

"Data science" became more widely used in the next few years: in 2002, the Committee on Data for Science and Technology launched the Data Science Journal. In a 2001 paper, he advocated an expansion of statistics beyond theory into technical areas; because this would significantly change the field, it warranted a new name. 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. Despite these differences, data science and data analysis are closely related fields and often require similar skill sets. Many statisticians, including Nate Silver, have argued that data science is not a new field, but rather another 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. In 2003, Columbia University launched The Journal of Data Science. Others argue that data science is distinct from statistics because it focuses on problems and techniques unique to digital data. For example, a data analyst might analyze sales data to identify trends in customer behavior and make recommendations for marketing strategies. During the 1990s, popular terms for the process of finding patterns in datasets (which were increasingly large) included "knowledge discovery" and "data mining". He reasoned that a new name would help statistics shed inaccurate stereotypes, such as being synonymous with accounting or limited to describing data. Statistician Nathan Yau, drawing on Ben Fry, also links data science to human-computer interaction: users should be able to intuitively control and explore data. In a 2001 paper, he advocated an expansion of statistics beyond theory into technical areas; because this would significantly change the field, it warranted a new name. Data scientists are often responsible for collecting and cleaning data, selecting appropriate analytical techniques, and deploying models in real-world scenarios. In addition to statistical analysis, data science often involves tasks such as data preprocessing, feature engineering, and model selection. They work at the intersection of mathematics, computer science, and domain expertise to solve complex problems and uncover hidden patterns in large datasets. In 2012, technologists Thomas H. "Data science" became more widely used in the next few years: in 2002, the Committee on Data for Science and Technology launched the Data Science Journal. The term "data science" has been traced back to 1974, when Peter Naur proposed it as an alternative name to computer science. Later, attendees at a 1992 statistics symposium at the University of Montpellier II acknowledged the emergence of a new discipline focused on data of various origins and forms, combining established concepts and principles of statistics and data analysis with computing. Data scientists often work with unstructured data such as text or images and use machine learning algorithms to build predictive models and make data-driven decisions. Jeff Wu used the term "data science" for the first time as an alternative name for statistics. Data science is multifaceted and can be described as a science, a research paradigm, a research method, a discipline, a workflow, and a profession. The modern conception of data science as an independent discipline is sometimes attributed to William S. In contrast, data science deals with quantitative and qualitative data (e.g., from images, text, sensors, transactions, customer information, etc.) and emphasizes prediction and action.