Payment Integrity Alert: The Use of Automation and Data Analytics
From the Office of Management and Budget (OMB) Office of Federal Financial Management and the Pandemic Response Accountability Committee (PRAC)

The intent of this joint alert is to raise awareness on areas where OMB and the PRAC believe there is opportunity to deploy strategies to mitigate payment integrity risks while promoting the equitable delivery of mission objectives, including managing privacy risks. This document does not constitute official guidance or require agencies to undertake specific tasks beyond consideration of appropriate steps to address ongoing or future issues related to payment integrity.

Good government serves the needs of the American people through the equitable delivery of a program's objectives, ensuring that Federal programs and services reach those they are intended to help while maintaining the integrity of those programs. Payment integrity, or reducing the risk of improper payments of government funds, is essential to effective stewardship of taxpayer resources and is vital to public trust in government institutions. Ensuring payment integrity relies heavily on the government's ability to quickly identify and mitigate payment integrity risks. While agencies have the overall responsibility for managing payment integrity risks, the Inspector General (IG) community has unique expertise to support the design and implementation of mitigation strategies for preventing improper payments and recovering overpayments as part of government-wide efforts. As such, information sharing between agencies and their IGs is very important, as agencies can provide proactive briefings to the IG and the PRAC on tracking, reporting, and financial controls in place, as well as request regular briefings from their IG on payment integrity risks and mitigation strategies relevant to their programs.[1]

Data management and automation are enabling capabilities of an agency's digital strategy and have the potential to support agency missions while mitigating payment integrity risks. Many agencies have at least foundational expertise in data-driven, evidenced-based decision making. The use of automation and data analytics are key elements of an effective multi-pronged approach for identifying potentially fraudulent payments. Agencies receiving funding for administration or technology modernization initiatives or activities should consider these elements for both short and long-term improvement.

Automation

OMB Memorandum 21-19, *Transmittal of Appendix C to OMB Circular A-123, Requirements for Payment Integrity Improvement* (March 5, 2021), defines automation as an automatically controlled operation, process, or system. Implementing intelligent automation and robotic process automation into a program's payment system could improve the government's ability to quickly mitigate payment integrity risk. While Federal agencies commonly use automation, the degree of data management and automation maturity and capabilities vary. Agencies are encouraged to continually evaluate their payment processes to identify where applying automation and data management techniques could provide opportunities for improvement without introducing bias into the process. In addition, IGs are encouraged to consider whether there are automation-related solutions that they could recommend as part of their oversight work that would be effective at improving the prevention of improper payments,

[1] See the first OMB-PRAC joint alert, Payment Integrity Alert: Risk Factors and Suggested Mitigating Strategies. Since March 2020, the Congress has enacted legislation providing more than $5 trillion in COVID-19 related reliefand recovery funding. This massive amount of funding has highlighted pre-existing issues with payment integrity,while also creating new opportunities and challenges.

while ensuring that these solutions are not exacting undue or unfair burden on recipients. In situations where automation needs to be implemented rapidly, programs can apply the following automation strategies.

Intelligent automation and robotic process automation (RPA) technologies are most efficient when usedto address a specific problem. To ensure smooth implementation of automation in the workplace and workforce, agencies may need to consider such issues as data management, process, and privacy policy.

Automation requires intimate knowledge of a process and allows for rethinking processes in a more holistic way. For example, understanding the application intake process, standardizing data elements, and subsequently digitizing the intake enables traceability and allows for better analytics in understanding payment integrity. Further, automation builds resiliency, improves efficiency, enhancesanalytic capability, and frees up staff to focus on more strategic aspects of their work.

Automation tools are only as good as the data they use and checking to ensure data are accurate shouldbe a priority. Agencies' considerations should also include how the use of automation might introduce unanticipated unfairness into the payment system, leading to disparate outcomes. Agencies should develop specific strategies to monitor potential bias and strategies for oversight of algorithms being deployed for automation and in data analysis. Agencies should engage stakeholders when making data-related decisions, tapping into the expertise of, for example, engineers, data scientists, ethicists, privacyand civil rights groups, security experts, and other specialists. Algorithms used in intelligent automationprocesses should be continually evaluated for fairness and accuracy.

Review Processes

Automated review processes can reduce improper payment risk when compared to those that rely on a manual- or labor-intensive review processes. For example, automated review processes can reduce the likelihood that a program will make improper payments when responding to a large increase in benefit program applications. Without deployment of automated capabilities, many agencies may need to process the increased workload with legacy systems not adequately prepared to process an increased volume of claims and, with the increased workload, the program's capacity to review and process the applications may be diminished. If an agency relies heavily on manually reviewing applications, programs will likely need to increase staff and resources to handle the higher application volume; new and inexperienced staff may be more prone to making errors in the manual review or verification process. However, agencies using automation and data analysis may also require the addition of staff with the expertise necessary to continually assess these processes. It is also important that agencies balance the usage of automation to confirm that there are appropriate opportunities for human intervention to mitigate errors in the overall decision-making process.

Agencies should consider the following when implementing or enhancing automation:

1. Collaborate with program managers, chief data officers (CDOs), chief information officers (CIOs),and information technology professionals in your agency to identify key stakeholders, select the process(es) to automate, assess impacts to processes and front and back-end systems, leverage subject matter experts on existing processes and train staff on new processes, identify potential solutions to aid in the automated review of collected information.
2. Leverage Application Programming Interfaces (APIs) for expert users to leverage in theirprocessing streams.

3. Work across program managers and CDOs to share data within agencies. Agencies are encouraged to address challenges to data sharing that include lack of agency will to participate driven by a lack of resources to meet demand, agency reluctance to deviate from historical postures of non-sharing, misinterpretations or information asymmetry on statutory authoritiesto disclose data, lack of consensus on ownership and provenance of data, and lack of standardson agreements and data quality.
4. Automate notifications to applicants to provide them with supplemental information.
5. Urge recipients of loans to leverage automated processes like repayment plans through pay.gov or online applications.

Automating the Verification Process for New Eligibility Criteria

Ensuring a beneficiary has fulfilled the eligibility criteria for a payment is necessary for maintaining payment integrity within a program. Programs often incorporate checks or other actions into their payment process to help mitigate the risk of issuing a payment to a beneficiary who is not eligible. When the eligibility rules of a program are changed, programs must interpret the changes and make anyneeded adjustments in a timely manner.

If processes are already automated, adapting the payment process to new legal provisions or new eligibility criteria may occur relatively quickly. For example, when existing financial systems automatically allow for ledgers, reports, data reviews, and oversight/monitoring to be performed, the payment integrity risk associated with new legal provisions being integrated into the payment process can be identified and mitigated. When a program has automated the validation of applicant eligibility criteria, modifications to the process to account for changes to the eligibility criteria can quickly be deployed across the entire payment system. For those processes that are not automated, agencies should consider the following actions to implement an automated process:

1. Identify other programs or agencies that automate checks on similar eligibility criteria and adopt leading practices from those programs or agencies or request technical support in program design.
2. Identify software solutions from existing vendors, for example, to use machine learning and algorithms for identifying and detecting potentially abnormal entries on applications, especially for new applications.
3. When using machine learning-based solutions, take steps to evaluate the flags associated with the technologies' usage to confirm that privacy, civil rights, and civil liberties are being preserved (with special attention to the data and processes used to build these solutions).
4. Identify any Federal databases helpful in determining eligibility that may be modernized to exchange data or confirmation of data existence through APIs.

Extending Existing Automated System Controls to New Programs

When a new program is created, agencies should consider whether the new program can adopt automation from existing programs. For example, agencies could leverage the automated controls in their existing electronic payment system for their new program if the new program will also make payments through the same electronic payment system. Given the likelihood that a new program mayserve a population unfamiliar with the application process, it is helpful to adopt existing automated controls that will reduce the risk that an incorrect or incomplete application will be paid. Automated system controls that may already be in place for existing programs, which may be particularly helpful tonew programs, include:

1. Automating calculations instead of relying on applicant or employee calculations.
2. Validating information against existing data sets instead of agency reviewer manually validating eligibility.

3. Checking applications for completion through automation instead of a visual review from agency staff. This will also include assessing the effectiveness of the automation to confirm that appropriate recipients are not disenfranchised.
4. Using APIs to send payment information from one system to another instead of the agency downloading from one system and then uploading into another system.

Applying Data Analytics to Payment Integrity Risks

Agencies should work closely with their CDO to establish robust data analytics capabilities that can move an agency from relying on a "pay-and-chase" approach to a preventative approach that allows the agency to identify potential improper payments before they occur. Data analytics can effectively identify indicators of fraud and improper payments by discerning trends, patterns, anomalies, and exceptions within data. Data analytics methods vary, and it is important to be aware of the various applications for identification of fraud and improper payments. Data analytics serves as a tool to improve all aspects of the payment processes from prevention to recovery.

Types of Data Analytic Techniques

A wide range of analytic techniques are available that can be used to improve payment integrity. The table below provides a high-level overview of common data analytic techniques that agencies should consider incorporating into their payment process.

	Description of Data Analytics Technique	How the Technique Can Improve PaymentIntegrity	Considerations for usage of the technique
Rule-Based	Focuses on transactional data; seeks to identity transactions that depart from expected procedures.	Can help isolate instances where a transaction departs from expected rules, including those governing use of purchase cards, procurement policies and applicants who may be on excluded parties lists, among other examples. For example, if a "rule" is that incarcerated individuals are not eligible for a benefit, a data match can be conducted to determine if the applicant is incarcerated before approving the transaction.	Low risk. Equity issues can arise when the deployed rules systematically flag practicesthat depart from expected procedures, but are not intentional payment integrity violations and can be corrected. Remedy: Ensure that rules are linked to standard procedures and that flagged transactions are verified for potential data errors.
Anomaly Detection	Focuses on investigating large sets of transactions, uses "unsupervised modeling" techniques to identify outliers compared to peer groups based on unknown patterns among common and individual fraudsters.	Can allow agencies to scan large datasets and quickly identity outliers that could indicate fraud. These outliers can then undergo further reviews.	Low risk. Equity issues can arise if applications from underrepresented groups (being small in number) are flagged as outliers by the unsupervised modeling system. Remedy: Ensure human oversight of identified outliers to verify that the identification of potential fraud is accurate.
Network/Link ~ *socialnetwork*	Looks at linked patterns, such as the social networks of individuals, to identify previously unknown bad actors.	Can be useful for uncovering organized fraud and associations between fraudsters. For example, an individual may not be suspicious based on their actions alone, yet suspicion may arise when their actions are connected to others through a set of commonalities based on associated attributes, revealing schemes that may have otherwise gone unnoticed. *Agencies should consult with their General Counsel in leveraging social networks and related data sources for automated evaluation of linked patterns.*	Medium-to-high risk. Equity issues are likely to occur with the 'guilt by association' mechanism used in these approaches, especially when using social networks where affinity groups associate strongly and historic or past disparate scrutiny has made individuals in certain networks more likely to be deemed bad actors than members of other networks. Caution: Social network-based identification is inherently risky due to the lack of reliable training data. Such methods should be used only with extensive validation and continuing human oversight.
Predictive	Uses known fraud or improper payment patterns to infer the existence of such patterns in data before a payment is made.	Can be joined with automation and help a payment system identify likely fraudulent transactions. Can also be used to automatically reject a payment when the existence of a number of known fraud or improper payment characteristics are present. *Agencies should take care to evaluate the outcomes of predictive analytics for fairness and potential bias.*	High risk. Equity issues are very likely to occur as such methods require extensive amounts of correctly identified training data and judgements about improper characteristics are subjective. Caution: Predictive identification should be only used with extreme care. If third-party vendors provide the solution, agencies are strongly recommendedto perform an independent evaluation/audit for equity concerns prior to and while using the method.
Text	Uses natural language processing tools that parse large text fields and pull out patterns or indicators, such as keywords that may indicate fraud or improper payments.	Can help review large amounts of textual data. Text analytics puts large amounts of textual data (such as from the internet) into a structured form and then can analyze strings of text to scan for red flags of fraud.	Medium-to-high risk. Equity issues are likely to arise when using text analytics (especially drawn from the internet) because of lack of true representation for groups using internet-based text data. Caution: Natural language processing methods (which are predictive methods as above) should be used with extreme care. If third-party vendors provide the solution, agencies are strongly recommended to perform an independent evaluation/audit for equity concerns prior to and while using the method.

Sources for data include, but are not limited to:

- Watchlists and excluded parties lists (e.g., Death Master File, Incarcerated Individuals, System for Award Management, etc.);
- Information obtained from meetings with investigative personnel from the IG;
- Federal and state law enforcement agencies, including Financial Crime Enforcement Network (FinCEN) data;
- Third-party data, such as credit data;
- Other Federal agency data for matching, such as employment records; and
- Agency's historic/prior improper payment data (payment mistakes and fraud) to model agency susceptibility to fraud and inaccurate agency payment processes.

Agencies can also leverage existing Federal resources for data analytics, including their agency's CDO, leading data analytics centers such as Health and Human Services (HHS) IG's Consolidated Data AnalysisCenter, the Payment Integrity Center of Excellence (PICOE),[2] and the Treasury Working System.[3] The IGcommunity should also consider leveraging the PRAC's Pandemic Analytics Center of Excellence (PACE) to augment their analytic capacities.[4]

To improve data analytics in the short-run, agencies can work with their CDO and/or CIO to:

1. Evaluate the data available and any data and data standards gaps, as well as the quality of the data collected. For example, programs can collect geographic data to discern fraud trends, identify which communities are served, and detect suspect Internet Protocol (IP) addresses.
2. Determine whether it is necessary to establish a centralized repository of data, which may include more than one program, such as information about identifiers that are fraud risks.
3. Ensure analysts have training and software needed to perform data analytics on large data sets (software ranges from basic tools such as Excel, to data visualization tools such as Power BI, to advanced regression analysis tools such as R and Python).
4. Identify staff with analytics and data science skills and prioritize these staff in hiring plans.
5. Identify clear objectives for the data searches such as eligibility criteria subject to the most errors or trends in fraud.
6. Test and validate/confirm new methods to identify risks such as fraud risks (for example, mining social media for potential fraud against a specific program) with the understanding that individual's representations on social media may be inaccurate or exaggerated.
7. Leverage existing public data that helps flag risks such as information on private identity fraud patterns collected by the Federal Trade Commission's (FTC) Consumer Sentinel Network.

[2] A community of experts within the Treasury, Bureau of the Fiscal Service, that assists agencies in reducing improper payments by creating solutions that proactively reduce improper payments, fraud risks, and fraud throughout the payment lifecycle.

[3] A centralized data and analytic service performed at Treasury as part of the Do Not Pay Initiative functions for all Federal payments. It allows agencies to perform pre-payment reviews as well as other activities such as investigation activities for fraud and systemic improper payment detection through analytic technologies and other techniques.

[4] The PACE was created by the PRAC to provide the enhanced capacity and scale necessary to oversee the $5 trillion in pandemic spending associated with COVID-19 relief. The PACE creates a leading-edge analytic platformto deliver analytic and investigation support, while leaving a long term and flexible platform for the Federal IG community when the PRAC sunsets.

8. Leverage talent in other agencies through detailees who have the technical background to analyze the data.
9. Use data visualizations of the data to improve display of the results, both positive and negative, so the public, decision makers, and stakeholders can better understand the program's efficacy and financial performance.
10. Perform periodic sampling of payments for comparison with historic/baseline improper payment types and to evaluate any improvements from automation and the data analytics components.