Crowe

March 2019

Testing the Effectiveness of Name Screening Solutions

An article by Kimberly A. Macadaeg, CAMS; Ralph D. Wright, CAMS; and Beatriz R. Young, CAMS

Economic sanctions programs continue to be an effective tool for managing national interests, but their successful execution requires that financial institutions have effective, risk-based systems in place. Sensitivity testing of sanctions compliance systems supports financial institutions in applying a risk-based approach to their sanctions programs. This approach mitigates the risk of noncompliance fines while reducing the time-consuming review of false positive results.

Sanctions environment

In the face of an ever-changing global economic landscape and major geopolitical challenges such as counterterrorism strategies, failing economies, and conflict resolution programs, sanctions have become the policy tool of choice for world governments, particularly in the West. The last two decades have seen a steady and substantial rise in sanctions imposed by the U.S. Department of the Treasury's Office of Financial Asset Control (OFAC).[1]

Institutions in the European Union (EU) and the United Kingdom (U.K.) have also seen increased sanctions-compliance demands, both due to their own jurisdictional regulations and as a result of their relationships with U.S. institutions. The latter requires the monitoring and execution of U.S. sanctions by foreign institutions partnered with U.S. institutions. Further, OFAC's 50-percent rule prohibits dealings with any entity of which half or more is owned by a sanctioned country or person.

Sanctions have also become more sophisticated and now target not only specific political and personal entities, but also particular activities within them. Entities in Russia, North Korea, Cuba, Venezuela, and Iran are of particular concern. These sanctions developments have resulted in an increase in the diversity of entity names within the sanctions lists, requiring greater precision from screening filters in order to comply with regulatory expectations.

Sanctions violations can be imposed regardless of the type of organization, the number of infractions, or the dollar amounts associated with restricted transactions. Failure to identify a sanctioned entity can lead to significant enforcement actions and pose reputational risk. OFAC levied $329 million in fines against Crédit Agricole in 2015 for violations of sanctions against Sudan and other countries.[2] Barclays Bank was the recipient of a comparatively minor $2.48 million fine in 2016 due to violations of sanctions against Zimbabwe.[3] And, in 2017, OFAC and the U.S. Department of Commerce assessed a record $1.19 billion in combined fines against ZTE Corporation for its violations of sanctions against Iran and its attempts to conceal those violations.[4]

While technically overseen by OFAC within Treasury, the execution of sanctions programs is largely the responsibility of financial institutions. Failure to comply with OFAC rules and other sanctions regulations could result in substantial fines, criminal and civil penalties, and enforcement actions. It is essential for financial institutions to test that their sanctions compliance systems are appropriately sensitive to evolving requirements while balancing the need to reduce the time-consuming and expensive review of false positive data.

What is sensitivity testing?

Sensitivity testing is an integral component of all sanctions compliance-monitoring systems. It evaluates a watchlist screening application's reaction to adverse data quality by assessing its ability to identify certain data transformations correctly, such as word truncation, concatenation, changes in word order, alternate spellings, and outright misspellings. This information is used to increase the effectiveness of the screening system by evaluating the calibration of the model and the model's outputs.

The testing assesses the model's ability to perform name matching for names degraded by a determined set of data quality business rules. That is, the model must be able to accurately identify true, or high-quality, matches from data that is known to be degraded, or unclear. This analysis aids in assessing the impact of identified model assumptions and limitations on match output and identifies weaknesses in the model's name-matching algorithm. From there, analysts can determine whether or not model adjustments or tuning exercises (or both) are the logical next steps in enhancing the monitoring system.

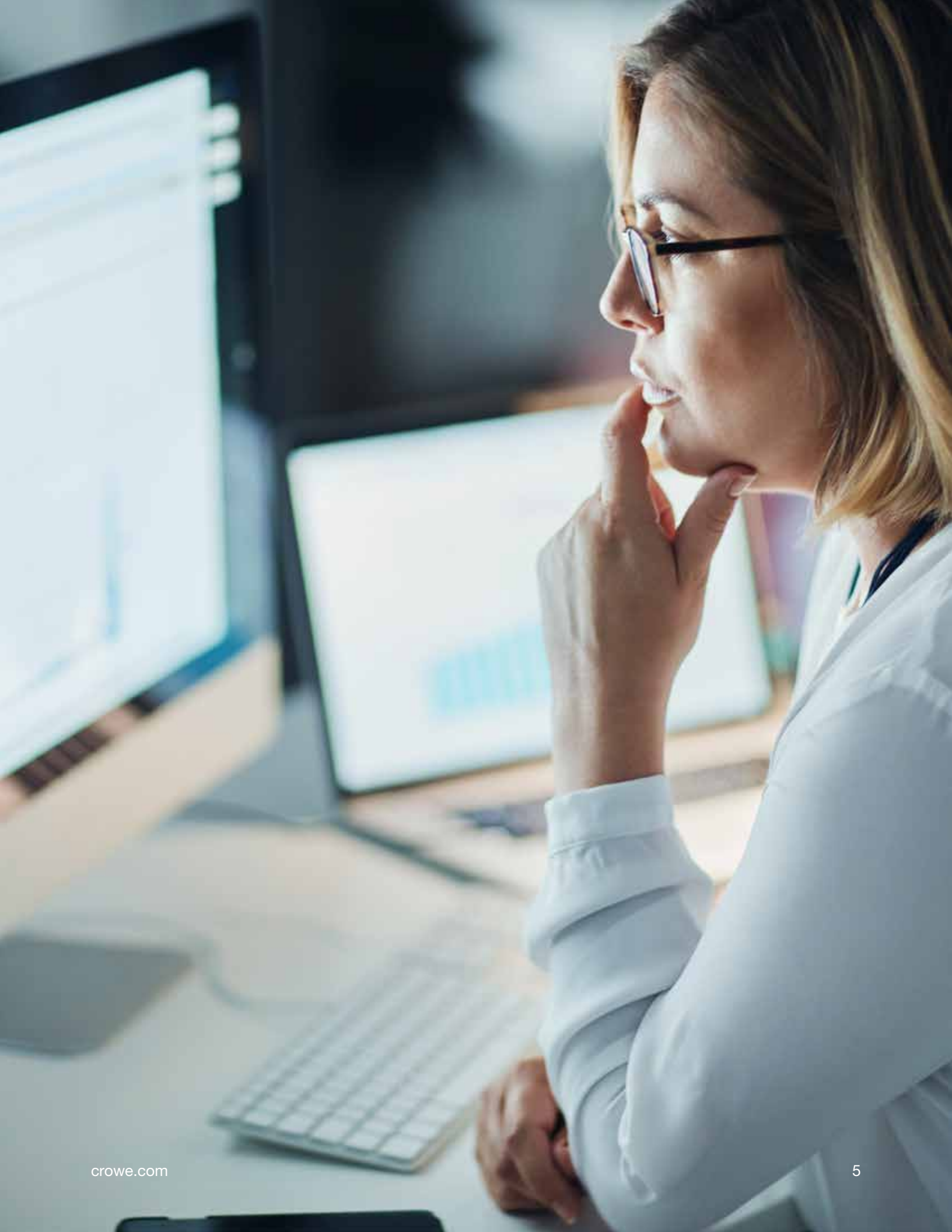
To conduct sensitivity testing, names are degraded using a degradation matrix and screened by the watchlist screening system. The results are then analyzed to see whether a match was generated. This step allows for the identification of any critical screening gaps that might exist and for the determination of which degradation rules led to the model's highest sensitivity.

Sensitivity testing objectives

The purpose of sensitivity testing is to identify the limitations of the matching logic of watchlist screening models. The matching logic – often referred to as "fuzzy logic" – generally refers to the set of algorithms, rules, synonym tables, foreign word transliterations, and other functionalities designed by the vendor to generate name matches between client data and watchlist content. The fuzziness of the logic creates space for the matching of inexact, but likely similar, data. For example, phonetic matching software, such as Soundex, matches words that are similar when said aloud. This probabilistic logic augments direct matching, or deterministic logic, allowing for a thorough evaluation of the performance of the watchlist screening model's match scoring algorithm and thresholds. The evaluation determines if they are calibrated effectively and optimized to generate as few false positives as possible while still generating high-quality alerts.

False positives are client records that are incorrectly identified as matches to a sanctions list entry. False positives might be generated as a result of a watchlist screening model's thresholds being set too conservatively or algorithms and parameters not aligning to the financial institution's customer base or data type. While false positives are an accepted reality in watchlist screening filters, analyzing the trends that cause false positives can provide opportunities for their reduction.

Sensitivity testing might also identify which, if any, quality issues inherent to the financial institution's data require immediate remediation so that the watchlist screening filter detects affected records. Alternatively, financial institutions might decide to adjust the model's configurations in an effort to compensate for those data issues.



Analyzing results

Sensitivity testing is guided by an institution's source system data quality assessment as well as general industry and regulator expectations. Some degradation rules are institution-specific and based on the type and quality of data that will be screened, while others are selected based on industry standards. Degradation rules that don't apply to a given institution will not result in effective recommendations.

The analytic techniques used to review the results of sensitivity testing are broad and rely on advanced data analytics software including data visualization tools. Typical analyses include assessment of the impacts of individual degradation business rules, data quality issues, threshold analysis, and degradation tolerance.

Sensitivity testing documents the filter's performance at various levels of degradation, which is calculated using statistical concepts such as edit distance, edit distance percentage, and degree of degradation. The goal of degradation tolerance analysis is to identify both the degree of degradation at which the system fails to return expected results and the degree of degradation at which there is a marked increase in false positives. These data points aid in the assessment of model performance, based on the data selected for testing. In addition, this information helps an institution understand the balance between false positives and false negatives. Further, it provides the analytic support for decisions regarding thresholds and configurations that align with their risk appetite.

Once the degraded names are matched against the appropriate watchlists, degradation rules that most usefully test the efficacy of the system must be identified. Overly relaxed degradation rules might capture too many false positives. But overly conservative degradation rules might allow false negatives to slip through, increasing the risk to a financial institution of missing the detection of a list entry.

Execution frequency

How and when should sensitivity testing be conducted? Reviewing the effectiveness of a model's filtering parameters and thresholds should be part of any financial institution's model risk management program. Depending on the institution's risk profile and specific model risk management testing principles, sensitivity testing should be conducted every one to three years. Changes to the institution's risk profile and other events might also trigger a review of the model's performance via a sensitivity testing exercise. These changes might include, but are not limited to, acquisitions, mergers, divestitures, joint ventures, reorganizations and the launch of new businesses or business lines. Regulatory environment changes might also warrant analysis of the watchlist screening model's current settings on an expedited timeline.

Sensitivity analysis can be conducted during tuning exercises or as part of model validation activities. Tuning is the process by which the thresholds and parameters within a model are analyzed and calibrated to improve its performance. As an important component of a tuning exercise, sensitivity testing assesses any potential negative, and unintended, impact to the system's effectiveness caused by the recommended threshold and parameter changes. Within a model validation, sensitivity testing is used as a performance testing strategy to determine if the model is able to effectively detect degraded data that might result from insufficient data quality or model limitations under the current configuration settings.

Sensitivity testing tests the robustness of a watchlist screening system's parameters to align coverage of name variations with an institution's risk appetite. By identifying weaknesses in the rules and algorithms used to screen for sanctioned entities and possible instances of money laundering, sensitivity testing can mitigate the risks of unidentified exact entry matches and unidentified potential matches due to adverse data quality issues.

Learn more

Kim Macadaeg
+1 415 946 7455
kim.macadaeg@crowe.com

Ralph Wright
Principal
+1 630 586 5203
ralph.wright@crowe.com

Bea Young
+1 202 779 9940
bea.young@crowe.com

---

[1] "United States Sanctions Tracker," Enigma Labs, 2018, https://labs.enigma.com/sanctions-tracker

[2] "Treasury Department Reaches Sanctions-Related Settlement With Crédit Agricole Corporation and Investment Bank for Approximately $329.5 Million," U.S. Department of the Treasury, Oct. 20, 2015, https://www.treasury.gov/press-center/press-releases/Pages/jl0223.aspx

[3] "Settlement Agreement Between the U.S. Department of the Treasury's Office of Foreign Assets Control and Barclays Bank PLC," U.S. Department of the Treasury, Feb. 8, 2016, https://www.treasury.gov/resource-center/sanctions/OFAC-Enforcement/Pages/20160208.aspx

[4] "ZTE Corporation Agrees to Plead Guilty and Pay Over $430.4 Million for Violating U.S. Sanctions by Sending U.S.-Origin Items to Iran," U.S. Department of Justice, March 7, 2017, https://www.justice.gov/opa/pr/zte-corporation-agrees-plead-guilty-and-pay-over-4304-million-violating-us-sanctions-sending

crowe.com

RISK-19003-007E