The term "data science" has been traced back to 1974, when Peter Naur proposed it as an alternative name to computer science. In addition to statistical analysis, data science often involves tasks such as data preprocessing, feature engineering, and model selection. Data science also integrates domain knowledge from the underlying application domain (e.g., natural sciences, information technology, and medicine). Data analysts typically use statistical methods to test these hypotheses and draw conclusions from the data. Vasant Dhar writes that statistics emphasizes quantitative data and description. Andrew Gelman of Columbia University has described statistics as a non-essential part of data science. In 2015, the American Statistical Association identified database management, statistics and machine learning, and distributed and parallel systems as the three emerging foundational professional communities. For instance, a data scientist might develop a recommendation system for an e-commerce platform by analyzing user behavior patterns and using machine learning algorithms to predict user preferences. 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. Data scientists are often responsible for collecting and cleaning data, selecting appropriate analytical techniques, and deploying models in real-world scenarios. 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. Data analysis focuses on extracting insights and drawing conclusions from structured data, while data science involves a more comprehensive approach that combines statistical analysis, computational methods, and machine learning to extract insights, build predictive models, and drive data-driven decision-making. Vasant Dhar writes that statistics emphasizes quantitative data and description. Vasant Dhar writes that statistics emphasizes quantitative data and description. 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. Both fields play vital roles in leveraging the power of data to understand patterns, make informed decisions, and solve complex problems across various domains. Data scientists are often responsible for collecting and cleaning data, selecting appropriate analytical techniques, and deploying models in real-world scenarios. 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. 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. For example, a data analyst might analyze sales data to identify trends in customer behavior and make recommendations for marketing strategies. Vasant Dhar writes that statistics emphasizes quantitative data and description. 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. Others argue that data science is distinct from statistics because it focuses on problems and techniques unique to digital 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.