FT07.05_2023 Data Validation Centre

IoT-driven business design – digitization of companies and society

Introductory information

Deployment area	IoT-driven business design – digitization of companies and society
Institute	FORCE Technology
Title	Data Validation Centre
Numbering	FT07.05_2023
Version	1.0
Period	Januar 2023 – december 2023
Contact person	Henrik Hassing (hnh@forcetechnology.com)

Changes

This is the first version of the activity description for 2023. It builds on activities and results completed in 2022.

Description

Objectives

Since 2022, the ambition with this activity has been to establish a centre for data validation and to develop the necessary methods, frameworks and related tools. The Data Validation Centre is a virtual centre of competences that will help both private companies and public institutions ensuring correct, accurate and useful data for their IoT applications – in other words assurance of data quality. The issue with quality of data is well-known in relation to IoT solutions where many different error sources can impact the data captured by sensors and the transportation of data in the network. These error sources can be both temporary and more permanent ones. The return on investment of an implemented IoT solution is very dependent on the quality of the data provided and used by the solution.

Establishing the Data Validation Centre is a significant part of the activity described herein, since the centre will provide expert knowledge within this area as well as establish a service suite that can be offered to support companies' development and implementation of IoT-based services. For instance in relation to the green transition, the digitalization and Industry 4.0, it is crucial to have reliable and high quality data, hence data validation is a central part of implementing an IoT solution. Through the 2021-2022 activities and projects with the SME Christonik and Frederiksberg municipality, it became clear that IoT data validation is a critical activity and must be carried out for all IoT applications. Furthermore, enterprises and organisations need effective tools to diagnose and fix data quality issues as data errors and anomalies can have a severe negative impact on services and operations employing IoT solutions. In addition, the Data Validation Centre within the framework of the Nordic IoT Centre will contribute to raising the awareness about the importance of data validations and how to systematically work with data quality which is essential for companies developing, delivering, procuring, and operating IoT applications.

The activity will also contribute to international standardization which will help promote and establish data validation as a key part of IoT solution development and operations. The concept of data validation and the process will be added to an IoT-related standard currently under development.

The activity plan for 2023 will continue the activities from 2022, focusing on:
- Further development of the Online Data Validation Service (ODVS) tools and applications
- Deployment of an ODVS on suitable infrastructure (e.g. cloud)
- Implementation of new functional features of the ODVS based on feed-back from testing with enterprises and organisations
- Contribution to an international standard for IoT data, to also include data validation

Content

Data itself can be in many formats. It can be sensor values, images, audio or video streams. It can be stored in a database or arrive as a continuous event stream. The primary focus of the ODVS is to validate IoT data received from SMEs as timeseries (repeated sensor values over time). There are only a few companies in the world offering services to validate IoT data.

The focus of the activities in year 2023 will be to further automate the ODVS and develop services, methods, and advance algorithms through which the process of manual IoT data validation can be reduced. The activity plan consists of the following activities that support the objectives above:

1. **Further development of the Online Data Validation Service (ODVS)**
   With the roadmap defined in 2022, the development of the OVDS in 2023 will focus on how to automate the service even further, (Phase 2). Please refer to the figure below, activities mentioned in 1, 2, 3, 4 and 7 were part of Phase 1 and they are already completed. Phase 2 comprise the activities 5, 6, 8, 9, 10, 11 and 12 will be the focus of the future development. If possible, the Phase 3 will be initiated as well, completing the automation of the services, which enables sending the automatically generated report to the customer. The Phase 2 activities in 2023 includes development and deployment of next generation ODVS.
   The increased automation and online execution of the data validation process will save cost and make data validation accessible to SMEs and enable efficient introduction and application of data validation as part of the overall operation of IoT solutions.

2. **Contribute to the development of an international standard within data validation**
   In 2022, several standards were reviewed to find out if they were candidates for including data validation, or if a new standard should be created. One international standard called "ISO/IEC 30178 – Data format, value, and coding" was identified as a candidate. It is under development and the scope has been expanded to include data validation. In 2023, work will be continued by contributing to the standard primarily within the area of data validation. The development is expected to be completed in 2024. The developed standards will help promote uniform processes and requirements to perform data validation across the industry, which is particularly important considering the explosive growth of IoT data across many different business segments, e.g. manufacturing, utilities, energy sector and Smart City applications.

3. **Execution of the data validation process from ODVS for further development**
   IoT data validation is a new concept on a national and international level. Feedback from the ecosystem is critical to ensure that the ODVS is being developed in the right direction to solve companies' needs. Therefore, the offered data validation process will be tested with 10 data sets received from companies via the online service, the results of which will provide input to further refinement and feature enhancement of the service. Furthermore, the test runs will ensure that the development focuses on the parts of the process with the largest optimization potential, e.g. data export, validation, visualization, processing and reporting process and algorithms.
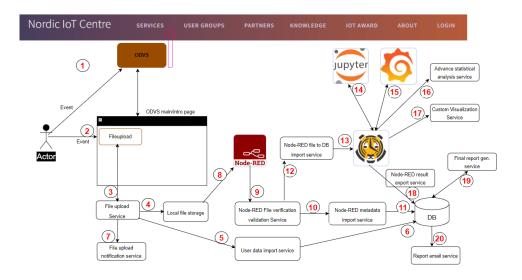
Figure 1 High level online data validation process and tools

Actors

The activity plan particularly involves FORCE Technology's and Alexandra Institute's competencies within IoT, software development, cloud computing and data analytics. It will be carried out in close collaboration between FORCE Technology and Alexandra Institute

Synergies/collaboration with other projects

As part of the activity plan, coordination is ensured with the following other efforts and projects, so that knowledge and services developed under the auspices of these become available to the target group.

- *"Fremtidens hybride testbeds"*, FORCE Technology: about the use and validation of data from tests
- *"Metrologi for digitalisering and datasikkerhed"*, DFM (lead) and FORCE Technology: The activity contributes knowledge about metrology within data validation of IoT data in particular
- *"Digital vandløsninger til den grønne omstilling"*, DHI (lead) and FORCE Technology: The activity contributes knowledge about IoT environmental sensors and IoT systems in harsh environments
- *"Digital sikkerhed, tillid og dataetik"*, Alexandra Institute (Lead) and FORCE Technology: Joint interest to find anomalies in datasets, more specifically for cybersecurity
- *"Digitale teknologier til datadrevet, bæredygtig vækst"*, Alexandra Institute (Lead) and FORCE Technology: Joint interests to improve the processes around data collection and transformation

Other R&D projects:

- This activity is expected to create synergies within the area of data validation together with projects from Energy Cluster Denmark (one project) and AI Denmark (one project).
- Also, FORCE Technology will apply for additional funding related to Data Validation in order to investigate further the synergies related to the national clusters such as AI Denmark, We Build Denmark and Energy Cluster Denmark

Advisory Board (Følgegruppe)

The activity plan has been presented at the Advisory Board meeting on Monday 28 November 2022.

Knowledge dissemination

Results developed under the activity plan are disseminated via the Nordic IoT Center (nordiciot.dk) and associated follow-up group and stakeholder groups. The activity includes communication of the results in at least

one related conference. The specific activities for knowledge dissemination are described in FT07.09_2023 Knowledge dissemination and ecosystem.