

In 1998, Hayashi Chikio argued for data science as a new, interdisciplinary concept, with three aspects: data design, collection, and analysis. Stanford professor David Donoho writes that data science is not distinguished from statistics by the size of datasets or use of computing and that many graduate programs misleadingly advertise their analytics and statistics training as the essence of a data-science program. In 1962, John Tukey described a field he called "data analysis", which resembles modern data science. Big data is a related marketing term. Data analysis typically involves working with smaller, structured datasets to answer specific questions or solve specific problems. However, data science is different from computer science and information science. In 1996, the International Federation of Classification Societies became the first conference to specifically feature data science as a topic. They work at the intersection of mathematics, computer science, and domain expertise to solve complex problems and uncover hidden patterns in large datasets. Stanford professor David Donoho writes that data science is not distinguished from statistics by the size of datasets or use of computing and that many graduate programs misleadingly advertise their analytics and statistics training as the essence of a data-science program. Jeff Wu used the term "data science" for the first time as an alternative name for statistics. Data science is an interdisciplinary field focused on extracting knowledge from typically large data sets and applying the knowledge and insights from that data to solve problems in a wide range of application domains. He reasoned that a new name would help statistics shed inaccurate stereotypes, such as being synonymous with accounting or limited to describing data. Turing Award winner Jim Gray imagined data science as a "fourth paradigm" of science (empirical, theoretical, computational, and now data-driven) and asserted that "everything about science is changing because of the impact of information technology" and the data deluge. Data science is an interdisciplinary academic field that uses statistics, scientific computing, scientific methods, processes, algorithms and systems to extract or extrapolate knowledge and insights from noisy, structured, and unstructured data. Others argue that data science is distinct from statistics because it focuses on problems and techniques unique to digital data. F. Big data is a related marketing term. Though it was used by the National Science Board in their 2005 report "Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century", it referred broadly to any key role in managing a digital data collection. Both fields require a solid foundation in statistics, programming, and data visualization, as well as the ability to communicate findings effectively to both technical and non-technical audiences. Moreover, both fields benefit from critical thinking and domain knowledge, as understanding the context and nuances of the data is essential for accurate analysis and modeling. In 1962, John Tukey described a field he called "data analysis", which resembles modern data science. However, the definition was still in flux. Data analysis typically involves working with smaller, structured datasets to answer specific questions or solve specific problems. The modern conception of data science as an independent discipline is sometimes attributed to William S. The professional title of "data scientist" has been attributed to DJ Patil and Jeff Hammerbacher in 2008.