CASE STUDY

Copyright Clearance Center

RightsDirect
A Copyright Clearance Center Subsidiary

FAIR Data in Action

Partnering with CCC to Support Contextualized Scientific Discovery

FAIR data principles — making data Findable, Accessible, Interoperable and Reusable — are essential elements that allow organizations to maximize the value of their digital assets. As knowledge workers within life sciences organizations know, for every scientific breakthrough, there are countless hours spent sifting through vast amounts of scientific literature, analyzing volumes of data, and creating linkages at the intersections of information.

But without the ability to make connections between multiple internal and external sources of structured, semi-structured, and unstructured data, the process of forming all relevant information pieces into a completed puzzle can be nearly impossible.

At a global biopharmaceutical company headquartered in the U.K., the Scientific Information Services team saw an opportunity to further the organization's mission of delivering value to current and future patients by transforming how they make digital assets available to the broader R&D organization. They concluded that if integrated content and data were more discoverable through a unified experience, researchers would be able enrich the R&D pipeline with new ideas more efficiently — resulting in more scientific breakthroughs.

Read on to learn how this biopharmaceutical company and Copyright Clearance Center (CCC) set off to develop a best-in-class solution to facilitate knowledge discovery and insights. Using CCC's aggregated search and personalization capabilities as a base, this partnership resulted in a solution that will enhance the research process by applying FAIR principles to digital assets.

THE FAIR OPPORTUNITY

The company's Scientific Information Services team identified several core themes as part of a content analytics and enrichment strategy, building from content unification through an aggregation platform, to data enrichment and search, to exploration and data analytics.

What is FAIR?[1]

Findable, accessible, interoperable, reusable.

These are the guiding principles that together add up to a "FAIR" way to manage scientific data.

This concept is not only important for the pharma industry, but for digital healthcare as a whole. FAIR began in the academic community and has since been embraced by scientific funding agencies such as the European Commission's ERC (European Research Council), as well as the National Institutes of Health in the US.

- **Findable** means that the data should be able to be found by an appropriate person at an appropriate time.
- **Accessible** means that it's accessible either internally through a license or publicly available.
- **Interoperable** in terms of how the data is formatted. It's standardized and annotated.
- And **reusable** for both people and machines.

[1] https://www.nature.com/articles/sdata201618
https://www.go-fair.org/fair-principles/

These themes would enable the company to achieve key overarching objectives such as:

- Realize the full value of digital assets
- Offer the organization best-in-class discovery and insight-generation technology
- Support multiple discovery workflows
- Leverage digital rights and licenses

However, their researchers faced three main obstacles in achieving these objectives:

- **Fear of missing information:** For researchers to be able to trust their results, they needed to have access to all relevant information.
- **Lack of unified data:** Researchers needed a wide spectrum of data at their disposal, from internal items such as market research reports and meeting posters produced by internal researchers, to external information including grant funding information, meeting abstracts, and published full-text journal articles. But often, this information was siloed in different departments and available through disparate products, making it difficult to obtain.
- **Synthesizing information to discover insights:** When dealing with disparate content and data sources, it's no surprise that their exploratory research was overwhelming. Not only was the ability to access information crucial, but researchers also needed to be able to easily separate the signal from the noise by identifying outliers that could yield scientific insights. "We need to constantly explore the fringes of our mainstream assets," was how one researcher expressed this sentiment.

Simply put: Researchers needed access to all relevant information, from a wide spectrum of sources, and the data environment needed to support computable workflows and advanced analytics.

"I'm a believer that 90% of what I need is published. I just need to find it. I want to create the story and then have all the references to prove it."

THE SOLUTION

CCC's platform capabilities in contextualized discovery, enhanced by machine learning, enable individuals to find the right content at the right time. Building on this foundation, the company and CCC collaborated to define workflow features based on user need and requests. Additionally, the custom branding throughout the tool conveyed to researchers that they were using a solution unique to their organization.

Today, these are some of the key features they utilize:

- **Unified search experience:** Using a scalable integration framework of data source connectors, researchers now have a single access point that brings together more than 10 data types, comprising more than 120 million records.
- **Personalized experience:** Through both explicit and implicit personalization, the solution displays the data and content most meaningful to each user. User topic and data source preferences, search activity, sharing, purchasing history, and more are taken into consideration.
- **Dynamic data visualizations:** Now, researchers can explore connections between scientific concepts more easily with dynamic visualizations of semantically enriched data.

WHAT'S IN A DATA SOURCE INTEGRATION?

THE UNIFIED DATA SOURCES NOW COMPRISE:

3 MILLION +
GRANTS

350k+
CLINICAL TRIALS

34 MILLION +
PATENTS

100 MILLION +
JOURNAL ARTICLES & BOOK CHAPTERS

100k+
DRUG RECORDS

9k+
COMPANY RECORDS

7 MILLION +
MEETING ABSTRACTS

100k+
NEWS BRIEFS

1000
MARKET RESEARCH REPORTS

How Do Dynamic Data Visualizations Work?

Researchers have access to search across indexes of full-text content semantically enriched with biomedical vocabularies. They also have access to other content and data that has disparate or no enrichment applied at the source. How can the two be brought together?

In real time, CCC processes the results data by scanning it with the chosen vocabulary (MeSH), and then we map that to the enriched data to create a tree map visualization. Presently, this allows researchers to see the evolution of research over time and to explore broader connected research areas.

NEXT STEPS

The Scientific Information Services team is now better equipped to engage with and support its scientific researchers. With FAIR data principles in place, the value of the digital information assets the company licenses are being maximized — and ultimately knowledge workers are empowered with a more efficient research process.

The roadmap for the future is to leverage the data integration work in two ways. First, because the value of data increases as connections can be made across larger data sets, they plan to incorporate additional third-party licensed data sets as well as internal, proprietary materials that may include content such as competitive analyses and scientific poster presentations.

Second, analytics capabilities being explored at the company will potentially take advantage of the unified data sets to yield scientific insights. For example, using CCC's RESTful APIs, scientific researchers within other workflow applications can benefit from information retrieval and aggregated data generated across the integrated resources. Additionally, knowledge graph applications can be extended to address particular use cases.

Stay tuned.

Copyright Clearance Center

Copyright Clearance Center (CCC) builds unique solutions that connect content and rights in contextually relevant ways through software and professional services. CCC helps people navigate vast amounts of data to discover actionable insights, enabling people to innovate and make informed decisions. CCC, with its subsidiaries RightsDirect and Ixxus, collaborates with customers to advance how data and information is integrated, accessed, and shared while setting the standard for effective copyright solutions that accelerate knowledge and power innovation. CCC is headquartered in Danvers, Massachusetts, with offices across North America, Europe and Asia.

LEARN MORE

Maximize the ROI of your digital information investments. Contact CCC today:

info@copyright.com

www.copyright.com/rightfind-navigate

+1.978.750.8400 (option 3)