d-fine

Open source IoT data pipelines

The digital shop floor – doing, not debating

analytical. quantitative. tech.

July 2022

Inhalt

Open source IoT data pipelines, July 2022

01. Going digital on the shop floor

Digitalising shop floor operations promises improved understanding of processes, shortened process times, reduced wastage, and cost savings. Yet, deploying the required digital systems requires expertise in system administration, software development and data analytics.

In our projects we accompany manufacturing clients through their digital journeys and help to deploy reliable IoT solutions that create valuable improvements. To speed things up, avoid lengthy vendor selection processes and ensure a steep learning curve right from the start we bring along flexible, modular open-source tools: The d-fine IoT platform d-scover@ is in active use at manufacturing clients and in research collaborations. We are releasing it under a permissive open-source license, so that businesses, institutions, or individuals can boost their IIoT projects, taking advantage of the experience we have accumulated over the years.

In this paper we explain the reasoning behind d-scover@, share best practices for most valuable IIoT use cases and provide some insight into challenges that projects will meet and how to address them.

02. Challenges in Capturing Value from Data

Although nearly all manufacturing sectors are looking for data-driven ways to improve their processes most companies find it hard to translate the great promise of modern digital technology into added value.

Serving clients from many industries we identify three main reasons why data-driven business transformation projects struggle or fail. All of them are related to an initial lack of practical experience that must be overcome to make informed decisions:

2.1 Unclear Business Cases

Business cases are only vaguely defined initially as they are designed to break new ground. Unfortunately, in many cases a clear business case fails to materialise as the project moves on. This is the most common failure scenario we encounter in digitalisation projects.

We provide advice on how to set realistic goals and formulate criteria for success before too much sunk cost has accumulated in section three.

2.2 Long Lead Times

Clients expect IT projects to move at much higher speeds than most other changes. Unfortunately, IIoT projects frequently get bogged down and take a long time to deliver until external advice is found. Most often this is not caused by the complexity of the problem but by overplanning, overengineering, or vendor selection.

Exchange with experienced practitioners speeds up identifying opportunities and modular, extensible tooling with right-by-default configuration accelerates initial implementations. In section four we explain how d-scover@ supports building real-world IIoT applications in the shortest possible time.

2.3 Transfer to Productive Environment Fails

The most frustrating failure is a failure to translate a prototype to productive use e.g., when security practices become too onerous, prototype code is hard to maintain, or when the system lacks stability and reliability. This is aggravated by pilot-development practices: Online tutorials, and guidelines often result in insecure, insufficiently isolated or partially outdated systems and leave "the details" to practitioners.

To avoid this we use modern infrastructure technologies – chiefly containerisation and infrastructure-as-code – that are maintainable and upgradeable in a straightforward way and provide a built-in security concept. d-scover@ is a ready-to use IIoT systems that runs on a local machine, on-premise or in the cloud and enables you to collect **real-world experience** in hours.

03. Building a Business Case

3.1 How can the business case be validated?

For an IIoT project to generate value, efforts for data collection, integration, analysis and maintainenance must be balanced against business benefits. This is difficult when it is not clear which data sources are available and how they can be accessed. At the same time users often either struggle to grasp the extent of possibilities inherent in new technology or expect silver bullets.

Rapid, iterative prototyping helps establish realistic expectations and allows stakeholders to clearly formulate expected benefits. This should take no more than a couple of weeks and deliver a first user experience within days.

3.2 How can practical experience be gained quickly?

Machine providers frequently promote proprietary IIoT platforms for a particular machine or process, Hyperscalers and automation equipment manufacturers provide "fully integrated" IIoT solutions to include almost arbitrary data into a digital twin of the value chain. Both require substantial preparation and upfront investment.

To understand benefits, challenges and scale of a "digital factory twin" it is however advisable to gain practical experience and insights before committing on a preferred vendor. An open-source solution like d-scover@ provides exactly this.

3.3 What does it take to make a start?

The market for IIoT software is diverse in terms of underlying infrastructure (SCADA-like on-prem systems vs fully integrated cloud solutions), service abstractions (XaaS) and technologies. Standardization efforts frequently do not deliver upon their promises or "outscope" issues arising from legacy infrastructure.

Together with the very real challenges of shop floor data integration, this often leads to IIoT projects being framed as hinging mostly on the choice of a software vendor. Inevitably disappointment ensues when rosy promises do not become reality. While many of the software packages offered do indeed provide impressive features, they do very little to answer the three main questions of IIoT:

- Why is it reasonable to assume that data will help to improve KPIs?
- Which data should be collected, stored, and analyzed?
- How does data get from the machine to the service backend?

These questions and real-world experience should inform the choice of vendor. d-scover@ provides first results and helps to answer the above questions within days without need for CAPEX.

04. Shortening Development Cycles for IIoT

4.1 Agile Practices for Shopfloor IT

Modern software development focuses on rapid implementation and short feedback cycles. Projects deliver user value at the earliest possible time and react flexibly to feedback. To do so, many development frameworks focus on reducing the effort for a minimal viable product: With a clear focus on providing functionality to users, both failure rates and product lead times reduce substantially.

Appliying this “agile” approach to enterprise software is somewhat harder, because of complex boundary conditions (such as regulatory concerns and company guidelines) and business requirements. Building MVPs can be challenging where interfaces to different systems are required and whenever security is a concern. Despite this, industrial manufacturing processes can benefit greatly by applying agile software practices to the shopfloor context. Two obstacles for rapid progress exist:

- Any useful setup involves hardware and connectivity considerations and aspects of legacy support.
- There are no “pre-live” environments for manufacturing facilities to experiment with.

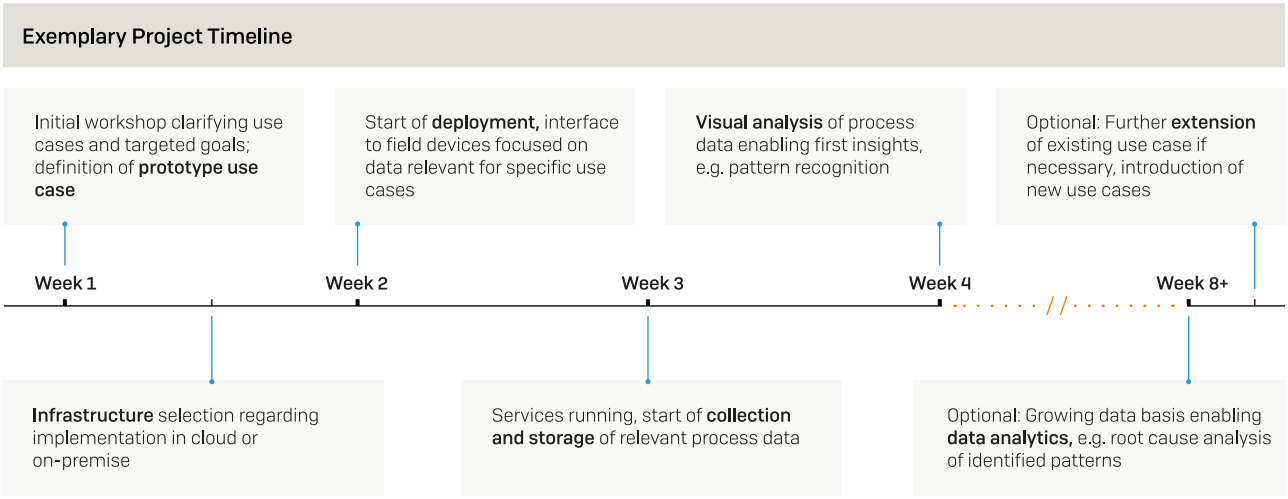
Consumer IoT devices have minimal computing power and are fully functional only when connected to cloud services with at most minimal interruption. Devices and functionalities are often tightly bound to a specific software and consumers give up control for improved usability.

For industrial applications, such an approach is unacceptable because it poses new, important, and unnecessary risks; It is also pointless because industrial equipment typically is installed and maintained in an organized fashion. Productive and sustainable software development has to follow a “correct by design” approach in which the standard configuration is production-ready, so development can focus on business needs.

4.2 Doing, Not Debating

Attempts at digitalisation can be bogged down by unclear objectives and overplanning. There is a perceived need to make strategic decisions, settle on a long-term plan and fully evaluate options and vendors. Because of that, required budgets and expected timelines quickly grow. The result is either inaction, or a big expenditure that is not informed by practical experience.

d-fine advocates for an exploratory project model, which delivers concrete results and generates expertise within weeks. This is the context in which an open-source platform with "ready to go" configuration is particularly important: a project can rely on this software without budget needs or vendor selection, flexibly integrate into an existing landscape, and begin to collect data without lock-in. All data stays under the full control of the data owner.



05. Being Ready for Production

5.1 How can systems be simple and flexible without being insecure?

A data integration platform provides a large attack surface to malicious actors. Deployment and test of updates require substantially larger effort than a desktop or back-office system does. Vulnerabilities are prevalent on all levels of connected systems ranging from insufficient security designs of low budget sensors to sloppy user behavior manifesting itself in default or trivial to guess passwords (See e.g. [1]).

Given the systemic degree of complexity, achieving short turnaround times without compromising security requires pre-configured structures and expert advice.

[1] https://www.trendmicro.com/vinfo/us/security/news/internet-of-things/smart-yet-flawed-iot-device-vulnerabilities-explained

5.2 Matching Software to Equipment Capabilities

Modern shop floor equipment often comes IIoT ready. Where time-honored "legacy" equipment is in use, additional sensors and other hardware are required to facilitate data acquistion.

Edge infrastructure provides networking capabilities, helps to isolate the machine from the shopfloor network and to "abstract" it by exposing standardized interfaces even where the actual equipment is highly divers. Its functions range from simple A/D conversion to edge analytics with specialized AI hardware. d-scover@ contains components which provide many of these capabilities out of the box for point-and-click configuration and are seamlessly extensible through Python scripts, Groovy scripts, or Java code.

5.3 Supporting Services and Deployment Strategies

Productive deployment involves integrating supporting services (most importantly identity management) and rolling out configurations reproducibly in different environments. Solutions need to run locally, on on-premise servers or in the cloud, depending on strategy and circumstances. Without a clear framework in place, this requires great discipline as well as substantial know-how in systems administration from developers.

These hidden requirements make up a good part of the complexity of going productive. Having them "built-in" from the start saves cost and allows to deploy and to collect feedback frequently.

d-scover@ is fully containerized with centralized, declarative configuration and can be deployed in a automatic fashion on nearly every Linux-capable device. We provide a Terraform configuration for deployment to the cloud provider of your choice. All installations provide https, network configuration and user rights management, with the option of integrating Active Directory.

06. A First IIoT Project

6.1 Getting Started Quickly

Every industrial corporate has unique machines and sensors and the most relevant data may vary even when using the same machine! Competing standards and protocols and the need to integrate external systems complicate this situation further.

Yet IoT projects must gain quick operational wins while being cost-effective, maintain a steep learning curve and adapt targets continuously. An "agile" IoT-project delivers small but frequent incremental value, demonstrates benefits and potential problems precisely and provides feedback into new increments.

d-scover@ allows experimentation in real-world settings with minimal effort because it is pre-packaged to be safe and secure, and provides ready-to-use connectors to sensors, controllers, APIs, and databases.

6.2 Tackling Real Problems Early

Every data integration project needs to solve tasks which can be structured as so called “increments”: measurable progress in the form of a releasable piece of software with business value. This is a challenge especially in early stages because some of the necessary infrastructure does not deliver immediate business benefit.

Support Infrastructure for IIoT-Applications

- Data integration / ETL pipelines to collect data from different sources
- Storage backend
- Data analytics and visualization
- Log integration
- System performance measurement
- Consistent user and privilege management
- Encryption

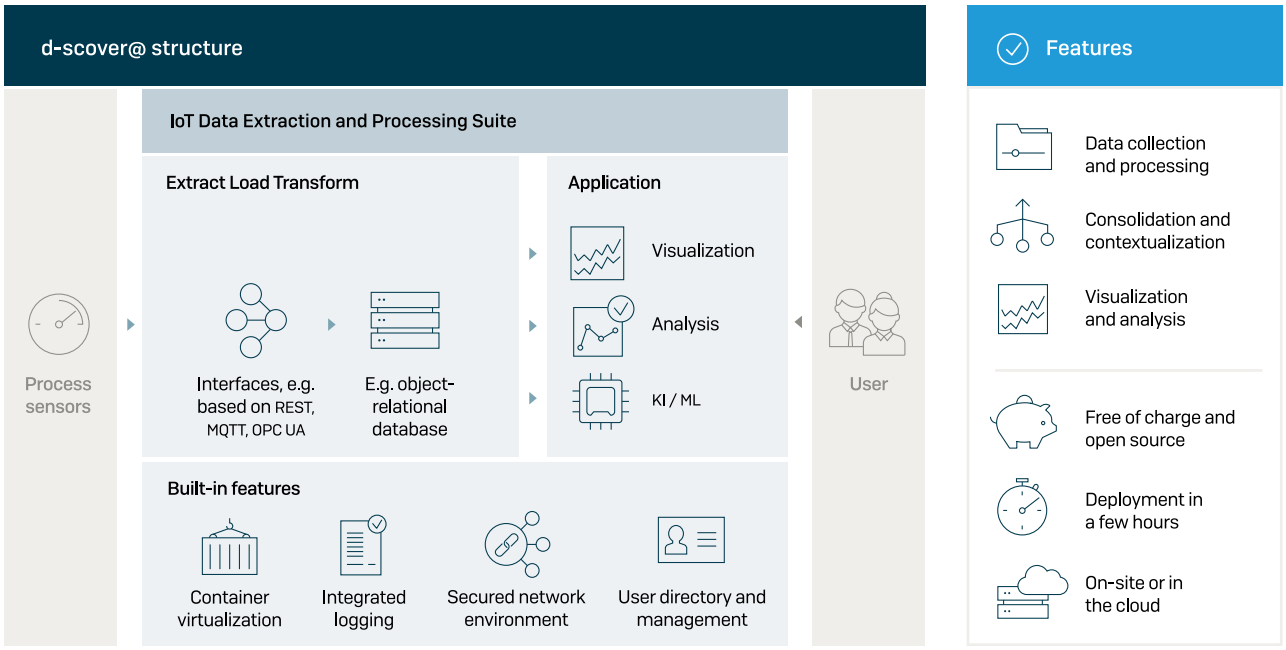
For each of these functional blocks a plethora of solutions exists. Their integration is either a (sub-)project itself or is provided by monolithic vendor-solutions. Selecting an IoT platform is a costly process with little tangible benefit and may easily result in a vendor lock-in. For companies in early “try & fail” stages this is particularly undesirable.

When developing the d-scover@ IoT platform we instead aimed at avoiding lengthy implementation phases and at enabling clients to gain practical IoT experience without substantial upfront investment.

6.3 A Running Start with d-scover@

d-scover@ facilitates reliable and secure DI-pipelines with minimal investment. It provides “snap-fit-components” for data bases, ETL, reverse proxy, web portal, user management, logging, analytics & network setup. The technology-stack is completely open-source based with container-based deployment, ready-to-use network configuration, CI-integration, and simple extensibility. It can be operated with standalone user management or integrated with Active Directory. An initial, preconfigured installation can be used to rapidly gather feedback and learn essential lessons.

It also includes modelling and data analysis capbilities via a secure execution environment for Python notebooks. With comprehensive documentation and correct-by-default setup, projects can move fast without breaking things and can focus on creating new insights from data. In both greenfield and brownfield approaches results can be presented to stakeholders after no more than a few weeks.



Scalability and adaptability enable a fast and simple implementation of prototypical and productive use cases.

6.4 Examples of Real World Usage

6.4.1 Manufacturing

A metal processing plant evaluates the influence of environment- and process parameters on product quality using d-scover@. d-fine supports the client with deploying the system on-site, connecting machine controllers and monitoring the acquisition of to process data.

In this way a full-featured analysis platform is available to engineers and plant managers, without running a full-scale project. On the basis of the collected data we analyse the impact of production settings and conditions upon the manufacturing process and product quality and provide input on the concrete benefits of a potential full-scale IIoT rollout.

6.4.2 Research

Within the BMWK research project “HyConnect”, d-fine and its industrial and academic partners research the potential of combining classic sheet metal forming with additive manufacturing to reduce resource consumption and ecological footprint.

d-scover@ is responsible for enabling the exchange of data across participating companies automatically and securely. To this end a blockchain-backend module has been developed and integrated. Mathematical and statistical analysis of collected data is automated through the extension points provided by d-scover@.

6.4.3 Education / Adademia

d-scover@ is in use at a number of academic institutions for research projects, digitalisation of lab environments and for teaching assignments. It is particularly valued for greatly simplifying end-to-end data collection and analysis pipelines. The effort saved thereby frees up scarce resources and allows researchers to focus upon their actual scientific interests.

Autoren

Christoph Charlet
Manager, d-fine GmbH, Frankfurt
christoph.charlet@d-fine.de

Dr. Tassilo Christ
Senior Manager, d-fine GmbH, München
tassilo.christ@d-fine.de

d-fine

Berlin
d-fine GmbH
Friedrichstraße 68
10117 Berlin
Germany
berlin@d-fine.de

Dusseldorf
d-fine GmbH
Dreischeibenhaus 1
40211 Dusseldorf
Germany
duesseldorf@d-fine.de

Frankfurt
d-fine GmbH
An der Hauptwache 7
60313 Frankfurt
Germany
frankfurt@d-fine.de

Hamburg
d-fine GmbH
Rödingsmarkt 9
20459 Hamburg
Germany
hamburg@d-fine.de

Munich
d-fine GmbH
Bavariafilmplatz 8
82031 Grünwald
Germany
muenchen@d-fine.de

London
d-fine Ltd
6-7 Queen Street
London, EC4N 1SP
United Kingdom
london@d-fine.co.uk

Vienna
d-fine Austria GmbH
Seilerstätte 13/25-26
1010 Vienna
Austria
wien@d-fine.at

Zurich
d-fine AG
Brandschenkestrasse 150
8002 Zurich
Switzerland
zuerich@d-fine.ch