MarkLogic®

IOP Publishing Cuts Costs and Increases Efficiency
How the developers taught themselves MarkLogic

IOP Publishing

The Institute of Physics (IOP) is a scientific society of over 40,000 physicists and physics enthusiasts. IOP advances physics research, applications, and education. As advocates for science, IOP educates policy makers and the public about the importance of physics. The institute's nonprofit publishing arm, IOP Publishing, produces its own scientific journals. Through its website, IOP Publishing also distributes a vast and evolving portfolio of content including journals, magazines, community websites, and conference proceedings. Always striving to increase accessibility to vital scientific content, IOP Publishing is a partner to the science community.

Summary

Customer
IOP Publishing

Industry
Science Publishing

Challenge
Replace relational database to shorten development cycle with a limited budget

Customer Use Case
Online platform of scientific content

Benefits
Agility, fast delivery of content, lower operational costs

> "With MarkLogic, our system is quicker because it cuts out the complexity of our technology stack – we don't need to convert to a relational database and then back again."
>
> Will Funnell
> Systems Developer,
> IOP Publishing

Highlights

- IOP Publishing is a global distributor of scientific content with more than 50,000 users across the world submitting one million page requests each day

- The IOP Publishing development team members taught themselves MarkLogic by attending MarkLogic events and by using tools available on the MarkLogic Community

- With MarkLogic, IOP Publishing drastically decreased its time to publish

- By migrating from Oracle to MarkLogic, the development team cut costs without any downtime

Challenge

As a global publisher and distributor of scientific publications, IOP Publishing customers expect to easily access the most recent content and to be able to do so at any moment. To meet this demand, IOP Publishing wanted a much more efficient way to deliver content. The legacy solution was complex, archaic, expensive to administer, and slow to deliver accurate results to customers. Built on Oracle and Hibernate, the eight-layer stack required a multi-step process to take unstructured content, like journal articles, from ingest to delivery.

To increase efficiency, IOP Publishing wanted to access content directly. To do this in Oracle, the IOP team had to copy new fields and rows in the database to represent the content. This process often resulted in multiple versions of content in different locations that created a mismatch between versions. Valuable time was lost tracing back to the source; this directly affected the scientists and other users who rely on the IOP platform as a vital research tool.

To provide the best service to its user community, IOP Publishing wanted to overhaul its biggest platform, IOPscience. With a limited budget and a mandate to switch systems with no downtime, the IOP Publishing team knew it couldn't rebuild the system from the ground up with traditional technology. The team members thought MarkLogic might be the solution because it is optimized for any data, at any volume, in any structure. However, the team was composed of five Java developers with little experience in non-relational databases. The team decided to teach itself how to develop with MarkLogic.

IOPscience Landing Page
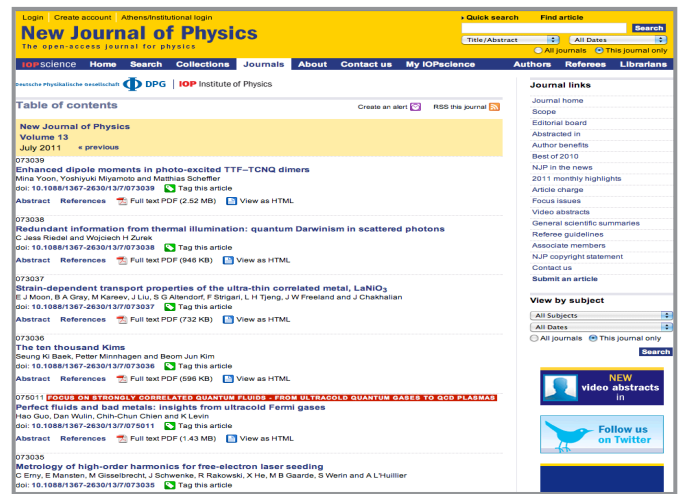
IOP Search Results Powered by MarkLogic

Solution

The IOP development team set about on its own to migrate from Oracle to MarkLogic section by section. Fortunately, rich training content is readily available to MarkLogic developers and was invaluable for this migration process.

Team members began by looking through the documentation, and experimented with small proof of concepts to boost their confidence. Then, by attending the MarkLogic User Conference, they learned how easy it would be to change schemas and add flexibility to their system. The team continued its education on the MarkLogic Community, a site for developers, students, and enthusiasts which includes free software, tutorials, screencasts, videos, demos, discussions, and contests. The team found the MarkLogic Community blog especially helpful with specific examples from seasoned developers. For example, a single blog post led the team to best practices for coding that drastically sped up IOP queries.

As the developers learned the system, they began migrating content from Oracle to MarkLogic. They started with the landing pages for articles, working on just the bibliographic information. From there, they took on the task of representing entire articles. Now, the team is moving entire journals over one by one.

IOPscience is a major source of scientific content with 450,000 articles dating back to 1874, more than 50,000 users across the world, and up to a million page requests a day. With MarkLogic, that content is efficiently delivered to end users. The team's self-taught approach to learning a new system paid off. By switching to MarkLogic, IOP Publishing is delivering content faster than ever while cutting costs. By eliminating the need for several database administrators, the team members are able to focus their time and energy on coding new design features to enhance the IOPscience user experience.

Next Steps

The IOP development team is in the process of moving all IOP journals to MarkLogic. Once the migration is complete, the IOP Publishing team plans to consider implementing MarkLogic search for all IOP content.

Conclusion

With its simplified architecture and revolutionary approach to handling Big Data and unstructured information, MarkLogic is a cost-effective, simple solution to the complexities and inefficiencies of Oracle. The IOP Publishing development team was not familiar with XQuery, but as good programmers tend to do, they took the initiative to learn a new language by themselves by using freely available MarkLogic resources. IOP Publishing completed the migration from Oracle to MarkLogic completely in house, under budget, and on time. Not only is the team saving money on administration, but it is also reducing its time to publish and add other features that enhance user experience.