



# 19 QUESTIONS TO ASK DATA QUALITY VENDORS

Over the years, we've filled out a lot of RFIs, RFPs, and RFQs.

And we couldn't help noticing that sometimes, while the questions would give the prospective buyer a lot of data about each vendor, it wasn't clear how that data was going to help the buyer make a decision.

Because many times the buyer would ask us questions like: Do you support address standardization? What algorithms can be used in your rules? Do you connect to Hadoop?

But the buyer wouldn't ask us questions like: How are you going to make me successful? What is my experience with you going to be like? How much of the heavy lifting does your solution handle and how much is left to me?

## FEATURES: NOT THE WHOLE STORY

A lot of data quality buyers are buying data quality for the first time, so it's understandable that they default to the feature questions—if you're new to the market, those can feel like complex, in-depth information. It's easy to compare vendor answers about features, which can lend a feeling of confidence to the decision.

But vendors are more familiar with the market than a first-time (or even a second-time, or third-time) data quality buyer is. To vendors, those feature questions are softballs: every vendor works hard to make sure they can check the same "boxes" as their competitors. So while it's important to get the answers to those feature questions, they alone don't tell you as much as you might hope.

## IT'S ALL ABOUT YOU: GET THE DETAILS

To get a real sense of what a vendor is offering, you need to force the vendor to explain themselves to you. Get them to go into detail about exactly what their solution is and what that means for you.

And remember to ask about the future. Your organization isn't going to get less data-driven than it is now; information about how your data quality will be able to grow is an important part of judging its long-term value to your company.

If a vendor answers these questions with generic statements instead of clearly articulating how its solution is relevant to your specific situation,\* tread carefully. Your data isn't generic, so a generic solution won't deliver maximum data quality.

*\* This assumes you've actually given the vendor specifics, which we strongly recommend—if you make vendors guess, their proposals just aren't going to be very valuable (or even comparable, necessarily). See our Vendor DQ Briefing template for an outline of the kind of information we suggest sharing.*

To get you started, here are our 19 essential questions for buying data quality.

## VENDOR APPROACH

1. Summarize Vendor's views about of the role and purpose of data quality within an enterprise IT and business ecosystem.

*Why ask this?: See if the vendor's approach to "why are we doing data quality" is broadly compatible with yours.*

2. Summarize the solution that Vendor recommends to achieve Buyer's stated goals.

*Why ask this?: Get a quick overview of the solution before you embark on more detailed questions.*

3. In order for Buyer to achieve Vendor's recommended solution, how many separate products are needed, including all named software products and modules from Vendor and from third parties?

*Why ask this?: Unfortunately, vendors who have many distinct product lines—and especially those who have expanded their offerings via acquisition—often don't put as much effort as you'd hope into making those lines interoperable. Ask the vendor to be explicit about how many moving parts their solution includes, so you know whether you need to push the vendor for details on how well they work together. The more parts there are, the more important it is to get those details.*

4. Describe the tangible outcome deliverables that Vendor's product will produce as part of the recommended solution for Buyer.

*Why ask this?: If the project assigned to you is supposed to have specific outcomes for the business, compare those with the outcomes that the vendor's solution will produce: Will the solution cover all your bases, or will you need additional software or work? Now is the time to make sure that the solution is going to deliver everything you need.*

## DATA SOURCES

5. Describe how the solution will enable Buyer to access data from its existing sources: [List your known sources and source types]

*Why ask this?: When you sit down and list these out, you might be surprised to find that some of your systems and datasets have been customized or are proprietary. So if you haven't already made this list, you need to do it now, so you can get information about how vendors will handle those specialized sources.*

## 19 QUESTIONS TO ASK DATA QUALITY VENDORS

6. Are there any sources or source types which, if added into the recommended solution at a later time, would require Buyer to purchase additional modules or functionality for the solution? If so, list them.

*Why ask this?: Depending on your current project, the answer to this question isn't going to necessarily be a dealbreaker, but it's a good thing to know ahead of time anyway. Because if your data quality effort goes well, you'll probably start thinking of how to bring those benefits to other data in your company.*

7. What is the marginal effort required to incorporate additional data sources beyond those specified for the original project?

*Why ask this?: A data quality solution that discourages you from adding new sources of data sets you up to encounter a dead end from its very first day—not ideal, to say the least. Look for solutions where adding a new data source requires about the same effort no matter how many sources you already have, and avoid ones where the more sources you have, the harder it is to add a new one. If the vendor starts talking about “complexity,” look closely: do they mean it's harder to add a source when the source itself is complex (reasonable), or that it's harder to add any source when your existing source ecosystem is complex (not reasonable)?*

### DATA

8. Describe how Vendor's solution will handle data in Buyer's systems that is unique to Buyer.

*Why ask this?: You probably have unique data and don't even know it. If you discover it during your data quality project, it's important to know how the vendor will respond to that.*

9. Is Vendor willing to perform a demo or proof of concept using Buyer's actual data? If yes, describe the parameters of the demo/POC.

*Why ask this?: Even if you're working with customer data--the most popular data domain for data quality--your business, systems, and operations combine to make something unique. The vendor should be willing to show you that their solution can handle your business's concept of a “customer.” This is even more important if you're working with NON-customer data: a vendor with a truly multi-domain solution shouldn't balk at showing you your own data in action. If the vendor's parameters try to dictate extremely narrow terms for what kind of data they're willing to include, that could be a warning flag—is their solution really flexible enough to cope with your data?*

10. Describe how Vendor's solution will remain relevant as Buyer's understanding of its needs for data and data quality evolve.

# 19 QUESTIONS TO ASK DATA QUALITY VENDORS

*Why ask this?: In data quality, you need to expect the unexpected. A solution that can't flex and grow as your needs and expectations change will quickly outlive its usefulness. The degree to which the vendor expects this and is prepared for it can give you information about how easy it will be to keep the solution current with your needs.*

## RULES

11. Describe how Vendor's data quality rule creation will allow Buyer to incorporate Buyer-specific needs and knowledge.

*Why ask this?: So many of the details of truly effective and meaningful data quality depend on things that only your business can know. A solution can't produce maximum value if it doesn't let you wrap those meaningful details in.*

## SOLUTION OUTPUTS

12. Describe what kind of outputs (statistics, data, worklists, etc.) Vendor's solution will produce.

*Why ask this?: Solutions that show basic statistics about records on a system-by-system basis and solutions that use machine learning to produce corrected golden records can both be marketed as a "data quality solution," but they meet very different needs and are generally positioned very differently within a company's flow of data. Use this question to draw out explicit information about what tangibles will be delivered by the vendor's solution, and make sure those tangibles meet your business and technical needs.*

## SOLUTION OPERATION

13. Describe how Vendor's solution supports and/or interacts with a [Name of project] initiative.

*Why ask this?: If your data quality is meant to support or work in conjunction with another data project, either immediately or in the future, it's important to know how the vendor sees that happening, so you can determine if that aligns with your needs and expectations.*

14. Describe how business users are expected to interact with the solution, including: What interfaces are available for business users? What functionality is accessible through these interfaces?

*Why ask this?: Business users are, of the groups that use a data quality solution, often the group with the least technical training and least comfort with learning new interfaces. But at the same time, their insights and knowledge are essential for a well-functioning data quality project. Weigh*

# 19 QUESTIONS TO ASK DATA QUALITY VENDORS

*the answer to this question against the needs and comfort of your specific business users.*

15. Describe Vendor's expectations for Buyer technical user interaction with the solution, including: What interfaces are available for technical users? What functionality is accessible through these interfaces? What degree of ongoing technical input from Buyer will continued operation of the solution require? If Vendor plans to provide technical user interaction to Buyer as a service, describe the scope of the recommended services.

*Why ask this?: It's important that you and the vendor be on the same page about who is going to do what. There's nothing wrong with a vendor-run solution, or with wanting to do it yourself; it's a question of whether the vendor and the solution can reach their maximum potential with the degree of interaction that you want.*

## INTEROPERABILITY AND SCALABILITY

16. Describe Buyer's options for extending the functionality of the recommended solution.

*Why ask this?: This question (and most of the ones in this section) are to provide you with insight into how your solution will be able to cope when the technology or needs around your data quality solution change, which is virtually inevitable. This question is to determine your options if your needs for the solution change: specifically, if you need to add functionality to the solution, like more rules, higher performance, or more output formats.*

17. Describe the options that Vendor's recommended solution provides for making the solution outcomes available to additional and downstream destinations, including analytics and BI tools. Address the following destinations in detail: [List of potential future outcome destinations]

*Why ask this?: New opportunities to use your quality-checked (and/or improved) data can happen at any time. You don't want to be held back from taking advantage of them because your DQ solution "wasn't designed for that."*

18. Describe the upward scalability of Vendor's recommended solution, including the factors (number of sources, number of users, volume of data, etc.) that determine when Buyer would be required to purchase additional software or technical resources.

*Why ask this?: This question is to determine your options if (or when) you want to expand the scope of your data quality... or even if you just want to give more resources to the existing solution (for example, if you want to have more users, do you need more licenses?).*

## TECHNICAL REQUIREMENTS

19. What are the minimum and recommended technical/hardware requirements for Vendor's solution?

*Why ask this?: This may seem like a softball question, but it's actually asking the vendor to consider a number of complex factors, and how the vendor responds will say a lot. Solution sizes will depend on the amount of data you have, types of analysis you intend to perform, and amount of impact on the performance of your existing systems that you're willing to tolerate.*