Progress®

Data Agility with a Semantic Knowledge Graph

WHITEPAPER

Data Agility with a Semantic Knowledge Graph

Organizations everywhere are investing heavily in changing data into knowledge, and insight into action. They are looking for better ways to encode organizational knowledge, and use it throughout their organization. A *semantic knowledge graph* (SKG) is a data structure that describes entities, their relationships and their meanings as well as the data itself.

When coupled with the active data used to create that knowledge, the result is *data agility*: the ability to quickly and easily make changes to how information is interpreted and acted upon. Leaders use data agility to change the way they work with information: **keeping data and everything that is known about it together at all times**.

They use this active, *reusable organizational knowledge* to achieve a variety of strategic objectives in any context where considerable human knowledge is required to evaluate information: research, operations, customer service, intelligence, complex manufacturing, compliance and more.

From an architectural perspective, ingested data flows from producer to consumer. The semantic knowledge graph is used to connect new data with current knowledge, taking action as necessary. Knowledge workers create new insights, and make them available to others as needed. Different parts of the organization consume the SKG and the data that created it in a variety of ways: whether that be using *informed search, contextual applications or grounded analytics*

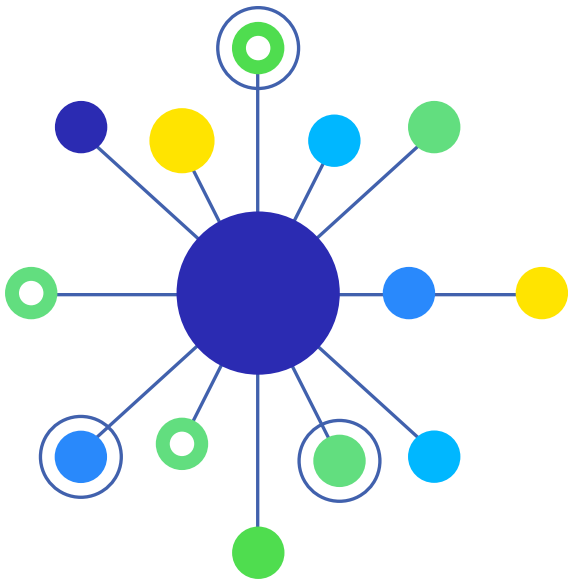


Figure 1: Data Agility: keeping data and everything that is known about it together at all times.

The Value of Data Agility

Many organizations invest in capturing and codifying organizational knowledge in an effort to improve a wide range of business processes that benefit from additional context.

Examples include optimizing global logistics in life sciences, corporate M&A data harmonization, and analyzing field intelligence in government security. The goal is to make a single set of coordinated investments that provide broad and *sustainable competitive advantage* as compared to their peers.

Advanced practitioners will point to *data agility* as the most important result of their investments. Data agility is the ability to make *simple and powerful changes to how any information is interpreted and acted upon*. Data agility creates active organizational knowledge, which delivers a rich and current context wherever important decisions are made.

To achieve this, the familiar model of separated data and metadata must change: *data and everything we know about it must be kept together at all times*. Much like packages separated from labels creates a problem, so does data separated from everything we know about it.

When a rich semantic knowledge graph is coupled with the data that created it, new advantages appear in any situation where improved human decisions make a significant impact:

- Research activities: biotherapies, industrial, intelligence, data science and more.
- Process coordination activities: manufacturing, logistics, human resources and more.
- Client-facing activities: knowledge-enhanced products and services, portals, high-value clients and more.

When data, metadata and meaning are connected and active, more benefits result:

- New information can be connected and interpreted in current grounded organizational knowledge, with appropriate action taken, e.g. upon receipt of an email from upset client, what does it mean and what should be done?
- Existing information can be interpreted in new ways, creating new insights that are available to all.
- Information can now be consumed in current context, with feedback loops as appropriate.

Informed search, contextual applications and grounded analytics all benefit from this innovative approach to complex data.

The Value of Informed Search

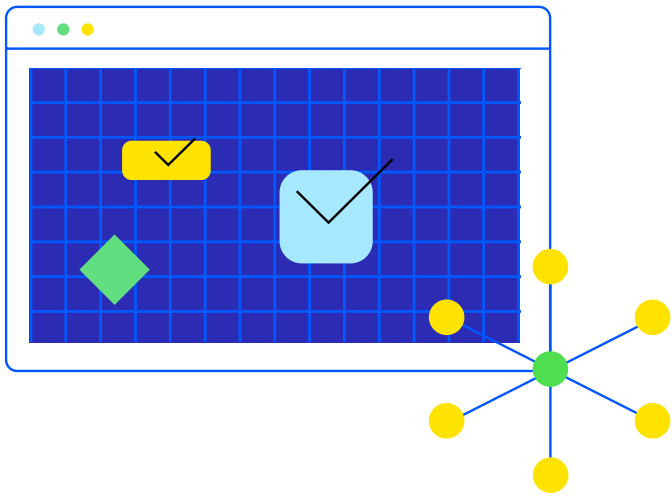
Familiar forms of search quickly run out of informed context: why are you searching for something, and what existing knowledge can be brought to bear? Life sciences may want a context of proteins: what is known about them and how they interact. Financial services may want context of complex financial products, associated risks and how they relate. Intelligence agencies want to search in a context of known actors and actions. Indeed, searching for information within a specialized context is seen everywhere.

Focused search means wanting to quickly connect new scraps of information to current knowledge, document those connections, infer new patterns, and continue. To the extent that "current knowledge" is codified and consumable, their work is greatly accelerated and their insights more quickly and broadly applied.

Encoding that knowledge in a consumable way is the role of the SKG and its data. Any new data artifact can be quickly filtered against what is already known. When new connections and insights are discovered, they are easily coded and thus shareable. Existing knowledge can be combed through, looking for new connections and insights. And when supporting downstream uses of research, the insights and the data that created them are documented and verifiable, hence trusted.

By connecting active data, active metadata and active meaning, learning happens faster, and is put to work faster.

The value of informed search using an active SKG with connected data has been well-documented in a broad range of industries: bioscience, financial services, logistics, manufacturing, intelligence and fraud, and more.



The Value of Contextual Applications

Many applications are intended to help people make informed decisions as part of a workflow: approving claims, coordinating logistics, assessing risk, providing superior client experiences and similar. A contextual application uses the SKG and the data that created it to help application users deliver decisions informed by current context.

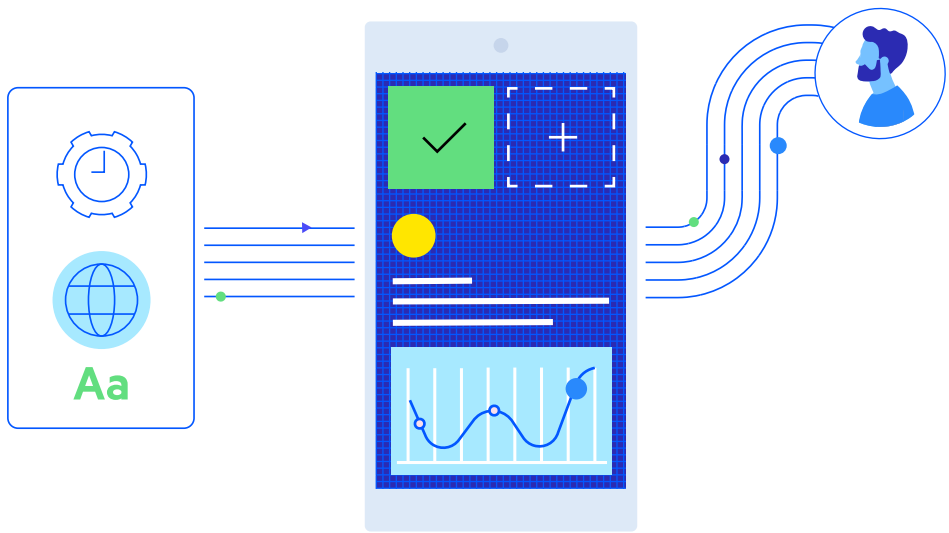
For insurance claims, this is being able to view a given claim in the context of everything that is known about claims, what things are being claimed as losses, what do we know about the claimant, what do we know from past experiences, and more. For client services, the person providing the service has immediate and current knowledge of the client, their current relationship and likely needs, the potential offerings that might fulfill those needs, likely concerns, and so on.

As a result, the person making the decision has an enormous advantage over an application user that is less informed. When done using an active SKG coupled with data, decision context is always current, and outcomes can easily be used to direct better decision making, as data plus derived knowledge is easily auditable and revisable.

As a result, the organization quickly learns to make better decisions at the point of engagement: the user application.

By connecting active data, active metadata and active meaning, it can do so faster.

The value of contextual applications using an active SKG with connected data has been well documented in a broad range of industries: financial services, healthcare, public sector, manufacturing, logistics and many more.



The Value of Grounded Analytics

Many organizations are substantially increasing their investment in analytics, and for good reason: as part of a digital transformation, analyzing and interpreting data becomes a required core competency. This takes the form of both enabling business users with the organization, as well as data research teams with deeper capabilities.

Both forms of analytical investment are well-complemented by grounded context of what is known and how it is known. Data researchers not only gain the insights of others, but also can easily "show their work" in the form of the source data and the reasoning that led to the conclusion.

Analytics users throughout the organization gain the benefits of an active SKG connected to the data that created it, resulting in better interpretations of new facts with supporting context. New insights can easily be added to the SKG in a shared and verifiable manner, creating a network effect of organizational learning.

As a result, the organization is able to reason over analytical insights more effectively, accelerating organizational learning and its application.

By connecting active data, active metadata and active meaning, it can do so faster.

The value of grounded analytics has been documented in a number of analytics-intensive environments, including intelligence, bioresearch, financial services and more.

Connecting Active Data, Active Metadata and Active Meaning

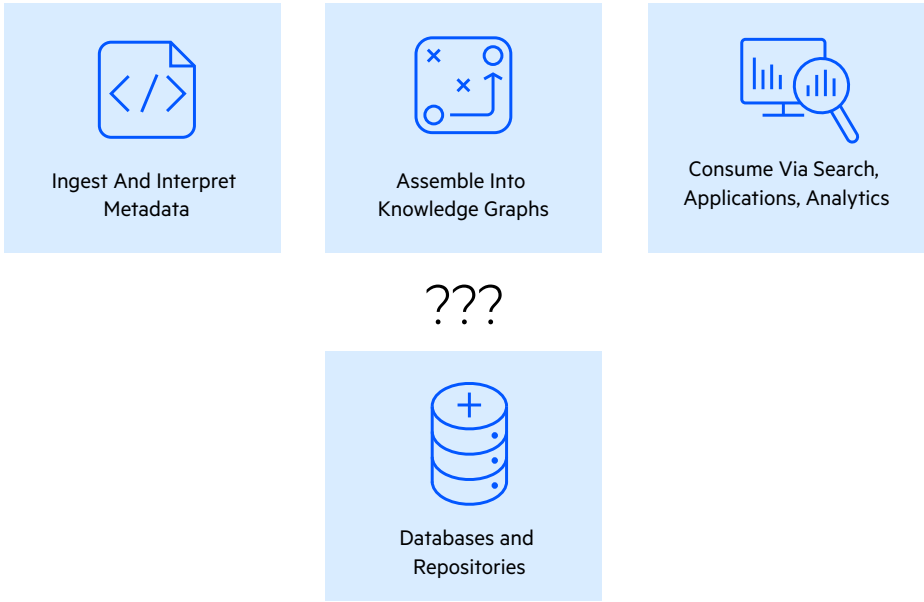
Previous approaches to these problems have been described as integration-centric. The solution to the problem is seen as integrating various components in the form of existing and potentially newer investments. The assumption is that, through clever integration, data and everything we know about it can be reconnected in a reusable and agile way.

Many practitioners will point to their experience that any such approach destroys data agility. As a result, the effort becomes increasingly unresponsive to the organizations that need it, and the effort is abandoned at some point.

By connecting active data, active metadata and active meaning, data agility results.

Any required change to how information is interpreted and acted upon is simple and powerful. Data – and what it means – can be quickly and easily tailored to the needs of any knowledge worker. This simple yet powerful capability turns out to be surprisingly useful in a number of contexts: informed search capabilities improve faster, contextual applications get smarter, analytical results can be better interpreted, and so on.

This same data agility also simplifies and streamlines many operational aspects of data management. Compliance, audit and other forensic activities can quickly get to "what did we know" and "when did we know it?" Security policy is driven by data and what we know about it, resulting in an unusually agile and powerful security model. Data management teams can more easily make better policy decisions now informed by how the organization values the data at hand, and why.



Connect, Create, Consume

As information enters, is analyzed and is consumed, the value of the SKG plus connected data provides different forms of value at each stage.

Upon data ingestion, semantic knowledge combined with encoded metadata can evaluate and connect new facts with existing ones. Important information can be acted upon, unimportant information can be safely ignored.

Many organizations invest in knowledge engineering or research to evaluate existing and new connections independently, to create new knowledge and the facts and reasoning behind their insights. Knowledge engineering and research workers have easy, convenient access to known facts, why they are known, their meaning in relation to other facts, how those decisions were made, and so on.

The metadata and its connections can be inspected as well, providing new value to familiar questions. For example, if the insurance question is "what is our exposure to tsunami risk?", an examination might quickly reveal incomplete knowledge of earthquakes or other subjects that support a more complete answer being required.

As different organizations want to consume the SKG and its data in different ways, data agility means the platform can keep up with their tempo of new and creative uses for information. Data agility also creates a powerful network effect, where different functions bring their problems – and their knowledge – to the commons. Data agility also enables innovation: as most friction has been removed from consuming knowledge, more ways of consuming it are now possible.

One advantage derives from all activities being connected to a single construct that contains data and everything that is known about it. In simple terms, *all participants are working off the exact same facts and their interpretations.* Another advantage results from any insight being able to be added anywhere, and used everywhere immediately. Any insight anywhere along the line is a candidate to be added to the current store of tribal wisdom.

Large scale policies become much easier to implement. A pragmatic example would be a new security policy that needed immediate and verifiable effect: essentially a new interpretation of information and what it means.

These patterns play out in powerful themes in any information-intensive organization that wishes to improve the effectiveness of their knowledge workers.

Barriers to Innovation

Successful practitioners will share their challenges with different constituencies when proposing a new way of doing things.

There is always a strong motivation to consider existing investments as part of the solution. They will tell you that, instead, they turn out to be part of the problem. They capture and use information for a different purpose, and are best considered as participants, not foundational components.

For example, data management teams will not be enthusiastic about the prospect of yet another unfamiliar technology to be responsible for, as well as the infrastructure required to support it. Once they understand the need to expose existing data assets in a new and more usable way, they better understand the motivation.

Similarly, data architects may propose an integration-centric solution, involving multiple familiar components. A proof-of-concept exercise involving transforming largely unknown data into something useful quickly exposes the unique value of data agility. Metadata-centric solutions that deliver data agility will often deliver usable results in hours or days. Integration-centric solutions may never be able to.

Creating the case for a new platform is never easy. Successful practitioners will point to targeting specific, difficult problems with high visible impact, then using the successful pattern elsewhere in the organization.

The Case for Data Agility

As a part of any digital transformation, organizations must learn to work with information in new ways. A metadata-centric platform that connects active data, active metadata and active meaning creates the opportunity to do just that: learn to work with information – and what it means – in entirely new ways.

By removing the traditional friction that separates data and what we know about it, we can more effectively turn data into knowledge, and insight into action. How thought leaders are using their new-found data agility will differ depending on what their goals are, but all will agree: it is a powerful organizational advantage.

About the Progress Data Platform

The Progress® MarkLogic® data management and Progress® Semaphore™ semantic AI platforms give Global 2000 and public sector organizations a faster, trusted way to unlock value from complex data and achieve data agility. Our unified, data-centric approach lets organizations securely connect data and metadata, create and interpret meaning, and consume high-quality contextualized data across the enterprise – enabling informed search, contextual applications, grounded data for analytics, and facts-based intelligence. Organizations can respond nimbly to business change while benefiting from rigorous data governance and transformational data security.

The platform combines a multi-model database, search, and semantic AI technology to couple data with its metadata and put information in context to its state, use, and audience. Organizations benefit from a single data resource that includes data and everything that is known about what the data means, eliminating the effect of data and knowledge silos and removing friction from any aspect of working with any information at any scale for any purpose. The platform enhances existing investments in data systems, delivering:

- **Data Agility** – the ability to quickly and easily make changes to any aspect of how information is interpreted and acted on

- **Ease of Connection** – integrate existing data sources, repositories, applications, and workflows

- **Semantic AI** – no-code software that uses machine learning and knowledge models to synthesize, enrich, extract, and harmonize all types of metadata

- **Trusted data** – traceable, transparent, repeatable, and auditable results, understandable by business users

- **Enterprise-grade capabilities** – security, scalability, availability, language support, temporality, interfaces, and standards-based architecture for both data and metadata using a single platform

Gain Agility with a Unified Data Platform

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

Rev 2023/01 RITM

Worldwide Headquarters

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

facebook.com/progresssw
twitter.com/progresssw
youtube.com/progresssw
linkedin.com/company/progress-software
progress_sw_

Progress