

4 Benefits of Organizing HR Data Consumption Around a Data Hub



Introduction

Human Resources (HR) data is generally managed within a dedicated system for the standard functions. HR managers need to have centralized information in order to develop skills while giving due consideration to each employee's personal situation and path within the company. However, they must also allow part of this data to be used in order to produce payroll, manage absences and assess the company's operating performance, while ensuring the confidentiality of sensitive data.

By organizing HR data around a data hub, you can provide – in a straightforward manner – the data your systems need to run without changing your existing data architecture, and ensure the all-important security of your HR data at the same time.

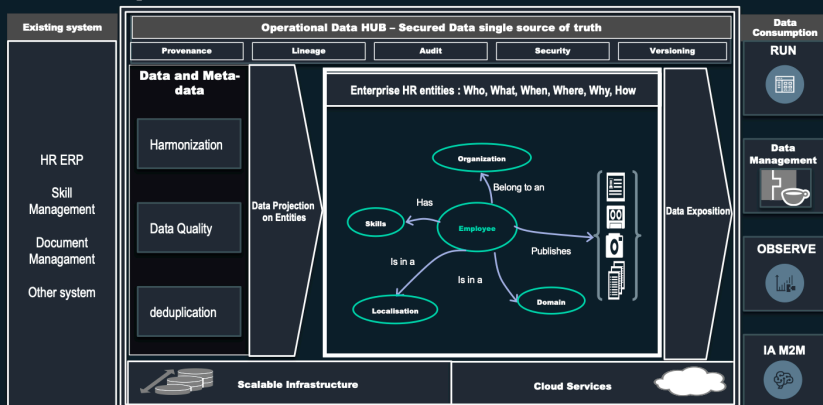
Agile Data Integration: The MarkLogic Solution

A data hub for HR data is a platform that allows Data Architects and Engineers to accelerate integration of all existing and new data types and models, and securely share the curated data needed to improve information discovery, insights, and applications development. Well-curated and governed data is foundational for successfully advancing the development of HR applications.

The MarkLogic Data Hub platform, powered by the most flexible and trusted NoSQL database in the market, includes a comprehensive set of capabilities to integrate, curate, and access all of your enterprise data and run operational and analytical applications at scale.

The data hub does not replace the company's data architecture; it integrates with the existing architecture to make it more flexible, by making optimal use of the various existing systems.

Operational Data Hub : A robust Architecture



Benefits of MarkLogic's Data Hub for HR data

#1 – Manage all data without any limitations thanks to a flexible data model

Most data integration approaches are based on the development of a single global model and the mapping (via ETL) of all the source data in the new model. This approach leads to a rigid, predefined organization in which any subsequent change proves costly.

MarkLogic's approach is different: the underlying principle is to describe a scalable data model, independent of the various systems, where each data item is characterized according to its context of use.



By separating the description of the data from its context of use, you can reconcile different models and upgrade the model in line with application needs, without undermining the initial organization.

This model is based on the W3C standard for representing concepts via RDF triples to describe the objects with their relations for each data use case.

Using semantics and semantic RDF triples to represent relations

The Resource Description Framework (RDF) is a set of W3C standards for representing machine-understandable concepts on objects, and the relations between these objects. It also forms the basis of the Semantic Web concept. The representation unit is called a triple, consisting of a subject, a predicate and an object, globally representing a fact/concept or a relationship (e.g. “Euro type Of Currency” or “Thor son Of Odin”). In a data hub, RDF triples provide a myriad of possibilities for managing data and the complexities of data integration.

This concept is very important since it serves to model the business data. As opposed to relational databases which are used to model physical data, the business data model, created using this semantic functionality is used to expose entities (employee, contract, leave, assignment, etc.) and their relations.

Each entity can be enhanced (by adding attributes, links, a contextual description, etc.) in relation to the businesses. These enhancements do not break the model; what worked prior to the enhancement will continue to work afterwards. This all-important characteristic of the data hub means that it can be used to iteratively respond to different business needs and upgrade the model in order to accommodate existing and future needs.



#2 – Integrate data and the data model in a data hub

Data integration traditionally separates data management solutions according to their use:

- Relational databases for processing and management;
- Data warehouses or datamarts for regular analysis and reporting.

Any change to the system or any new analysis requirement leads to long, often complex updates to the overall model, and which have an impact on the existing systems.

Running counter to this approach, MarkLogic's Data Hub for HR data compiles – in the same repository – the data and its representation, which can then be used directly for all purposes.

The HR data is immediately available for any new use, without calling into question the model's organization. The operational data provides a base for the real-time analysis of your data.

Your management systems can be upgraded simply and independently of the data hub's organization. Whatever the functional changes your businesses require for their applications, or the technological upgrades of the management systems, your data repository remains coherent and available for any new use, without the need for long and costly upgrades to the data model.

Rapid implementation

Considering that its data model is independent of the source data models, you can implement a data hub rapidly, without having to first design all the use cases to standardize all the company's HR data. Accordingly, in the space of 3–4 weeks, you can build a simple solution tailored to a given use case, while creating the foundations to use the repository for any other project.

With each new use case, you enhance the model by defining the necessary elements, without undermining the models defined for the previous use.

This brings all the flexibility and agility required by modern organizations.

#3 – Improved data governance

More than ever before, increasingly stringent personal data regulations, such as GDPR require businesses to put additional data governance measures in place to ensure the source of all data.

MarkLogic's data hub does not replace the original data models - it enhances them, and retains trace and record of all data movements.

All the original source data is thus preserved (which is essential for governance) without the need for prior decisions as to which data source is important or not.

Consequently, with data provenance and lineage, you can guarantee the origin and the source of all data, and retain the trace of all its transformations.

Whatever the data's use, you keep a trace of the transformations made to the data, from its sources through to its consumption, to ensure that the transformation has not denatured the source data.

You can audit the data from its source to its consumption.

#4 – A powerful security model

Data about people is very sensitive and requires control over access. While management applications must be able to utilize the useful data in accordance with functional needs, certain analysis must be possible on anonymized data that protects confidentiality.

The data hub model serves to define access rules in line with the data consumption context.

You can:

- Retain the source data security model;
- Define a security model tailored to each use case, and notably define a different access model for the consumption of consolidated data;
- Define transformation rules that render certain information inaccessible, if necessary.

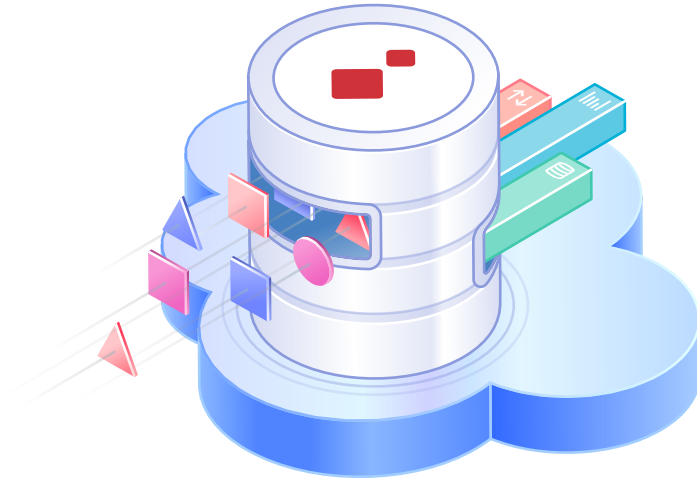
You thus guarantee the confidentiality of your sensitive HR data.

A central hub for a new form of HR intelligence

The MarkLogic Data Hub is flexible - it easily integrates all data formats, structured and unstructured, and real data.

By providing the most flexible approach for ingesting and accessing multi-structured data, the data hub constitutes a unique base for data utilization and analysis.

It can thus provide unified and reliable data for use in AI engines, among others. As such, it forms the reference base that serves to build the predictive models needed to anticipate the company's HR changes.



Conclusion

The data hub for HR data offered by MarkLogic provides a simple, flexible solution that can be easily adapted to any data architecture used by enterprises. Thanks to an independent, scalable model, it provides the framework you need to utilize your HR data in complete security, without the need for long and costly development cycles.

You can use it to set up an agile data architecture in which your HR data is available on request or in self-service mode.

About MarkLogic

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.

999 Skyway Road, Suite 200
San Carlos, CA 94070

+1 650 655 2300
+1 877 992 8885
www.marklogic.com
sales@marklogic.com

© 2021 MarkLogic Corporation.

MarkLogic and the MarkLogic logo are trademarks or registered trademarks of MarkLogic Corporation in the United States and other countries. All other trademarks are the property of their respective owners.

 MarkLogic®