From Raw Sensor Data to Trustworthy Digital Twins

Vaishnavi Mall, Frank Leinenbach, Georg Frey, Antonio Krüger, Hans-Georg Herrmann
(Chair of Automation and Energy Systems (AES), Saarland University, DFKI, Fraunhofer IZFP)

Industrial sensor data often leaves the lab as bare numbers in different proprietary vendor formats. This work presents a nine-stage pipeline transforming raw sensor measurement data from three sensor modalities (Eddy Current, Ultrasound, and Thermography) into standard Asset Administration Shells (AAS). Each shell carries the measurement's meaning, quality information, and usage terms, with FAIR-related validation and independent tools used to verify the results. Building on this work within EnFoSaar, these trusted measurements can be used to build digital twins of multi-energy systems, where electricity, heat, and hydrogen data must be reliable before AI can plan and control them.

A number without a story

An ultrasonic scan of an iron plate fills a 62 MB file with 15 million values. What reaches the next lab is often just numbers and files:

.nde
62 MB
3D volume

→ 58.37

What a machine cannot tell, and the standard that can answer it:

- **Meaning** — What is it? How was it measured? — AAS + RDF
- **Correctness** — Is it complete and unchanged? — SHACL
- **Permission** — Who may use it, and how? — ODRL
- **Findability** — Can anyone find it next year? — FAIR

Three instruments, three file formats, no shared vocabulary. The standards exist; this pipeline joins all of them in one run.

Now the number has a story

58.37 The same peak amplitude, now with everything a machine needs to use it.

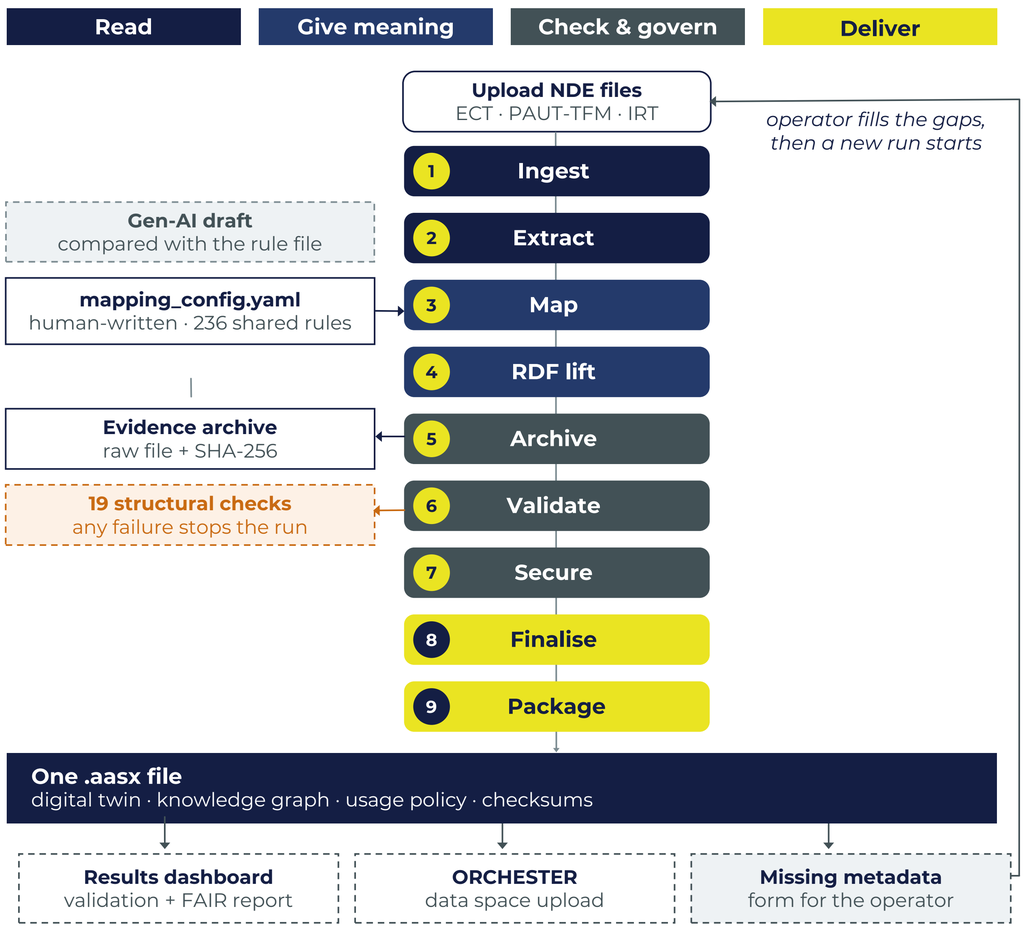
Specimen - AAS	M36_SJF_B_71
• **Nameplate**	Evident MXU · Serial QC-0078656 · SW 5.16.0
• **TechnicalData**	5L16-A1 · 5 MHz · L-L path
• **MeasurementValue**	40 dB gain · gate 1 at 10% threshold
• **InspectionResult**	Peak Depth 2.2491 mm · **Peak Amp 58.37 dB**
• **TimeSeries**	381 rows · SHA-256 verified

What a machine can tell now, and the standard that answers it:

- **Meaning** — AAS + RDF
  What is it? How was it measured?
  Peak ultrasonic echo at **58.37** dB (2.2491 mm depth) captured via 5L16-A1 5 MHz probe at 40 dB hardware gain.
- **Correctness** — SHACL
  Is it complete and unchanged?
  0 standard violations across 2 SHACL engines; payload locked with SHA-256 verification
- **Permission** — ODRL
  Who may use it, and how?
  Governed by ODRL Policy: Internal read-only access with explicit restrictions on third-party redistribution.
- **Findability** — FAIR
  Can anyone find it next year?
  Passes 13–14 of 17 FAIR indicators; queryable via persistent ID & SPARQL graph endpoints.

One pipeline, nine stages

The core contribution: one run from raw file to validated AAS package.



Next: from scan to system

How can digital twins of multi-energy systems stay trustworthy when their data comes from many sensors, vendors and partners?

One AAS per measurement → **Digital twins of energy systems, built on trusted measurements**

What carries over

AAS standard · SHACL checks · FAIR data · Usage terms

Possible directions, open for discussion

1. Automatic data connection from energy sensors to digital twins
2. Quality checks on data before AI forecasts and anomaly detection
3. Plain-language questions to twins, with answers traced to the data
4. Digital calibration certificates (DCC) inside the AAS

Working with energy data? Let's talk!

Gefördert aus dem Transformationsprogramm Forschung und Wissenstransfer Saar durch:

Ministerium der Finanzen und für Wissenschaft SAARLAND