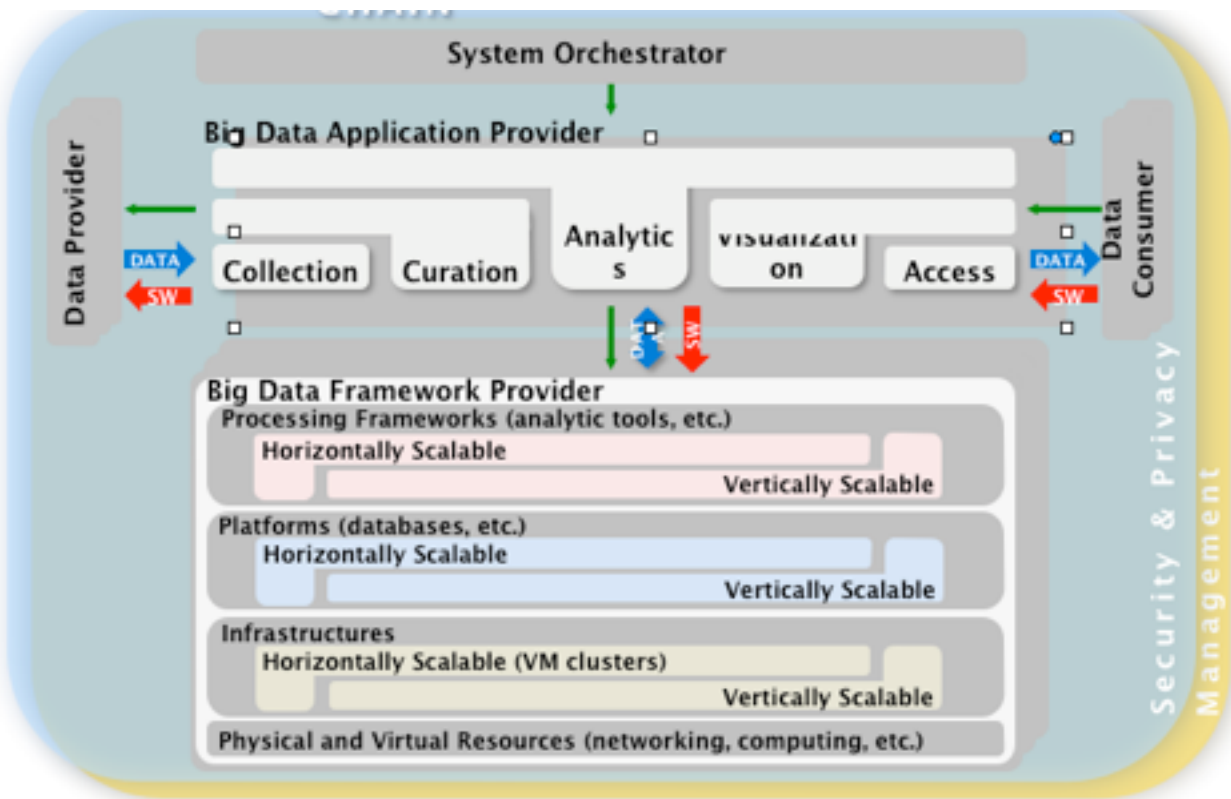


Artificial Data Scientist and Human Analytics Processing

Current NIST Reference Architecture

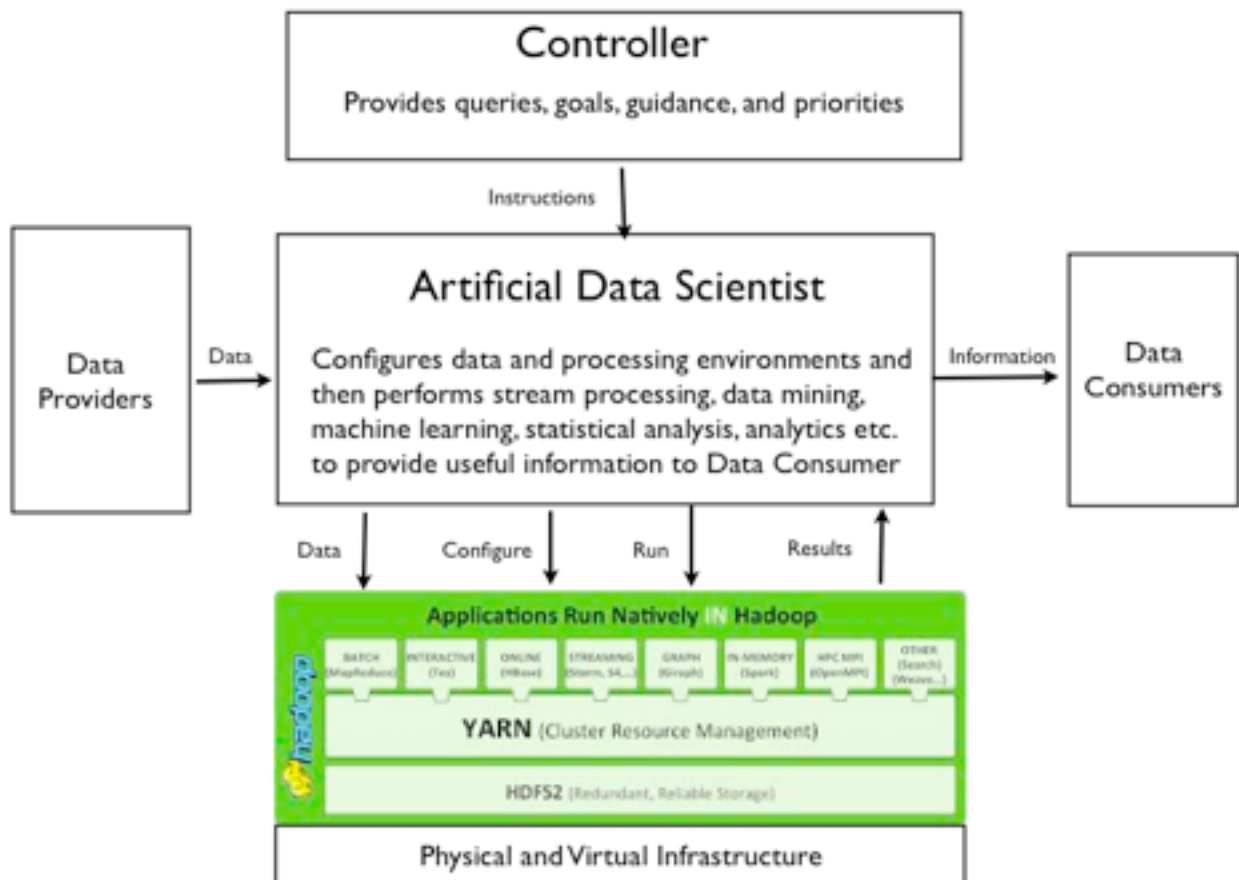
In the NIST Reference Architecture, the Providers can be multiple automated programs and/or manually controlled operations. This reflects the current state of Big Data processing and architectures. For analytics, the Big Data Application Provider could consist of a very skilled team of Data Scientists who know how to configure, run, orchestrate, Big Data analytics applications and post-process for delivery to Data Consumers.



Artificial Data Scientist Architecture

The Artificial Data Scientist is a software package that automates the role of human Data Scientists. This component will probably be needed in the future due to the shortage of data scientists. (Apache Yarn is used as an example in this diagram because it enables multiple data science to run in the same data and process environment.) This simplifies the work of the Artificial Data Scientist. IBM's [Watson](#) and Narrative Systems [Quill Artificial Intelligence Data Engine](#) are early examples of an Artificial Data Scientists.

The Controller will probably be a manual process for the foreseeable future. It will be hard to automate goal selection and query creation.



Brain Architecture for Analytics

The analytics processing within the human brain can be represented in a similar architecture to the Artificial Data Scientist. The role of the Controller is assumed by the conscious mind setting goals and creating queries. There are processes similar to stream processing, data mining, machine learning, and statistical analysis that the brain can perform. (Mechanisms unknown). The Analytic Thinking Process Control is the intermediary step that decides what approach to take to achieving the goal or responding to a query. It also converts the results into information to drive future actions. All of the components of the architecture are potential research topics. See [IBM's Cognitive Computing](#) for an attempt to simulate brain processing.

