Quality control in SimaPro Database implementation

Parse

Check data format valid

Check naming and processes and substances for duplicates

Update

Store data in software

Pre-process data for better performance

Regionalize substances

Cross-check

Calculation results

Uncertainty values

Comments

Hyperlinked data

Synchronize

Consistency with existing data

Available across all software we provide

Correspondence for new/removed processes

Flag/Remove obsolete processes

Release

Inform users

Release documentation of changes

Conveying quality-related information to users

- Data Quality Indicators exist for some time, but seem not being used by most users
  - Users tend to pick based on names

As software developers,
should we provide info on quality of data, or
help find what is fit for purpose?

Conveying quality-related information to users

Fit for purpose

"Is the data I have close enough to reality, so it doesn't distort my model?"

- For ISO compliance - One of 1st tasks in data collection step is data quality requirements

    → Software could guide users into setting the needs and propose relevant data

Conveying quality-related information to users

Example of a guided process

1. User starts new LCA project

2. Guided definition of data requirements

3. Search for data with feedback on alignment to requirements

4. Data quality analysis step with help from the software

Conveying quality-related information to users

Guided definition the data quality requirements

- Example:
  - Type: Energy
  - Age: Less than 2 years old
  - Technology: Market average
  - ...

- Add a set of requirement per data type (e.g. energy, infrastructure...)

Conveying quality-related information to users

Guided search for data

- Each dataset found is rated against requirements (age, technology, geography…)

- Weighting can be calculated and a simple score or colour code shown to check overall matching
  - User defines under which score data are not worth to be shown

- Help find proxies
  - e.g. if users look for a chemical for which no data is available, propose one with similar production process
  - Links to online repositories like GLAD, or libraries like the Community-Driven Life Cycle Inventory and NorEnviro

Quality assessment tools in LCA software

- Also in ISO, assessment of the data quality of complete model is mandatory, but again often overlooked
  - → Too complex for average practitioner?

- Software tools can help, by analysing big data fast
  - Each datapoint can be listed with colour codes to give an overview
  - Overall model can be rated against criteria set at start

  - Even possibly link assessment with hotspots detection to guide user where effort is actually needed

Questions?

Jori Coustillas
SimaPro Product Owner
coustillas@pre-sustainability.com

Artur Donaldson
LCA Data Scientist
donaldson@pre-sustainability.com

PRé

pre-sustainability.com   PReSustainablty   PRé Sustainability