The Evolution of Metadata Management

REPORT

As a data architect, you probably already understand the importance of metadata management. By defining and leveraging information about the data itself (descriptions, definitions, relationships, etc.) you can achieve better data quality, governance, unity, reliability, and security, among other benefits.

But in 2021 Gartner pronounced that traditional metadata practices were now "insufficient" to meet the needs of the enterprise. In its Market Guide for Active Metadata Management, it heralded the transformational capabilities of "active metadata." This evolved type of metadata powers some of the most innovative data architectures, from data fabrics and digital twins to semantic knowledge graphs, and enables powerful new business use cases.

Read on for an overview of this metadata evolution.

Passive vs Active Metadata

To grasp the concept of active metadata, it helps to contrast it with traditional, passive metadata. If your data is a digital file, then the (passive) metadata is the author, file type, date created or modified, and so on. While these types of metadata fields are basic and static, they still provide some elements of context and meaning to data.

For example, at a life sciences company, standardized metadata fields for clinical trials (e.g., agent dose, agent administered time, disease recurrence type, etc.) help ensure that data is consistent, shareable, and compliant with regulations. This gives pharmaceutical team members greater transparency into data, facilitating better communication, deeper insights, and faster results.

What Is Active Metadata?

What makes metadata "active" rather than static or passive? It changes as the data changes. More specifically, it immediately reflects a change in state, its use, or its intended audience of its associated data as it moves through the information supply chain. Active metadata creates new meaning for existing data, making it more insightful and actionable to improve business operations.

Think of this metadata as a backpack attached to your data, which travels everywhere your data goes. Inside that backpack is everything you know about the data – for example, its meaning, its relationships with other data elements, its process journey, its governance record of how it changed over time – which means that as data flows through the supply chain, you're continually adding new items to the backpack. All of this is then available for use in analytics, process management, and for contextual applications.

Here's a summary of metadata's evolution from passive to semantic, active, and augmented:

Metadata Type	Description	Examples
Passive metadata	Information about data	System-applied dates, creator, source
Semantic metadata	Describes the meaning of data	Topics, product, geography, audience, concepts, and relationships
Active and augmented metadata	Intelligent and dynamic data	Facts, status, PII, protective and retentio markers, data orchestration, ML analytics

To take advantage of active metadata, you typically deploy machine learning (ML) and/or business oriented AI algorithms to automate metadata management. An AI/ML-powered data platform or tool can continuously collect, create and analyze metadata, then trigger automated actions or proactive recommendations such as:

- Detecting missing or incorrect data, then correcting the issue
- Protecting or masking data when it contains personal information
- Adding context to dashboards to help leaders make more informed decisions

Most architectures have metadata management (active or otherwise) as a distinct layer in the data stack. But this approach is sub-optimal. Metadata IS data (a different type of data, but data nonetheless) and it should be addressable as data from a single data platform without compromise. Separate layers suggest custom integration between different vendor products with all the consequential risk of versioning, compatibility and loss of capability.

In summary, effective metadata management tightly coupled with its associated data eliminates the effect of both data and knowledge silos, delivers data agility and accelerates insightful business decision making.

Metadata Challenges for the Enterprise

Here's what's holding back most organizations from leveraging the full power of their metadata:

- Their metadata is passive only, with rigid structures and schema that is rarely well maintained.
- Users input metadata manually—a time-consuming process that restricts the size of the metadata set, limits the ability to add or change metadata in real time. Or, metadata is system-generated and rife with errors, as there is no applied context.
- Metadata is stored separately from the data itself, eliminating the possibility of actively updating it to reflect how its related data changes, how it was used, and prevents joined up analytics.
- Integrating the metadata layer requires a third-party tool, adding further complexity to the system.

These limitations severely hamper an organization's data agility—defined as the ability to quickly and easily make changes to how any information is interpreted and acted upon, at any scale, for any purpose.

Put Metadata at the Center of Your Strategy

How can you achieve data agility? By putting metadata at the center of your data strategy. That means connecting active data, active metadata, and active meaning; and keeping data and everything that is known about it together at all times.

That's the power of the Progress® MarkLogic® data management and Progress® Semaphore™ semantic AI platforms.

- Couple data with its metadata to establish a single data resource
- Create and manage active metadata, allowing you to integrate, store, manage, govern, contextualize, and surface data regardless of format, schema, or type
- Employ semantic AI to synthesize, enrich, extract, and harmonize all types of metadata
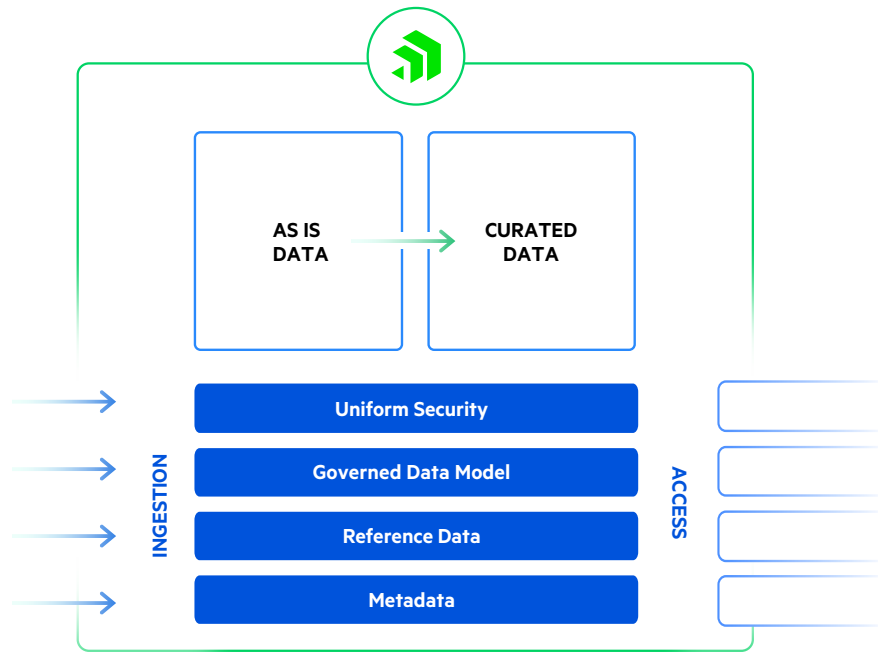
Unlock the Power of Metadata.
Request a demo at progress.com

About Progress

Dedicated to propelling business forward in a technology-driven world, Progress (NASDAQ: PRGS) helps businesses drive faster cycles of innovation, fuel momentum and accelerate their path to success. As the trusted provider of the best products to develop, deploy and manage high-impact applications, Progress enables customers to build the applications and experiences they need, deploy where and how they want and manage it all safely and securely. Hundreds of thousands of enterprises, including 1,700 software companies and 3.5 million developers, depend on Progress to achieve their goals—with confidence. Learn more at www.progress.com

Worldwide Headquarters

Progress Software Corporation
15 Wayside Rd, Suite 400, Burlington, MA01803, USA
Tel: +1-800-477-6473

f   facebook.com/progresssw
🐦 twitter.com/progresssw
▶ youtube.com/progresssw
in linkedin.com/company/progress-software
○ progress_sw_