



Case Study

AbeBooks, a global new and used book retailer and subsidiary of Amazon.com, migrates to the cloud and increases ETL literacy within the company by 150%



www.matillion.com

AbeBooks, with Amazon Redshift, has been able to upgrade to a comprehensive data warehouse with the enlistment of Matillion ETL for Amazon Redshift. In this case study, we share AbeBooks' data warehouse success story.

The Challenge

There are some database challenges that are universal; primarily performance and high costs of ownership. This was no different for AbeBooks, so they looked to Amazon Redshift for the solution. The cloud-based database allows users to quickly and simply launch fully managed environments that scale up/down based on workload needs. Most noteworthy, Amazon Redshift is secure and compatible with a host of other solutions, such as Matillion. In addition to these features, Amazon Redshift is a cost-effective data warehouse solution, you can run 1TB a year for less than \$1000. For these reasons "Amazon Redshift was the natural choice ... It offered better performance and lower costs than the database platform we were previously using ... In addition, with Athena and Spectrum, we can leverage big data and query this cheaply," explains Dmitry Anoshin, Data/BI engineer at AbeBooks.

With a platform selected AbeBooks wanted to look at new ETL tools. Their ETL solution at the time was inherently linked to their legacy on-premise database. This made it "complicated and fragile and essentially a black box to everyone except for a few DBAs." A new ETL/ELT would be needed to migrate their database, connect with third party data sources such as AdWords and Facebook to continually load new data, transform data, work with Tableau to visualize that data and provide enhancements such as notifications and performance tracking. AbeBooks wanted "ETL to be more accessible to business users," added Dmitry. Based on their previous experience of ETL, they were expecting to incur significant problems loading data into the cloud. They looked at both on-premise and cloud-based ETL/ELT tools to evaluate their options.

About AbeBooks

AbeBooks, a subsidiary of Amazon, is an online marketplace for books, fine art and collectibles. Trusted independent sellers from around the world offer for sale millions of new, used and rare books, as well as art and collectibles, such as antique maps and vintage magazines, through the AbeBooks websites.

www.abebooks.com



The Solution

As part of their migration project, AbeBooks selected their new technology stack comprising of **Matillion ETL for Amazon Redshift**, **Amazon Redshift** and Tableau which can be used alongside AWS Big Data tools such as Athena and Spectrum.

Amazon Redshift

AbeBooks was already using Amazon Redshift in a limited capacity until this year when it became the core database for their data warehouse. Dmitry went on to say that “For us Redshift offered better performance and lower costs than the database platform we were previously using.”

Matillion

Matillion ETL for Amazon Redshift was recommended to AbeBooks by Solution Architects who had previously supported customers on AWS cloud deployments. Dmitry admitted: “I was very skeptical about Matillion—I had never heard of it and honestly thought that this is just a buzzword, but we took the recommendation and put it on our list of tools to evaluate.” After undergoing a technology options assessment, AbeBooks progressed with a PoC through the 14-day free trial offered on the AWS Marketplace. Matillion ETL for Amazon Redshift is available exclusively on the AWS Marketplace as an AMI (Amazon Machine Image) that is deployed inside your VPC (Virtual Private Cloud). Therefore, a customer has full control over the network security. Moreover, Matillion is accessed via a web browser so there is no software installations or hardware configurations on client machines.

AbeBooks tested their ETL options against their various use cases and found Matillion to be the “fastest and easiest” solution. To procure Matillion, AbeBooks went to the AWS Marketplace, spun up an instance and started a 14-day free trial within minutes.

“If you are doing ETL to Redshift, I think you absolutely have to look at Matillion because it really is that “best of breed” tool on the Marketplace right now.”

Dmitry Anoshin
Data/Business Intelligence
Engineer, AbeBooks

“Matillion ETL for Amazon Redshift is purpose-built for Amazon Redshift, and gives customers the tooling required to quickly and capably deliver analytics projects using AWS and Redshift.”

Matthew Scullion
CEO, Matillion

The Benefits

By implementing Matillion ETL for Amazon Redshift, AbeBooks has overcome several huge business challenges. First, they were able to quickly assess and procure the right tools for the job in a retail-like transaction on the AWS Marketplace. Next, they seamlessly migrated their database to Amazon Redshift, ahead of schedule, with the help of Matillion ETL. Finally, they were able to build a number of orchestration and transformation jobs to bring data in and transform it, leveraging the power of Amazon Redshift and Spectrum.

Their AWS compatible technology solution has addressed the many challenges AbeBooks was facing with their on-premise database and legacy ETL tool. ETL is no longer a black box. All jobs are self-described with notes meaning, "if I show a colleague stored procedure code from our old ETL processes I'll likely just get a pained stare. But if I show them a Matillion orchestration I'll get a completely different reaction—the visual processes are naturally intuitive," says Dmitry. Shining a light on the ETL black box has increased the number of ETL users within AbeBooks by 150%. Since users can self-serve by looking up existing jobs to inform new jobs, DBAs and software engineers' time is safeguarded, thus leading to faster development cycles.

Summary

Since we couldn't write a better summary than our customer already did, we'll let them have the last word.

"Anybody that has worked in tech for some time knows there are good software tools and not-so-good software tools out there. Both might do about the same thing, but your choice of the good tool means you will do it with less pain and grief, and maybe even a bit of joy. And because of this you will achieve your goal faster, better, cheaper. If you are doing ETL to Redshift, I think you absolutely have to look at Matillion because it really is that 'best of breed' tool on the marketplace right now."

Benefits of using Matillion ETL for Amazon Redshift

- ✓ Built specifically for AWS and Amazon Redshift
- ✓ Intuitive browser-based user experience – easy on-boarding and powerful
- ✓ Push-down ELT architecture – simplified infrastructure, fast performance
- ✓ Powerful feature set
- ✓ Retail-like acquisition through AWS Marketplace
- ✓ Affordable pricing for everyone, from small startups to Fortune 500 companies
- ✓ Wide range of data source connectors, all included
- ✓ A fully-integrated, data-integration tool that requires no additional development or maintenance staff

About Matillion

Founded in 2011, Matillion has offices in Manchester, UK and New York City. Matillion delivers technology that helps companies exploit their data using the Cloud. Matillion is one of a very small number of Amazon AWS Big Data Competency holders worldwide. Matillion ETL for Amazon Redshift is available in all regions on the AWS Marketplace, and Matillion is an AWS Advanced Technology Partner. Learn more at www.matillion.com.



Matillion is an AWS Advanced Technology Partner and an AWS Big Data Competency holder.
Matillion ETL for Amazon Redshift is available worldwide via the AWS Marketplace.