

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. 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 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. This can involve tasks such as data cleaning, data visualization, and exploratory data analysis to gain insights into the data and develop hypotheses about relationships between variables. Others argue that data science is distinct from statistics because it focuses on problems and techniques unique to digital data. Big data is a related marketing term. "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. Davenport and DJ Patil declared "Data Scientist: The Sexiest Job of the 21st Century", a catchphrase that was picked up even by major-city newspapers like the New York Times and the Boston Globe. He describes data science as an applied field growing out of traditional statistics. 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. Big data is a related marketing term. Others argue that data science is distinct from statistics because it focuses on problems and techniques unique to digital data. The term "data science" has been traced back to 1974, when Peter Naur proposed it as an alternative name to computer science. In 1996, the International Federation of Classification Societies became the first conference to specifically feature data science as a topic. In summary, data analysis and data science are distinct yet interconnected disciplines within the broader field of data management and analysis. Data analysts typically use statistical methods to test these hypotheses and draw conclusions from the data. However, the definition was still in flux. Data scientists are often responsible for collecting and cleaning data, selecting appropriate analytical techniques, and deploying models in real-world scenarios. Others argue that data science is distinct from statistics because it focuses on problems and techniques unique to digital 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. However, data science is different from computer science and information science. 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. Jeff Wu again suggested that statistics should be renamed data science. However, the definition was still in flux. 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.