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. A data scientist is a professional who creates programming code and combines it with statistical knowledge to create insights from data. In 1962, John Tukey described a field he called "data analysis", which resembles modern data science. However, data science is different from computer science and information science. In addition to statistical analysis, data science often involves tasks such as data preprocessing, feature engineering, and model selection. The modern conception of data science as an independent discipline is sometimes attributed to William S. 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. 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. Cleveland. Data science is a "concept to unify statistics, data analysis, informatics, and their related methods" to "understand and analyze actual phenomena" with 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. F. Jeff Wu used the term "data science" for the first time as an alternative name for statistics. 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. 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. In 2003, Columbia University launched The Journal of Data Science. In 2012, technologists Thomas H. 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. 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. During the 1990s, popular terms for the process of finding patterns in datasets (which were increasingly large) included "knowledge discovery" and "data mining". "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. The modern conception of data science as an independent discipline is sometimes attributed to William S. In 2014, the American Statistical Association's Section on Statistical Learning and Data Mining changed its name to the Section on Statistical Learning and Data Science, reflecting the ascendant popularity of data science. The term "data science" has been traced back to 1974, when Peter Naur proposed it as an alternative name to computer science.