VTT Technical Research Centre of Finland

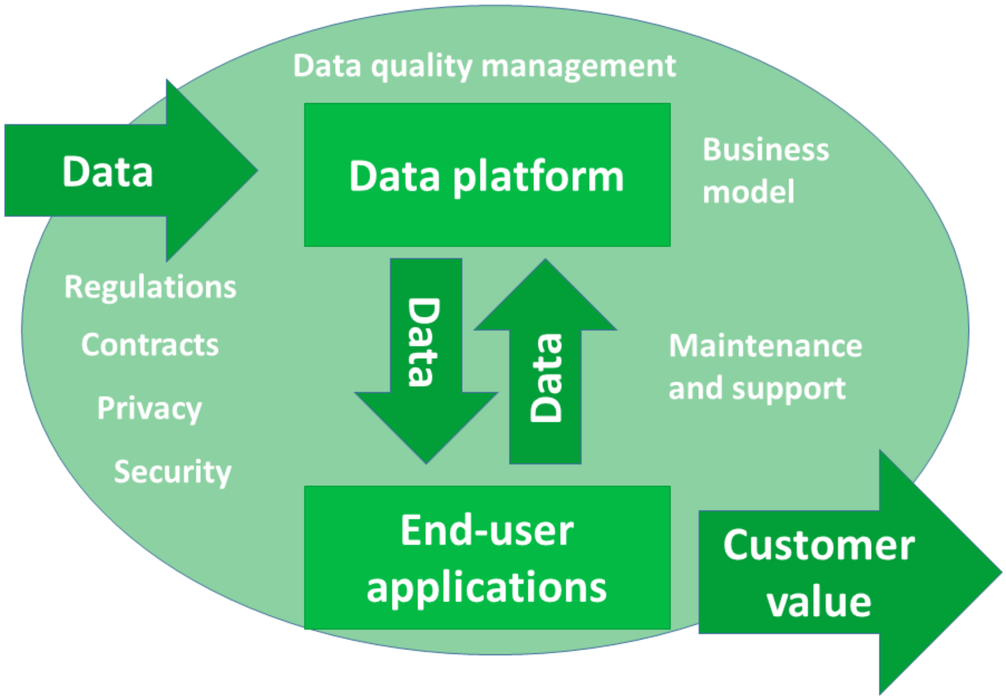
Data Platform for Smart Otaniemi Ecosystem

Vesanen, Teemu; Piira, Kalevi; Lavikka, Rita; Piippo, Jouko; Biström, Henri; Kannari, Lotta

Published: 14/01/2021

Document Version
Publisher's final version

RESEARCH REPORT VTT-R-00041-21



Data Platform for Smart Otaniemi Ecosystem - Requirements, Specifications, and Usage

Authors: VTT Technology Research Centre of Finland: Teemu Vesanen, Kalevi Piira, Rita Lavikka, Jouko Piippo, Henri Biström, Lotta Kannari

Confidentiality: Public

Version: 14.01.2021

beyond the obvious

Data Platform for Smart Otaniemi Ecosystem - Requirements, Specifications, and Usage	
Customer, contact person, address Smart Otaniemi ecosystem, https://smartotaniemi.fi/	**Order reference** -
Project name Smart Otaniemi (first pilots) WP 2 Platform and Connectivity	**Project number/Short name** 118621
Author(s) Teemu Vesanen, Kalevi Piira, Rita Lavikka, Jouko Piippo, Henri Biström, Lotta Kannari	**Pages** 36
Keywords Smart Otaniemi, data platform, ecosystem	**Report identification code** VTT-R-00041-21

Summary

This report focuses on the platform part of WP 2, Platforms and Connectivity in Smart Otaniemi. The connectivity part of the WP 2 has been reported in a separate report. Smart Otaniemi is an innovation ecosystem that connects experts, organisations, technologies and pilot projects. The project focuses on smart energy to promote sustainability and cost-effectiveness. The ecosystem is exploiting and experimenting with new technologies through pilots. In this ecosystem, connectivity and data platforms are one of the key enablers towards service chains across industries enabling large-scale use of renewable energy sources in energy systems, and reducing energy consumption and CO2 emissions.

The report describes how the concept of a data platform was further developed in WP 2: first by continuing the implementation of the existing VTT research platform to enable use case trials in Smart Otaniemi and then by starting the planning of a commercial platform. The report begins by discussing data markets in general and how to increase data utilisation in the built environment. After this, the report presents the current VTT research platform in detail, for example, the multiple sources of data collection, services on data and interfaces to other systems. Then, the report presents end-user applications of the VTT research platform. Finally, the report discusses current challenges in digital data sharing and reports requirements for a commercial data platform.

The report finds that inter-organisational data sharing takes time and several challenges exist. For example, data sharing often necessitates the implementation of an API infrastructure, which incurs costs. Also, companies still lack use cases and capabilities for using and acquiring data. Lack of common rules for data sharing also slows down the development of success stories for data sharing. The report discusses ten requirements for a commercial platform: Platform owner, data ownership, data availability, data integrity, data's metadata, data security, obedience to GDPR, data access control, varying performance requirements and data sharing contracts.

Confidentiality	Public

Espoo 14.1.2021

Signed by	Reviewed by	Accepted by
Senior Scientist Teemu Vesanen on behalf of all authors	Senior Scientist Esa Nykänen	Senior Scientist Ismo Heimonen

VTT's contact address
VTT, P.O.Box 1000, 02044 VTT, Finland

Distribution (customer and VTT)
Printed to VTT Archive.
Available online in VTT's Research Information Portal, https://cris.vtt.fi/

Preface

Smart Otaniemi is an innovation ecosystem that connects experts, organisations, technologies and pilot projects. The project focuses on smart energy to promote sustainability and cost-effectiveness. The ecosystem is exploiting and experimenting with new technologies through pilots. In this ecosystem, connectivity and data platforms are key enablers towards service chains across industries enabling large-scale use of renewable energy sources in energy systems, and reducing energy consumption and CO2 emissions.

This report describes how the concept of a data platform was further developed in this project: first by continuing the implementation of the existing VTT research platform to enable use case trials in Smart Otaniemi and then by starting the planning of a commercial platform. The report focuses on the platform part of WP 2 Data Platforms and Connectivity of Smart Otaniemi. The WP's connectivity part has been reported in a separate, confidential report (Horsmanheimo *et al.*, 2020).

The report begins by discussing data markets in general and how to increase data utilisation in the built environment. After this, the report presents the current VTT research platform in detail, for example, the multiple sources of data collection, services on data and interfaces to other systems. Then, the report presents end-user applications of the VTT research platform. Finally, the report discusses current challenges in digital data sharing and reports requirements for a commercial data platform, based on interviews conducted with the Smart Otaniemi companies participating in WP 2.

The report finds that inter-organisational data sharing takes time and several challenges exist. For example, data sharing often necessitates the implementation of an API infrastructure, which incurs costs. Also, companies still lack use cases and capabilities for using and acquiring data. Lack of common rules for data sharing also slows down the development of success stories for data sharing. The report discusses ten requirements for a commercial platform: Platform owner, data ownership, data availability, data integrity, data’s metadata, data security, obedience to GDPR, data access control, varying performance requirements and data sharing contracts.

We want to thank Business Finland for financing the project and all companies involved in the WP 2 research: ABB, E2M, Eneron, Ensto, Fidelix, Fourdeg, Gravito, Nokia, Parkkisähkö, Process Genius, Tieto, UPM Energy, Väre, VTT and Wirepas.

Espoo 14.1.2021

Authors

Contents

1 Introduction

Smart Otaniemi is an innovation ecosystem that focuses on smart energy to promote sustainability and cost-efficiency. Smart energy services necessitate data sharing between companies. At the moment, companies collect data with several company-specific data platforms operating in silos, which makes it difficult or even impossible to combine data from different sources. In Smart Otaniemi, the aim is to enable easy data sharing between companies to enable the collaborative development of data-enabled services.

Smart energy pilots have used a VTT research platform for sharing data between organisations. VTT has developed the VTT research platform over the past ten years. The research platform has been designed to create easy access to data from different sources to enable the development of new types of aggregated services and applications for energy, building, and communication market sectors. However, the platform was not designed as a market place for commercial activities, such as selling data-enabled services and applications. The platform focuses on the technical side of sharing data. It can read data from original systems via interfaces, perform analytics and offer the results

via interfaces. However, the support for increasing trust is mostly missing.

The role of the VTT research platform is to enable testing of technical solutions for data sharing. A commercial platform trial is needed to get valid information about the business side of sharing data between companies, such as costs related to a system that provides the contractual support for scalability and trust to the system.

The initial questions that the companies of WP 2 had in the beginning were:

1. What is a commercial data platform business model that reasons the costs?
   a. How to fund the startup of the commercial platform?
   b. Who owns the platform?
   c. Who owns the solutions/apps in and around the platform?
   d. Who is operating the platform?
   e. What are the vision and strategy?
2. How to find and implement the right technology?
   a. How to make it secure enough?
   b. How to balance between open data sharing and restricted rights?
   c. How to separate monitoring and control parts with different requirements?
   d. Should we have a network of cloud services or a single server?
3. What kinds of contracts are needed?
   a. How to ensure that participants keep the agreed level of data quality?
   b. What data is available and with what terms?

In the beginning, the companies agreed that the data platform used in Smart Otaniemi should enable support for technical data quality management. The platform would also need to provide the contractual basis and rules of the ecosystem, ensuring that participants keep the agreed level of data quality. The business model must provide value for every stakeholder in the data management process. Otherwise, if an organisation sees no value in sharing data, it will not be interested in using resources to maintain the data quality. Thus, the quality chain will break. For example, a use case of real-time occupancy data handling was drafted in the first business model workshop (Figure 1). The use case was developed based on the idea that each organisation would create and capture value.

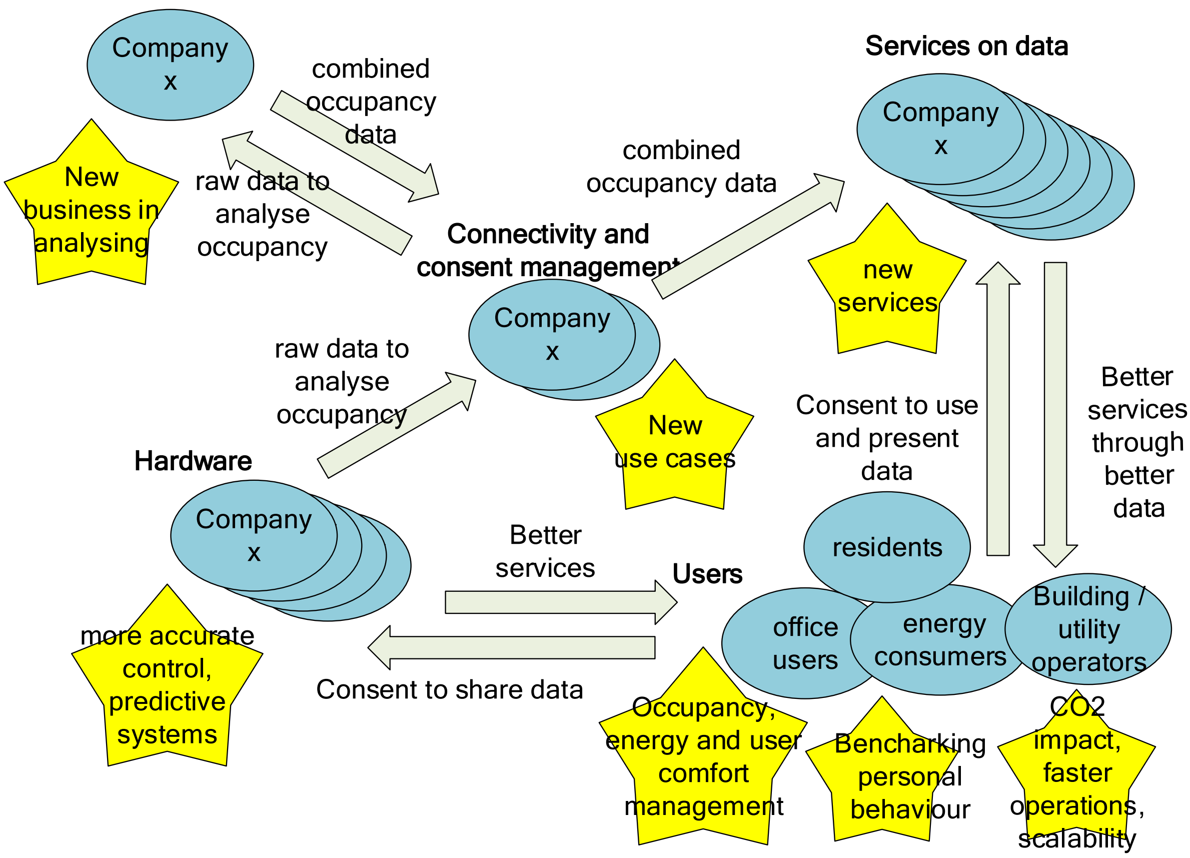


Figure 1. New business in analysing occupancy

2 Towards data markets

2.1 Data markets and data value chain

Digital raw data, products and services derived from raw data are sold and bought in data markets. The size of the data market in the EU-28[1] in 2015 was estimated at 54 billion euros (IDC, 2017), and it has steadily grown ever since. The global data market is currently estimated to grow at an annual rate of around 10,6%, and the market size is estimated to be worth 229 billion dollars by 2025[2].

Data markets create opportunities for new business, but they also reshape existing businesses. However, many companies still seem lulled into the idea that data markets are not related to their business. Nowadays, though, data markets affect almost all companies, not just ICT companies with the know-how to handle large volumes of data.

Data can be used to make traditional industries smarter. IoT market analysis in the Finnish estate and construction sector (Granlund 2017) divides the IoT development path into four different steps: operational efficiency, new products and services, outcome economy and autonomous economy (Dooley

[1] The EU-28 in 2015 consisted of 28 European Union countries (Belgium, Bulgaria, Czech Republic, Denmark, Germany, Estonia, Ireland, Greece, Spain, France, Croatia, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Malta, Netherlands, Austria, Poland, Portugal, Romania, Slovenia, Slovakia, Finland, Sweden, United Kingdom) which operate as an economic and political block. https://www.igi-global.com/dictionary/eu-28/58384
[2] https://www.marketsandmarkets.com/PressReleases/big-data.asp

et al., 2017). For example, in the real estate and construction industry, data use helps provide better indoor air environments in an energy-efficient manner (operational efficiency). Data can also help in building high-quality and cost-effective facilities with rethought services (new products and services). It can also enable moving from fixed rents and service costs to the dynamic outcome-based pricing (outcome economy) and enable building services to respond to the user needs automatically (autonomous economy).

While most of the data utilisation seems to lay on the first two steps, i.e., efficiency and new services, some services have features of the autonomous economy. As an example, Senate Properties has created a service that allows its property users to provide feedback on the indoor conditions and send a message or a service request using a mobile phone (VastuuGroup, 2019). In the case, humans can be seen as sensors that detect the working environment's quality and create data that is used to evaluate how well the asset management team has performed. If this is not “outcome economy” yet, at least it is much closer to that, than just monitoring, e.g., indoor temperatures or control signals of the building system.

Data markets in the built environment are currently being created. For example, data related to land use and construction are continuously and increasingly generated. These data could be used to support better decision making. For example, in the fight against climate change, it would make sense to know where and how to build so that the carbon footprint is kept to a minimum and the diversity of regions can be preserved.

The data value chain starts with data collection and creation and continues with the storage, aggregation, analysis and finally, distribution. The impacts of the European data economy can be divided into two: Macroeconomic impacts concern GDP growth, new SMEs and job opportunities, and the data-driven competitiveness of the EU industries. The microeconomic impacts relate to cost savings, increased flexibility, new products and services, and improved customer services. The elements leading to these impacts are shown in Figure 2.

In the figure, the framework conditions have been divided into regulatory and non-regulatory ones. Most of the conditions obviously require both sides. Regulations can set the minimum level of requirements, but companies willing to offer the highest quality services will certainly do more than the minimum. On the other hand, in many countries, the regulations will also cover the part that is classified here as non-regulatory. The point in the original report, where the figure has been adapted from, was to define framework conditions that identify the main factors, which will enable or prevent the development of the European data market and economy. It is good to note that the report is based on a project that started in 2013. Already at that time, it was recognised that privacy, ownership, copyrights and security are the fields that should be regulated to promote the data market development. Especially the data ownership is still a challenge that seems hard to solve with the non-regulatory measures.

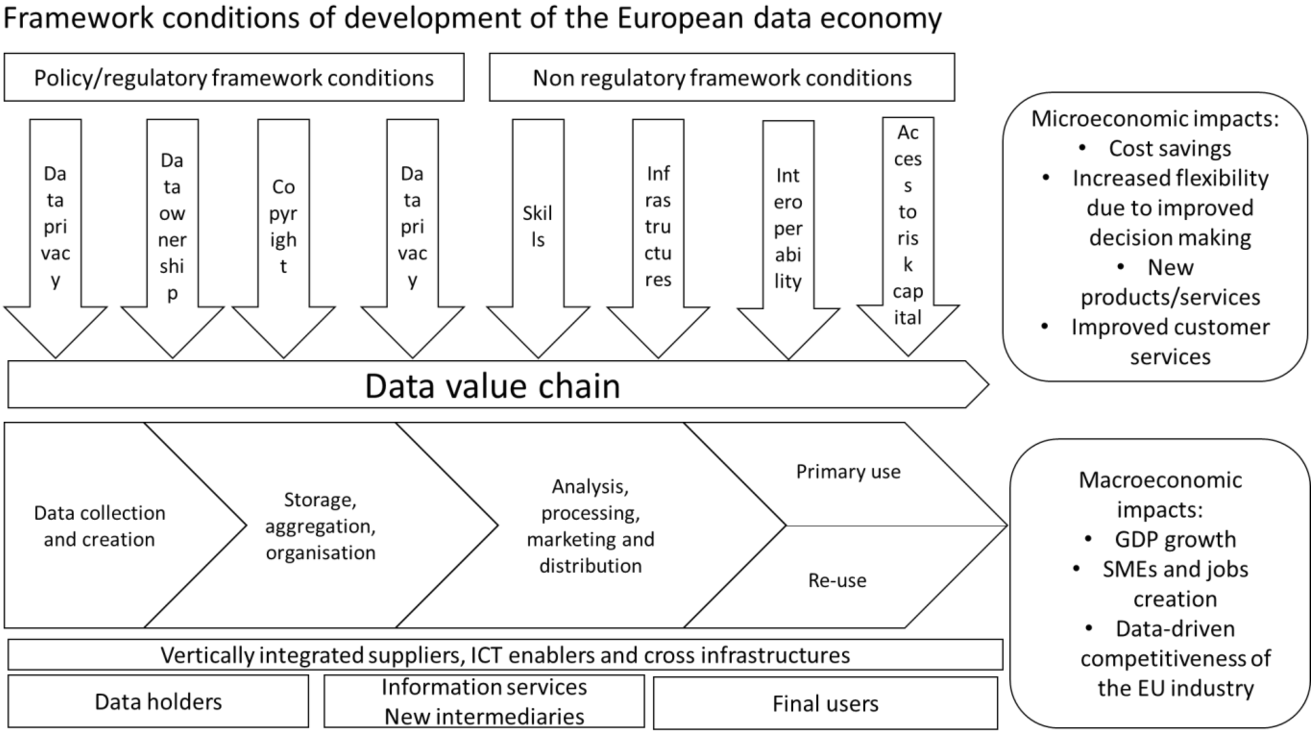


Figure 2. Data value chain (Modified from IDC, 2017, p. 28)

2.2 Towards greater data utilisation in the built environment

Recently, there have been a few initiatives that are aiming to increase data utilisation in the built environment. For example, the Finnish Ministry of the Environment launched a development project that will develop an information system for the built environment (Ympäristöministeriö, 2020). If the project is successful, the built environment's data can be reliably used to make decisions and provide data-driven services.

Another initiative is the model terms for data sharing, provided by the Technology Industries of Finland. This is an important effort since established practices for agreeing on data access rights do not yet exist (Teknologiateollisuus, 2019). The model terms can be used by companies to define the origin, purpose, and ownership of data. Transparency in the use of data should be a high priority, as it increases trust between actors.

The recent launch of the Finnish Center for Artificial Intelligence (FCAI) has also increased the awareness of possibilities for utilising data.[3] FCAI aims to bring together researchers and industry and public sector actors to solve real-life problems using both existing artificial intelligence expertise and entirely new artificial intelligence methods. The Big Data Value Association (BDVA), on the other hand, is an international, industry-driven organisation that aims to increase data-driven collaboration among actors by providing tools for co-creation that should lead to new services and know-how exchange.[4] A third organisation, worth mentioning, is the International Data Spaces Association (IDSA), which aims to design a reliable data economy architecture with the help of business and the research community. IDSA

[3] Finnish Center for Artificial Intelligence, 26.8.2020, available at: https://fcai.fi/
[4] The Big Data Value Association (BDVA), 26.8.2020, available at: https://www.bdva.eu/about

strives to ensure data management through an open, vendor-independent peer-to-peer network architecture, providing access control for data coming from different regions.[5]

A quick data experiment through a hackathon provides one method to gather greater courage in utilising data. Some smaller, agile ICT-companies are usually more than willing to experiment with incumbents. In the best scenario, a quick experiment increases the experience-based understanding of utilising data in business collaboration.

It is not too late to get involved. If a company lacks the courage to experiment, it is essential to make a data inventory. For example, the structure, quality, sensitivity, and ownership of data provide an understanding of the potential for exploiting data resources. The data inventory helps determine whether additional help is needed from companies providing data services related to integration, integrity, reliability and traceability. For example, the Finnish financial company, OP group, has started publishing annual data balance sheets. *“In business, data is not just a raw material or a by-product of service production. It is a new and valuable capital that needs to be nurtured and which return possibilities need to be identified and exploited. Knowledge is a resource for those in possession of it.”,* it states. [translated from OP data balance sheet[6]]

3 Current VTT research platform

3.1 Data collection from multiple sources

VTT has a long background in developing shared data use and platforms for building energy data in Otaniemi. In earlier activities, VTT has developed a research platform based on oBixHistory server. It contains the following data from Otaniemi buildings. This data is ready to be shared and used in the pilots in the authorisation scope by the building or open data owners.

- Building data & Control
    - ~ 3 000 data points from BACSs in two buildings in Otaniemi
    - VTT Office in Vuorimiehentie 3 and Otakaari 4
    - There are points such as
        - air pressures
        - air CO2 concentrations
        - variable air volume control unit flow measurements
        - freezing alerts
        - light intensities
        - humidities
        - district heat energy, water flow and temperature measurements
        - heating network equipment statuses, temperatures, pressures, pressure differences
        - water leakage alerts
        - service water flow and temperature measurements, equipment statuses
        - rainwater pumping station status
        - ventilation damper control and flow measurements, time extension buttons
        - ventilation CO2 measurements and setpoints for VAV control (typically from exhaust air ducts)

[5] The International Data Spaces Association (IDSA), 26.8.2020, available at: https://www.internationaldataspaces.org/our-approach/
[6] OP Financial Group, OP Data Balance Sheet 2018, 18.3.2019, available at https://www.op.fi/op-financial-group/media/bulletins?id=3237960_PRC

- frequency controller status for ventilation fans
- ventilation fan status
- ventilation fan on-off permits
- ventilation air filtering pressure difference
- heat recovery unit statuses
- room temperatures
- cooling network pump statuses
- cooling network pressures
- cooling control setpoint temperatures
- cooling unit statuses
- lighting controls
- rain and snow sensors
- emergency lighting service request
- fire alarm
- fire alarm service request
- outdoor anti-ice heatings
- access control
- access control service request

- Energy data
  - district heating, electricity, service water
  - ~ 50 buildings in Otaniemi, see the preliminary list in Figure 3
- FMI open weather data and forecast
  - Measurements from all stations: temperature, humidity
  - Measurements whenever available in the weather station: global radiation, diffuse radiation
  - Locations:
    Alajärvi Möksy (101533), Asikkala Pulkkilanharju (101185), Enontekiö Kilpisjärvi kyläkeskus (102016), Enontekiö Kilpisjärvi Saana (102017), Enontekiö Näkkälä (102019), Espoo Nuuksio (852678), Espoo Tapiola (874863), Haapavesi Mustikkamäki (101695), Hailuoto Keskikylä (101776), Hailuoto Marjaniemi (101784), Halsua Purola (101528), Hammarland Märket (100919), Hanko Russarö (100932), Hanko Tulliniemi (100946), Hanko Tvärminne (100953), Hattula Lepaa (101151), Heinola Asemantaus (101196), Helsinki Harmaja (100996), Helsinki Helsingin majakka (101003), Helsinki Kaisaniemi (100971), Helsinki Kumpula (101004), Helsinki Vuosaari satama (151028), Hyvinkää Hyvinkäänkylä (101130), Hämeenlinna Katinen (101150), Hämeenlinna Lammi Evo (101189), Hämeenlinna Lammi Pappila (101154), Ilomantsi Mekrijärvi (101651), Ilomantsi Pötsönvaara (101649), Inari Angeli Lintupuoliselkä (102026), Inari Ivalo lentoasema (102033), Inari Kaamanen (102047), Inari Kirakkajärvi (102055), Inari Nellim (102052), Inari Raja-Jooseppi (102009), Inari Saariselkä Kaunispää (102006), Inari Saariselkä matkailukeskus (102005), Inari Seitalaassa (129963), Inari Väylä (102042), Inkoo Bågaskär (100969), Joensuu Linnunlahti (101632), Jokioinen Ilmala (101104), Jomala Jomalaby (100917), Jomala Maarianhamina lentoasema (100907), Joroinen Varkaus lentoasema (101420), Joutsa Savenaho (101367), Juuka Niemelä (101609), Juupajoki Hyytiälä (101317), Juva Partala (101418), Jyväskylä lentoasema (101339), Jämsä Halli lentoasema (101315), Jämsä Halli Lentoasemantie (101338), Kaarina Yltöinen (100934), Kajaani lentoasema (101725), Kajaani Petäisenniska (126736), Kalajoki Ulkokalla (101673), Kankaanpää Niinisalo lentokenttä (101291), Karvia Alkkia (101272), Kaskinen Sälgrund (101256), Kauhajoki Kuja-Kokko (101289), Kauhava lentokenttä (101503), Kemi Ajos (101846), Kemi I majakka (101783), Kemi Kemi-Tornio lentoasema (101840), Kemijärvi lentokenttä (101950), Kemiönsaari Kemiö (100951), Kemiönsaari Vänö (100945), Kirkkonummi Mäkiluoto (100997), Kittilä Kenttärova (101987), Kittilä kirkonkylä (101990), Kittilä lentoasema (101986), Kittilä Levitunturi (101981), Kittilä Lompolonvuoma (778135), Kittilä Matorova (101985), Kittilä Pokka (101994), Kokemäki Tulkkila (101103), Kokkola Hollihaka (101675), Kokkola Tankar (101661), Korsnäs Bredskäret (101479), Kotka Haapasaari (101042), Kotka Rankki (101030), Kouvola Anjala (101194), Kouvola Utti lentoasema (101191), Kouvola Utti Lentoportintie (101219), Kristiinankaupunki Majakka (101268), Kruunupyy Kokkola-Pietarsaari lentoasema (101662), Kuhmo Kalliojoki (101773), Kumlinge kirkonkylä (100928), Kuopio Maaninka (101572), Kuopio Ritoniemi (101580), Kuopio Savilahti (101586), Kustavi Isokari (101059), Kuusamo Juuma (101899), Kuusamo Kiutaköngäs (101887), Kuusamo lentoasema (101886), Kuusamo Ruka Talvijärvi (806428), Kuusamo Rukatunturi (101897), Kökar Bogskär (100921), Lahti Laune (101152), Lappeenranta Hiekkapakka (101252), Lappeenranta Konnunsuo (101246), Lappeenranta lentoasema (101237), Lappeenranta Lepola (101247), Lemland Nyhamn (100909), Lieksa Lampela (101636), Liperi Joensuu lentoasema (101608), Liperi Tuiskavanluoto (101628), Lohja Porla (100974), Loviisa Orrengrund (101039), Luhanka Judinsalo (101362), Lumparland Långnäs satama (151048), Maalahti Strömmingsbådan (101481), Mikkeli lentoasema (101398), Multia Karhila (101536), Muonio kirkonkylä (106435), Muonio Laukukero (101982), Muonio Sammaltunturi (101983), Mustasaari Valassaaret (101464), Mäntsälä Hirvihaara (103794), Nurmijärvi Röykkä (101149), Oulu Oulunsalo Pellonpää (101799), Oulu

Vihreäsaari satama (101794), Parainen Fagerholm (100924), Parainen Utö (100908), Parikkala Koitsanlahti (101254), Pelkosenniemi Pyhätunturi (101958), Pello kirkonkylä (101914), Pietarsaari Kallan (101660), Pirkkala Tampere-Pirkkala lentoasema (101118), Pori rautatieasema (101064), Pori Tahkoluoto satama (101267), Porvoo Emäsalo (101023), Porvoo Harabacka (101028), Porvoo Kalbådagrund (101022), Porvoo Kilpilahti satama (100683), Pudasjärvi lentokenttä (101805), Puolanka Paljakka (101831), Puumala kirkonkylä (150168), Pyhäjärvi Ojakylä (101705), Raahe Lapaluoto satama (101785), Raahe Nahkiainen (101775), Raasepori Jussarö (100965), Rantasalmi Rukkasluoto (101436), Ranua lentokenttä (101873), Rauma Kylmäpihlaja (101061), Rautavaara Ylä-Luosta (101603), Rovaniemi Apukka (101933), Rovaniemi lentoasema (101920), Rovaniemi rautatieasema (101928), Salla kirkonkylä (101959), Salla Naruska (101966), Salla Värriötunturi (102012), Salo Kiikala lentokenttä (100967), Salo Kärkkä (100955), Savonlinna lentoasema (101430), Savonlinna Punkaharju Laukansaari (101441), Savukoski kirkonkylä (101952), Seinäjoki Pelmaa (101486), Siikajoki Ruukki (101787), Siilinjärvi Kuopio lentoasema (101570), Sodankylä Lokka (102000), Sodankylä Tähtelä (101932), Sodankylä Vuotso (102001), Somero Salkola (101128), Sotkamo Kuolaniemi (101756), Suomussalmi Pesiö (101826), Taivalkoski kirkonkylä (101885), Tampere Härmälä (101124), Tampere Siilinkari (101311), Tampere Tampella (151049), Tohmajärvi Kemie (101459), Toholampi Laitala (101689), Tornio Torppi (101851), Turku Artukainen (100949), Turku Rajakari (100947), Utsjoki Kevo (102035), Utsjoki Kevo Kevojärvi (126737), Utsjoki Nuorgam (102036), Vaasa Klemettilä (101485), Valtimo kirkonkylä (101743), Vantaa Helsinki-Vantaan lentoasema (100968), Varkaus Kosulanniemi (101421), Vesanto Sonkari (101555), Vieremä Kaarakkala (101726), Vihti Maasoja (100976), Viitasaari Haapaniemi (101537), Virolahti Koivuniemi (101231), Virrat Äijänneva (101310), Ylitornio Meltosjärvi (101908), Ylivieska lentokenttä (101690), Ähtäri Myllymäki (101520)

- Fingrid open data
  - Down-regulating price in the Balancing energy market
  - Electricity consumption forecast
  - Frequency containment reserve for normal operation
  - Solar power generation forecast
  - Temperature in Tammisto
  - Up-regulating price in the Balancing energy market
  - *also, others are easily connected to the platform*
- Wireless energy harvesting sensors
  - EnOcean, controllable radiator valves, air temperatures & humidity, windows, radiator and wall surface temperatures, CO2, etc
- ABB CMS-700 electricity measurement (active power, energy) via Modbus TCP field bus at one second time interval
- GEF PV panel active power generation an average 15 second time interval
- MQTT based measurements (Wirepas mesh technology connected sensors via MQTT broker, PV panels via MQTT broker)
- Besides a lot of file-based measure data sets (cases where an online connection is missing)

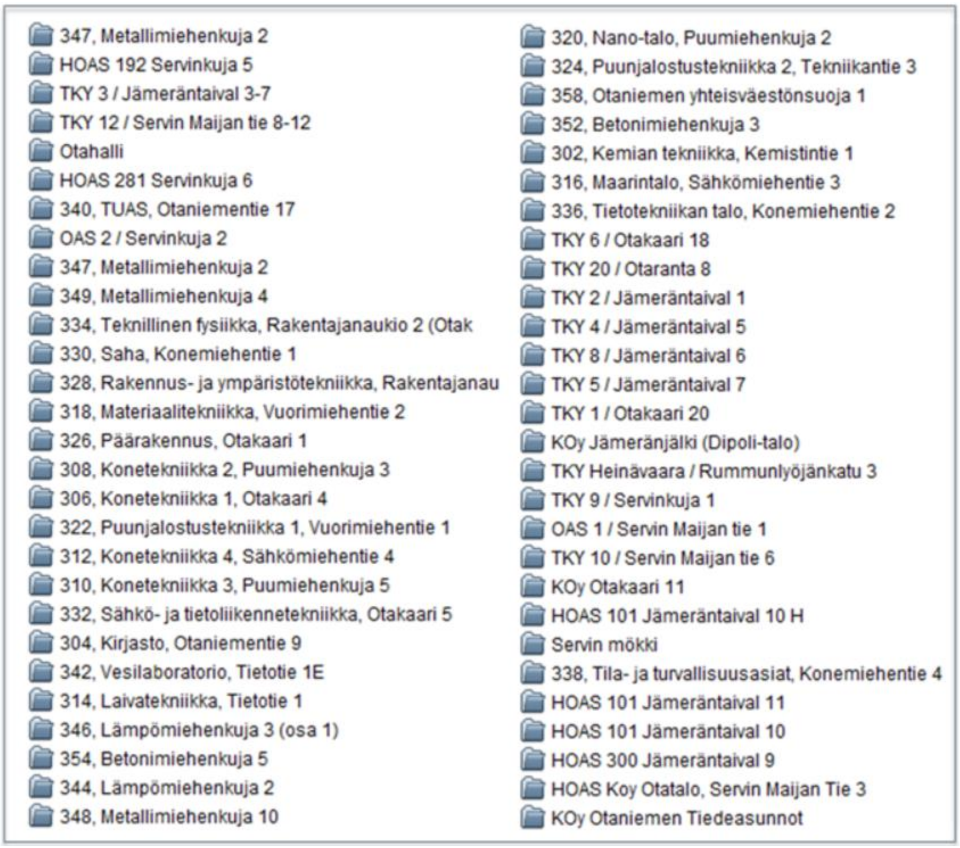


Figure 3. Preliminary list of buildings, which district heating, electricity and water consumption data are collected via EnerKey data API to VTT research platform.

3.2 Services on data

In addition to the data, the platform also offers some services:

- The data can be connected to the building information models stored either to VTT BIMServer installation or Drumbeat BIM-storage in the web.
- Heat consumption can be connected to forecasting and anomaly detection service utilising genetic algorithms boosted machine learning of calculation model parameters based on available online and history data.
- Simulation service can be used to get comparison data for measured values.
- Machine learning neural network models, e.g., forecasting electricity consumption.
- Thermal comfort indexes can be calculated based on building automation data.
- Individual optimised thermal comfort can be controlled via existing building automation systems or energy harvesting controllable radiator valves.
- Digital twins of the pilot buildings can be studied.
- Virtual data points (aggregated and post-processed data points) can be calculated and stored.

3.3 Interfaces to other systems

The platform includes interfaces to read or subscribe to the data listed in the previous chapter. There are also to-way interfaces, e.g., to several building automation systems.

An architecture overview of the VTT research platform is depicted in Figure 4.

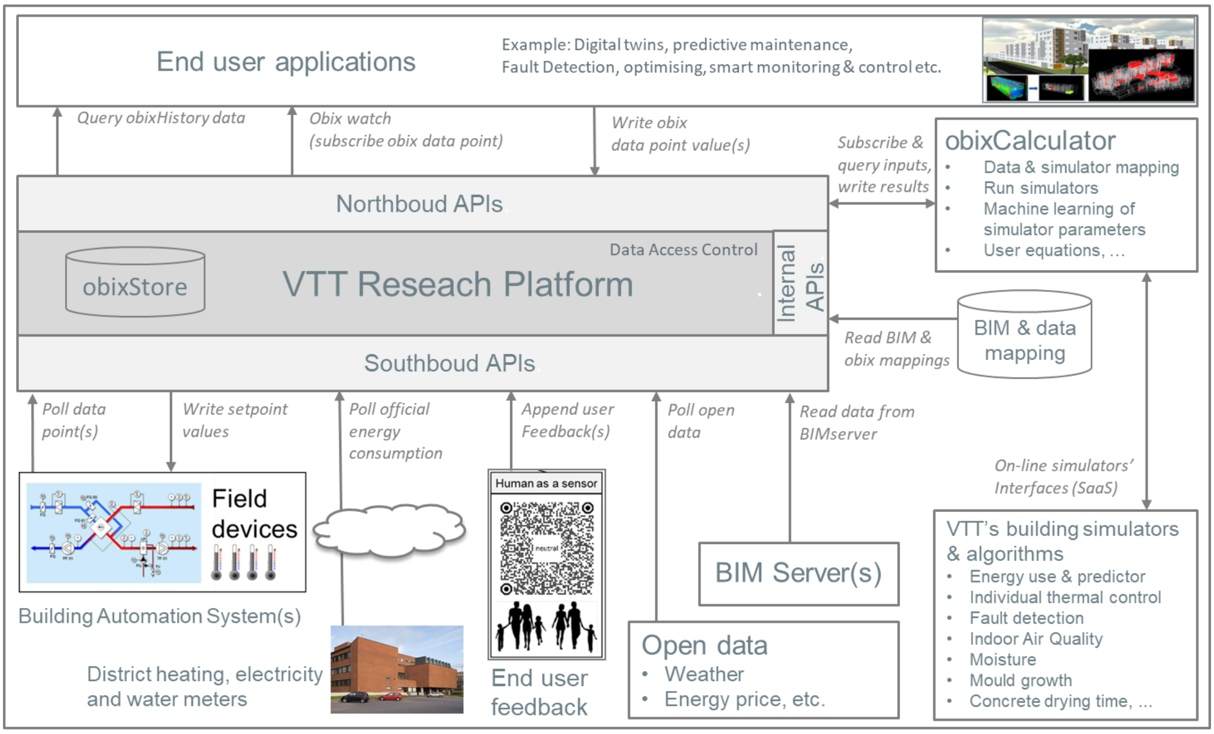


Figure 4. An architecture overview of the VTT research platform.

The main components of the VTT research platform are

- Platform's core services
  - obixHistory database (obixStore) for storing timestamped data point values
  - data point level data access management for different users
- Platform's southbound interfaces (SBI, see Figure 5) for collecting data from partners systems
  - Poll (regular reading of data from partners' systems or cloud to obixStore):
    - Siemens Desigo CC, Siemens Navigator, Caverion Niagara/oBIX, AtmosCare, Fourdeg, Modbus and related EnOcean, BACnet BACrest, KNX (BAOS), Fidelix, Granlund Metrix, EnerKey, EDI Messages (Electric, district heating & water meters), Helen District Heating, Helen District Heating Demand Response, GEF, ABB CMS-700, Telia, Nuuka, ZigBee, Savox, SmartWatcher, MeshWorks Wireless, Schneider PME (SFTP server related polling), Loimua, NNE, third party cloud platform, …
    - Open data services like FMI Weather data & forecast, Fingrid, …
  - Append (receive and store data to obixStore which other system or apps are sending to platform):
    - VTT's QR code-based end-user feedback (thermal sensation, indoor air quality, etc.), indoor positioning, Suunto Movesense, IoT / MQTT, external simulators, third party cloud platform, …
  - Write (regular reading data from obixStore and writing it to a partner system; e.g., change building automation related setpoint values):
    - Siemens Desigo CC, Caverion Niagara/oBIX, Fourdeg, Modbus and related EnOcean
  - BIMserver interface for reading Building Information Models (BIM) from BIMserver.org (or drumbeat) BIM server
- Platform's northbound interfaces (NBI) for an application developer and end-user apps access to the platform's data

        - Query obixHistory data is the most important API for end users. This API is REST-based API using basic HTTP authentication and HTTP POST messages. The result values are XML coded obixHistories () including timestamped values of the data point between given start time and end time.
        - oBIX Watch (subscribe obix datapoint) API makes it possible to subscribe to the obixStore data point and get a notification when the value is changed.
        - Write data point value is the same API as southbound API Append (see the platform southbound API description).
- Internal interfaces
    - Read BIM & obix mappings, which makes it possible, e.g., to visualise and analyse data in the 3D building model
- obixCalculator for running online connected simulators and used defined equations in the platform
    - Mapping obixStore data with online connected simulators inputs and outputs
    - Run obixCalculator connected simulators
    - Genetic algorithms boosted machine learning of simulator parameters, e.g., online connected digital twins
    - Possible to write user equations using basic mathematical notations
    - Write simulator results to obixStore
- VTT's platform connected (building) simulators and algorithms as a service (SaaS)
    - Energy consumption & load predictor
    - Individual thermal control
    - Fault detection
    - Indoor Air Quality
    - Moisture
    - Mould growth
    - Concrete drying time

The VTT research platform southbound interfaces are shown in Figure 5. From Smart Otaniemi point of view only EnerKey, Schneider AtmosCare, GEF, ABB CMS-700, Fidelix FX-3000-C, Modbus, EnOcean, Granlund Metrix, MQTT based measurements (e.g., Micronova: Wirepas sensors and PV panel), EDI messages and open data via FMI's and Fingrid's APIs have been utilised in current Smart Otaniemi pilots. On the other hand, the technical interoperability with the other southbound APIs exists if some pilot building has related systems.

Figure 5. VTT research platform southbound interfaces.

4 VTT research platform use cases and end-user applications

Figure 6 illustrates the overall concept of the VTT research platform. The basic idea is first to collect data from different data sources and then harmonise the collected data. After this, the data are connected with different kinds of online simulators, smart algorithms and machine learning digital twins to the platform for research related monitoring, analysing, optimising, predicting, smart control, fault detection, fault diagnosis, predictive maintenance and visualisation.

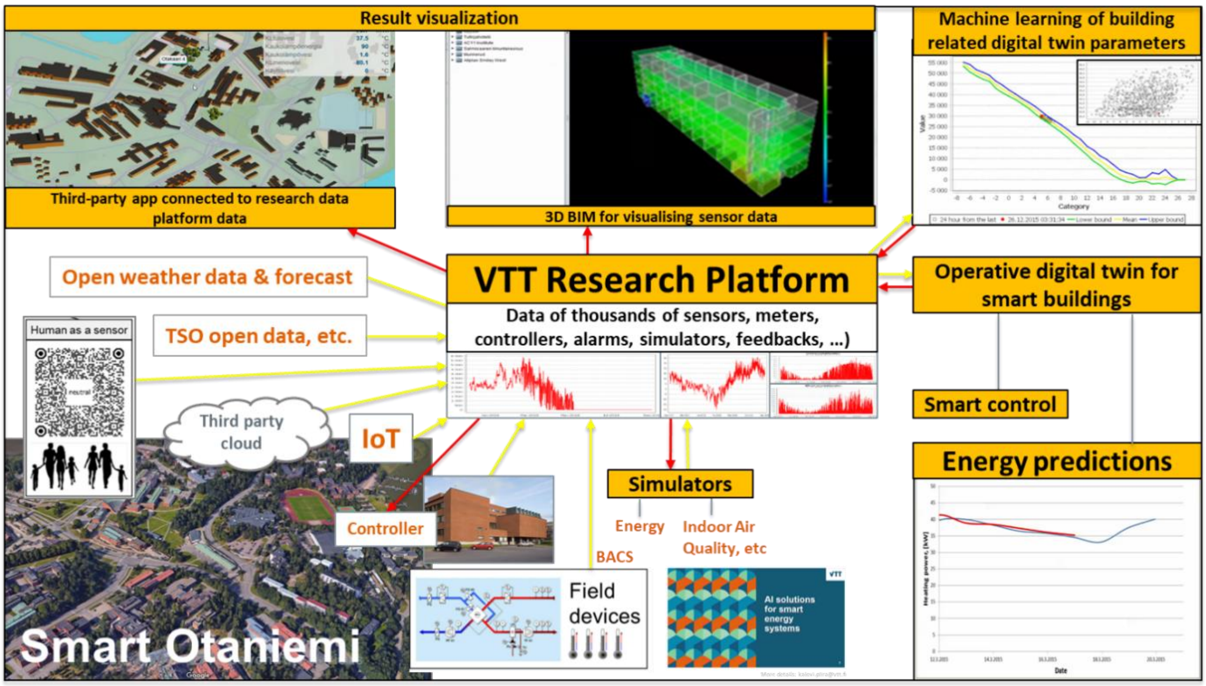


Figure 6. The overall concept of the VTT research platform.

Some examples of the VTT research platform connected demonstration and end-user applications are described as follows. These applications read data by VTT research platform harmonised APIs, and the online dashboards are implemented using open source based Grafana application.

4.1 Otaniemi district heating and electricity consumption dashboards

The official district heating and electricity-related energy consumption data (used for billing) are collected via EnerKey Energy Data Platform (Figure 7).

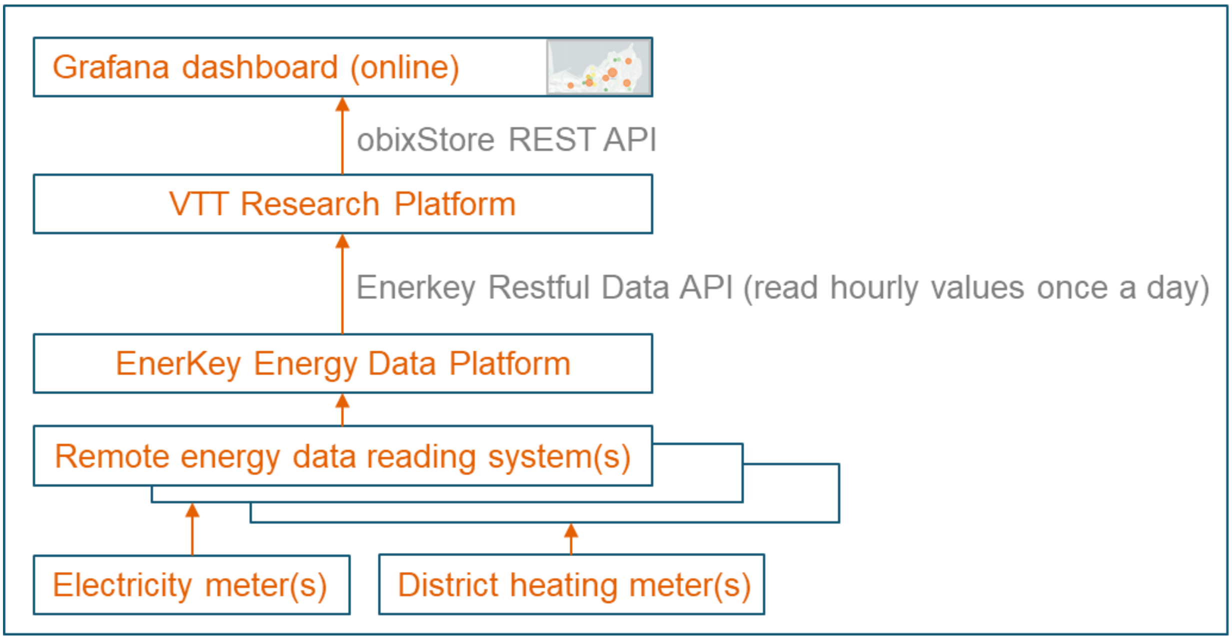


Figure 7. Data collection from EnerKey Energy Data Platform to VTT research platform.

An example of Otaniemi district heating and electricity consumption dashboards are shown in Figure 8. These dashboards are utilising VTT research platform API for data collected by Enerkey REST Data API and visualised by research platform connected open source Grafana dashboard utilising Leaflet (JS Library for interactive maps), OpenStreetMap and CartoDB.

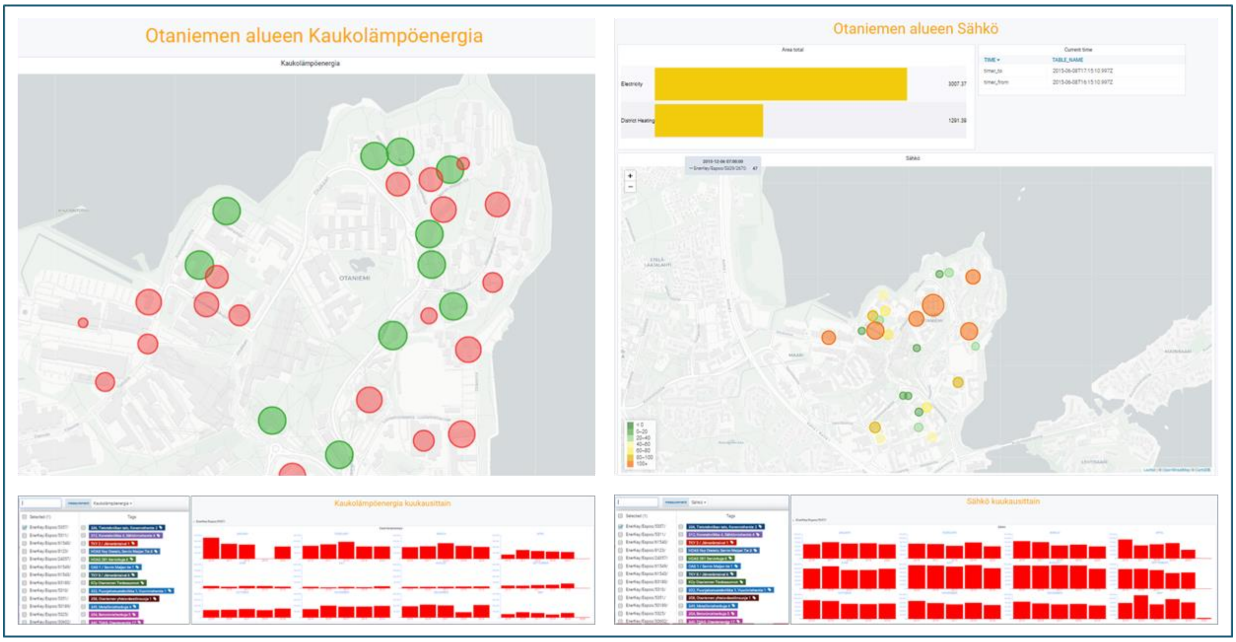


Figure 8. An example of Otaniemi area buildings district heating and electricity consumption dashboards (left district heating, right electricity and below selected buildings monthly consumptions).

4.2 PV panel real-time monitoring dashboard (case GEF)

The PV panel's cumulative electricity generation values (by GEF inverters) are collected an average 15 seconds interval to VTT research platform via GEF vision API. The power values are calculated online using VTT's obixCalculator, triggered when GEF API data is changed, as shown in Figure 9.

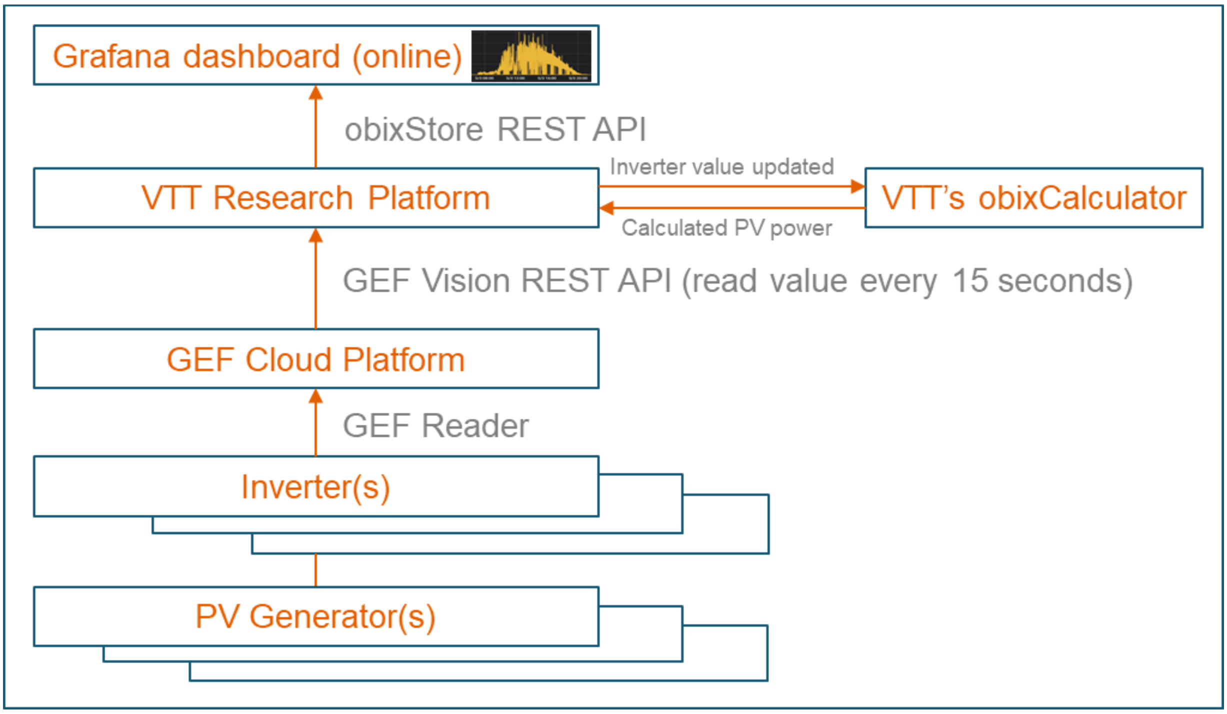


Figure 9. Data collection from GEF's PV panel related inverters via GEF Cloud Platform to VTT research platform.

An example of Grafana based PV panel real-time monitoring dashboard for cumulative and active electricity generation (case GEF) is shown in Figure 10. The PV panel monitoring dashboard allows end-users to select the measured and calculated values from the last five minutes to collected history trends.

Figure 10. An example of a PV panel real-time monitoring dashboard for cumulative and active electricity generation (case GEF inverters in Väre building, presented time resolutions are 30 days, seven days, manually zoomed and 15 min).

4.3 Electric power real-time monitoring dashboard (case ABB CMS-700)

The connection of Micronova building-related ABB CMS-700 to VTT research platform is shown in Figure 11. Measurement values (active power and energy) are read in one-second interval via Modbus TCP field bus and stored to the platform.

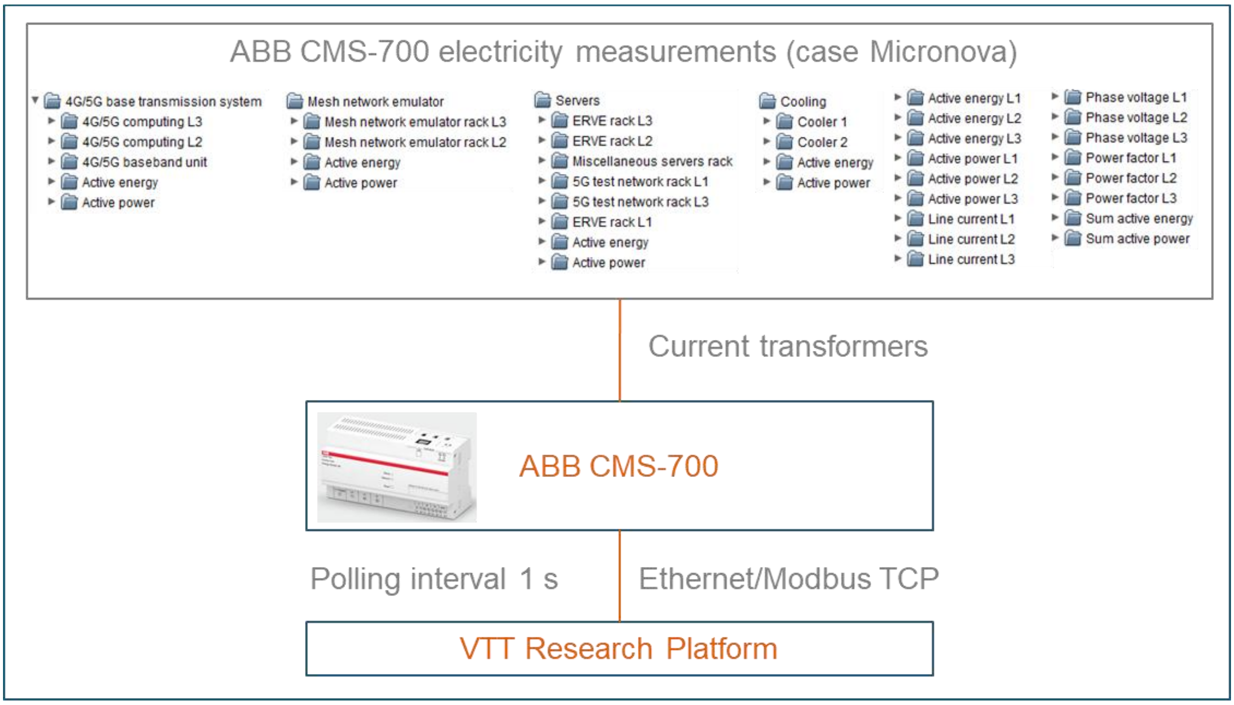


Figure 11. ABB CMS-700 connection to VTT research platform.

An example of an electric power real-time monitoring dashboard (case ABB CMS-700) is shown in Figure 12.

Figure 12. An example of an electric power real-time monitoring dashboard (case ABB CMS-700).

4.4 Clivet chiller unit real-time monitoring dashboard (case e2m Fingrid market access in Kemistintie 3)

The access of Kemistintie 3 chiller (Clivet) to Fingrid energy marked was done using Energy2market (e2m) technology. This solution was also connected to VTT research platform via Fidelix FX-3000-C (Figure 13).

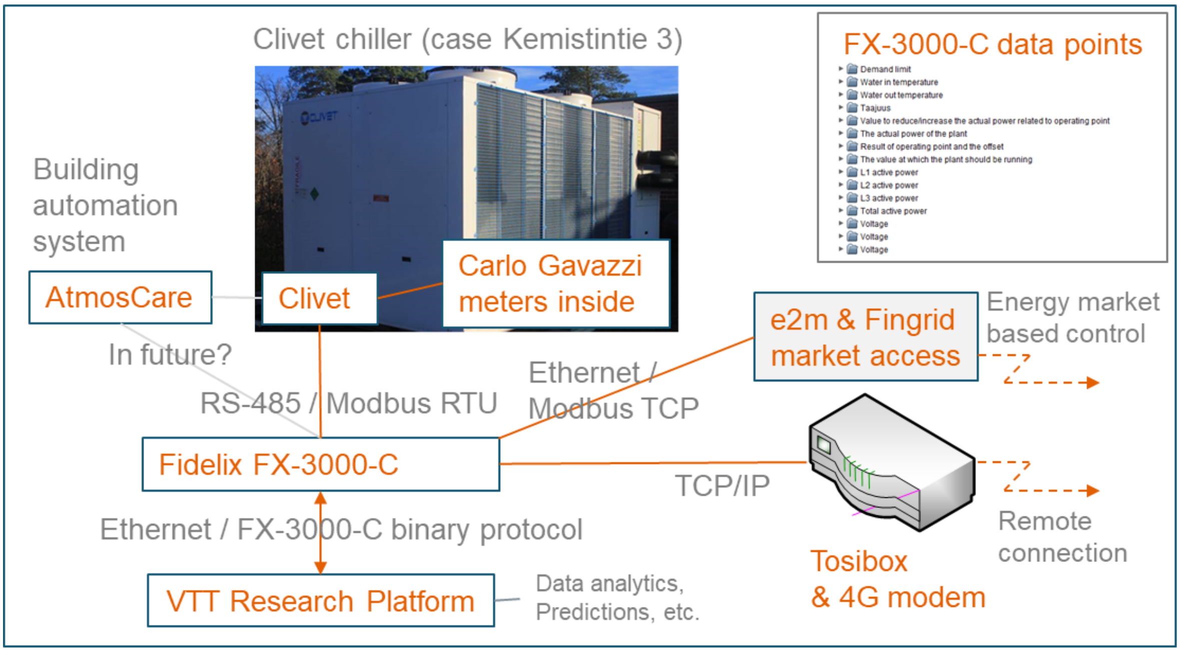


Figure 13. Access of Kemistintie 3 chiller to Fingrid energy market by e2m technology and related VTT research platform connection by Fidelix FX-3000-C.

An example of the VTT research platform based chiller unit real-time monitoring dashboard (case Kemistintie 3 Cliver chiller unit) is shown in Figure 14.

Figure 14. An example of a VTT research platform based chiller unit real-time monitoring dashboard (case Kemistintie 3 Cliver chiller unit via Fidelix FX-3000-C).

4.5 Dashboard for digital twin based forecasting electricity consumption - case Kemistintie 3

Digital Twin for Kemistintie 3 electricity consumption and related forecasting has been tested using Gated Recurrent Units (GRU, a gating mechanism in recurrent neural networks) RNN neural network by VTT research platform data.

An example of the platform connected Grafana dashboard for digital twin based forecasting electricity consumption (case Kemistintie 3 main electricity meters 1 & 2) is shown in Figure 15.

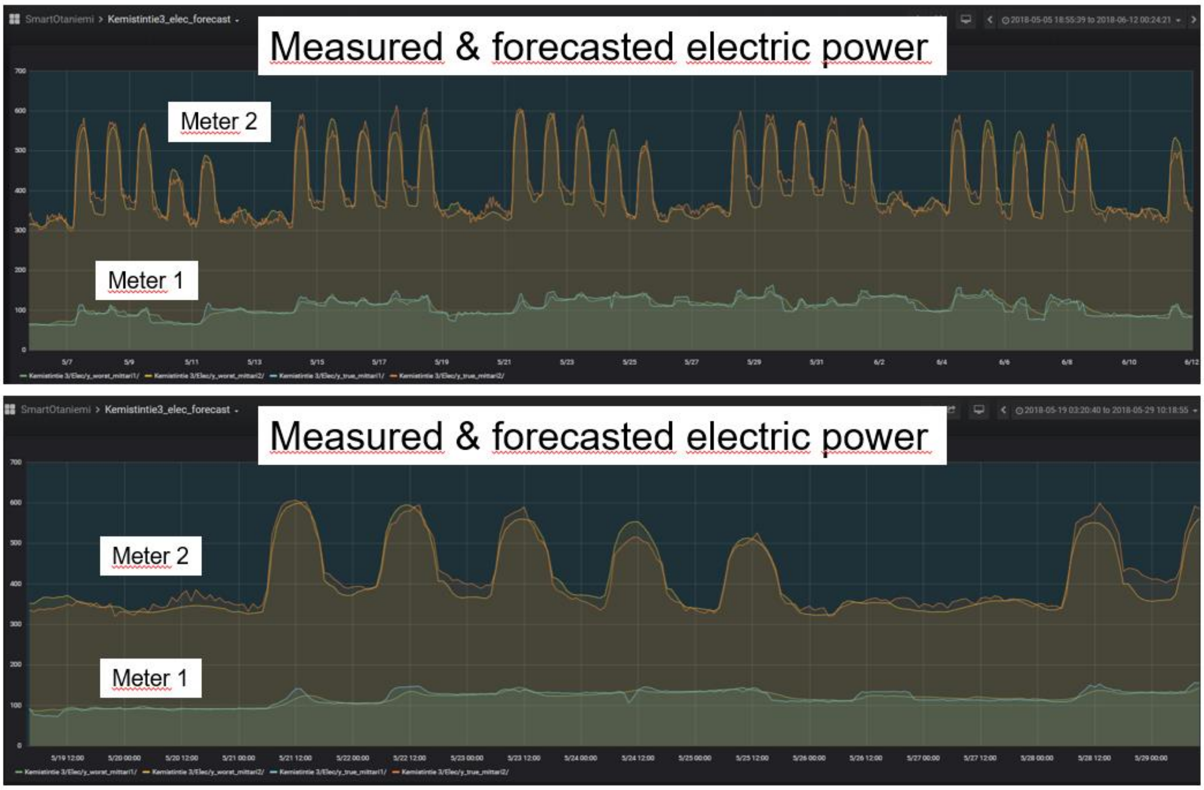


Figure 15. An example of the VTT research platform connected Grafana dashboard for digital twin based forecasting electricity consumption (case Kemistintie 3 main electricity meters 1 & 2).

4.6 Space monitoring dashboard (case Vuorimiehentie 3, space A105)

Vuorimiehentie 5 space monitoring (and control) use two different sources;

1) EnOcean based energy harvesting sensors (temperature, humidity, etc.) and energy harvesting radiator thermostats (space heating control) and

2) Indirect Computec Building Automation System data points by Granlund Metrix REST API (Figure 16).

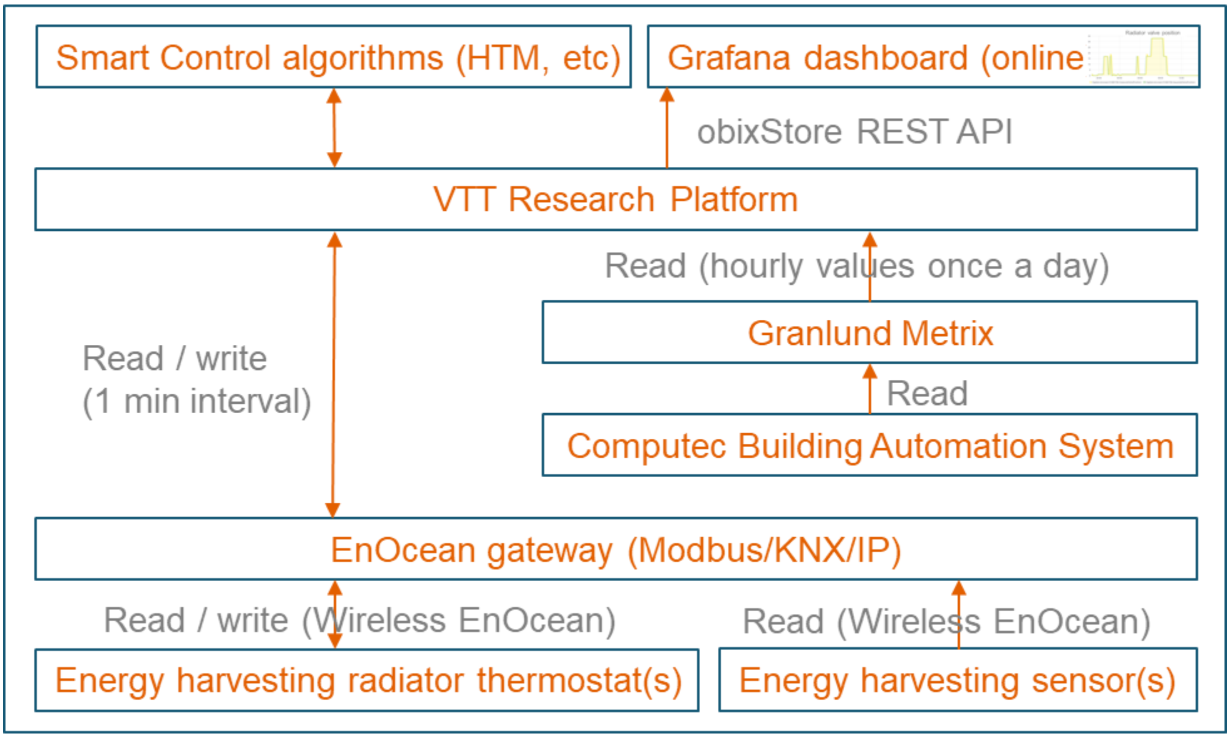


Figure 16. Data transfer between VTT research platform and energy harvesting sensors (space monitoring) and energy harvesting radiator thermostats (space heating control), case selected rooms in VTT main building at Vuorimiehentie 3, Espoo Otaniemi.

An example of the VTT research platform connected Grafana dashboard for Vuorimiehentie 3 space measurement data is shown in Figure 17, for the last two years data. Figure 18 zooms to one selected day data, whereas Figure 19 illustrates data for the last two days.

Figure 17. An example of the VTT research platform connected Grafana dashboard for monitoring space data (case Vuorimiehentie 3, Space A105, last two years of data).

Figure 18. An example of the VTT research platform connected Grafana dashboard for monitoring space data (case Vuorimiehentie 3, Space A105, zoomed to one-day data).

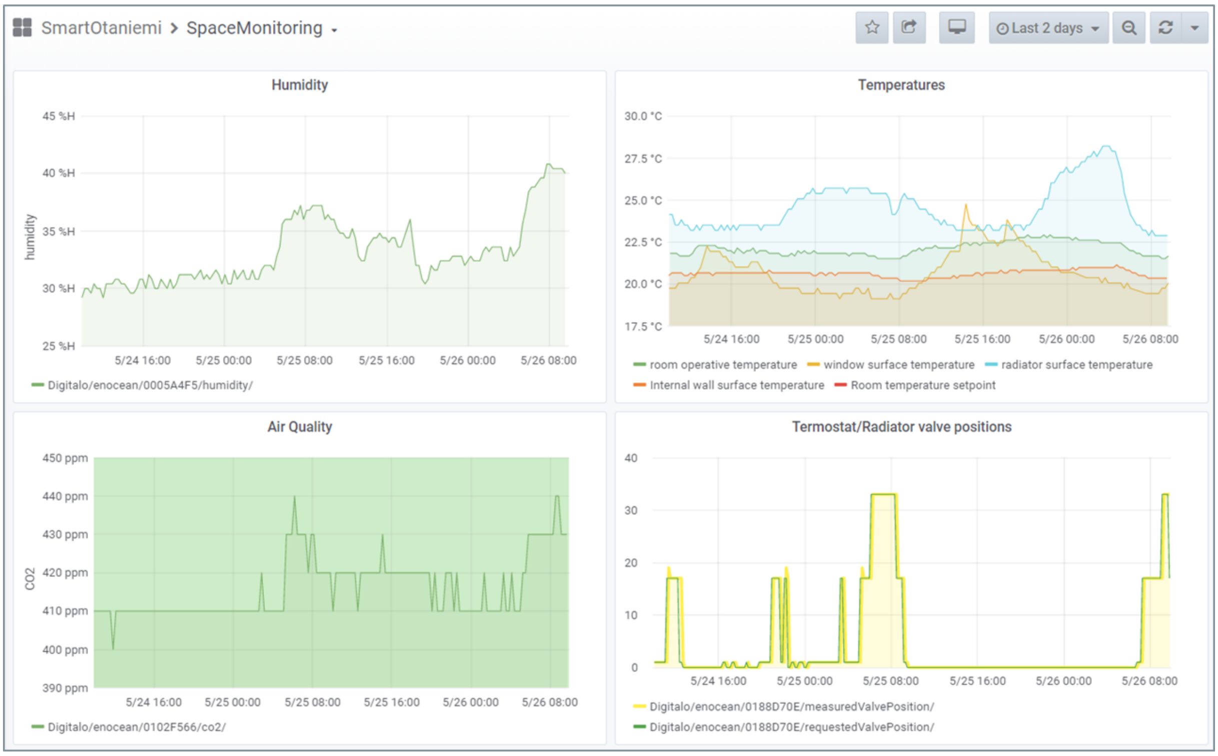


Figure 19. An example of the VTT research platform connected Grafana dashboard for monitoring space data (case Vuorimiehentie 3, Space A105, data for the last two days).

The same kind of data (thousands of data points per building) has been collected, monitored and analysed from many Otaniemi area building via building automation system, e.g., by Schneider AtmosCare building automation system in Otakaari 4 and Fidelix building automation system in Aalto Inn at Otaranta 4. However, lately, some GDPR and data access right issues have emerged.

4.7 Dashboard for digital twin based control of individual indoor thermal comfort (case studied human in Vuorimiehentie 3 space A105)

The used digital twin solution is utilising several data sources:

- space model (e.g. Vuorimiehentie 3 space A105) and its
- related measurements (energy harvesting sensors supported by building automation measurements),
- energy harvesting and wireless controllable radiator thermostat,
- VTT's human thermal model and
- the end-user QR code-based (smartphone application) related feedback of their thermal sensation (Figure 20).

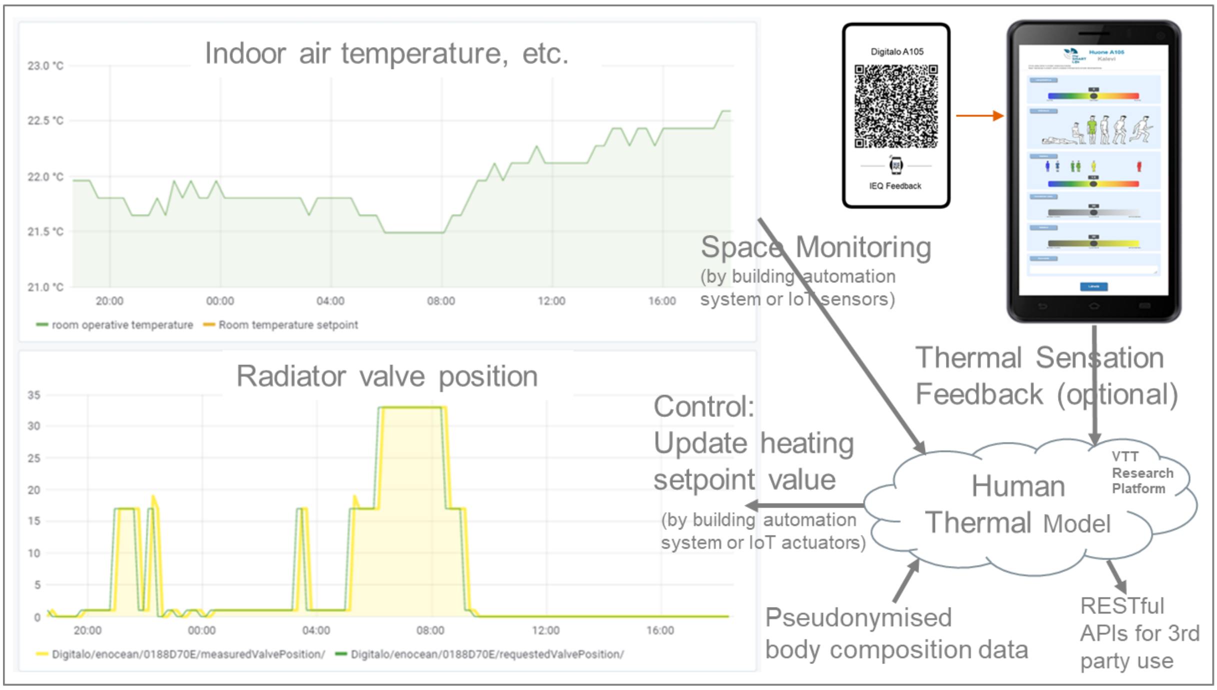


Figure 20. An example of the VTT research platform connected control application concept (HTM based digital twin for individual thermal comfort control).

An example of a dashboard for digital twin based control of individual indoor thermal comfort (case studied human in Vuorimiehentie 3 space A105) is shown in Figure 21.

Figure 21. An example of a dashboard for digital twin based control of individual indoor thermal comfort (case studied human in Vuorimiehentie 3 space A104).

4.8 Smartphone and QR codes based monitoring

An example of QR code-based smartphone application related monitoring of VTT research platform data (QR codes in space) is shown in Figure 22. The QR codes are typically placed near the studied device or sensor.

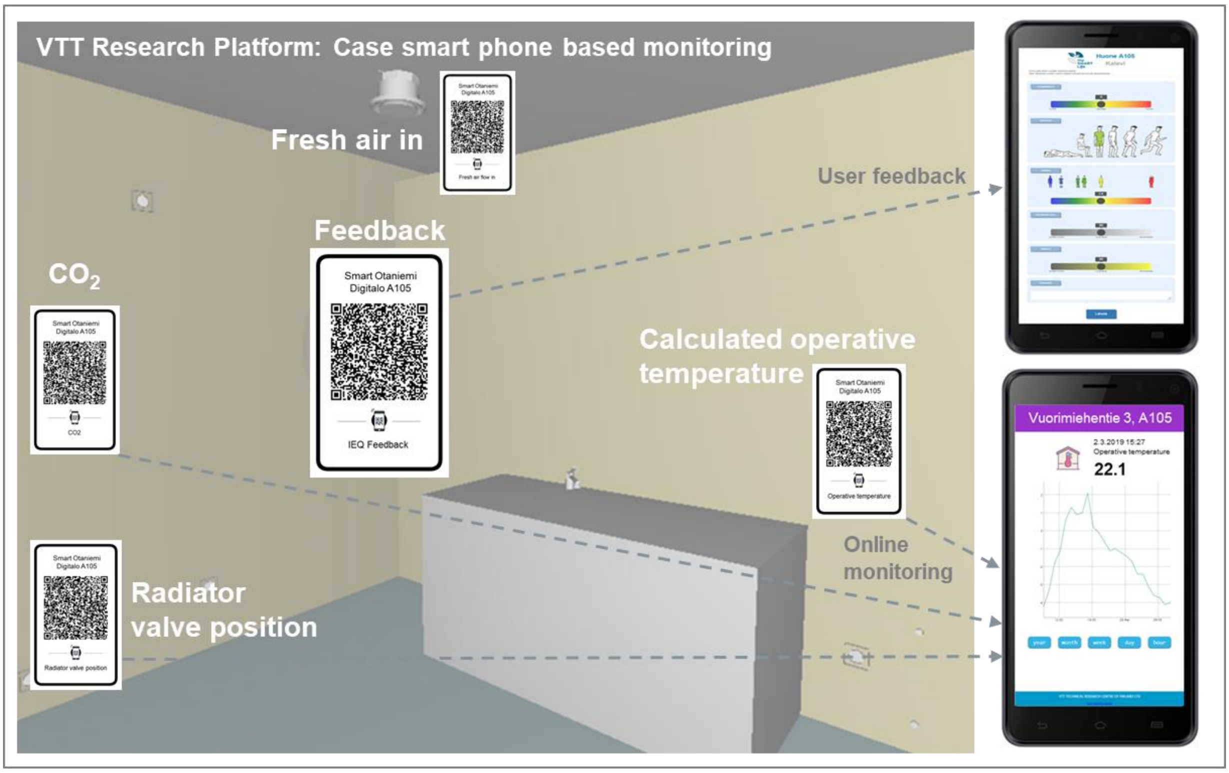


Figure 22. An example of smartphone-based monitoring of VTT research platform data.

The same kind of monitoring (by smartphone or web browser) is available also by browsing all (if authorised) thousands of VTT research platform obix data points (measurements, control signals, feedbacks, simulation results, data analytics results, etc.).

4.9 3D BIM-based visualising and analysing Smart Otaniemi data

A solution for mapping, visualising and analysing Smart Otaniemi data by 3D BIM Models via VTT research platform is shown in Figure 23.

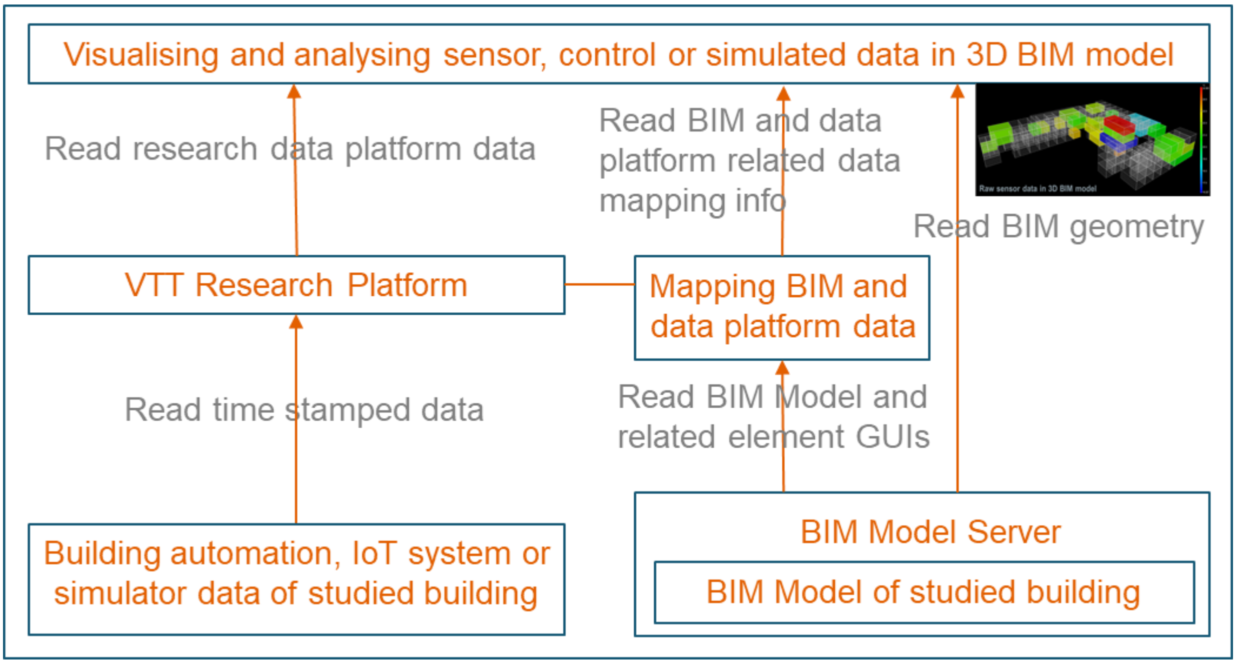


Figure 23. Solution for mapping, visualising and analysing Smart Otaniemi data by 3D BIM Models via VTT research platform.

An example of 3D BIM-based visualising and analysing Smart Otaniemi data is shown in Figure 24. The studied building was Otakaari 4, and the used building automation system was Schneider's AtmosCare. The used BIM Model server was based on open-source BIMserver.org via VTT's BIM Server client, VTT's BIM and BACS mapping solution and VTT's BimZone visualising and analysing client application.

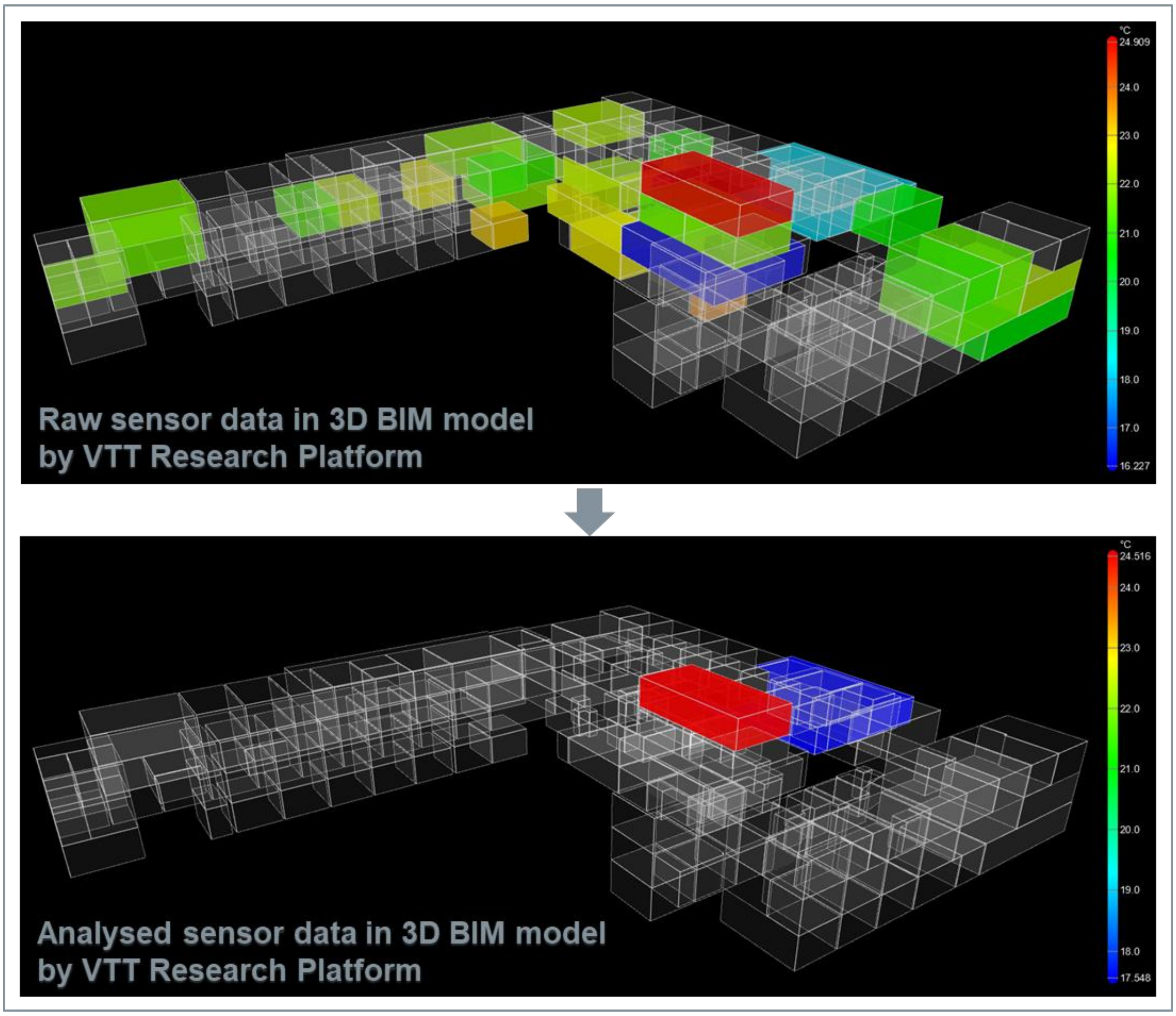


Figure 24. An example of 3D BIM-based visualising and analysing Smart Otaniemi data.

This 3D BIM model-based tool (VTT's BimZone) make it possible to analyse and visualise VTT research platform data. For example, we can study, in which place (e.g., space) and in which time (e.g., last summer) the selected data point value (measurement, a control signal, feedback, simulation result, data analytics result, costs, etc.) is/was higher or lower than selected value.

In other projects, also the Unity game engine based 3D model has been used to visualise the platform data.

4.10 File imported datasets

Data files (CSV or Excel) related timestamped measurement data can be imported to VTT research platform (Figure 25).

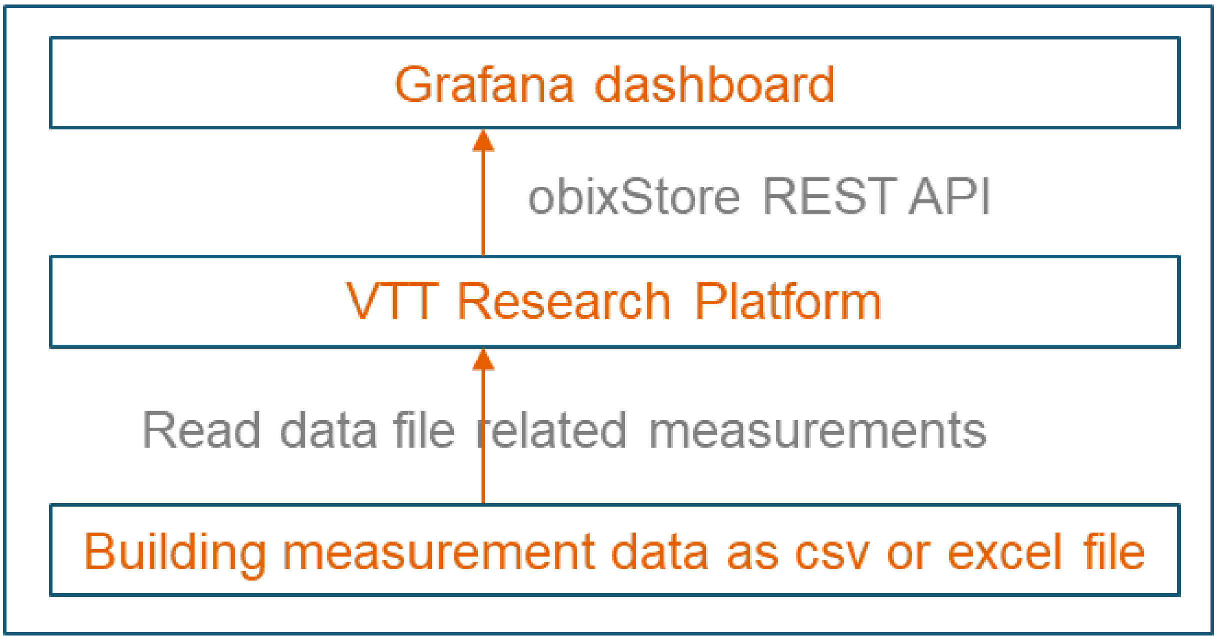


Figure 25. The concept for importing timestamped measurement data from files to VTT research platform.

An example of an Excel file that is imported to the VTT research platform (visualised by Grafana dashboard) is shown in Figure 26. Data platform file import is used when an online connection is not possible to use.

Figure 26. An example of Excel file imported data to VTT research platform visualised by Grafana dashboard.

4.11 Open data on VTT research platform

Open data can be read and saved to VTT research platform, e.g. by REST APIs (Figure 27).

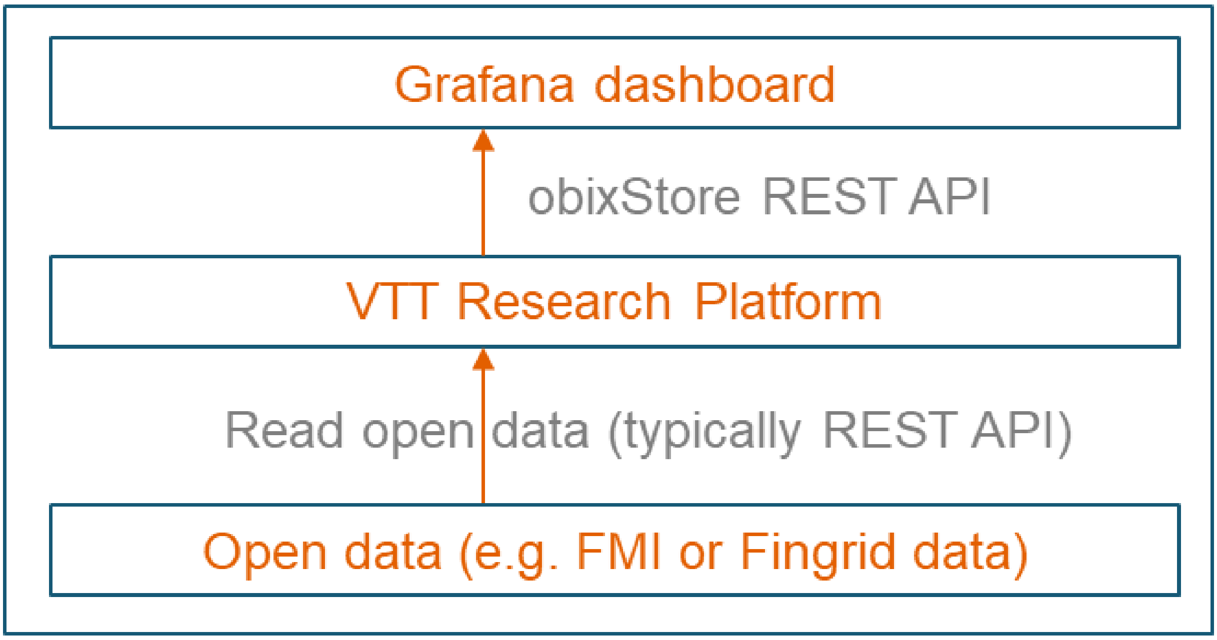


Figure 27. The concept for reading open data to VTT research platform.

An example of open data on the VTT research platform (case FMI's weather data) is shown in Figure 28. This data from over 200 FMI's weather stations are an important input for different simulation, forecast and analytic models. Open data is also collected continuously by Fingrid's open data APIs and City of Helsinki public building-related energy consumption (district heating, electricity) by Nuuka open data APIs.

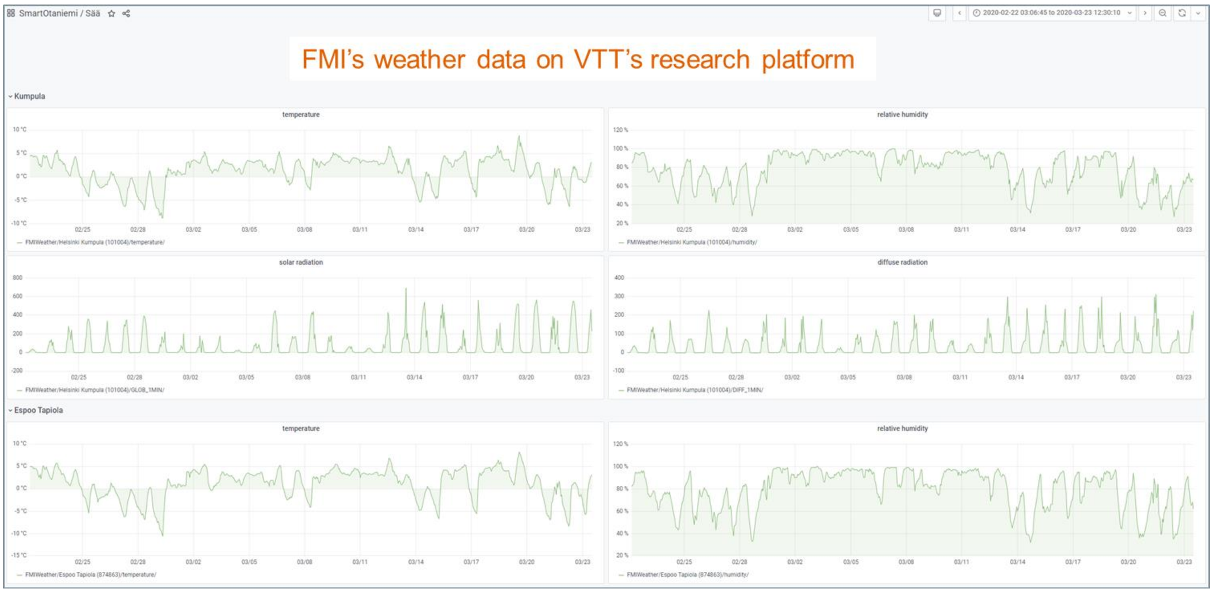


Figure 28. An example of open data on the VTT research platform (case FMI's weather data).

4.12 Third-party application (case Process Genius)

Third-party applications can read, subscribe and write data to VTT research platform via harmonised research platform REST APIs (Figure 29).

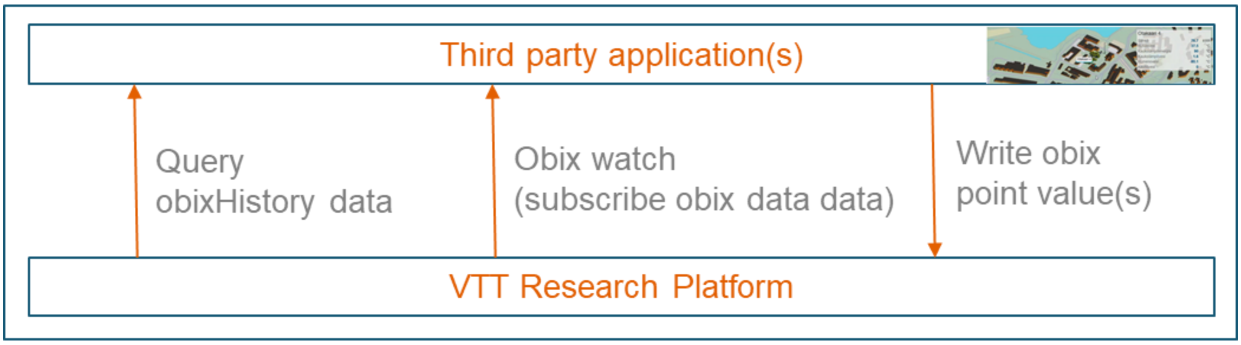


Figure 29. The concept for utilising (reading, subscribing and writing) data from VTT research platform by third-party applications via harmonised REST APIs.

An example of the VTT research platform connected third party application (case Process Genius) is shown in Figure 30.

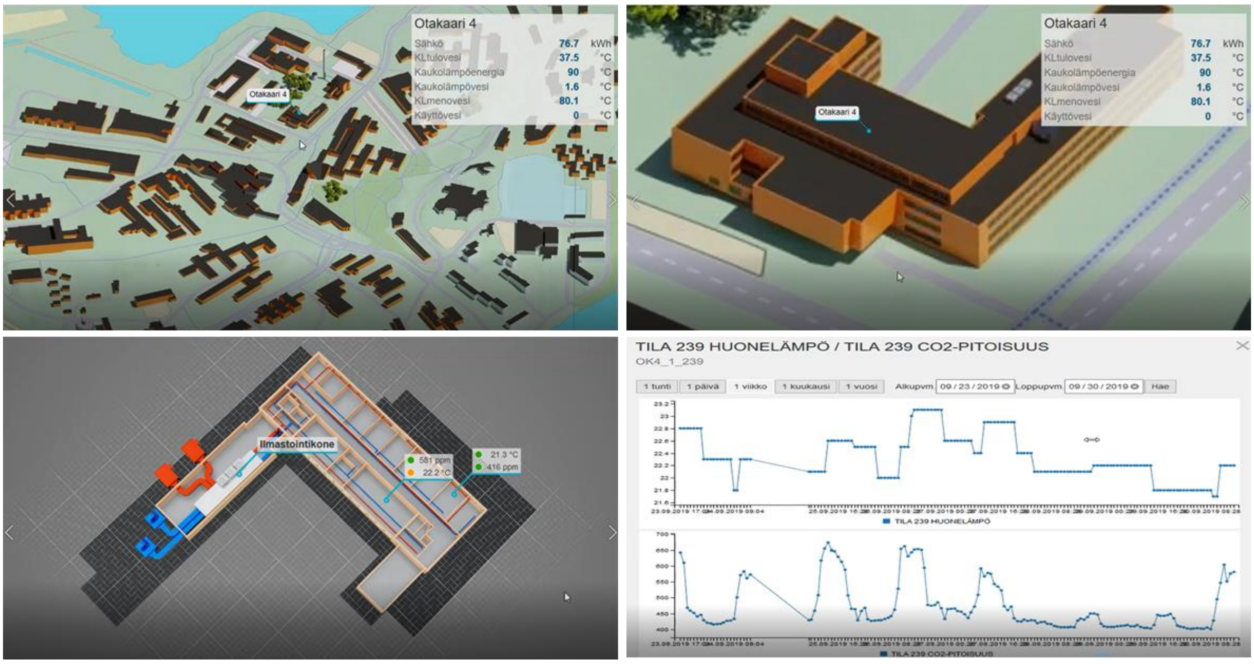


Figure 30. An example of the VTT research platform connected third party application (case Process Genius).

The Process Genius application includes the Otaniemi area 3D model, including Otakaari 4 BIM model and an online connection to VTT research platform by platform REST API (query obixHistory data).

5 From VTT research platform to a commercial platform

5.1 Companies' current challenges in sharing data

In general, the challenges in utilising data related to the built environment have been the same as in many other industries. Data is stored in various formats and distributed across multiple information systems that are not interoperable (Halttula, 2020). Besides, there may be deficiencies in the reliability of the data. These challenges are not a surprise, since the use of data is complex, and includes several issues that can be categorised into technical, security, openness, availability, processes and politics (Kortelainen *et al.*, 2019).

VTT researchers investigated the current challenges in inter-company data sharing by conducting nine interviews with the Smart Otaniemi WP2 consortium companies. The interviews took place between September and December of 2019.

The current challenges were grouped into the following categories:

1. **Problems in data acquisition and harmonisation**

   Some companies that are willing to use data for creating data-based services are encountering challenges in data acquisition. For example, data are hard to collect, or other companies are unwilling to share data since data can potentially be valuable, but the value of data is challenging to determine. Another challenge is that data are in the wrong format and need to be harmonised before usage. A couple of companies mentioned that they would need help cleaning up their data into a standardised format. An interviewee mentioned that if a company does not have the skills needed to use its data, one option is to sell data for other companies that have capabilities to create data-based services. However, this option necessitates understanding the value of data.

2. **Lack of use cases and capabilities for using data**

   Some companies lack use cases for their data. These companies are unsure of how they could use their data, whereas some other companies do not yet possess the capabilities to use their data. One interviewee suggested that the requirements for sustainability might provide the needed urge for using data. Customers are already looking for products and services that can be proven to be sustainable, and data can help the companies in this task. In addition to sustainability thinking, service thinking and service design might be the approaches that could help in creating customer-oriented and data-driven services.

   Some companies were willing to start experimenting with inter-company data transfer based on available data concerning energy and electric vehicle charging. The interviewees mentioned three possible use cases for inter-company data sharing: 1) smart electric vehicle charging, 2) remote monitoring, and 3) load control and aggregation. However, the interviewees raised concerns about the yet unclear energy aggregation business models on the consumer side. There is a lack of knowledge of what kind of data consumers consider worth paying for.

3. **Lack of common rules for data sharing**

   Many companies mentioned that no common and shared rules for inter-organisational data sharing exist. Companies want to be sure their data is not misused and that the agreed rules are followed.

In summary, inter-company data sharing is still in its infancy in the interviewed companies. Data is mainly used internally to make companies' business more efficient. Many companies have not identified commercially viable use cases for data, even though many companies recognise the potential for new business. This is in line with Gartners' study that found one-third of organisations struggling with finding suitable use cases for data (Andrews, 2018). Some companies fear that they do not understand the value of their data, and are thus hesitant to share it with other companies, especially as rules for data sharing have been missing.

All companies recognised the potential value of data to provide better services to their customers or develop their internal processes. For example, some companies would like to share their massive data sets to develop their equipment or provide better service to their building management service customers. Some of the recognised use cases for data would have necessitated inter-organisational data sharing. For this reason, the project tried to identify potential data-sharing platforms.

DigiPLACE project, for example, has listed 308 existing digital platforms, most of them specific to the construction sector (DigiPLACE, 2019). Most platforms are developed in the USA, but several platforms also originate from Europe. DigiPLACE project categorised most of the platforms as ‘collaborative platforms’ used by different project parties to exchange data and documents for better communication and more effective project management. However, commercial platforms providing marketplaces for digital, machine-to-machine readable, data and services are less developed. In this report, a data-sharing platform is defined as a multi-layer cloud service or network of services that can enable, e.g., connecting, storage, analysis and services for different data sources. This definition entails the ideas of machine-to-machine readable data but also a ‘data model’, which enables the user to understand the data flows in a human-readable format. In general, a data model illustrates the organisation of data elements and how they relate to one another and real-world entities' properties[7].

However, the consortium companies could not agree on an existing commercial platform that could be used for inter-company data sharing and service creation. First, the companies pondered whether they could collaboratively create such a commercial platform and later start a new company. However, building a platform from scratch takes time and requires large amounts of resources and knowledge on how the data platform markets function.

According to a study, two trade-offs are encountered when building a digital platform (Karhu *et al.*, 2020). The first trade-off is ‘variety versus unity’, where the platform owner faces the trade-off between attracting a variety of complements, while simultaneously preserving the platform’s unity. The second trade-off is ‘open versus closed’, where the owner faces the trade-off between the openness of the platform to attract collaborators, while at the same time avoiding exploitation by competing platforms. (Karhu *et al.*, 2020)

The same study found four digital platform tactics — leverage, control, exploit, and defence — that can be used to balance the two trade-offs, and boundary resources, which are the platform owner’s methods to execute the tactics. *Leverage tactic* is used to grow the platform user and complementor bases. Application Programming Interfaces (APIs), software development kits (SDKs) and open-source licenses are examples of boundary resources for opening the platform for complementors. (Karhu *et al.*, 2020)

Control tactic refers to platform rules that make sure complementors innovate within the set constraints. Guidelines, publisher terms and conditions, and client library are examples of boundary resources of the control tactic. *Exploit tactic* is used by competitors to create a competing new and separate platform. For example, a competitor may copy the platform resources if they are published under an open-source license, thus avoiding costs in creating their platform. (Karhu *et al.*, 2020)

Finally, the *defence tactic* is used to balance having an open platform that attracts collaborators and a platform that prevents competitors from exploiting the platform. Here the boundary resources are open-source licenses, such as GNU's General Public License (GPL) that forces the license user to share and license changes under the same terms, which forces the user to open source their modified platform. (Karhu *et al.*, 2020)

Eventually, the idea of creating an own platform was not continued. The use cases in different companies were very different, and the value of data was hard to determine, the path forward was not clear enough for that.

---

[7] See data model definition, e.g. in Wikipedia, available from: https://en.wikipedia.org/wiki/Data_model#:~:text=A%20data%20model%20(or%20datamodel,properties%20of%20real%2Dworld%20entities.

5.2 General business requirements for commercial big data platforms

Energy-related use cases often include large amounts of data that need to be processed and possibly shared between companies. Thus, in these types of use cases, a big data platform is needed. These platforms can handle huge amounts of data cost-efficiently. Several commercial big data platforms exist, for example, SelectHub[8] lists and ranks 51 big data commercial platforms. SelectHub lists ten key business requirements for big data platforms, which are next shortly presented.

1. Variety, velocity, veracity, volume and value of data
    - Variety (data types): support various forms of data. For example, big data technology should analyse and bring together data of different types, such as messages, social media conversations, photos, sensor data, and video.
    - Velocity (stream analysis): refers to the speed at which new data is generated and the speed at which data moves around. The platform can analyse data, while it is being generated (in-memory analytics) without putting it into databases.
    - Veracity (data uncertainty): refers to the trustworthiness of the data, i.e., how accurate the data is.
    - Volume (amount of data): refers to the amount of data generated every second. Big datasets tend to be nowadays so large that they cannot be stored and analysed using traditional database technology.
    - Value (business leverage): the platform should turn data into valuable information.
2. Database management
3. Data warehousing: an integrated, granular, historical single point of reference for data.
4. Data analytics
    - Big-data analytics for e-commerce: Data is often categorised into structured (name, address, preferences, sex, and age) and unstructured (likes, tweets, clicks, videos). How to turn unstructured data into meaningful insights.
    - In-memory analytics: In-memory analytics give concurrent, in-memory and multiuse access to structured and unstructured data to solve complex problems quickly. It should divide analytic processes into easily manageable pieces with computations distributed in parallel across a dedicated set of blades.
    - Predictive analytics: Predictive analytics is the practice of extracting information from existing data sets in order to determine patterns and predict future outcomes and trends.
    - Text analytics: Text analytics capabilities should capture the knowledge of domain experts into dictionaries and semantic rules for re-use. It should allow customisable information extraction for logical reasoning to draw inferences from natural, unstructured communications. It should recognise the entity and relationships to classify words that can be analysed for business meaning.
5. Hadoop integration
    - Data viewer for Hadoop: data viewer aids in consuming information from a Hadoop store.
6. Social media integration
    - Social analytics capabilities to analyse the data coming from social media streams.

---

[8] SelectHub, available from: https://www.selecthub.com/. Accessed 6.8.2020.

7. Reporting and portal
   - Report import/export capabilities to aid in visualising and consuming data.

8. Scalability
   - Ability to scale data through scale-in, scale-up and scale-out techniques. Optimise data's distributed storage layer and retune and rebalance data workloads regularly.

9. Technical requirements
   - On-premise installation: system should be able to be installed on-premise to provide a privately hosted system
   - Hosted installation: the system is available as a hosted, software-as-a-service (SAAS) offering

10. Functional requirements
    - Database software as a service (dbSAAS): capabilities to be licensed on a subscription basis and can be centrally hosted.
    - Unstructured text management and search: effectively manage unstructured text and enable search capabilities on such data.
    - Dashboard: a single page report, which is displayed on a web page that is linked to a database, which allows the report to be constantly updated.

5.3 Smart Otaniemi WP2 companies' technical requirements for a commercial platform

The interview analysis, concerning Smart Otaniemi WP2 companies' expectations for a commercial platform, revealed ten factors that need to be provided by the commercial platform:

1. **Platform owner** – The platform needs to have an owner, with whom the contract for sharing data is signed. Data sharing is executed for commercial purposes.

2. **Data ownership** – Data should stay in the control of the initial owner if it is so decided. In practice, this means that the owner should be able to determine to whom s/he wants to share the data.

3. **Data availability** - Some companies were concerned about how to ensure that data is all the time available. For example, contracts concerning data are often written so that data can be used only for a certain period of time, such as during a research project. Sitra's IHAN programme's Rulebook workgroup has written the first version of the IHAN Rulebook for Data Networks to facilitate the creation of many-to-many agreements needed in many-sided data ecosystems (Pitkänen *et al.*, 2019). In addition, a commercial platform should not lock-in their users in the spirit of 'fair data economy' (Ilves and Osimo, 2018).

4. **Data integrity** - Data should be accurate, truthful, complete, retrievable and verifiable throughout its lifecycle.

5. **Data's metadata** – Metadata provides information about the data, which makes finding relevant data easier. Examples of metadata are author, date created, location, security class, and units

6. **Data security** – Data security refers to the protection of data from unauthorised access. Security issues can be relevant in the context of sharing energy data since energy data is part of Finland's crucial infrastructure.

7. **GDPR** – General data protection regulation (GDPR) gives control to individuals over their personal data by forcing organisations handling their data to have appropriate technical and organisational measures to protect data. GDPR applies to organisations operating in the EU or offering good or services to individuals in the EU. GDPR may be relevant in the context of charging an electric vehicle if the data related to it is used and the charger is revealed. Thus, data should be anonymised.

8. **Data access control** – Data access control should guarantee that only the authorised users are given access to appropriate data. Data access control is also connected to the MyData approach, which aims at providing individuals with control over their data (Kuikkaniemi *et al.*, 2015). Data sovereignty is a similar term to MyData. International Data Spaces (IDS) is an international 'virtual' organisation aiming to ensure that persons and companies are capable of self-determining their data (Otto *et al.*, 2019). For this purpose, IDS has created a reference architecture model, which states the requirements for a secure and trusted data exchange in Europe.

9. **Performance requirements** – Performance requirements vary between use cases. As a rule of thumb, the higher the performance requirements, as it is in sharing energy-related data, the more specialised the database system should be.

10. **Contracts** – Contracts should specify how the data is used in the receiving organisation, how accurate and real-time the sending parties' data should be, how long the data should be available, and what the cost of using data is.

Several commercial big data platforms can meet these requirements. However, the use of these platforms brings costs, and in the early phases of inter-company data sharing, the commercial value of data is still blurred. This complicates the selection of a commercial data platform.

6 Conclusion and next steps

Data markets are emerging, and also the built environment actors want to start utilising their data for new types of data-driven services. Data are being collected more and more by both private and public organisations. Smart Otaniemi innovation ecosystem, specifically the VTT research platform, has enabled companies to develop and test data-driven use cases, such as remote inspection with drones and smart EV charging. However, the commercialisation of the use cases necessitates a data-sharing platform that allows the sharing of linked data, structured machine-to-machine readable data, for commercial purposes.

One major challenge with the term *data-sharing platform* has been that different actors understand it very differently. For example, large online cloud storage services such as OneDrive or Dropbox are often considered as platforms in general terms. However, these platforms do not enable the sharing of machine-to-machine readable data and automatic services, maybe also automatic payments with those services, which depict Smart Otaniemi ecosystem players' needs. Hence, in this report, a data-sharing platform has been defined as "a multi-layer cloud service or network of services that can enable, e.g., connecting, storage, analysis and services for different data sources".

The list of expectations for a commercial platform, presented by Smart Otaniemi WP2 companies can be met by an already existing commercial big data platform. The challenge is to agree on a platform that can be flexibly taken into use for inter-company data sharing, which is still in its infancy and does not yet promise commercial success. For this reason, the next research step is to try the data-sharing platform provided by the Platform of Trust (PoT). The preparations for the trial are ongoing. The use case has been named as 'better indoor conditions energy-efficiently using up-to-date data'. The data will come from a Finnish public school operating in Espoo. The idea is to combine indoor air data with energy data and apply an algorithm to modify parameters. As a result, better indoor conditions can be produced energy-efficiently.

References

Andrews, W. (2018), *Build the AI Business Case - A CIO's Guide to Building the Strategy and Business Case to Implement AI in the Enterprise*, available at: https://aiuserforum.com/wp-content/uploads/2019/02/ai-business-case-ebook.pdf.

DigiPLACE. (2019), *D3.1 - Level of Implementation of Digital Technologies and Digital Industrial Hubs/Platforms in the European Construction Sector and in Other Industrial Sectors*.

Dooley, K., Ihasalo, H., Jylhä, T. and Sairanen, S. (2017), *IoT Market Analysis in the Finnish Estate and Construction Sector*, available at: https://issuu.com/granlundoy/docs/iot-raportti_en_lores.

Halttula, H. (2020), *Enhancing Data Utilisation in the Construction Project Lifecycle through Early Involvement and Integration*, University of Oulu.

Horsmanheimo, S., Mehnert, S., Tuomimäki, L., Lembo, S., Hongisto, M., Huilla, S., Ahola, K., *et al.* (2020), *Technical Report on Connectivity Pilots*.

IDC. (2017), *European Data Market SMART 2013 / 0063 Final Report*.

Ilves, L. and Osimo, D. (2018), *Policy Brief - A Roadmap for a Fair Data Economy*, Helsinki.

Karhu, K., Gustafsson, R., Eaton, B., Henridsson, O. and Sørensen, C. (2020), "Four tactics for implementing a balanced digital platform strategy", *MIS Quarterly Executive*, Vol. 16 No. 3, pp. 197–213.

Kortelainen, H., Kortelainen, J., Pulkkinen, A., Juhola, A., Hemming, B., Kantorovitch, J., Heikkilä, E., *et al.* (2019), *Data Typology in Manufacturing Industries*.

Kuikkaniemi, K., Poikola, A. and Honko, H. (2015), *MyData – A Nordic Model for Human-Centered Personal Data Management and Processing*, available at: http://urn.fi/URN:ISBN:978-952-243-455-5 %5Cnhttps://github.com/okffi/mydata.

Otto, B., Steinbuß, S., Teuscher, A. and Lohmann, S. (2019), *Reference Architecture Model, Version 3.0*, available at: https://www.internationaldataspaces.org/wp-content/uploads/2019/03/IDS-Reference-Architecture-Model-3.0.pdf.

Pitkänen, O., Yli-Jaakkola, J., Mäkinen, J., Kuohuva, K., Hiekkanen, K., Koskinen, J., Malkamäki, S., *et al.* (2019), *Rulebook for Fair Data Economy - Rulebook Template for Data Networks*.

Teknologiateollisuus. (2019), *Model Terms of the Technology Industries for Data Sharing*, available at: https://teknologiainfo.net/sites/teknologiainfo.net/files/documents/pdf/Introduction to Model terms of the Technology industries for data sharing.pdf.

VastuuGroup. (2019), "Up-to-date conditions management with the help of data - Case Senate Properties", *Case*, available at: https://platformoftrust.net/en/cases/up-to-date-conditions-management-help-of-data-case-senate-properties/ (accessed 17 June 2020).

Ympäristöministeriö. (2020), "Rakennetun ympäristön tietojärjestelmä", available at: https://www.ym.fi/fi-FI/Maankaytto_ja_rakentaminen/Ohjelmat_ja_strategiat/Rakennetun_ympariston_tietojarjestelma (accessed 12 June 2020).