

## REPORT REPRINT

# Unifi adds AI-driven recommendations to Data as a Service software

**MATT ASLETT**

**23 OCT 2017**

Unifi Version 2.0 adds an artificial intelligence-powered recommendation engine, among other things, to the company's combination of data governance, catalog, discovery, preparation and collaboration functionality.

---

THIS REPORT, LICENSED TO UNIFI SOFTWARE, DEVELOPED AND AS PROVIDED BY 451 RESEARCH, LLC, WAS PUBLISHED AS PART OF OUR SYNDICATED MARKET INSIGHT SUBSCRIPTION SERVICE. IT SHALL BE OWNED IN ITS ENTIRETY BY 451 RESEARCH, LLC. THIS REPORT IS SOLELY INTENDED FOR USE BY THE RECIPIENT AND MAY NOT BE REPRODUCED OR RE-POSTED, IN WHOLE OR IN PART, BY THE RECIPIENT WITHOUT EXPRESS PERMISSION FROM 451 RESEARCH.



©2017 451 Research, LLC | [WWW.451RESEARCH.COM](http://WWW.451RESEARCH.COM)

Unifi Software's combination of data governance, catalog, discovery, preparation and collaboration technology is now described as a Data as a Service platform. The latest version, 2.0, adds an artificial intelligence (AI)-powered recommendation engine, among other things, as the company plans to hit triple figures in terms of customer adoption by the end of the year.

---

## THE 451 TAKE

While we might not be convinced by the term 'data as a service' given the potential for confusion, Unifi Software has clearly grown its customer base rapidly in 2017 and the launch of Unifi Version 2.0 has the potential to drive further enterprise adoption with enhanced recommendation, lineage, metadata and collaboration functionality, as well as a new cost-based optimizer to ensure the most appropriate use of MapReduce or Spark for data-preparation execution. This market is still emerging, but Unifi is well-placed to take advantage of growing interest in strategic self-service data platform initiatives.

---

## CONTEXT

Unifi Software was founded in 2014 by former executives of analytics database firm Greenplum and advanced analytics software provider Alpine Data (which was incubated with Greenplum). The company launched the first version of its software, then described as a data-integration offering, in early 2015. Unifi was always more than a simple data-integration product, however. Version 1.0 provided a combination of self-service preparation and integration of data stored in relational databases and Hadoop, based on the company's automated parsing functionality, which creates a search-based catalog of enterprise data as it is ingested into Hadoop.

This combination of self-service data preparation and catalog-based data governance has emerged as a fundamental enabler of multi-function 'data lake' environments. There is less agreement on what to call it, however. Unifi has opted for the term Data as a Service platform to describe its combination of data governance, catalog, discovery, preparation and collaboration functionality, which is positioned to address the needs of both business analysts and IT professionals. The term data as a service is already used elsewhere in the industry, however – specifically to describe public cloud services (also called data markets or data clouds) through which enterprises can subscribe to public and private data sets (typically behavioral, demographic used to identify and target specific groups of users/customers).

Unifi's Data as a Service platform is clearly different, but overlaps – a public data cloud (data as a service) offering could be a source of data, alongside data from internal enterprise applications made available to data analysts using Unifi's Data as a Service platform. There is clearly some potential for confusion here, although it doesn't appear to have inhibited the company's customer traction to date.

When we last spoke with Unifi, in December 2016, it had doubled its customer count from the 10 it had when we initiated coverage in March 2016, and was aiming for 25 by the end of 2016. The company is now aiming for over 100 paying customers by the end of 2017, and boasts 70PB of data under management and 1,200 daily individual users. Its list of reference customers includes Havas Media, A.T. Kearny, Boston Biomedical, Compassion-First Pet Hospitals, Essentra and Sabre Holdings. The rapid growth in customer accounts this year is all the more impressive since Unifi's focus has been on 'strategic' adoption driven by data stewards and IT professionals, which necessarily results in a longer adoption cycle than stand-alone data-preparation tools, which tend to be targeted initially at business users and departmental use cases.

Unifi launched Version 2.0 of its Data as a Service platform in September. The latest version added a number of new capabilities to its existing combination of data governance, catalog, discovery, preparation and collaboration functionality, not least an AI-powered recommendation engine designed to reduce the time taken for users to prepare and analyze data. Also new is a single integrated console that provides search-based discovery of all data assets, as well as the ability to expose lineage information to trace the origins of a particular metric or data set. Additionally, Unifi Version 2.0 features a metadata business glossary, collaborative commenting, support for both MapReduce and Spark as data-preparation execution engines, and cost-based optimization functionality to select the most appropriate execution engine for an individual task.

Previously, in August, Unifi introduced a new service called RegAlert! designed to monitor data assets to ensure compliance with the European Union's General Data Protection Regulation related to the protection of personal data. Unifi has also raised a fresh round of funding since we last spoke with the company. In March, it announced that it had raised a \$17.5m series B round led by Scale Venture Partners with participation from existing investors Canaan Partners and Pelion Venture Partners. The round brought the total raised by Unifi to \$32m.

## COMPETITION

The closest competition to Unifi Software comes from other vendors offering a combination of self-service data preparation and data management/governance, especially those providing it in a single product, as well as others with niche offerings in each category. We believe that Podium Data and Zaloni are the most directly comparable with Unifi. Podium describes its combination of self-service data preparation and data management as a Data Marketplace, while Zaloni used to offer two separate components but has now combined them in its Data Lake Management Platform.

Cask is also a player in this space with its Cask Data Application Platform, as is Cambridge Semantics with its Anzo Smart Data Platform. Immuta, meanwhile, is focused on data management specifically for data science. Data-integration heavyweight Informatica also offers a combination of self-service data integration and data management/governance via Informatica Big Data Management, which includes Intelligent Data Lake for self-service data preparation and Enterprise Information Catalog.

IBM is a key vendor in this space as well with its combination of IBM Bluemix Data Connect and InfoSphere Information Governance Catalog, while the data-cataloging opportunity is also a target for the likes of Alation, Collibra and Waterline Data as well as Amazon Web Services and Microsoft via Azure Data Catalog. There are still further contenders more heavily focused on self-service data preparation, including Datawatch, Oracle (Big Data Preparation), Paxata, Pentaho, SAP (Agile Data Preparation), SAS Institute, Talend and Trifacta, while there are also a variety of self-service analytics products with data-preparation capabilities from the likes of Alteryx, Birst, ClearStory Data, Datameer, Looker, TIBCO Spotfire and Yellowfin.

## SWOT ANALYSIS

### STRENGTHS

Unifi Software was founded by a management team with data management and analytics experience and has gained traction this year, generating a healthy number of paying customers.

### WEAKNESSES

We are not convinced by the term data as a service, believing it has the potential to create confusion given its usage for cloud services through which enterprises can subscribe to public and private data sets.

### OPPORTUNITIES

We are seeing a growing interest in data management and governance as enablers of strategic self-service data-preparation and data lake initiatives.

### THREATS

The incumbent data management and integration vendors already have supplier relationships with established enterprises and will look to squeeze out emerging data lake specialists.