Enabling Data for Secondary Use

Eighty percent of healthcare data is unstructured. This data is critical to deriving new insights, innovation and knowledge for research hospitals, insurance and medical claims providers - among other industries. Accessing and analyzing this data, while ensuring it is properly protected, will improve healthcare outcomes and drive operational efficiencies.

Privacy Analytics’ software provides customers with the flexibility to anonymize structured and unstructured data. This can be done either as an integrated solution or as a standalone text de-identification product called TEXT. This software enables statisticians and data analysts to conduct critical analyses on valuable unstructured data, while allowing data managers to safeguard personal information used for secondary purposes.

Safeguard Your Data

The growing volumes of unstructured data from multiple sources heighten organizations’ susceptibility to potential data breaches. With TEXT, privacy and compliance officers strengthen the protection of their organizations' data assets by extending de-identification to unstructured data. With a standardized, enterprise-wide approach, privacy and compliance officers can:

- Ensure company-wide de-identification practices that are compliant with HIPAA and other legal requirements, as well as internal organizational policies and procedures governing privacy and personal information;

- Extend the practice of de-identification to unstructured formats residing in electronic health records, medical devices, clinical notes, discharge summaries and other data sources, to optimize protection and coverage of data for secondary purposes; and,

- Increase the value of underlying data assets with the ability to capitalize on de-identified unstructured information for analysis.

Gain Analytic Utility and Insight

TEXT redacts personal information, such as names and dates of birth found in physician notations in structured databases, medical devices’ text fields, or XML. It can de-identify MS Office files, text fields in unstructured databases, PDF files and more.
TEXT allows analytic professionals to configure the redaction of unstructured data, ultimately preparing it for secondary use and analysis, by:

- Discovering and annotating personal information residing in multiple text formats and data sources, as well as filter IDs, including credit cards, driver licenses and medical codes;

- Improving the quality of de-identified data by identifying misspellings and text errors through fuzzy matching;

- Allowing for better insight into anonymized geospatial data by maintaining the first three digits of postal and zip codes; and,

- Evaluating the measurement and tuning of precision and recall by comparing pre-determined samples of a dataset.

**Integrated Solution**

TEXT provides users with a highly scalable solution that, through its API, automatically integrates with customers' IT environments. It can extract, de-identify and index personal information in text formats residing in standard database tables.

TEXT integration with the rich capabilities of CORE enables analysis of structured and unstructured data values, matching the corresponding de-identified unstructured text to ensure accurate analyses. For statisticians and data analysts, TEXT enables them to:

- Automate the de-identification of unstructured data from multiple sources to gain richer analytic value and insight;

- Mitigate the risk of re-identification by detecting exposure of personal information, while also determining its relative analytic quality; and,

- Enhance the value of data assets by maintaining the relationship of masked and anonymized values for more granular, higher quality analyses.

**Leveraging Your Data Assets**

Enabling personal information for secondary use is critical to driving innovation, deriving new insights, and gaining new knowledge for your organization. TEXT is the industry’s most comprehensive de-identification solution when dealing with unstructured data.

TEXT: *the proven, responsible way to unlock personal health data.*

If you would like to learn more, please visit [www.privacy-analytics.com](http://www.privacy-analytics.com)