Market Insight Report Reprint

Coverage Initiation: Entity resolution specialist Senzing focuses on a persistent data quality pain point

July 15 2022

by Paige Bartley

The company is focused on offering an entity resolution API with a number of deployment options that can facilitate either commercial or enterprise use cases. It is hoping to make entity resolution easy for developers, and easy and cost-effective for organizations.

451 Research

S&P Global
Market Intelligence

This report, licensed to Senzing, developed and as provided by S&P Global Market Intelligence (S&P), was published as part of S&P's syndicated market insight subscription service. It shall be owned in its entirety by S&P. This report is solely intended for use by the recipient and may not be reproduced or re-posted, in whole or in part, by the recipient without express permission from S&P.

Introduction

Data quality challenges become more complex with evolving data types and increasing volumes of enterprise data. It is therefore not surprising that data quality remains a top enterprise pain point. Based on 451 Research's Voice of the Enterprise: Data & Analytics, Data Management & Analytics 2021 survey, 47.7% of respondents report "data quality" as the biggest challenge in relation to enterprise analytics initiatives. And one of the most fundamental technical challenges in data quality remains entity resolution: the task of recognizing when records are the same, even when they have been described differently.

Senzing is an entity resolution specialist, focused on offering an entity resolution API with a number of deployment options that can facilitate either commercial or enterprise use cases. By focusing on high accuracy, near real-time (rather than batch) functionality, ease of deployment and low cost of use, Senzing hopes to make entity resolution easy for developers and as an extension, easy and cost-effective for organizations.

THE TAKE

In the data management market at large, it seems as if there is a pendulum swing by technology providers to provide broader and more overlapping data management platforms. Our Data Management Market Map 2022 supports this notion, with "convergence" of market segments as a key theme. Yet amid this vast data management nexus, opportunity emerges for specialist providers. For purchasers that have a well-defined pain point or problem set, sometimes a specialized and cost-effective tool is exactly what is desired. Sweeping or integrated platforms can potentially be excessive in these cases, especially if they overlap with existing IT investments.

Senzing knows precisely where it fits and precisely what its primary data quality use case is. By courting enterprise DIY users of technology, as well as original equipment manufacturer (OEM) and value-added reseller (VAR) arrangements, it has multiple go-to-market options. Its challenge, in catering to a developer end-user audience, is that it may sometimes miss out on higher-level purchase-influencer discussions, particularly with IT and data leadership.

Context

The history of the Senzing technology is intertwined with IBM Corp.'s acquisition, research and development efforts. In 2005, IBM acquired the Nevada-based Systems Research & Development (SRD) for its real-time entity resolution technology. Its core technology, known as Non-Obvious Relationship Awareness (NORA), was essentially rebranded and then sold as IBM InfoSphere Identity Insight. In mid-2009, the underlying technology for Senzing was initiated at IBM as a new entity resolution project code-named "G2," focused on technical aspects such as domain agnosticism, self-tuning and self-correcting capabilities, privacy by design, relationship awareness, real-time capabilities, scalability and ease of deployment and operation.

IBM's G2 entity resolution project technology became commercially available in 2012, with thousands of copies shipped. Senzing was formally spun off as an independent company from IBM in 2016 by co-founders Jeff Jonas (CEO and chief scientist), Jeff Butcher (chief architect) and Greg Adair (chief engineer). Senzing had its first product release in 2017, gained customer and partner momentum between 2018-2020, and later signed its first large OEM agreement in 2021.

To date, Senzing has raised a seed round of funding, although the company does not currently disclose funding totals. Reporting a capital-efficient business model, Senzing has not yet needed to pursue a series A round or further institutional funding at this time. The company is based in Las Vegas, and reports roughly 25 employees with a distributed/remote team model.

©Copyright 2022 S&P Global Market Intelligence. All rights reserved.

Product

The Senzing product is a C library accessible from Python, Java and REST using a locally hosted API that runs either on-premises or in the cloud. No data flows to the Senzing company. Because many entity resolution use cases involve customer data and personally identifiable information (PII), this design aids in security and privacy by limiting the exposure of data to an additional business party.

Senzing is designed specifically for developer use, as opposed to many master data management (MDM) or data quality products on the market that are designed more for line-of-business roles. The architecture enables real-time use cases, such as entity resolution in banking, because it offers very low latency and does not use batch processing. Built to be highly accurate out of the box, the technology continues to improve in accuracy with use and exposure to data. It can scale to tens of billions of records, handling thousands of transactions per second.

The Senzing product is also designed to be low maintenance, with minimal advance data preparation required. Active maintenance allows developers to perform updates and deletes while loading data and servicing queries, with the product online 24/7.

Privacy by design was also a key consideration in the development of the Senzing product. Aside from data not flowing back to the Senzing company, the entity resolution methodology also favors false negatives. This means that in efforts to pinpoint specific people or entities, the wrong people aren't "scooped" up by association or similarity. The product also has data protection features and features for explainability of how, why or why not a specific match occurred.

Out-of-the-box functionality which the product specializes in includes people and organization entity types; complicated name, address, and identifier data; as well as global scripts.

Strategy

Senzing's selling motion operates primarily via three routes: direct sales, OEM partners and VARs. While the company first focused on direct sales, it has more recently expanded into OEM deals, including those with firms such as NICE Actimize, IBM, Black Ice, Siren and Redhorse. OEM and VAR deals in the past year have represented about 20% of ARR, with this number steadily increasing and projected to be closer to 40% in 2023.

Packaging for the product underscores low total cost of ownership (TCO) and currently hinges on a model of subscription pricing based on data source records. A data source record is defined as a single record processed for loading into Senzing, and updates to records and searches aren't counted toward the total. Deletes of records generally reduce the count of data source records. There are incremental tiers of yearly pricing based on number of data source records, with options going all the way up to unlimited data source records for a fixed yearly cost.

©Copyright 2022 S&P Global Market Intelligence. All rights reserved.

Competition

Senzing is a technology component maker, not a platform provider. In this sense, some of its most direct competition comes from direct development and data science efforts rather than commercialized products. Libraries such as Dedupe exist for Python users seeking entity resolution use cases, but commercial products such as Quantexa also exist for entity resolution.

The competitive market for broader data quality providers is more complex, because some large data quality vendors provide or compete in entity resolution but also might eventually be attractive partners or even OEM customers for Senzing. Precisely (fka Syncsort) is worth noting in this market for its use of location enrichment data and unique ID approach to deduplicating and resolving entities. However, other providers that are well known for data quality that may include entity resolution include Experian Data Quality, Hitachi Vantara, Informatica Inc., SAS Institute, Semarchy, Syniti, Talend, Tibco and Zaloni.

MDM platforms are also both potential competitors and potential consumers of the Senzing product. Examples not already listed above include companies such as Ataccama, Boomi, insightsoftware (via Magnitude assets), Mastech (via InfoTrellis assets), Reltio, SAP SE, Stibo Systems, Syndigo and Tamr.

Entity resolution can also occur in other technologies, such as those used to discover data and associate data with specific defined individual entities. Often these products have data privacy use cases. BigID is prominent example of technology underpinned by data discovery and entity resolution, but others such as 1touch.io, OneTrust and Spirion also have capabilities in this space.

SWOT Analysis

STRENGTHS

Senzing has the benefit of being extremely focused; the company is not trying to boil the ocean in terms of data management, or even data quality. The entity resolution technology has been developed to be highly automated, highly accurate, cost-effective, privacy-aware and real time. By striving to be the best at what it does, the company hopes to carve out a niche for both direct and OEM/VAR sales.

WEAKNESSES

Senzing is unapologetically an entity resolution specialist, and further, its API-driven technology is geared toward a developer audience, potentially limiting its enterprise reach with higher decision-makers. It is also lean in terms of funding relative to many other specialists in the data quality space. The company needs to be strategic with its partnerships and go-to-market in order to maximize opportunities.

OPPORTUNITIES

As a technical specialist, focused on an underlying data quality issue, Senzing has room to grow with its OEM relationships. If the company were to embed its technology within more of the larger and better-known data quality platforms, it could reap the financial benefits without having to spend much on direct marketing and sales. Senzing is a component more than a platform, and OEM strategy is ideal.

THREATS

In the data management market, the broad platform players are always looking for strategic technical acquisitions that could potentially complement existing capabilities or fill gaps in the portfolio. Senzing, as a small company with limited funding, could possibly become an attractive target for M&A. Such a turn of events could undermine the company's existing partnership ecosystem and inherent agnosticism.

©Copyright 2022 S&P Global Market Intelligence. All rights reserved.

CONTACTS

The Americas
+1 877 863 1306
market.intelligence@spglobal.com

Europe, Middle East & Africa
+44 20 7176 1234
market.intelligence@spglobal.com

Asia-Pacific
+852 2533 3565
market.intelligence@spglobal.com

www.spglobal.com/marketintelligence

Copyright © 2022 by S&P Global Market Intelligence, a division of S&P Global Inc. All rights reserved.

These materials have been prepared solely for information purposes based upon information generally available to the public and from sources believed to be reliable. No content (including index data, ratings, credit-related analyses and data, research, model, software or other application or output therefrom) or any part thereof (Content) may be modified, reverse engineered, reproduced or distributed in any form by any means, or stored in a database or retrieval system, without the prior written permission of S&P Global Market Intelligence or its affiliates (collectively, S&P Global). The Content shall not be used for any unlawful or unauthorized purposes. S&P Global and any third-party providers, (collectively S&P Global Parties) do not guarantee the accuracy, completeness, timeliness or availability of the Content. S&P Global Parties are not responsible for any errors or omissions, regardless of the cause, for the results obtained from the use of the Content. THE CONTENT IS PROVIDED ON "AS IS" BASIS. S&P GLOBAL PARTIES DISCLAIM ANY AND ALL EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, ANY WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE OR USE, FREEDOM FROM BUGS, SOFTWARE ERRORS OR DEFECTS, THAT THE CONTENT'S FUNCTIONING WILL BE UNINTERRUPTED OR THAT THE CONTENT WILL OPERATE WITH ANY SOFTWARE OR HARDWARE CONFIGURATION. In no event shall S&P Global Parties be liable to any party for any direct, indirect, incidental, exemplary, compensatory, punitive, special or consequential damages, costs, expenses, legal fees, or losses (including, without limitation, lost income or lost profits and opportunity costs or losses caused by negligence) in connection with any use of the Content even if advised of the possibility of such damages.

S&P Global Market Intelligence's opinions, quotes and credit-related and other analyses are statements of opinion as of the date they are expressed and not statements of fact or recommendations to purchase, hold, or sell any securities or to make any investment decisions, and do not address the suitability of any security. S&P Global Market Intelligence may provide index data. Direct investment in an index is not possible. Exposure to an asset class represented by an index is available through investable instruments based on that index. S&P Global Market Intelligence assumes no obligation to update the Content following publication in any form or format. The Content should not be relied on and is not a substitute for the skill, judgment and experience of the user, its management, employees, advisors and/or clients when making investment and other business decisions. S&P Global Market Intelligence does not endorse companies, technologies, products, services, or solutions.

S&P Global keeps certain activities of its divisions separate from each other in order to preserve the independence and objectivity of their respective activities. As a result, certain divisions of S&P Global may have information that is not available to other S&P Global divisions. S&P Global has established policies and procedures to maintain the confidentiality of certain non-public information received in connection with each analytical process.

S&P Global may receive compensation for its ratings and certain analyses, normally from issuers or underwriters of securities or from obligors. S&P Global reserves the right to disseminate its opinions and analyses. S&P Global's public ratings and analyses are made available on its websites, www.standardandpoors.com (free of charge) and www.ratingsdirect.com (subscription), and may be distributed through other means, including via S&P Global publications and third-party redistributors. Additional information about our ratings fees is available at www.standardandpoors.com/usratingsfees.