

DATA QUALITY SELF-EVALUATION WORKSHEET

This self-evaluation will walk you through some of the most important questions to consider about your data quality project and your organization. With your answers in hand, you'll be better positioned to articulate your needs to vendors and determine who can do the best job of meeting your requirements.

Tip: Pair this self-evaluation with our Data Quality Vendor Checklist when it's time to create a vendor evaluation shortlist.

PROJECT BACKING

How invested is this project's leadership in the need for data quality?

1	2	3	4	5
Not at all				Extremely

How willing is this project's leadership to embrace new technology and technology vendors?

1	2	3	4	5
Not at all				Extremely

How willing is your organization to embrace new technology and technology vendors?

1	2	3	4	5
Not at all				Extremely

Tip: If your organization is conservative regarding new technology, is skeptical of introducing data quality, or both, you might need to stick with big-name vendors in your first search—because it doesn't matter how much you like a more innovative vendor if you can't get the purchase approved. Once data quality has proven itself (or if the big-name software can't deliver) you can try for more room to look further afield.

If your organization is skeptical of data quality for the opposite reason—because previous solutions from big-name vendors have failed to deliver on their promises—focus on vendors with more innovative and flexible offerings. To build your case for the purchase, give vendors details about your needs and ask them to demonstrate specific examples of how they can produce value for you.

DATA QUALITY EXPERIENCE

Do you think that your organization as a whole has a strong understanding of what data quality is and how to achieve it?

1	2	3	4	5
Not at all				Extremely

Do you feel that your team, specifically has a strong understanding of how data quality will achieve the goals of this project?

1	2	3	4	5
Not at all				Extremely

DATA QUALITY SELF-EVALUATION WORKSHEET

Tip: Customer data is the most popular, so every vendor will say they're ready for it. Your business is unique though, so don't take it for granted that the vendor's concept of "customer data" matches up exactly with yours. Get a custom demo or proof of concept that uses your data before you make any long-term commitments—and that goes double if you're using any non-customer data.

What kinds of data sources are you planning to involve in this project? Select all that apply.

- | | |
|---|---------------------------------------|
| <input type="checkbox"/> SQL | <input type="checkbox"/> Streaming |
| <input type="checkbox"/> RDBMS | <input type="checkbox"/> Custom |
| <input type="checkbox"/> Flat file (Excel, Word, PDF, etc). | <input type="checkbox"/> Other |
| <input type="checkbox"/> Hadoop, Spark, etc. | <input type="checkbox"/> I don't know |
| <input type="checkbox"/> Cloud (AWS, Azure, etc.) | |

Tip: If you don't know or if you have custom systems, focus on vendors who are eager to help you figure out your sources and ready to adapt to whatever they find.

About how many data sources (separate systems and files) do you plan to involve in this project?

- ☐ 10s
- ☐ 100s
- ☐ 1,000s
- ☐ 10,000+
- ☐ I don't know

Tip: Also consider how varied your sources are. Connecting to 4,000 Excel files, formatted identically, takes overall less work than connecting to 400 Excel files that each arrange the data differently... and some vendors will be better at dealing with those 400 files than others. Vendors will consider the number and variety of your sources when scoping your project, so getting as much information as possible—and being up-front about your knowledge gaps—will get you the most realistic proposals from your shortlist.

PROJECT USERS

In your current vision of the project, how often would the following groups of users from your organization regularly interact with the solution?

Business users, such as analysts

- ☐ Daily ☐ 1-3 times/week ☐ 1-2 times/month ☐ 1-5 times/year ☐ Never

Technical users, such as developers

- ☐ Daily ☐ 1-3 times/week ☐ 1-2 times/month ☐ 1-5 times/year ☐ Never

IT users, such as operations/backup support

- ☐ Daily ☐ 1-3 times/week ☐ 1-2 times/month ☐ 1-5 times/year ☐ Never

Tip: A data quality solution will need all these users to be involved in some way, so knowing who will be using the solution internally gives you an idea of what support (if any) you'll need from a vendor.

DATA QUALITY SELF-EVALUATION WORKSHEET

PROJECT GOALS

What do you want this project to produce? Select all that apply.

- ☐ Quality statistics (about data)
- ☐ Quality statistics (about a business process)
- ☐ Worklists of records for manual correction
- ☐ Automatic record corrections
- ☐ Pushed corrections
- ☐ Golden records
- ☐ Other _____

Have you been asked to deliver specific information, statistics, analytics, or data sets as a result of this project? If so, what are they?

How will this project affect or interact with your company's high-value assets and/or competitive differentiators?

Tip: Think broadly when you answer this question—if your project didn't affect either of these things, it wouldn't be worth the company's time to pursue it. If you improve analytics, does that drive better product innovation? If you repair customer data, can you start (or continue!) delivering the industry's best customer experience?

What questions are you hoping to answer with this project?

Tip: Again, think broadly, and include questions that data quality answers indirectly, through its support for other projects.