ESS Control System Data Lab - Executive Summary

Runeson, Per; Andersson, Per; Hall, Anna; Larsson, Jan Eric; Rathsman, Karin; Söderberg, Emma

2021

Document Version:
Publisher's PDF, also known as Version of record

ESS Control System Data Lab

Executive Summary

Per Runeson[1], Per Andersson[2], Anna Hall[3], Jan Eric Larsson[2], Karin Rathsman[4], Emma Söderberg[1]

[1]Lund University, [2]GoalArt, [3]Big Science Sweden, [4]ESS

February 19, 2021

Driven by the idea to use alarm data to explore machine learning across Industry 4.0 applications, the goal of this pilot study was to explore how to collect, store, manage and share data from the ESS Control System. Generally, we seek to make any control system data available for research and innovation but started with alarms as a feasible domain in which to explore machine learning. The goals were threefold, each explored in a work package:

1. How to govern a data ecosystem, and which tools are needed to support it?
2. How can alarm data be interpreted across industrial contexts, i.e., which meta data and reference models are needed?
3. How can data sharing be practically and legally handled at ESS?

In summary, we identify a set of potential alleys for continued work to foster industrial innovation and collaboration in a control system data ecosystem with ESS as a catalyst.

Project results

The governance topic was explored, comparing ESS with two other emerging data ecosystems (RoDL in automotive, and JobTech in the job market domain), and we conclude that there are potential, mutual benefits in data sharing and collaboration for ESS and other research and industry actors. However, as the keystone actor, initiating the ecosystem and providing the initial data, ESS has to continue to advance the data ecosystem and ensure that there is a common interest among the stakeholders. Further, a technical solution to support access to data and meta data, while preserving the integrity and security of the ESS facility should be established.

The alarm data at ESS is well structured and managed, and the needed meta data exists or can be created, although it has to be collected into a common, shareable infrastructure and format, to support proper interpretation of the primary data. Meta data are of three types, 1) pointwise (existing annotations to primary data), 2) operational modes and states (well defined and extractable from data), and 3) meta data model of the facility (a digital twin instance creatable from existing information). The project has elaborated a practical example of available data and meta data for a vacuum gauge and proposed methods to create any missing meta data, thereby demonstrating the feasibility of the approach.

Sharing data is an alien phenomenon to many organizations' legal entities, and the field is generally less mature than, e.g., open source licenses. Thus, there are no standard and generally adopted license models. However, as ESS is operating to support open science, ESS has a policy for intellectual property rights and inventions. The policy explicitly addresses the scientific data but may easily be extended to also comprise the control system data. This is sufficient to continue in the specific case, although the general licensing issues remain a research and practice challenge.

In addition to producing these technical results, a reference group was set up to establish a network of stakeholders, to validate the work, and form future collaboration. Industry (ABB, Perstorp, Tetra Pak, Hitachi, and Sinch), research facility (DESY) and university (Lund university, innovation and collaboration office) representatives make up the reference group, which has gathered in advisory meetings and collaboration workshops and seminars. The reference group has confirmed the potential value of collaboration around ESS control system data, both in terms of future knowledge and solutions, and in the established network as such.

Next steps

Based on the prestudy, we conclude that ESS may function as a catalyst for Industry 4.0 digitalization, both in industry and other research facilities. We identify needs for advancing general technical, governance and legal support for data ecosystems. However, for the nearest steps, the more specific machine learning issues related to alarms have higher priority. They would also integrate the interests and potentials for the industry and research facility network of national and international partners, established in the pilot project.

Proposed continuation projects include: 1) Creating a digital twin instance by machine learning, 2) ML-based operation state identification, 3) ML-based fault detection, 4) ML-based root-cause analysis of alarm cascades, and 5) Model-based root-cause analysis of alarm cascades.

These topics are relevant for research facilities and Industry 4.0 digitalization efforts, which we propose being explored in collaboration projects to shape a future data ecosystem. Thus, the innovation potential of the large-scale research facilities can be exploited also indirectly, by acting as a catalyst for collaboration.

Reports from and related to the project

Per Andersson, Jan Eric Larsson, and Karin Rathsman. ESS control system data lab – work package 2 final report. Technical report, ESS-3216740, 2021.

Anton Engström. Towards agile data engineering for small scale team. Master thesis LU-CS-EX 2020-40, Lund University, Department of Computer Science, 2020.

Karin Rathsman. ESS Control System Data Licenses, Policy and Strategy for the future. Technical report, ESS-3192568, 2021.

Per Runeson, Johan Linåker, and Thomas Olsson. Open data ecosystems – an empirical investigation into an emerging industry collaboration concept. Journal of Systems and Software, Submitted, 2021

Per Runeson and Emma Söderberg. Tools and Ecosystems for Open Control Systems Data at ESS, Technical report, Lund University, Department of Computer Science, ISSN 1404-1200, Technical Report 106, 2021.