

Product Report: Triple Boost for Data Consistency and Increased User-Friendliness (Part 2 of 3)

Data input validation with Deep Qualicision AI

In the last issue of PRODUCTION manager, we discussed how auto-completion based on the Deep Qualicision AI Framework can achieve measurable improvements in data consistency and user friendliness during data collection using input forms. However, auto-completion does not yet fully guarantee correctness of an entire data set. This is because the entered data would be transferred to a database for further processing without additional verification. As a result, cells may be swapped or incorrect units interchanged. This is where data input validation based on the Deep Qualicision AI Framework is used to provide both syntactic and semantic validation of the data in a fully automated manner.

Data input forms, some with many available fields, are well known as an integral part of a wide range of business processes. Some fields are mandatory, whereas others are optional. In addition, syntax and semantics

primarily depend on the respective context. Constantly keeping track of this during day-to-day work is a big challenge for those processing such forms. In the course of time, errors consequently can occur in entered data, which may have ex-

tensive consequences for subsequent processes.

A practical example: Entering a sales contract

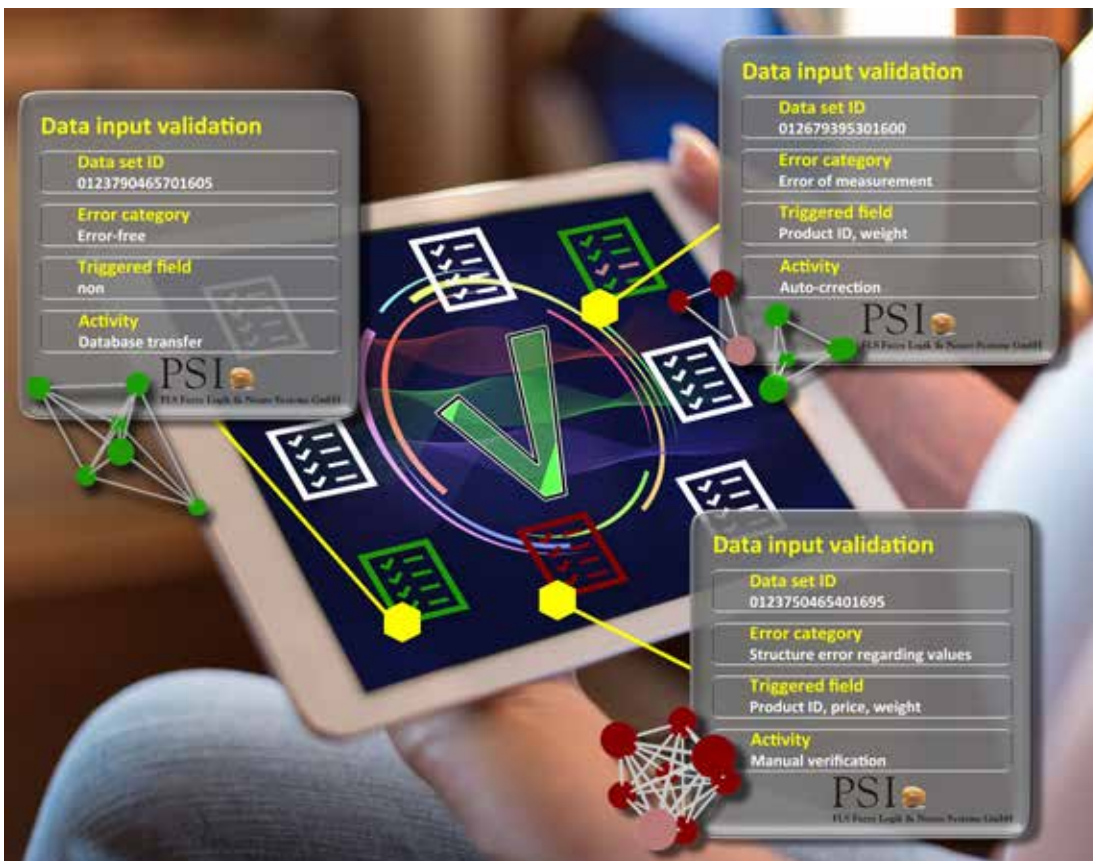
When entering a sales contract, the unit tons is selected for the agreed quantity instead of kilograms and the decimal point is moved one digit forward in the price. If the entered contract is later executed without previous data validation, far too large a quantity (1000x) is sold at much too low a price (0.1x).

As a result, most companies use control offices to check the correctness of entire forms before being transferred to a database. But this process takes a long time and is still error-prone

for several reasons. It is therefore necessary to achieve the highest possible degree of automation in order to guarantee the accuracy of the data and thus minimize the risk of inconsistencies.

Validation of input data based on rules

The manual inspection of all entries by a control office is not economically feasible in most companies. The data sets of form entries to be checked must therefore be initially pre-filtered for auditors so that they



Data input validation with Deep Qualicision AI.

can focus on errors with potentially significant consequences. Rule-based systems are often used in this regard. They search for fixed anomalies in entry data sets, mostly by performing threshold checks. For example, if an entered weight exceeds the value of one ton, the control office will conduct a check of the input data.

This ensures that large orders are always checked for correctness. However, if there is no additional checking rule for the price, the extent of the damage can still be large if, for example, several hundred underpriced data sets are transferred to the database with a weight just under one ton, bypassing the control office. In addition, such strict control systems remain rigid in processes that change over time. The codebase must be continuously adapted by appropriately qualified developers in order to continue supplying the control office with relevant data sets. Rather, a mechanism is needed that automatically detects anomalies in the structures of the full data sets and continuously adapts to the current situation.

Data-driven input validation by means of Qualitative Labeling combined with machine learning

For most business processes with form-based data collection, there is already a broad basis of historicized data. By using Qualitative Labeling and machine learning from the Deep Qualicision AI Framework, process-specific structures in input data sets can be learned from past data. This can be done both universally and per user, in order to guarantee ideal adaptability to any process. Data-based approaches offer a wide range of benefits, especially for identifying multi-stage rela-

tions in data, such as a correct ratio between the entered weight and the reported price.

KPI-based self-learning input validation as part of a Deep Qualicision AI complete system

Qualitative Labeling along with a knowledge base of historicized data trained by machine learning form the basis for input validation adapted from the Deep Qualicision AI Framework. However, the connection to a control system should not be neglected, as these are still suitable for fixed dependencies between attributes. Deep Qualicision AI Framework also enables decision support by simply preferring different evaluation KPIs. In this way, significant deviations from forecasts can be evaluated in a comprehensible manner.

In addition, the control mechanism is kept constantly up to date as the knowledge base is continually adapted through fully automated learning of new data. A KPI-based, self-learning AI system such as this offers the ability to provide an automated system for continually monitoring the data acquisition process based on historicized data and a constantly growing knowledge base. This ensures that only those data sets that contain certain anomalies are filtered for manual reviews without determining on fixed properties such as a minimum weight.


Deep Qualicision-based data input validation as an extension of auto-completion

An auto-completion system that is already in operation can be fully extended with data input validation functionality using the common Deep Qualicision AI Framework. In this

- + Detection of input errors as anomalies in data collection
- + Automated validation of all data sets entered
- + Significant time savings in subsequent data processing
- + Consistency across the entire database
- + Qualitative standardization and plausibility analyses
- + Continuous learning of the knowledge base to sustain status up to date

way, a further measurable increase in user friendliness and data consistency can be achieved.

Automated duplicate detection as the final component of the AI complete system

If the syntax and semantics of the data sets have been learned from historicized data as well as during data input and validation, this can be used immediately to search for duplicates in existing databases. In the final step, a knowledge base expanded in this way serves to complete the AI system by supplementing it with automated duplicate recognition. This provides a further boost in data consistency and user friendliness based on the Deep Qualicision AI Framework (Part 3 continues in the next edition of PRODUCTION manager). 

PSI FLS

Fuzzy Logik & Neuro Systeme GmbH

Dr. Jonas Ostmeyer

ostmeyer@psi.de

www.deepqualicision.ai