

Enterprise Analytics Modernization

9/25/2021

Enterprise-level business intelligence, data management, and analytics have evolved materially over the last few years. However, many organizations still maintain highly siloed data with limited self-service and enterprise-level data management and analytics capabilities. Analytica was engaged by a federal government client, the Department of Homeland Security (DHS), to help modernize their enterprise analytics capabilities and support the analysis of hundreds of millions of records from across DHS agencies. The client faced difficulty linking immigrant identities from the disparate records stored in more than four dozen separate data silos. Downstream, this hindered their ability to perform timely mission critical enterprise-level planning, prescriptive, and predictive analytics.



The Challenge

The client maintained over four dozen separate data systems. This siloed data ecosystem limited the client's ability to conduct enterprise-level planning and analysis and to report quickly and comprehensively to Congress and the public. In 2016, DHS launched the Immigration Data Integration Initiative (IDII) to tackle these challenges. The IDII seeks to establish a single, authoritative source of immigration data that is accessible to stakeholders from across the Department. The initiative encompasses three main lines of work: implementing uniform enterprise-wide immigration data standards; linking immigration records across DHS data systems; and establishing an IT environment that provides stakeholders real or near real-time access to validated data. Analytica, as part of their engagement, supported downstream initiatives around IDII in performing immigrant person-matching in a timely and accurate manner. DHS had developed an initial entity resolution engine in SAS but needed support to make significant improvements in performance and person-level matching confidence.



Our Approach

Analytica built an entity resolution engine that used existing SAS as well as ported functionalities into Python to support DHS's transition to a scalable, modern analytics environment, and a suite of Tableau dashboards to provide DHS Executives with timely and up-to-date reporting and analysis of enterprise immigration statistics. Our work is building a high-confidence matching engine that will allow the DHS to establish a single source of accurate data across hundreds of millions of event records sourced from over a dozen DHS agencies and in different data formats. Analytica is using sophisticated fuzzy matching and phonetic encoding technologies to populate a single source of record to be used for anonymous immigration statistics within the DHS. The system applies matching business rules that are tailored to DHS immigration data and are capable of ingesting diverse subject matter records. Analytica is also providing a suite of executive dashboards in Tableau for DHS HQ to view complex visualizations securely.



The Solution

Analytica's matching engine and Tableau dashboards provide the client with the key foundational elements in the modern enterprise analytics environment needed to support mission critical scalable statistics, analysis, and reporting capabilities. Our use of advanced, modern phonetic encoding and fuzzy matching allowed DHS to increase record matching by 60% allowing DHS to address long-standing challenges of establishing and updating immigrant identities from across hundreds of millions of records with high confidence and performance. This periodically refreshed, single source of anonymized immigrant records will yield ongoing benefits for DHS as this serves as a critical data source in many future use cases.

Contact Us

Phone: 202.470.4806

Fax: 1.888.265.4689

Email: bufederal@analytica.net

www.analytica.net