

They work at the intersection of mathematics, computer science, and domain expertise to solve complex problems and uncover hidden patterns in large datasets. F. 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. "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 1996, the International Federation of Classification Societies became the first conference to specifically feature data science as a topic. 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. While data analysis focuses on extracting insights from existing data, data science goes beyond that by incorporating the development and implementation of predictive models to make informed decisions. A decade later, they reaffirmed it, stating that "the job is more in demand than ever with employers". While both fields involve working with data, data science is more of an interdisciplinary field that involves the application of statistical, computational, and machine learning methods to extract insights from data and make predictions, while data analysis is more focused on the examination and interpretation of data to identify patterns and trends. Data analysts typically use statistical methods to test these hypotheses and draw conclusions from the data. "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. Data science is a "concept to unify statistics, data analysis, informatics, and their related methods" to "understand and analyze actual phenomena" with data. During the 1990s, popular terms for the process of finding patterns in datasets (which were increasingly large) included "knowledge discovery" and "data mining". A data scientist is a professional who creates programming code and combines it with statistical knowledge to create insights from data. Data science, on the other hand, is a more complex and iterative process that involves working with larger, more complex datasets that often require advanced computational and statistical methods to analyze. Jeff Wu again suggested that statistics should be renamed data science. Data science is a "concept to unify statistics, data analysis, informatics, and their related methods" to "understand and analyze actual phenomena" with data. F. Others argue that data science is distinct from statistics because it focuses on problems and techniques unique to digital data. He reasoned that a new name would help statistics shed inaccurate stereotypes, such as being synonymous with accounting or limited to describing data. Jeff Wu used the term "data science" for the first time as an alternative name for statistics. As such, it incorporates skills from computer science, statistics, information science, mathematics, data visualization, information visualization, data sonification, data integration, graphic design, complex systems, communication and business. Both fields play vital roles in leveraging the power of data to understand patterns, make informed decisions, and solve complex problems across various domains. 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 2003, Columbia University launched The Journal of Data Science.