SOLUTION BRIEF

The Problem

Organizations today must comply with a number of changing regulations and ensure their data is secure. But the critical task of accurately identifying sensitive data remains a challenge. While many solutions on the market today claim they can identify organizational data types, one of the top concerns remains how exactly these solutions are doing it.

The Solution

Identifying sensitive data requires multiple data identification methodologies. Only a combination of disparate technologies can help you fully understand not only specific data but also its context. This is done by, for example, evaluating sentences surrounding a word, as well as the type of document the data is contained within. Only a multi-faceted data protection solution is sufficiently capable of identifying and protecting organizational data.

How Titus can help

While there are several methods of classifying and identifying data within the data protection industry, Titus provides the most robust and flexible data detection tools available on the market today. Titus incorporates advanced Machine Learning capabilities, enabling customers to quickly and accurately detect and protect their sensitive data. The combination of Titus data identification tools is called our "**Data Detection Engine.**"

The Titus Data Detection Engine provides a range of data detection tools, which, when used in conjunction, build the most robust security programs available.

RegEx, a.k.a. regular expressions, is a sequence of characters that define a search pattern. This method of search inspects the data character-by-character looking for a pattern that matches the search pattern. Many data protection organizations stop here when it comes to data detection methodology.
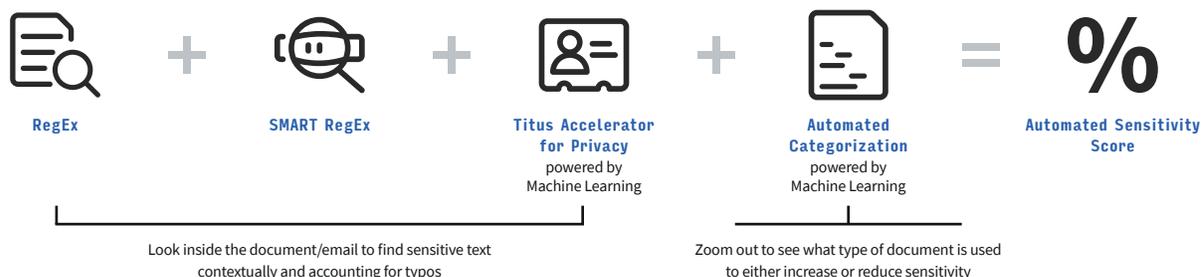
Smart RegEx pre-processes data into tokens such as words or numbers which allows for more targeted scans of well-formed sensitive data in large documents with repeated search queries. The pre-processing helps the solution focus and can improve performance dramatically.

Automated Categorization *powered by machine learning* leverages machine learning to recognize a topic by correlating words in a document that are indicative of said topic. This depends on a model that was trained from a corpus of documents, sorted by topic.

Titus Accelerator for Privacy *powered by machine learning,* also uses a type of machine learning called deep learning to categorize data. This tool doesn't consider the topic, but rather scans through content focusing on finding PII (personally identifiable information). Titus Accelerator for Privacy uses both words and the context of those words surrounding it to reduce false positives, and focuses on PII, PHI (personal health information), and PFI (personal financial information). Below, the multiple methodologies within the Titus Data Detection Engine work together to provide a better understanding of the sensitivity of data.

Identifying data **with Titus**

Below, the multiple methodologies within the Titus Data Detection Engine work together to provide a better understanding of the sensitivity of data.
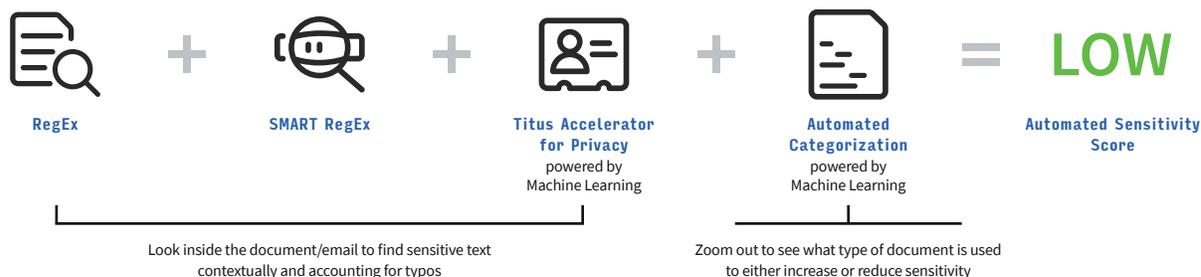

| RegEx | + | SMART RegEx | + | Titus Accelerator for Privacy powered by Machine Learning | + | Automated Categorization powered by Machine Learning | = | Automated Sensitivity Score % |

Look inside the document/email to find sensitive text contextually and accounting for typos

Zoom out to see what type of document is used to either increase or reduce sensitivity

To provide an example, if an end user is saving a document, the Titus Data Detection Engine will leverage RegEx, Smart RegEx, and Titus Accelerator for Privacy to discover a number of items within a file, including any names, email addresses, phone numbers, and credit card numbers.

While the combination of this information within a document would be considered highly sensitive, the Automated Categorization feature indicates that there is a very high probability that this document is, in fact, an admin guide, which contains only sample data. With that in mind, the context related to the data discovered within the document takes on a different meaning, and the document itself deserves a low sensitivity score.

| RegEx | + | SMART RegEx | + | Titus Accelerator for Privacy powered by Machine Learning | + | Automated Categorization powered by Machine Learning | = | Automated Sensitivity Score LOW |

Look inside the document/email to find sensitive text contextually and accounting for typos

Zoom out to see what type of document is used to either increase or reduce sensitivity

Most data detection solutions on the market today use a single methodology to identify and protect organizational data. Unlike other solutions available, Titus solutions leverage the Titus Data Detection Engine, comprised of multiple unique data detection methodologies that best ensure data is correctly identified and handled, ultimately creating the most robust privacy and security ecosystem for our customers.

---

titus
by HelpSystems

www.titus.com

About **HelpSystems**

HelpSystems is a people-first software company focused on helping exceptional organizations Build a Better IT™. Our holistic suite of security and automation solutions create a simpler, smarter, and more powerful IT. With customers in over 100 countries and across all industries, organizations everywhere trust HelpSystems to provide peace of mind. Learn more at www.helpsystems.com.