

Breakout Session E-2: Information

Lead: Jelmer Pier Van der Gaast; **Scribe:** Moritz Roidl

Participants: Leon McGinnis, Debjit Roy, Sadan Kulturel-Konak, Anike Murrenhoff, David Porter, David Šourek, Sunderesh Heragu, Layek Abdel-Malek, Fabiana Tornese

This session addressed the key issues of data usage in material handling. The balance between data quantity and quality, the impact of data sharing, and the differences between data and information was debated. The session also discussed semantics, metamodels, and standardization, emphasizing the need for consistent data handling for future technological advances in material handling. Finally, the importance of information sharing in the supply chain was addressed, including digital twins and cybersecurity.

The session began by challenging the prevailing assumption of data volume supremacy and raised the question of the proper use of available data as a more important issue. Data sharing was identified as a key challenge, particularly the frequent tendency of researchers to publish results without releasing the associated data sets. Leon McGinnis' anecdote about a company that is able to collect terabytes of daily data but has problems with efficient data use and storage, highlighted this issue. The group acknowledged the problems of data availability and searchability in academic institutions, where data dashboards often scatter information. Finally, the discussion highlighted the important distinction between data and information and the role of semantic models in transforming raw data into meaningful, useful information.

The discussion shifted to semantics, metamodels, and standardization in data management. There was agreement on the potential benefits of a common semantic metamodel to facilitate tools that can translate between data sets and applications. Despite current challenges such as different terminologies from different domains, initiatives such as the *International Data Space Organization* (<https://internationaldataspaces.org>) are working on such solutions. The group acknowledged the need for a semantic model, such as those used in the CAD industry or electronic circuit design, and discussed the possibility of using open-source software development models to create de facto standards. The importance of a unified data model for the effective functioning of digital twins was emphasized, highlighting the need for structured and standardized data.

In the context of supply chain information sharing, the importance of plug-and-play models was highlighted, and the discussion focused on data warehousing. The need for a unified data model for digital twins and the associated cybersecurity challenges were discussed. The lack of a common (simulation) object library was seen as a missed opportunity to streamline processes. The idea of an open source software development model to create de facto standards was again proposed. Finally, David Porter emphasized the importance of sharing the right information to foster collaboration, citing the success of *General Transit Feed Specification* (GTFS) streams for public transit schedules as a model worth emulating.