

Developing Data Capacity

Questions to Ask a Contractor

If you are preparing to hire an outside company or contractor to support your data analysis efforts, below are a few recommended questions to ask regarding the use of data. The way the contractor or company answers these questions should help you identify whether they are the right fit to work with your program.

How do they get the data they will use?

- Do they expect you to enter the data into their system manually?
- Is an electronic transfer possible? Who is responsible for managing the transfer?
- Who is responsible for the export and import of the raw data?

What data systems do they work with?

- Do they work only with ChildPlus? MyHeadStart.com? COPA?
- Do they have the ability to integrate data across multiple systems?

Do they have a HIPAA- and FERPA-compliant process for exchanging and storing data?

- Do they meet any other relevant data protection standards (eg. state requirements)?

What is their policy on data ownership?

- Does your organization always own its data?
- How can they ensure your organization always have access to the data it owns?



Developing Data Capacity

Glossary

Business Intelligence (BI):

A term used to describe the combination of technologies and practices for data collection, integration, analysis, and visualization in order to support data-informed decision-making.

Continuous Improvement:

The use of data and evidence to improve practice.

Culture of Data:

An environment in which data is consistently used to help guide key decisions in an organization. Example: Teachers have ready access to school readiness data about their students and plan curriculum to incorporate overall goals of their program with individual student needs. At the end of the year, teachers use their student records to see if their curriculum affected student outcomes. The grantee as a whole can examine patterns across all student records to see if program outcome goals were met.

Data Analysis:

The process by which raw data is evaluated in order to provide usable information. This can take many different forms in different contexts.

Data Architecture:

A set of rules, policies, procedures, and standards for the collection, storage, management, and integration of data within an organization. Example: Personally identifiable information (PII) is collected and stored and accessed only in one system and is never shared with anyone via email or any other platform.

Data Collection:

The process by which information is acquired, input, and stored. This includes any and all information used by an organization.

Data Integration:

The process by which data from different sources is brought together to be used in combination. Example: Attendance records from one system being aligned to demographic and assessment data.

Data Literacy:

The ability to find useful information in data. It is the knowledge and application of data as appropriate for a person's role in an organization.

Data Quality:

The perception that the data collected is "fit for purpose." The data accurately tells you the information you need to know and is available in a timely and usable way.

Data Transparency:

The ability to easily access the data necessary to perform one's duties no matter where it is located. The knowledge of how the data that is collected is being used by the organization. Example: A classroom teacher takes attendance daily and can look up student attendance records. The same teacher knows that attendance data is used to generate daily call lists and to look for longer-term patterns like chronic absence.

Data Visualization:

The process by which data is displayed so that it is understandable to the viewer. Data visualization allows people to see patterns, trends, and other relationships that are not be easily seen in text displays (or raw spreadsheets) of data.

Data Warehouse:

A means of storing data, often having pulled it from a variety of different, isolated systems into one comprehensive location.

Garbage In – Garbage Out (GIGO):

The principle that the quality of the output is dependent on the quality of the input. If the data entered is low quality, you cannot make high-quality decisions from it.