WHITEPAPER

Simplify and Accelerate Drug R&D

with the MarkLogic Data Hub Service for Pharma R&D

PRESENTED BY: PUBLISHED BY:

MarkLogic® | **FiercePharma**

CONTENTS

INTRODUCTION

Pharmaceutical research and development is complex and fraught with challenges. From drug discovery to clinical trials, throughout every stage of the R&D process, collaborating groups run the risk of duplicating—or even ignoring—the work of their peers in other parts of the organization. For example, biologists may be unaware of critical findings from toxicologists, and chemists may not know about assays already performed. Competitive information such as competitor performance and regulatory filings is also difficult to gather and analyze, slowing down investment decisions from the executive committees, such as approving a Phase III trial.

The technology root cause of this inefficiency is a patchwork infrastructure of disconnected data silos that restrict access to information, hamper collaboration, create data quality issues and drive up costs across all segments of the business.

Data silos create unnecessary roadblocks, leaving researchers in the dark to potential complications and gaps in their research. Researchers spend countless hours searching for information, both inside and outside of their organizations. But,

even when they do have access to information, it is often limited and grueling to synthesize as data of varying origins and formats is difficult— even impossible—to combine for analysis. For researchers, it's always a challenge to know if they have all of the information they need.

In addition to data sources, there are new and numerous changing ontologies. An ontology is helpful in classifying the concepts and their relationships used across pharma. The problem is, multiple ontologies are relevant to a single concept, and the "right" ontology for each data source changes over time. If researchers outside the organization use different terminology and methods of classification, it is difficult to find critical information about the genes, proteins, pathways, small molecules, papers, authors, drugs, and conditions being researched.

All of this results in unnecessary frustration, wasted time and exorbitant costs. Getting new drugs to market can now cost up to $2.6 billion and take over 10 years.* Without proper information access and management, drug companies will find it difficult to accelerate results and lower development costs.

Bringing a new drug to market can now require up to

$2.6 billion and
10-15 years

*Source: IFPMA, Facts and Figures 2017

PUBLISHED BY:
FiercePharma | ■■MarkLogic®

A BETTER WAY TO CONDUCT RESEARCH

What if your research teams had a single platform for sharing and searching for information? MarkLogic's Data Hub Service for pharma R&D provides exactly that, delivering a better way to share, access and synthesize data across teams.

The Data Hub Service for pharma R&D (the hub) is a single pane of glass that enables researchers to load their teams' key information and combine it with public knowledge bases to allow easy access to the widest possible array of information. Whether you load R&D data, competitive information, public or licensed third-party data sets, or internal proprietary data from assays to small molecules or trip reports—the hub allows your team to conduct research with confidence and accelerate innovation.

Using advanced search capabilities, users can leverage massive data assets—including decades of research and clinical trial data—faster and more efficiently than with data lakes or other custom-developed IT solutions.

As users push deeper into their search, they can gather, annotate and package related pieces of information, which they can then securely share with colleagues, both inside and outside their organizations. When new, relevant information is added, users are immediately notified via real-time alerts. The hub can easily export this newly linked data for enhanced machine learning and AI processing, allowing your company to accelerate its entire R&D pipeline.

With the hub's innovative approach and agile, next-generation data technology, data challenges no longer hinder drug discovery and development.

> "Using advanced search capabilities, users can leverage massive data assets—including decades of research and clinical trial data—faster and more efficiently than with data lakes or other custom-developed IT solutions."

BUILT ON A TRUSTED ENTERPRISE PLATFORM

The technology is built on MarkLogic's Data Hub Platform and runs as a cloud solution. The whole platform is backed by the MarkLogic multi-model database, which has been confidently used by enterprises for nearly two decades. The platform is more agile, governed and secure than relying on a data lake, which many organizations spend years building with disappointing results. Through the hub, your organization gets immediate value, and MarkLogic's expertise in managing cloud infrastructure means your IT department is free to focus more energy on solving your business challenges.
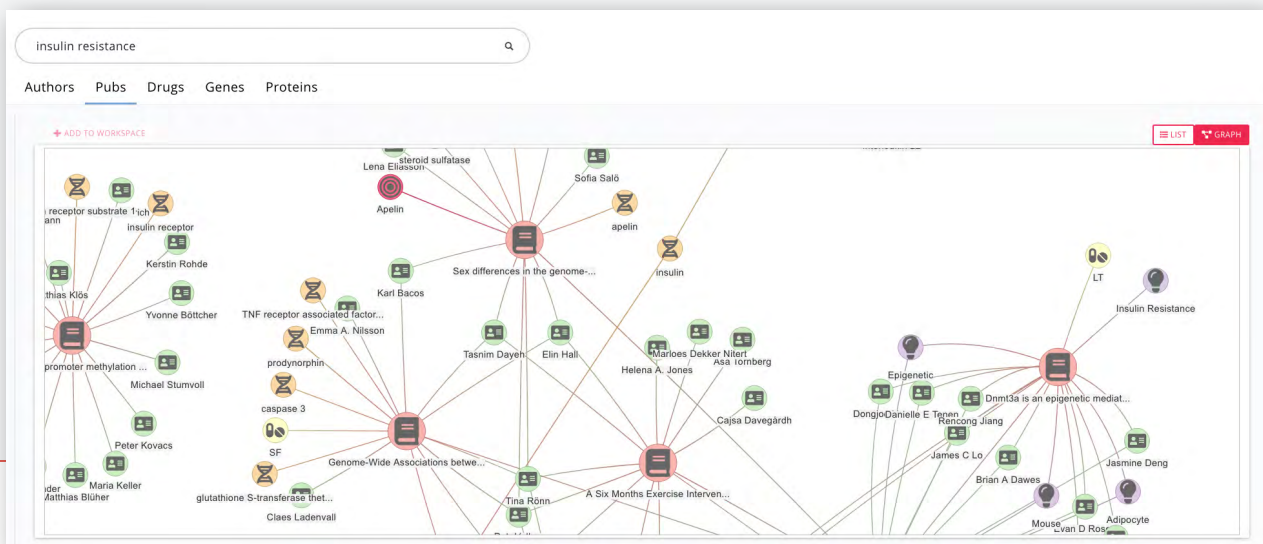
KEY FEATURES

The hub deploys in minutes, allowing any research team to immediately gain access to any pharma data set, uncover hidden relationships between data sets, and discover potential collaborators for their work. With data barriers out of the way, pharma teams are freed up to concentrate their efforts more fully on the research.

Quickly Load Any Pharma Data Set

The platform easily scales so that groups can start with their most important data—including publications, authors, drugs, compounds, genes etc.—make it discoverable, and then view it within context of data provided by other groups to drive collaboration. Both structured database information and unstructured publication and report data is managed and can be linked to a person, drug, or compound, etc. based on user preference.

Visualize Relationships

With the hub's underlying multi-model database, users can visualize any existing relationships between data and even discover new relationships. Users can view, navigate and search the robust graph of connections to see the structured data that is related to each entity and view how researchers are connected to institutions, publications and other peers.

Customize Search Results

The hub's advanced search capabilities allow users to customize results quickly and easily, such as adjusting facets to drill down into particular search results. Users can perform these customizations immediately without submitting an IT ticket, allowing them to accelerate results and tailor them to their unique specifications.

MarkLogic PHARMA RESEARCH HUB

💼 ▾ Likely experts for insulin resistance excercise

insulin resistance 🔍

Authors **Pubs** Drugs Genes Proteins

CURRENT FILTERS
🔖 ☑ Author: Jun Wang

REVISED DATE
Last Year
2 Years Ago
3 Years Ago
4 Years Ago
6 Years Ago
7 Years Ago
0 1 2 3 4 5

AUTHOR
☑ Jun Wang (15)
☐ Hui Li (2)
☐ Kerrin S. Small (2)
☐ Yingrui Li (2)
☐ A. Katrina Loomis (1)
☐ A. Reghan Foley (1)
☐ Aarno Palotie (1)
☐ Aaron Day-Williams (1)
☐ Adam S. Butterworth (1)
☐ Adam Shaw (1)

CITATIONS
☐ 201+ (1)
☐ 101-200 (2)
☐ 51-100 (2)
☐ 26-50 (3)
☐ 0 (7)

TOPICS
☐ Brain tumor stem cells (1)
☐ Cancer stem cells (1)
☐ Colon adenocarcinoma (1)
☐ EGFR (1)
☐ Gene expression signature (1)
☐ Glioblastoma multiforme (1)
☐ Immune system regulation (1)
☐ NSCLC (1)

+ ADD TO WORKSPACE
Found 15 results in 0.431 seconds.

📋 LIST GRAPH

📖 **Involvement of Insulin Signaling Disturbances in Bisphenol A-Induced Alzheimer's Disease-like Neurotoxicity**

2017-08-08

Bisphenol A (BPA), a member of the environmental endocrine disruptors (EDCs), has recently received increased attention because of its effects on brain insulin resistance. Available data have indicated that brain insulin resistance may contribute to neurodegenerative diseases. However, the associated mechanisms that underlie BPA-induced brain-related outcomes remain largely unknown. In the present study, we identified significant insulin signaling disturbances in the SH-SY5Y cell line that were mediated by BPA, including the inhibition of physiological p-IR Tyr1355 tyrosine, p-IRS1 tyrosine 896, p-AKT serine 473 and p-GSK3α/β serine 21/9 phosphorylation, as well as the enhancement of IRS1 Ser307 phosphorylation; these effects were clearly attenuated by insulin and rosiglitazone. Intriguingly, Alzheimer's disease (AD)-associated pathological proteins, such as BACE-1, APP, β-CTF, α-CTF, Aβ 1–42 and phosphorylated tau proteins (S199, S396, T205, S214 and S404), were substantially increased after BPA exposure, and these effects were abrogated by insulin and rosiglitazone treatment; these findings underscore the specific roles of insulin signaling in BPA-mediated AD-like neurotoxicity. Thus, an understanding of the regulation of insulin signaling may provide novel insights into potential therapeutic targets for BPA-mediated AD-like neurotoxicity.

...recently received increased attention because of its effects on brain insulin resistance Available data have indicated that brain insulin resistance may...

📖 **The Noncoding RNA Expression Profile and the Effect of lncRNA AK126698 on Cisplatin Resistance in Non-Small-Cell Lung Cancer Cell**

2013-05-05

Cisplatin resistance in non-small-cell lung cancer cells may relate to the changes in noncoding RNAs. Among these, AK126698 appears to confer cisplatin resistance by targeting the Wnt pathway.

...small- cell lung cancer is limited by the acquired drug resistance Identification the RNAs related to the cisplatin resistance may help...

📖 **The goose genome sequence leads to insights into the evolution of waterfowl and susceptibility to fatty liver**

2015-05-05

The online version of this article (doi:10.1186/s13059-015-0652-y) contains supplementary material, which is available to authorized users.

...first domesticated poultry. Geese are capable of rapid growth, disease resistance and high liver lipid storage capacity, and can be easily...

📖 **Regulatory Roles of Non-Coding RNAs in Colorectal Cancer**

2015-08-08

Non-coding RNAs (ncRNAs) have recently gained attention because of their involvement in different biological processes. An increasing number of studies have demonstrated that mutations or abnormal expression of ncRNAs are closely associated with various diseases including cancer. The present review is a comprehensive examination of the aberrant regulation of ncRNAs in colorectal cancer (CRC) and a summary of the current findings on ncRNAs, including long ncRNAs, microRNAs, small interfering RNAs, small nucleolar RNAs, small nuclear RNAs, Piwi-interacting RNAs, and circular RNAs. These ncRNAs might become novel biomarkers and targets as well as potential therapeutic tools for the treatment of CRC in the near future and this review may provide important clues for further research on CRC and for the selection of effective therapeutic targets.

Topics

Instantly Receive Notifications of New Information

Users can designate custom notifications and alerts when information is updated in the hub or when new information is added. Any search query can be turned into an alert, so users are consistently informed of new information and updates that impact their research.

Save and Share Workspaces

Users can aggregate search results, such as key publications, experts, compounds and more, in one place for future reference and sharing with colleagues. As users gather information and add it to a workspace, they can easily return to this information, view their history and see collections that others have shared with them. When key employees leave the company, their workspaces and collections stay behind, remaining discoverable and relevant.

Exercise Epigenetics Experts

DESCRIPTION

This workspace gathers information on the relationship between exercise and epigenetics, with a focus on exercise-related impacts on DNA methylation.

AUTHORS (3)

Gerwyn Morris

Tir Na Nog, Bryn Road seaside 87, Llanelli, Cardiff, Wales SA152LW UK

Autism Bipolar disorder Schizophrenia Chronic fatigue syndrome Cytokines Depression Immune dysfunction Inflammatory Mitochondrial dysfunction Multiple sclerosis Nitric oxide Oxidative stress Parkinson's disease Peroxynitrite Psychiatry Neurology

AUTHOR

Alfonso Varela-López

Department of Physiology, Institute of Nutrition and Food Technology "José Mataix", Biomedical Research Center (CIBM), University of Granada, Avda. del Conocimiento s.n., Armilla, Granada 18100, Spain;

aging antioxidants diet mitochondria nutrition oxidative stress ubiquinone

AUTHOR

Abdullah Aldehaiman

(E.D.-G.)

dimerization miRNA motility anoikis chaperon PTEN FIP200 LKB1 PI3K regulation

AUTHOR

PUBLICATIONS (1)

Coenzyme Q and Its Role in the Dietary Therapy against Aging

Molecules; 2016-3-18

Coenzyme Q (CoQ) is a naturally occurring molecule located in the hydrophobic domain of the phospholipid bilayer of all biological membranes. Shortly after being discovered, it was recognized as an essential electron

PUBLICATION

Improve Results with AI and Machine Learning

MarkLogic's Smart Mastering feature delivers better search results on higher-quality data. Built-in machine learning helps users find the most relevant data and recommends additional and better search queries to bridge knowledge gaps. Users can also apply quality rules to data exports sent to bioinformatics and AI systems outside the hub.

HOW IT WORKS

Samantha is a researcher at a large pharmaceutical company. She has been tasked with finding more information about a particular gene of interest. The company needs to know what has been previously published about the gene, what molecules or drugs are already being worked on that target proteins related to that gene, who within the research community has studied it, whether they can collaborate with those researchers, which other companies are also researching this gene, and whether there are any known patents.

Samantha logs into the hub and performs a general search for the gene. She receives a list of results and uses custom search filters to narrow the information down to a specific date range and list of key publications.

She finds a relevant article written by a research scientist at a university and saves the article to her workspace for future reference. The knowledge graph displays the author's connections to several other researchers whom she adds to her workspace.

Samantha then dives deeper into the connections to the university to determine whether they have published more articles on the subject or filed relevant patents (which often give a preview of future activities not yet made public). After adding a few more results, her workspace includes about three patents, 10 articles, and 20 research scientists. She notices that two researchers were recently hired by a competing company, so she removes them from her workspace.

She then shares the work package with a colleague who has studied a related topic. Her colleague further refines the list of results and adds another researcher who is working on a similar topic, before sending the package back to Samantha.

What's the value of this work?

In a matter of minutes, Samantha has completed work that would ordinarily take days and has taken a proactive approach to getting new information when it becomes available. She has compiled a thorough and highly-relevant set of people, institutions and publications to help decide who to work with, recruit or track as competitors. Samantha can now focus her efforts on the most important results from her query. She turns the query into an alert so she will be notified of any new information on the subject as her research progresses.

> " In a matter of minutes, Samantha has completed work that would ordinarily take days and has taken a proactive approach to getting new information when it becomes available. "

PUBLISHED BY: FiercePharma | MarkLogic®

ACCESS THE RIGHT INFORMATION WITHOUT IT BURDEN

With the Data Hub Service for pharma R&D, your IT department will be completely relieved of the burden of data integration complexity. The Hub is agile and secure, and it frees organizations from the expense and hassle of building other custom IT solutions that may be orphaned or abandoned when key personnel leave or as business needs change.

The technology is deployed using the MarkLogic Data Hub Service, which allows it to run in the cloud without purchasing servers or standing up a software development effort. Under the hood is a multi-model database that allows many data models and types to be handled natively, queried, searched and run at large scale without IT intervention. MarkLogic's Data Hubs are reliable, highly available (HA) and trusted by thousands of organizations around the world to provide 100% data consistency and availability.

With the hub, drug companies can bring their data to the cloud in minutes with no infrastructure to buy or manage—freeing up their IT departments to focus on other concerns. And, the hub provides a level of security and scalability unmatched by other solutions.

STREAMLINE DRUG R&D PROCESSES

MarkLogic®

Pharma R&D is complex. Accessing the data you need should not be. Simplify your R&D cycle with MarkLogic's Data Hub Service for pharma R&D. Within minutes, your research team can gain access to the widest possible array of R&D data available and have relevant information right at their fingertips.

When your researchers are empowered to:

- Connect all data to discover hidden relationships and new avenues for research
- Leverage massive data sets up to 10 times faster than with custom-developed solutions
- Quickly and securely find, synthesize, and share information
- Stay current on new research as it becomes available

Then, your organization can:

- Move life-changing drugs to market faster
- Reduce drug trial abandonment
- Improve drug quality and safety
- Lower drug development costs
- Decrease risk and regulatory compliance burden

ABOUT MARKLOGIC

Data integration is one of the most complex IT challenges, and our mission is to simplify it. The MarkLogic Data Hub is a highly differentiated data platform that eliminates friction at every step of the data integration process, enabling organizations to achieve a 360° view faster than ever. By simplifying data integration, MarkLogic helps organizations gain agility, lower IT costs, and safely share their data. Headquartered in Silicon Valley, MarkLogic has offices throughout the U.S., Europe, Asia, and Australia.

Visit **www.marklogic.com** to learn more.