

= Paxata =

Paxata is a privately owned software company headquartered in Redwood City , California . It develops self @-@ service data preparation software that gets data ready for data analytics software . Paxata 's software is intended for business analysts , as opposed to technical staff . It is used to combine data from different sources , then check it for data quality issues , such as duplicates and outliers . Algorithms and machine learning automate certain aspects of data preparation and users work with the software through a user @-@ interface similar to Excel spreadsheets .

The company was founded in January 2012 and operated in stealth mode until October 2013 . It has received more than \$ 10 million in venture funding . Analysts have praised Paxata for creating software that is user @-@ friendly for non @-@ technical business users , but caution that they are in a noisy marketplace .

= = History = =

Paxata was founded in January 2012 . It initially raised \$ 2 million in venture capital . The company came out of stealth mode in October 2013 . Simultaneously with its public release , Paxata announced an \$ 8 million funding round led by Accel Partners . Adoption of the software grew quickly . In March 2014 , In @-@ Q @-@ Tel acquired an interest in the startup . It raised an additional \$ 18 million in funding in September 2015 . It also began working with Cisco to jointly develop the Cisco Data Preparation suite of software and services .

= = Software = =

Paxata refers to its suite of cloud @-@ based data quality , integration , enrichment and governance products as " Adaptive Data Preparation . " The software is intended for business analysts , who need to combine data from a variety of sources , then check the data for duplicates , empty fields , outliers , trends and integrity issues before conducting analysis or visualization in a third @-@ party software tool . It uses algorithms and machine @-@ learning to automate certain aspects of data preparation . For example , it may automatically detect records belonging to the same person or address , even if the information is formatted differently in each record in different data sets .

The software has a spreadsheet @-@ based user interface . Patterns and anomalies in the data are color @-@ coded in the spreadsheet . Then users are provided with instructions on how to resolve data quality issues or to supplement the data with contextual information . Data sets and related quality issues can also be addressed in a collaborative environment through the " Paxata Share " feature . It runs on Apache Spark .

According to analyst firm Ovum , the software is made possible through advances in predictive analytics , machine learning and the NoSQL data caching methodology . The software uses semantic algorithms to understand the meaning of a data table 's columns and pattern recognition algorithms to find potential duplicates in a data @-@ set . It also uses indexing , text pattern recognition and other technologies traditionally found in social media and search software . One of the software 's users is dairy producer Dannon , which uses the software so that business staff can create their own reports on merchandising , supply chain and product data , without the IT department .

= = Reception = =

In its 2014 report " Cool Vendors in Data Integration and Data Quality " , Gartner praised Paxata for developing a " business @-@ user @-@ friendly " data quality product that does not use code . Ventana Research said its spreadsheet @-@ based user interface " should resonate well with business analysts , " who are resistant to move away from familiar Excel @-@ like programs .

Gartner also said Paxata was recognized in the report due to its automated , algorithm @-@ based features and how it tracks any changes made to the data .

Ventana Research said Paxata was in a " noisy marketplace " . According to Gartner , while Paxata is an early entrant into the market , many startups and large corporations are making investments in developing similar competing products . According to Gigaom and IT Business Edge , one way Paxata differs is that it automatically merges multiple data @-@ sets into a single table , so it can be easily imported into a visualization or analysis tool .

Gartner said Paxata will have a difficult time finding a compelling pricing model , when many data discovery tools that it supplements provide some similar features . In contrast , Ventana said Paxata 's pricing was " a pretty small amount " compared to the amount of time users can save .