EDM Webinar

Data Management in an AI-Driven Era:
Why MDM is No Longer Sufficient

Live Date: January 24, 2024

Featuring:

Dr. Michael Stonebraker, Adjunct Professor, Massachusetts Institute of Technology

Anthony Deighton, Data Products General Manager, Tamr

Moderator: **Eric Bigelsen,** Head of Industry Engagement, EDM Council

Webinar Resources:	Additional Links:
Recording: View webinar **Presentation: View slide deck**	**EDM Council website:** **www.edmcouncil.org** **Tamr website: www.tamr.com**

WEBINAR Q&A:

Thank you to the panelists for providing the below answers to all questions posed during the live webinar. For more information or additional questions, contact us.

Moving to organizing data by entity; do you need to go further and organize at enterprise level in a holding company environment or where there are multiple entities within an enterprise that can be grouped based on data sets used?

In a holding company or multi-entity enterprise environment, it's beneficial to organize data at both the entity and enterprise levels. By organizing data at the entity level, you address the specific needs and context of each entity. However, consolidating and organizing at the enterprise level is crucial for a holistic view and strategic decision-making. This approach allows for cross-entity analysis, identifying synergies, and leveraging economies of scale. Technologies like Tamr's facilitate this by unifying disparate data sources, ensuring consistency, and maintaining the integrity of data across the enterprise.

How can advances in AI/ML help solve the data quality problem?

AI and machine learning offer significant advancements in addressing data quality issues. These technologies can automate the identification and correction of data inaccuracies, inconsistencies, and duplications. AI algorithms can learn from patterns and anomalies in data, enabling them to predict and rectify errors proactively. Machine learning models, especially when coupled with human feedback loops, continually improve data quality over time. This results in more reliable, accurate, and actionable data for business insights.

Certain countries still have addresses i.e. in front of ABC college or opposite XYZ hospital etc. even for global organizations. How do we resolve such a scenario even using AI?

In countries with non-standard addressing systems, AI can be instrumental. AI algorithms, particularly those in natural language processing (NLP), can interpret and standardize these addresses by contextual understanding of landmarks and local references. By integrating geospatial data and with ML, AI systems can convert vague addresses into precise geolocations. This process enhances the accuracy of data for global organizations operating in regions with such addressing challenges.

We are challenging some requests re what is analytics vs what is an AI use case, various definitions for AI and what counts as AI vs iterative modeling and predictive analytics. Is what we do today already AI we just have not boasted or branded it as AI?

The line between analytics, predictive modeling, and AI is often blurred. Traditional analytics and iterative modeling are precursors to what we now term AI. AI encompasses a broader scope,

including machine learning, deep learning, and other advanced algorithms that can learn, adapt, and make decisions. If your current processes involve learning from data, making predictions, or automating decision-making, they could indeed be considered forms of AI. The key distinction is the level of sophistication and autonomy in the algorithms used.

Solving DQ at the source level could still be challenging while consolidating / mastering data. Any quick comments on improving DQ at the source level using AI ahead of mastering?

Addressing data quality at the source is crucial for any data management strategy. AI can significantly aid in this by implementing real-time analysis and correction algorithms at data entry points. Machine learning models can predict common errors and inconsistencies based on historical data, prompting corrections early in the data lifecycle. Additionally, AI can assist in standardizing and cleansing data from various sources before it enters the master data management system, ensuring a higher quality of data for unification and analysis.