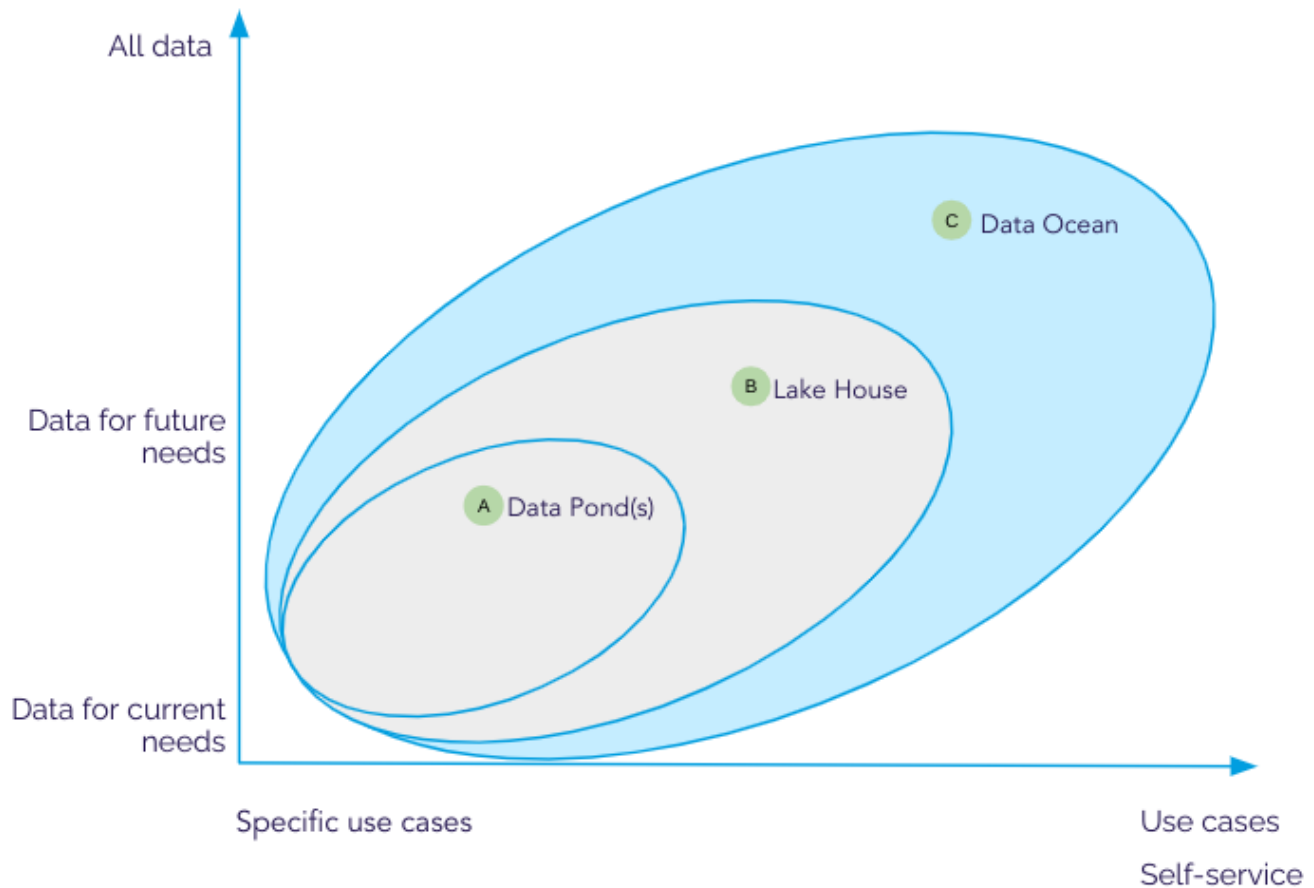


Data Ponds, Data Lakes and Data Ocean



Data Ponds

Data Ponds serve as the starting point within the Data Ocean.

They contain data that is limited to what is specifically needed by individual products or teams. These ponds act as repositories for product-specific data, ensuring that relevant data is readily available and easily accessible to the teams working on specific projects or products.

Data Lakes

Data Lakes represent the next level of data storage and accessibility within the Data Ocean.

They contain a broader range of data (structured and unstructured) that is required by various business units and support self-service capabilities.

Business units can tap into the Data Lakes to access and analyze data that aligns with their specific needs, promoting a more agile and data-driven approach.

Business units can access and analyze data from the Data Lakes that is relevant to their own needs, enabling a more flexible, agile and data-driven strategy.

Data Ocean

The Data Ocean, which extends self-service data and data-driven decision-making capabilities to all enterprise data and is the focal point of the system. Data from diverse sources, including Data Ponds and Data Lakes, are combined and made available to the entire organization through this central repository. The Data Ocean enables local autonomy while promoting global sharing by allowing both the vertical and horizontal use of data.