



What ASSISTANT brings you: Interview with Dr. P-O Östberg on the ASSISTANT project and Data Fabric

Dr. P-O Östberg is the founder and CTO of BiTi Innovations, a spin-off from Umeå University, Sweden, focused on commercialisation of research outcomes from systems research in autonomous distributed systems. In addition to his role at BiTi he is also an associated professor at Umeå University and has held researcher positions at several universities including Uppsala University, Sweden, Karolinska Institutet, Sweden, Ulm University, Germany, and the Lawrence Berkeley National Laboratory at the University of California, Berkeley, USA. His research interests are mainly oriented around the use of machine learning, simulation, and optimization techniques to construct AI systems for planning, and in ASSISTANT he and BiTi are leading the technical work on development of a data fabric for adaptive manufacturing.

Félicien: I really like to shorten your name if you do not mind. P-O, what is a data fabric?

P-O: A data fabric is a system that provides a unified architecture for management and provisioning of data. Data fabrics are typically realized as sets of services providing consistent access and interfaces for data, and parts of the systems are often deployed across several types of resources and networks to combine the use of on-premise and cloud resources.

Félicien: Sound excellent. So, what will the data fabric be used for ASSISTANT or how is it relevant to have a data fabric for complex projects like ASSISTANT?

P-O: The ASSISTANT data fabric will be used to coordinate and provide for the data needs of the ASSISTANT digital twins, combined simulation-AI systems that optimize the operations of various types of manufacturing systems. Manufacturing data are typically complex and diverse, and data fabrics are essential to complex research projects as ASSISTANT as they allow systems and people to focus on the research / tasks at hand and consider data as utilities - particularly so for the ASSISTANT digital twins that combine real world data with advanced simulations to make predictions and evaluate planning strategies for complex manufacturing operations. In ASSISTANT, the data fabric will also abstract the use of different types of (cloud, edge, and on-premise) computing resources and will house mechanisms for intelligent planning and control of the fabric itself.

Félicien: Interesting, we will cover the topic of the digital twin in the next interview. What are the technologies and methodologies that will be used for the implementation of this data fabric?

P-O: The ASSISTANT data fabric is designed as a layered architecture of data services where individual services are implemented using well established technologies such as Java, Python, REST and Google protocol buffers, and all services are abstracted behind well documented APIs for ease of integration. The architecture defines a

set of core services with exchangeable reference implementations and has multiple customization points in the form of plug-ins for adaptations and integrations. The data fabric also supports the use of domain models and has built-in metadata structures enabling, e.g., reactive data curation, advanced search, and data federation; and supports streaming data analysis through integration with the Intrasoft streaming platform. The platform builds on open, freely available technologies and is designed for flexibility to promote adoption in manufacturing industry.

Félicien: Amazing, tell me P-O, how should a data fabric be maintained?

P-O: The value of a data fabric lies in its ability to harness and refine value from data for its host organization. For this reason, it is important to realize that data fabrics are not products you install once as a turn-key solution but rather platforms you develop over time and continually adapt to new needs and opportunities in your business. In short, data fabrics should be designed, operated, and maintained with a data-centric view to fully realize their potential.

Félicien: Well, but hard to discuss the structure of the data fabric without talking about security and maybe also safety. So, what are the safety and security aspects related to the implementation and exploitation of a data fabric?

P-O: In contrast to other ASSISTANT systems, where cyberphysical systems such as production cells with both human operators and assembly robots are not uncommon, the data fabric is a digital system with few (physical) safety issues. In security however, the data fabric finds several challenges, e.g., compliance with GDPR regulations for the localization, migration, and application of data; as well as privacy and integrity aspects for data storage and transmission. An interesting feature of the ASSISTANT data fabric research is that we are developing algorithms that based on operator guidance rules can

control and restrict the deployment of services and movement of data to provide a framework for controlling the behavior of the fabric subsystems. A similar system is also being developed to provide a basis for ethical reasoning on the same, to allow for, e.g., selection of green power resources for intensive computations and storage of large data sets.

Félicien: Fantastic, but what is the complexity in terms of data access, storage, and manipulation of an operational data fabric?

P-O: Data fabrics are by their nature large, complex software stacks that typically evolve over time to meet the needs of the host organization. The ASSISTANT data fabric is specialized for adaptive manufacturing and provides a flexible architecture coupled with well documented easy to use tools for management and provisioning of data. Use of the ASSISTANT data fabric systems is not trivial and in certain situations requires specific manufacturing domain knowledge, e.g., for effective use of domain models and metadata-based graph searches, but we are making every effort to make it less complex than many other data fabric system. Storage, access, and manipulation of data should not be more complex than using a distributed file system, and we are placing focus on automating both the internal operations as well as advanced features such as semantic interpretation and analysis of data as much as possible.

Félicien: Last question, what do you think of the ASSISTANT project?

P-O: The ASSISTANT project is a very interesting research project that combines the extensive manufacturing knowledge of large organizations such as Siemens, Atlas Copco, and the PSA group with advanced research skills of several high profile European universities. I personally like the opportunity to learn more about how planning and scheduling is done in manufacturing environments (my personal experience has so far mainly been based on supercomputing and cloud environment research), and I think access to the types of competence

and data we have in the project are key enablers for good research and innovation. It is also a very nice project to work in, full of skilled and positive people. The project is still new, but I am very optimistic about the outcomes of the project and would recommend anyone interested to check out results as they are published.

Félicien: Thank you P-O for your time.

P-O: You are welcome and thank you for having me.