

Quality Checker DPUs

Quality Checker is a new type of DPU introduced recently, to raise the significance of quality assurance during data processing.

When you develop a data processing pipeline with strict data governance policies, the output data is required to conform to certain rules and constraints. These rules do not only enforce at data model and data schema level, but also inspect data at content level. Therefore quality checkers are introduced to inspect the data and react depending on constraint satisfactions. They directly forward the data from input channel to output channel after any inspection operation without modifying the original data. They then stop the pipeline execution or produce an exception report as additional output when constraints are not satisfied.

At this time, one RDF data checker exists, driven by SPARQL ask queries.