovative solutions. We also make building simulations to guide decision makers of the project like the building owner, Client and other stakeholders for efficient design, code compliance and work on research development projects. I do a little bit of consultancy as well".
		(2) Can you briefly tell us about the current work model and the participants working in your group and their competencies/roles?	(2) "Most of my colleagues at MOE are from MSc architectural engineering background. My colleague Steffen has a background in building energy design and is also involved in making guidelines and standards for the Danish building industry. We do a lot of building simulations using Bsim. Also heat calculations for buildings, For CFD simulations for HVAC and DGNB certification. We also have some

Primary questions	Supporting questions	Interview questions	Answers
			HVAC engineers in the group working on Revit for modelling HVAC systems. And members in the group that make dynamo scripts for automation. But I focus more on sustainability and indoor environment collaboratively working with them. We are also working on making innovative solution based on research".
		(3) Who are the interested stakeholders or participants who need to see BIM, sensor, and simulation data visualizations? What type of meetings would you define these as? Who are the participants?	(3) "Interested stakeholders can be ourselves as well, to help us analyze the data and understand faster. But we get money from the building owner who wants us to make this analysis. We also have Real Dania which does research analysis on building sustainable and healthy buildings with the building owners in Denmark. In this case they are the client that had a hard time to analyze and measure sensor data from one of our buildings represented with charts. That's why we are exploring the ways to analyze this senor data better. Other stakeholder would be facility managers and make it easier for them to see where problems in the building are there and it's hard for them to see that with plots, graphs so better visualizations for energy data would help detect

Primary questions	Supporting questions	Interview questions	Answers
			malfunctions in HVAC systems. It would be better if they can see it with a spatial recognition of the areas with issues. We are trying to evolve in the facility management arena as well. So, working with PowerBi was one of the first try to visualize sensor data but there are still lot of challenges with PowerBi, with making web applications would do so much better. We don't have lot of software developers, but we are still developing in that area. These meetings are with Real Dania, we were showing them this Power BI graphs how the rooms perform and where there is overheating. So, they wanted tools with better visualizations with the spatial elements of the buildings. So, our clients have sensors and they also must communicate this analysis of data to the building owners. There could be variety of stakeholders, like there can also be building end users who want to see the indoor environment quality data."
		(7) What exactly is the purpose of the visualization? What needs to be communicated through these	(7) "The purpose of the visualization is show to the discrepancies in the sensor data of temperature or Co2 values to look at the critical scenarios in the buildings. They can see which rooms are critical use

Primary questions	Supporting questions	Interview questions	Answers
		visualizations? What is usually the later stage after data representation in the meetings that you mention in the brief?	that to focus on the discrepancies and improve physically. The goal of the meeting is that we should be able to find a way to decrease something that is not efficient. Some extra analysis after for future designs to tweak the data and make optimizations and would be helpful as well. After the meeting we try to find some understanding of the building with sensitivity analysis, what is accepted and what should be immediately rectified. They lack visualization of sensor data is time consuming. Post meeting, we make a summary on what has being decided upon, main an extra visualization for a specific design how certain factors relate. There is a lot of reporting to document it what was agreed upon basis on what data. So, we don't have a visualization for documenting this right now."
		(9) Which type of Simulations is the data generated from and from what software tools?	"Traditionally we have simulation data, normally the data we look at is artificial data. Energy frame simulation for the DGNB OR BSIM software calculation software or daylight simulation data. There is difficulty in selecting simulations. Also, there is CFD simulation data sometimes."

Appendix II –Dynamo for Revit script



Appendix III –Python script

```
import glob
import os
import pandas as pd
import numpy as np
import json

#relative path
rel_path = os.path.realpath(r'.\*')
print(rel_path)
all_files = glob.glob(rel_path + "/*.csv")

#create singular data frames
temp = []
dfs = []
room_name = []
sensor_id = []
room_number = []
frames = []
#check and remove "empty" csvs from the list
invalid_csv = []
i = 0

for file in all_files:
    temp = pd.read_csv(file, sep=';')
    if temp.shape[0] < 3:
        invalid_csv.append(file)
        all_files.remove(file)
        #dfs.remove(files)
        print(' not valid csv', file)
   # info = temp.info(verbose=True)
    temp_column_name = list(temp.columns)

    room_name = temp_column_name[2]
    room_split = room_name.split()
    room_connect = room_split[0] + ' ' + room_split[1]
    room_number.append(room_connect)

    sensor_id.append(temp_column_name[0])

for file in all_files:
    dfs= pd.read_csv(file, sep=';', header=1, na_filter=False,
                     skip_blank_lines=False, skiprows=0, usecols=[0,1,2,3,4,5,6])
    #print('new DATA FRAMES', file, dfs)

    #column insert
    temp_room_number = room_number[i]
    temp_sensor_id = sensor_id[i]
    dfs.insert(0, 'id', temp_room_number)
    dfs.insert(1, 'sensor_id', temp_sensor_id)
    i += 1
```

```python
    #column rename
    column_names = list(dfs.columns)
    new_column_names = [
                        'id',
                        'sensor_id',
                        'timestamp',
                        'temperature',
                        'humidity',
                        'CO2',
                        'noise average',
                        'noise peak',
                        'particle sensor']
    column_dictionary = dict(zip(column_names, new_column_names))
    new_column_name_mapping = dfs.rename(columns=column_dictionary,
                                         inplace=True, errors='raise')
    new_cn = list(dfs.columns)
    dfs = dfs.iloc[:-2]
    dfs = dfs.replace([np.nan, '', ' '], 0)

    dfs = dfs.astype({
                    'id' : 'string',
                    'sensor_id' : 'int16',
                    'timestamp' : 'datetime64[ns]',
                    'temperature' : 'float16',
                    'humidity': 'float16',
                    'CO2': 'float16',
                    'noise average': 'float16',
                    'noise peak': 'float16',
                    'particle sensor': 'float16'
                    })

    dfs = dfs[::-1]

    frames.append(dfs)

result = pd.concat(frames)

tojson = result.to_json(orient="records", date_format="iso")
parsed = json.loads(tojson)
tojson_file = result.to_json(path_or_buf="to_print.json", orient="records",
                             date_format="iso", lines=True)
print(json.dumps(parsed, indent=4))
```