Query ID: test2115

Query Text: true or false the term big data refers only to the amount of data being generated

Retrieved Documents:

0	doc73572	Big data. The term has been in use since the 1990s, with some giving credit to John Mashey for coining or at least making it popular.[18][19][20] Big data usually includes data sets with sizes beyond
0	doc73637	Big data. Big data analysis is often shallow compared to analysis of smaller data sets.[186] In many big data projects, there is no large data analysis happening, but the challenge is the extract, tra
1	doc73569	Big data. Lately, the term "big data" tends to refer to the use of predictive analytics, user behavior analytics, or certain other advanced data analytics methods that extract value from data, and sel
0	doc2401202	Data (Star Trek). In "The Measure of a Man", a Starfleet judge rules that Data is not Starfleet property. The episode establishes that Data has a storage capacity of 800 quadrillion bits, (100 PB or 8
0	doc73628	Big data. Critiques of the big data paradigm come in two flavors, those that question the implications of the approach itself, and those that question the way it is currently done.[164] One approach t
0	doc2422578	Data quality. The Data QC process uses the information from the QA process to decide to use the data for analysis or in an application or business process. For example, if a Data QC process finds that
0	doc73585	Big data. 2012 studies showed that a multiple-layer architecture is one option to address the issues that big data presents. A distributed parallel architecture distributes data across multiple server
0	doc2149058	Information society. The growth of technologically mediated information has been quantified in different ways, including society's technological capacity to store information, to communicate informati
0	doc1112100	Data type. Almost all programming languages explicitly include the notion of data type, though different languages may use different terminology. Common data types include:
0	doc2422577	Data quality. Data QA processes provides following information to Data Quality Control (QC):
0	doc631440	Column (database). Also, as the data type of each column is alike, better compression occurs when running compression algorithms on each column, which will help queries churn results more quickly.[8]
0	doc1055301	Data validation. Definition and design contexts:
0	doc415415	Law of large numbers. That is,
0	doc126153	Array data structure. Some array data structures do not reallocate storage, but do store a count of the number of elements of the array in use, called the count or size. This effectively makes the arr
0	doc2291413	Logical data model. A logical data model is sometimes incorrectly called a physical data model, which is not what the ANSI people had in mind. The physical design of a database involves deep use of pa
0	doc1761256	Database. Database storage is the container of the physical materialization of a database. It comprises the internal (physical) level in the database architecture. It also contains all the information
0	doc790887	Internet access. The download (to the user) and upload (to the Internet) data rates given above are peak or maximum rates and end users will typically experience lower data rates.
0	doc657565	Equation. which are both true for all values of I .
0	doc1442258	Long and short scales. It should be remembered that "billion" does not mean in American use (which follows the French) what it means in British. For to us it means the second power of a million, i.e.
0	doc574386	Data analysis. Once the data is cleaned, it can be analyzed. Analysts may apply a variety of techniques referred to as exploratory data analysis to begin understanding the messages contained in the da
0	doc1792302	Information privacy law. Personal data covers both facts and opinions about the individual.[8] It also includes information regarding the intentions of the data controller towards the individual, alth

0	doc1837378	Financial data vendor. According to the 2009 Burton-Taylor report, the Market Data industry exited 2009 at US$22.68 billion after closing 2008 at US$23.01 billion. In 2009, Thomson Reuters and Bloombe
0	doc1755111	Digital forensics. (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and met
0	doc2257709	Boolean data type. In programming languages with a built-in Boolean data type, such as Pascal and Java, the comparison operators such as > and a are usually defined to return a Boolean value. Condit
0	doc1866060	Statement (logic). In logic, the term statement is variously understood to mean either:
0	doc1076788	Database model. Common logical data models for databases include:
0	doc1561697	Syllogism. If a statement includes a term such that the statement is false if the term has no instances, then the statement is said to have existential import with respect to that term. It is ambiguou
0	doc1789858	Backup. As long as new data are being created and changes are being made, backups will need to be performed at frequent intervals. Individuals and organizations with anything from one computer to thou
0	doc389241	Median. If there is an even number of observations, then there is no single middle value; the median is then usually defined to be the mean of the two middle values.[1][2] For example, in the data set
0	doc957824	Sample size determination. Sample sizes may be chosen in several different ways:
0	doc470902	Marketing information system. Kotler and Philip have said that "both primary and secondary researches offer loads of the data and information needed for the marketers, whereas the secondary data sourc
0	doc2626930	Missing data. Data often are missing in research in economics, sociology, and political science because governments choose not to, or fail to, report critical statistics.[1] Sometimes missing values a
0	doc1959490	Regression analysis. The performance of regression analysis methods in practice depends on the form of the data generating process, and how it relates to the regression approach being used. Since the
0	doc2067501	Smart growth. Growth is "smart growth", to the extent that it includes the elements listed below.[1][15][16]
0	doc2221771	Social alienation. Geyer (1996: xxiii) believes meaninglessness should be reinterpreted for postmodern times: "With the accelerating throughput of information [...] meaningless is not a matter anymore
0	doc1253689	Star schema. In computing, the star schema is the simplest style of data mart schema and is the approach most widely used to develop data warehouses and dimensional data marts.[1] The star schema cons
0	doc213941	Cloud storage. Cloud storage typically refers to a hosted object storage service, but the term has broadened to include other types of data storage that are now available as a service, like block stor
0	doc1985493	Histogram. It implicitly bases the bin sizes on the range of the data and can perform poorly if n < 30, because the number of bins will be smallless than sevenand unlikely to show trends in the data
0	doc1639474	Prescriptive analytics. In addition to this variety of data types and growing data volume, incoming data can also evolve with respect to velocity, that is, more data being generated at a faster or a v
0	doc415390	Law of large numbers. According to the law of large numbers, if a large number of six-sided dice are rolled, the average of their values (sometimes called the sample mean) is likely to be close to 3.5
0	doc2071577	Data cleansing. Administratively, incorrect or inconsistent data can lead to false conclusions and misdirected investments on both public and private scales. For instance, the government may want to a
0	doc1088598	Array data type. Depending on the language, array types may overlap (or be identified with) other data types that describe aggregates of values, such as lists and strings. Array types are often implem
0	doc2102932	Scalability. While DBMS vendors debate the relative merits of their favored designs, some companies and researchers question the inherent limitations of relational database management systems. GigaSpa
0	doc1043468	Paging. Although the processor in this example cannot address RAM beyond 4A GB, the operating system may provide services to programs that envision a larger memory, such as files that can grow beyond
0	doc2503863	Web Ontology Language. [The closed] world assumption implies that everything we dont know is false,

		while the open world assumption states that everything we dont know is undefined.
0	doc89353	Computer data storage. Main memory is directly or indirectly connected to the central processing unit via a memory bus. It is actually two buses (not on the diagram): an address bus and a data bus. Th
0	doc1051555	Image scanner. The size of the file created increases with the square of the resolution; doubling the resolution quadruples the file size. A resolution must be chosen that is within the capabilities o
0	doc2480666	Object-relational database. An RDBMS might commonly involve SQL statements such as these:
0	doc2407624	SQL. SQL's controversial "null" value is neither true nor false (predicates with terms that return a null value return null rather than true or false). Features such as outer-join depend on null value
0	doc2168340	NTFS. Sparse files are files interspersed with empty segments for which no actual storage space is used. To the applications, the file looks like an ordinary file with empty regions seen as regions fi
0	doc1249362	Statistical inference. This paradigm calibrates the plausibility of propositions by considering (notional) repeated sampling of a population distribution to produce datasets similar to the one at hand
0	doc2504429	Observations and Measurements. The core of the standard provides the observation schema. An observation is an act that results in the estimation of the value of a feature property, and involves applic
0	doc284861	Significant figures. For quantities created from measured quantities by multiplication and division, the calculated result should have as many significant figures as the measured number with the least
0	doc1189622	Positivism. Verified data (positive facts) received from the senses are known as empirical evidence; thus positivism is based on empiricism.[1]
0	doc1328446	Primary source. "The definition of a primary source varies depending upon the academic discipline and the context in which it is used.
0	doc1512458	False color. In addition, variants of false color such as pseudocolor, density slicing, and choropleths are used for information visualization of either data gathered by a single grayscale channel or
0	doc18490	Null hypothesis. The concept of a null hypothesis is used differently in two approaches to statistical inference. In the significance testing approach of Ronald Fisher, a null hypothesis is rejected i
0	doc1543324	Census in the United Kingdom. Traditionally, outputs are released in the form of tables of counts at various levels of geography. However, microdata, known Samples of Anonymised Records (SARs) are UK
0	doc436985	Megabyte. The megabyte is commonly used to measure either 10002 bytes or 10242 bytes. The interpretation of using base 1024 originated as a compromise technical jargon for the byte multiples that need
0	doc2308734	k-anonymity. There are 6 attributes and 10 records in this data. There are two common methods for achieving k-anonymity for some value of k.

Non-retrieved Relevant Documents: