NUCLEUS
RESEARCH

INFORMATICA ACQUIRES GREENBAY TECHNOLOGIES

ANALYSTS

Daniel Elman, Nick Grizzell

THE BOTTOM LINE

Informatica acquired GreenBay Technologies to support the vision for Data 4.0 and bring AI-powered intelligent data management and digital transformation to support data-driven, cloud-native, and innovative technologies. GreenBay Technologies will support Informatica data management tools through advanced capabilities in AI and machine learning. Informatica expects the MDM capabilities to make headlines first but also expect GreenBay's capabilities to surface in multiple products across the product portfolio in the long term. The acquisition of GreenBay Technologies is the second major acquisition of 2020 to augment the Informatica platform.
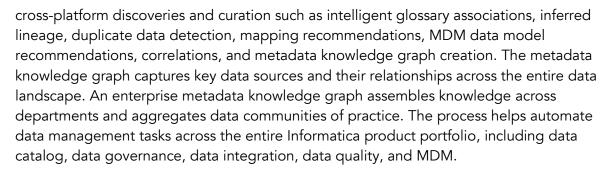
THE ANNOUNCEMENT

On August 18, 2020, Informatica announced the acquisition of GreenBay Technologies. GreenBay Technologies is a Wisconsin-based data management startup that develops AI and machine learning data management tools, with ties to the University of Wisconsin-Madison.

Informatica is a software development company covering a wide range of solutions, including data integration, ETL processes, cloud computing, data quality, data governance, and master data management (MDM), among others. The acquisition of GreenBay Technologies aims to strengthen Informatica's Data 4.0 strategy of AI-powered intelligent data management and digital transformation tools to support data-driven, cloud-native, and innovative technologies. GreenBay Technologies was co-founded by Dr. AnHai Doan, UW's Vilas Distinguished Achievement Professor, who oversees multiple data management research projects at the University's Department of Computer Science. Dr. AnHai Doan and his staff are working closely with Informatica's R&D team to enhance Informatica's CLAIRE engine and power industry-leading data management solutions. Informatica's acquisition of GreenBay Technologies is the first to result from a partnership with a university, and Informatica plans to deploy this strategy in other critical areas of data management. For example, Informatica R&D is collaborating with the ADAPT Research Centre in Dublin to leverage its applied research in AI and natural language processing.

GreenBay Technologies supports several areas surrounding AI-based approaches for entity and schema matching, deep learning, augmented AI using embeddings, and active learning-based technology. Many of GreenBay's innovations started with research work on CloudMatcher, which is a cloud service pioneered by GreenBay Technologies for hands-off entity matching to make matching tasks deployable by business-level users like domain experts. The systems use labeled pairs created by users to infer blocking rules, perform blocking, learn a matcher model, and perform entity matching.

GreenBay Technologies provides three core capabilities to augment Informatica solutions in areas such as entity matching, schema matching, and metadata knowledge graph. Entity matching is the process of finding a matched set of entities from two sets or a list of entities and will support data integration, MDM, data governance, risk, and compliance. The GreenBay solution excels at batch matching, delta matching, and real-time matching to support non-uniform data and a broader set of match fields when compared to declarative fields. Schema matching supports the process of identifying objects which are semantically related to find correspondence among concepts of different distributed, heterogeneous data sources. GreenBay will accelerate Informatica's roadmap for schema matching for data integration, data quality, and data privacy products to extend current capabilities, enable automated processes, and identify sensitive data. With GreenBay, schema matching becomes a core technique to drive automation across the Informatica product portfolio with

cross-platform discoveries and curation such as intelligent glossary associations, inferred lineage, duplicate data detection, mapping recommendations, MDM data model recommendations, correlations, and metadata knowledge graph creation. The metadata knowledge graph captures key data sources and their relationships across the entire data landscape. An enterprise metadata knowledge graph assembles knowledge across departments and aggregates data communities of practice. The process helps automate data management tasks across the entire Informatica product portfolio, including data catalog, data governance, data integration, data quality, and MDM.

GreenBay Technologies is providing ML-based innovations for entity matching to improve the accuracy of master data domains such as customer, product, and supplier. The solutions deep learning approach can be used to match long texts and relatively low-quality data. GreenBay will accelerate Informatica's data catalog and governance roadmap with ML-based inferred lineage to discover data flow relationships where scanner-based or manual lineage discovery is not possible and provide a shared data language with a metadata knowledge graph spanning multiple subject areas and organizations. The capabilities of GreenBay Technologies will provide an estimated 12-month advantage over competitors surrounding metadata solutions.

BENEFIT TO INFORMATICA

The key differentiator as a result of acquiring GreenBay Technologies comes from providing additional capabilities to further Informatica's AI-powered intelligent data management and digital transformation vision. GreenBay Technologies will help Informatica enhance the 360-degree view of business operations through greater accuracy of matching master data to customers, products, suppliers, and other domains. The complete view continues through a greater understanding of enterprise data with improved capabilities to infer data lineage and relationships. GreenBay will also increase data trust with improved capabilities to auto-generate and apply data quality rules based on concept schema matching. Additionally, Informatica and GreenBay Technologies can reduce the risk of a data breach by increasing the accuracy of identifying sensitive data across the enterprise data landscape.

Furthermore, the acquisition will strengthen the core capabilities of Informatica's CLAIRE engine across its Intelligent Data Platform, giving companies the ability to identify, access, and derive insights from organizational data to make informed business decisions. With GreenBay Technologies AI and machine learning capabilities already embedded in the CLAIRE engine, Informatica can accelerate the vision for self-integrating systems that automatically infer and link target schemas to source data, improve capabilities to infer data lineage and relationships, auto-generate and apply data quality rules based on concept

schema matching, and increase the accuracy of identifying sensitive data across the enterprise data landscape.

CONCLUSION

The next few years should be a time of rapid growth for a few dominant players like Informatica with further consolidation as smaller companies with niche capabilities and specializations are acquired or outcompeted. With COVID-19 forcing companies to integrate remote work and more companies adopting cloud-based solutions as security, performance, and availability increase, we will see a surge in demand for integration solutions. Additionally, more companies than ever before are migrating to cloud-based solutions, and with that comes a significant increase in the scale and complexity of application and data ecosystems to manage. As we move forward, the companies that consolidate solutions will come out on top as customers look towards simplified processes and automation tools that support a diverse set of business applications.

New companies will struggle to find a balance between CRM, ERP, and BI analytics data along with third-party solutions to integrate these platforms and gather insights. As customers add more applications and data sources to their stacks, the need for integration technology will only grow. Informatica currently provides a wide range of integration solutions for automated data management processes, data quality, and intelligent cloud services. With the addition of GreenBay Technologies, Informatica gains a major accelerator to its own AI and metadata management capabilities, allowing it to further separate from competitors and deliver greater value to customers. The GreenBay acquisition will greatly improve Informatica's AI capabilities and help accelerate its roadmap for intelligent automation of data management and integration tasks. We see automation becoming more and more important to customers across areas like infrastructure as this frees up developers from repetitive tasks for more value-add projects and eliminates or reduces the potential for crippling human errors with negative implications that can cascade to affect or even halt multiple different areas of business.