

## Case Study

# How Dorling Kindersley enhanced content discoverability, improved workflows and controlled costs



### Company name

Dorling Kindersley

### Location

Publishing

### Industry

Global

### About

Established in 1974, Dorling Kindersley (DK) publishes bestselling, award-winning and visually stunning content for all audiences, in which the unrivalled clarity of the words and pictures come together to spectacular effect.

Dorling Kindersley (DK) is undergoing a period of significant business change following the Penguin Random House merger, as control over processes and technology is transferred from Pearson Technology to DK. Previously existing plans to migrate away from DK's legacy content store — an inefficient system with poor usability, built upon Documentum — suddenly became business-critical and highly time-sensitive.

Copyright Clearance Center (CCC) developed a custom content platform and search interface which would support DK's wider strategic goals, allowing powerful granular discoverability across DK's internal product store and offering more efficient content reuse, new products and additional revenue streams.

As a partner working with DK since 2012, CCC was engaged to provide a phased approach for speedily and reliably delivering this solution, which would ultimately bring together a single product archive of up to 5000 unique products.

Following successful delivery of the platform in September 2014 on time and budget, CCC has signed a managed service contract for the platform.

### Technologies used

- Alfresco Web Framework
- ElasticSearch
- Amazon Web Services incl. EC2, RDS, S3, Cloudfront and SQS
- OWLIM RDF framework
- Mongo DB





We're in transition of thinking about content in addition to print," said Russell Watts, Program Manager at DK, "CCC are helping us in that transitional stage and have been able to offer some great expertise in terms of change management and transformation projects."

**Russell Watts**  
Program Manager at DK

### Significant business change

DK's Product store was owned by Pearson Technology and built upon the Document Management System, Documentum. This product was slow to use and not user-intuitive, with limited search capabilities and disordered relationships between products. It was also difficult and time-consuming for DK users to locate a particular piece of content or put in a request for it. A lack of controlled, reliable workflow processes and little auditing or tracking in place meant that users struggled to follow the progress of their content requests. Once requests had been put in, it could take up to a day for a product to be turned around.

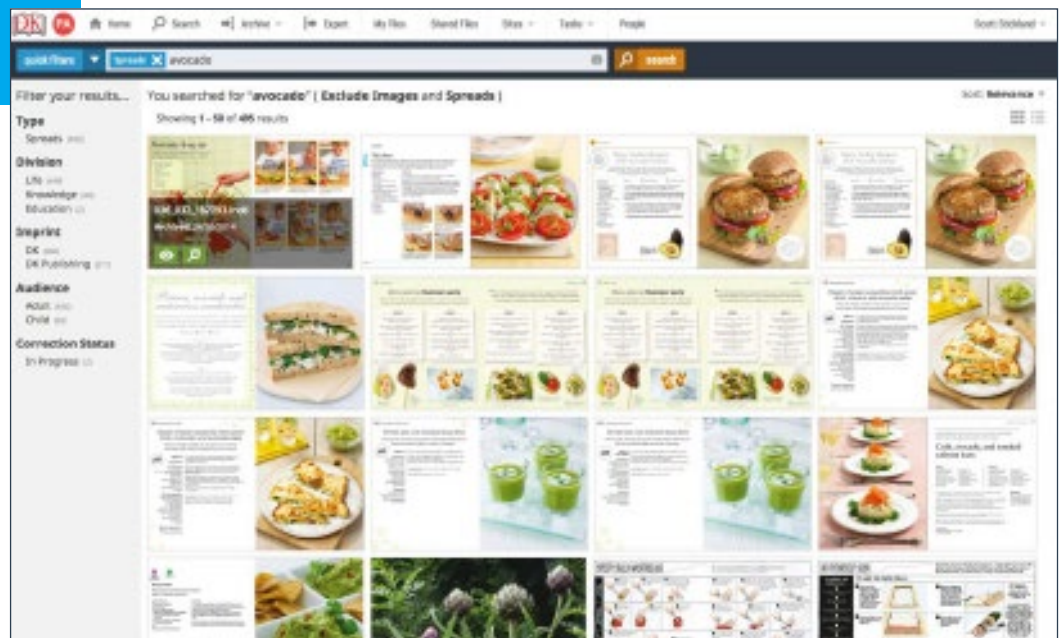
This difficulty in accessing DK's most valuable asset — its unique content — was leading to major inefficiencies in product assembly, which was impacting upon speed-to-market, cost-effectiveness and ability to fulfil consumer demand.

### Rapidly expanding digital output needs

Following the Penguin Random House (PRH) merger of 2013, DK decided to migrate all its content from Pearson Technology systems into its own product archive solution that could also support the organization's rapidly expanding digital output needs. With 90% of DK's current revenue being print-led, it was key that DK transformed its systems and processes to enable digitally-driven content. For a number of strategic reasons, it was essential for DK to have this change fully effected by the end of 2014, a significant challenge given their solution requirements and the large amounts of content that needed to be transitioned.

### 5000 unique products

As a partner working with DK since 2012, CCC was engaged to provide a phased approach for delivering this solution, which would ultimately bring together a single product archive of up to 5000 unique products, totalling around 70TB.





## Solution

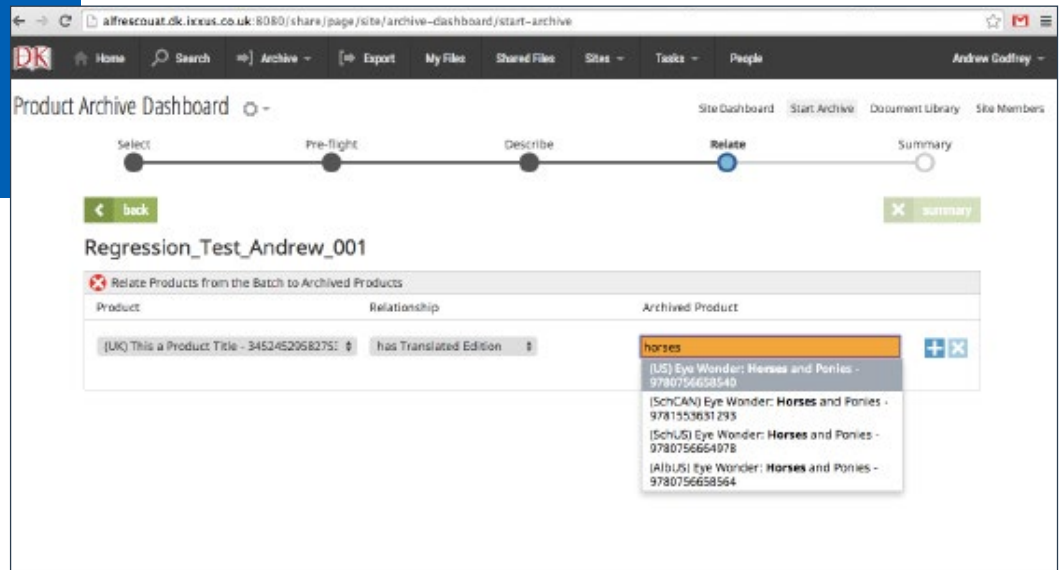
The DK Product Archive is a cloud-based custom application used to manage archived content (~5000 Products), allowing production and creative teams to effectively archive, ensure updates and corrections are stored with the source material and reuse content for new products, all from within a consistent user interface specifically tailored to DK requirements.

The solution is hosted on Amazon Web Services (AWS) within a secure Virtual Private Cloud (VPC). It also uses a set of key technologies including:

- **Alfresco's** web framework and Workflow engine for archive management, notifications, corrections, content requests and export.
- **ElasticSearch** as the primary search service.
- **OWLIM** as the semantic repository, combining RDF triple stores, inference engine and SPARQL query engine to form meaningful relationships between different products, their content and assets.
- **MongoDB** as a source of audit and user activity information.
- **Amazon's SQS** service as the robust message queue, managing communication between the technical components.
- **Amazon's S3** for scalable cloud storage of physical files (IDML files, etc)

It is also integrated into several PRH services including AD for user authentication, SMTP relay for email notifications and Local Network for file transfers.

CCC provided tooling to support a mass content migration into the new solution persisting DK's content structure and adding additional information to drive content discoverability and user experience



## Case Study

### Sleek & Visual User Interface

As a visually-led, high-design organization, DK required a very visual interface allowing at-a-glance access to images and full page spreads. By incorporating thumbnail display, multi-page layouts and easily navigable product ontologies, the solution allows a user to get high-level, image-driven views across DK's product archive.

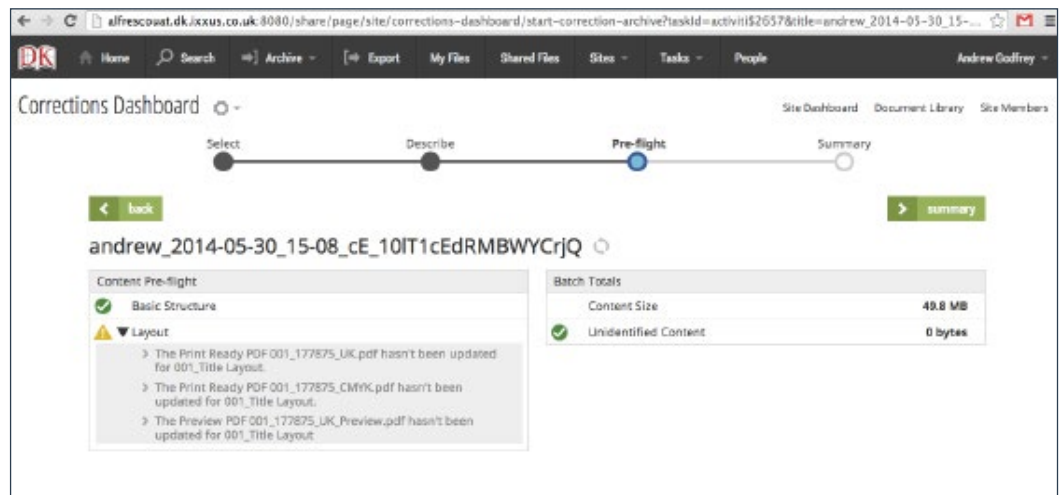
### Enhanced discoverability

With metadata enrichment, semantic services and a Big Data store, the platform delivers scalable discoverability and management of DK's print content, whilst enabling a foundation for future digital content. The award-winning enhanced faceted search expertise from CCC has been incorporated to deliver powerful search, including custom filters, facets and ISBN search, making it simple for DK users to surface the most relevant content to their needs and ensuring that full value is derived from each content asset.



### Full content lifecycle

The solution also allows control of the full renditions and corrections lifecycle; for example, highlighting whether a particular asset has been overused, or is only licensed for particular types of use, or has been translated into linguistic variants (e.g. US/UK editions). With an integrated Corrections dashboard, DK can minimize data quality issues and ensure that all publications meet the high standards which consumers expect from them.



## Case Study

These are some internal DK stakeholder comments during user acceptance testing:



"I've been working at DK for 8 years and this is the first thing that has made me go wow"

"Can I start using it today please?"

"It's like being let out of prison"

"It's fantastic. I can't wait to see more products"

"Don't you think it's amazing?"

"If you see me crying, it's tears of joy"

"This is genius"

### Image transformation for thumbnail preview

By integrating automated image transformation in the background, the platform renders large files such as .IDML, .AI, .EPS and .PSDs as smaller, more accessible PDFs, JPGs and PNGs to allow for in-platform thumbnail view. Users can view high quality thumbnail images and multi-page spread previews as part of search results, allowing them to make informed editorial and production decisions without having to first access large, multi-asset files. This accelerates product assembly and curation, driving editorial efficiencies and speeding time to market.

The screenshot shows the 'Legacy Import Dashboard' interface. It features a navigation bar with 'Home', 'Search', 'Archive', 'Export', 'My Files', 'Shared Files', 'Sites', 'Tasks', and 'People'. The main content area is divided into several sections:

- Legacy Import Job Queue:** A table with columns: Title, Started, Progress, ETA, Status, and Actions. It lists jobs like '178510\_Disney\_Fairies\_ULT\_GUIDE\_to...', '187216\_Little\_Hide\_and\_Seek\_Farm', '187501\_Pocket\_Eyewitness\_Science', and '195566\_Mindbenders'.
- Legacy Import Stats:** A summary table with columns: Status, Amount, and Actions. It shows counts for Available (17), Queued (4), Failed (0), and Successful (20).
- Legacy Titles Available For Import:** A table with columns: Title, Status, and Actions. It lists titles like '175432\_EWG\_Mythology', '175607\_Baby\_Bee\_Bee', '175749\_LEGO\_Star\_Wars\_Vis\_Dict', etc.
- Site Members:** A section showing '1 - 1 of 1' members, including 'Administrator Manager'.

### Results

Over a four phase project, CCC successfully delivered a fully functioning Product Archive solution as defined by DK's stakeholders. Each development phase including AWS infrastructure setup has been delivered on time and on budget, consistently augmenting the original scope with key improvements in the true sense of Agile development. All key stakeholders have been quick to compliment the intuitive user interface and functionality.

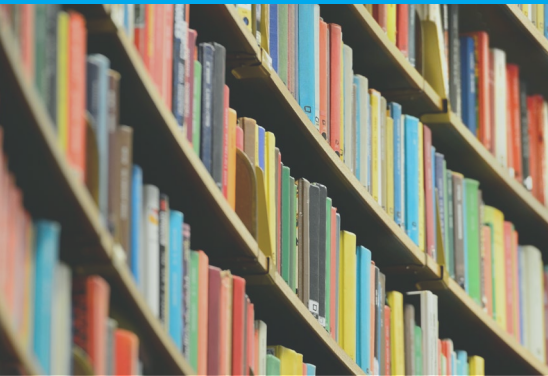
The Product Archive went live in September 2014 to a core set of DK users, and will be rolled out to the wider creative teams across the globe (a user-base of around 700) in early 2015. These are some internal DK stakeholder comments during user acceptance testing:



Not only did CCC deliver on time, they went above and beyond the initial scope of the project."

**Peter Buckley**

Technology Director at DK



## Benefits

### 1 Content Discoverability and Reuse

Enhancing the discoverability and accessibility of content and assets allows archived products to find new relevancy and new leases of life. Once discovered, legacy content can also be repurposed, remixed and reused to create entirely new product, opening up potential revenue opportunities.

### 2 Consistently Managed Workflow

All content processing in and out of the solution is wrapped with a consistent workflow definition which mirrors DK's processes, allowing a centralized overview for content-processing teams and enabling efficient handovers of information. The improved, highly discoverable content store has streamlined production, enabling major efficiency gains and an intuitive, user-driven experience.

### 3 Improved analytics capability

By building a foundation for future analytics and reporting, the platform will deliver critical business intelligence within 1-2 years, allowing DK to discover things like: how many products they have on X subject, how many versions exist of X publication, how many times has X image been used cross-publication, etc.

### 4 Cost and Control

Cost and control was a huge issue for DK, with little capability to fulfil requests when made; especially when the system was under heavy load. By deploying the entire architecture out to AWS managed services, DK now has direct control over scalability and development of the product store, allowing them to own their content like never before.



## Case Study

DK Publishing is world renowned for its distinctive, highly visual books that inform, inspire, and entertain readers of all ages. Publisher of New York Times bestselling LEGO® and Star Wars® non-fiction books, Smithsonian titles, the award-winning Eyewitness series for children and Eyewitness Travel Guides, in addition to a wide selection of other books for adults and children. BradyGames, Alpha (Idiot's Guides) and Rough Guides are also available from DK, a division of Penguin Group.

## The future

DK and CCC have recently signed a managed service agreement for the solution and are formulating plans to extend the current feature set, incorporate new content types and expand the solution's scope to include Work in Progress (WIP)/Authoring tools in order to achieve an end-to-end publishing and editorial solution.

Under the managed services agreement, CCC is providing application support, full systems management of the AWS environment and a proactive monitoring service.

### Digital transformation

DK and CCC are currently working together in partnership to set out an ongoing content roadmap and digital transformation strategy for the organization.

"With pressure to get content out of Documentum and under our own control following the PRH merger, time was of the essence," said Peter Buckley, Technology Director at DK, "Our users and stakeholders have felt involved in the process since day one, building a close working relationship with the CCC team — something which shows in their eagerness to get their hands on the finished product!"



### About CCC

A pioneer in voluntary collective licensing, Copyright Clearance Center (CCC) helps organizations integrate, access, and share information through licensing, content, software, and professional services. With expertise in copyright and information management, CCC and its subsidiary RightsDirect collaborate with stakeholders to design and deliver innovative information solutions that power decision-making by helping people integrate and navigate data sources and content assets.

© 2021 Copyright Clearance Center, Inc. All rights reserved.



### Learn more

Learn more about our licensing, content, and data solutions:

[copyright.com](https://copyright.com)

[solutions@copyright.com](mailto:solutions@copyright.com)