

SODA 

Data Quality Manifesto

01

All data teams should be equipped and empowered to manage the quality of their data.

It's our role as the data management team to make sure that all users are equipped to do the right thing when it comes to data. That's what creates trust.

02

Good data starts with good engineering.

Data should be at the center of the architecture of your systems and processes and should never be treated as “exhaust”.

03

Data quality is ensured by consistent process execution.

No single business is 100% automated. Every hand-off between people in a business process can result in bad data, so process owners play a key role in ensuring data quality.

04

Data quality should be measured by consumer expectations.

Data quality is all about what people need and want to do with the data. You need an effective way to crowdsource and validate current and future data consumer expectations.

05

Structural data quality fixes should be the norm.

Most data issues can be traced to a structural root cause. Rather than fixing all the downstream problems, you should find and fix the source.

06

The business impact should drive which data to fix first.

When looking at data problems, you should take into account all downstream usages of the data, and the criticality of those to the business processes.

07

Data validation and correction should happen at entry.

We all know those “lazy” business people that don’t fill-in the data, or we’ve come across faulty firmware and resulting sensor data. Sometimes validation at correction can’t be done at source, and that’s the reality that we have to live with...

08

Data quality management should be transparent to everyone.

Transparency is foundational to trust. If data teams can see who's managing the data they use, and what's been done to ensure its reliability, then they'll be more likely to use that data.

Questions? Feedback?

hello@sodadata.io

SODA 

Special Thanks To:

Thijs Nagel
Frank Gresnight

SODA 