DAY ONE PROJECT

Strengthening the Integrity of Government Payments Using Artificial Intelligence

James E. Cook, Gordon C. Milbourn III, and Chuck Howell

November 2020

Summary

Tens of billions of taxpayer dollars are lost every year due to improper federal payments. These improper payments arise from agency and claimant errors as well as outright fraud. Data analytics can help identify errors and fraud, but often only identify improper payments after they have already been issued.

Artificial intelligence (AI) in general—and machine learning (ML) in particular (AI/ML)—could substantially improve the accuracy of federal payment systems. The next administration should launch an initiative to integrate AI/ML into federal agencies' payment processes. As part of this initiative, the federal government should work extensively with non-federal entities—including commercial firms, nonprofits, and academic institutions—to address major enablers and barriers pertaining to applications of AI/ML in federal payment systems. These include the incidence of false positives and negatives, perceived and actual fairness and bias issues, privacy and security concerns, and the use of ML for predicting the likelihood of future errors and fraud.

Challenge and Opportunity

Improper federal payments—i.e., payments that should not have been made or were made in the wrong amount—were estimated at approximately $175 billion in fiscal year (FY) 2019, representing about 4% of total outlays.[1] Since not all federal programs are required to prepare improper payment estimates, actual improper payments are likely even higher.[2] Improper payments are likely to be especially high in FY 2020 due to the approximately $3 trillion of emergency federal spending motivated by COVID-19.

Data analytics play an important role in reducing improper payments. However, many data analytic approaches only detect potential errors or fraud after improper payments have already been made. This means that many federal programs operate in "pay & chase" mode, spending considerable time and resources to recover funds that have already been paid out in error or due to fraud. This is a fundamentally flawed approach that typically results in recovery of only a portion of the misallocated funds. The challenge for agencies in efficiently mitigating improper payments is to rely more heavily on pre-pay analytics that can be embedded in payment processes as well as other predictive or prescriptive approaches.

Proactively preventing errors and deterring fraud is much better than detecting them after the fact. Artificial intelligence (AI) in general and machine learning (ML) in particular (AI/ML) can help facilitate the move "left of check". ML is especially valuable for analyzing significant amounts of data moving with high velocity through agencies' payment systems. ML can be applied to transaction data in real time to identify subtle patterns in the behavior of individuals who are

[1] Government Accountability Office (2020). *Financial Audit: FY 2019 and FY 2018 Consolidated Financial Statements of the U.S. Government*. GAO-20-315R.

[2] Office of Management and Budget (n.d.). Annual Improper Payments Datasets.

making honest mistakes or, in the case of fraudsters, trying to evade detection. ML can also be combined with modeling to predict the occurrence of systemic problems that could produce improper payments.

Plan of Action

The next administration should launch an initiative to integrate AI/ML into federal agencies' payment processes. Successful development and implementation of ML algorithms for widespread use in government payment processes will require White House leadership and coordination, collaboration across virtually all federal agencies, and mobilization of non-federal actors (e.g., state and local governments, the private sector, nonprofits, and academia). Key considerations are outlined below.

White House leadership and coordination

Integrating AI/ML into the payment processes of most or all federal agencies will require a great deal of interagency coordination and leadership from the Executive Office of the President. As such, the next administration should develop a White House strategy for advanced and innovative uses of data analytics (particularly ML) to strengthen the integrity of government payments. Further, the Office of Management and Budget should ensure that the Chief Information Officers Council's new cross-government data science training program emphasizes use of ML to address payment integrity challenges and interfaces.

Agency participation

All agencies that process payment requests should assess their payment processes for opportunities to add or expand the use of AI, particularly ML. Specifically, agencies should undertake the following activities either individually or in collaboration with other agencies facing common challenges (e.g., using analytic cells):

- Define desired results. To guide effective deployment of AI/ML in payment processes, agencies should develop a clear sense of "what success looks like". This may involve, for individual programs, factoring in the biggest root causes of improper payments as well as establishing criteria to decide if pursuing an ML solution has adequate return on investment (i.e., if the cost and risks of the novel solution are worth the potential benefits). Agencies must also make informed tradeoffs among various performance metrics (e.g., cost of false positives vs. false negatives).
- Bring to bear appropriate knowledge and skillsets. Effective integration of AI/ML solutions into government payment processes will require skilled developers, IT specialists, data wranglers, and data scientists to provide technical expertise as well as staff with programmatic and payment integrity knowledge to guide development of tailored solutions.
- Identify relevant dataset(s). An agency's payment dataset alone may sometimes be sufficient to support development of AI-/ML-based solutions to limit improper payments. More frequently, payment datasets will need to be combined with other relevant

datasets, such as beneficiary data, do-not-pay lists, watch lists, data from other government agencies, and data from commercial or open sources. Agencies should establish processes for identifying and incorporating new datasets of interest over time.

- Select analytic tools to be used. After key problems have been identified, agencies can determine what tools or IT approaches offer the best solutions. Agencies should assess whether their existing computing infrastructures are capable of implementing such solutions or whether new, specialized computing environments are needed. Agencies should establish processes for identifying and incorporating new, leading-edge tools and methods over time.
- Consider user and stakeholder needs. Agencies should determine how to validate, enhance, and communicate findings from AI/ML in a usable form. Agencies should also consider whether proposed solutions could have disparate impacts on and/or benefits for different stakeholder populations.
- Establish feedback loops. Strong feedback loops are needed to continuously improve AI/ML solutions: for instance, to identify new or evolving errors and fraud indicators as well as to assess solution performance. Strong feedback loops also allow operators to revisit and revise initial assumptions as experience is gained from operating in the real world.
- Share best practices. Agencies should routinely exchange best practices and lessons learned from efforts to integrate AI/ML in payment processes.

<u>Mobilization of non-federal actors</u>

Using policy levers such as grant programs and cost-sharing opportunities, the next administration should mobilize non-federal actors (including commercial firms, nonprofits, and academic institutions) to conduct research into payment integrity in the broader context of AI. In particular, the next administration should promote collaborative research with government entities such as the National Institute of Standards and Technology to address the following four major enablers and barriers to expanding use of ML for payment integrity:

1. Managing false positives and negatives when applying ML algorithms.
2. Addressing perceived and actual fairness and bias issues inadvertently built into algorithms.
3. Addressing privacy and security concerns about gathering and sharing sensitive financial data.
4. Increasing use of ML for predicting interactions, events, and other occurrences that are likely indicators or triggers of improper payments.

Frequently Asked Questions

What causes improper payments?

The vast majority of improper payments are attributable to agency and claimant errors. Examples of agency errors include computer programming problems that result in miscalculated payment amounts, or human misinterpretation of claim documentation. Examples of claimant errors include an individual who mistakenly believes that they are eligible for a government benefit; a taxpayer who miscalculates a refundable credit on their tax return; or a contractor who accidentally submits a duplicate invoice.

Claimant "errors" cross the line into fraud when an individual willfully does something to secure a benefit to which they are not entitled. Fraud may also be committed based on the use of stolen or synthetic identities. It is difficult to determine the precise amount of federal payments lost to fraud. Fraudsters are adept at concealing the nature of their transactions, meaning that significant amounts of fraud likely go undetected. Moreover, even improper payments that are identified as likely fraudulent are not usually labeled by agencies as fraud until the potentially fraudulent activity has been investigated and adjudicated.

What are some examples of processes where AI/ML could reduce improper payments?

Examples include:

- Mortgage loans. Lenders use ML to identify individuals who are likely to be a credit risk even though they do not have a long history of making payments. Agencies that administer government-funded housing programs could use a similar approach to proactively identify and mitigate risks.
- Credit cards. Card issuers use ML to identify potentially anomalous transactions for verification by the cardholder. Agencies that provide benefits via debit or EBT cards could similarly use ML to identify potentially erroneous or fraudulent transactions before honoring them.
- Debt collection. Collection agencies use ML to inform their strategies for determining which individuals are likely to pay and which individuals require follow-up. Agencies that use Recovery Audit Contractors to identify errors and recover improper payments might adopt similar approaches.
- Healthcare fraud. One insurance organization is using advanced AI tools to detect indicators of fraudulent activity much faster than before. Agencies that administer healthcare programs could adopt similar approaches.
- Cybersecurity. Cyber risks can manifest in high-volume "noisy" threats as well as "low and slow" campaigns that are very hard for humans to detect. Fraudsters engage in both

types of campaigns. ML is adept at identifying both activity patterns, enabling early detection and neutralization of threats.

What are the biggest concerns with leveraging AI/ML to counter improper payments?

Developing and deploying advanced ML algorithms across government payment processes would enable agencies to make better use of the massive amounts of data they already collect. However, it is important to be aware of the following concerns:

- Like other data analytic approaches, ML algorithms are not infallible. False positives and false negatives are both problematic possibilities. Identification of false positives wastes resources in the follow-up that must be performed. Identification of false negatives means that payment integrity issues may be missed.
- Unknown bias can be embedded in ML algorithms.
- Many agencies may not have the breadth and depth of expertise needed to implement ML successfully.
- Agency and programmatic "silos" may limit exchange of analytic approaches, including ML algorithms. Poor communication may result in duplicative efforts that waste time and resources.
- Agencies acting individually cannot maximize the effectiveness of using ML in predictive payment integrity analytics. Cross-agency coordination is essential.

What is the federal government already doing to ensure payment integrity and leverage AI? How can the next administration build on these efforts?

In support of the President's Management Agenda,[3] the current administration has created an interagency team to meet objectives within the Cross-Agency Priority Goal "Getting Payments Right".[4] There is an Executive Steering Committee consisting of the Office of Management and Budget and four Departments, which leads a much larger interagency team.

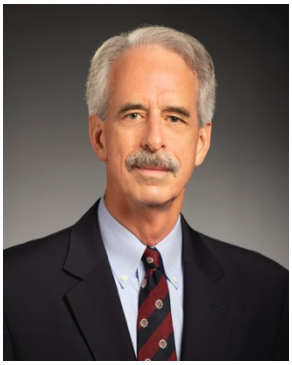
Since May 2016, the National Science and Technology Council (NSTC) has been coordinating AI research and development at the interagency level. In 2019, President Trump issued an Executive Order[5] establishing the American AI Initiative[6] to promote and protect national AI technology and innovation. A new NSTC Select Committee on AI[7] coordinates supporting activities.

[3] The White House (2018). *President's Management Agenda*.
[4] General Services Administration & the Office of Management and Budget (2020). Getting Payments Right.
[5] The White House (2019). Executive Order on Maintaining American Leadership in Artificial Intelligence.
[6] The White House (2020). Artificial Intelligence for the American People.
[7] Office of Science and Technology Policy. NSTC.

The next administration could link these existing interagency efforts to achieve the goals outlined in this paper. A simple but important first step would be a public statement that explains how AI/ML could help address pervasive payment integrity problems, and explicitly states the administration's commitment to using cutting-edge solutions to reduce improper payments.

About the Authors

Jim Cook is an advocate for data-driven, outcome-focused government. He has spent his career helping to advance improvements in Federal and State government performance, focusing on better use of technology, acquisition reform, benefits administration, and effective stewardship of public funds through high-integrity payment systems. In his current role, he leads MITRE's strategic corporate interactions with policymakers in the Executive Office of The President and Congress; and promotes development of new strategic partnerships with the non-profit policy community through MITRE's Center for Data-Driven Policy. He also fosters partnerships with private sector, academia, and other non-profit associations and foundations to address public interest challenges at the federal, state, and municipal levels. Since 1983, Jim has led major programs at the Federal and State levels focused on modernizing business and technology operations and disbursing large volumes of payments to the public. He has experienced firsthand the importance of ensuring integrity in payment programs and the processes and systems that support them. From 2007 to 2017, Cook was vice president and director for MITRE's Center for Enterprise Modernization (CEM), the federally funded research and development center sponsored by the Department of the Treasury and co-sponsored by the Department of Veterans Affairs (VA) and the Social Security Administration (SSA). He has sponsored ongoing efforts to research causes of improper payments and financial fraud and advance innovation, collaborative approaches to address them through data sharing, new public-private partnerships, and application of technology to advance prediction, detection and decision-making.



Gordon Milbourn III is MITRE's Policy Leader for Payment Integrity, focusing on fraud and other improper payment issues across the government. He began his career in 1974 with the IRS Inspection Service, serving in various audit positions nationwide for 12 years before moving to the Naval Audit Service and then the EPA Office of Inspector General. In 1999 Milbourn returned to the IRS as an audit executive with the newly formed Treasury Inspector General for Tax Administration and acted for over 2 years as the Deputy IG for Audit, before completing his federal career as the Assistant IG

for Audit at the U.S. Postal Service. Throughout his audit career he concentrated on improving government efficiency and effectiveness and combating fraud, waste and abuse, and in 2008 received the David M. Walker Excellence in Government Performance and Accountability Award. Upon retiring from federal service, Milbourn came to MITRE where he has worked with numerous federal sponsors on systems, accountability, and Payment Integrity challenges. Milbourn received his bachelor's degree in accounting from the University of Virginia and completed the DOD Program Manager's Course in leadership and systems engineering. He is a Certified Internal Auditor, a Certified Fraud Examiner, and a Certified Financial Crime Specialist.

Chuck Howell is focused on adapting tools and techniques from high-assurance systems engineering and from various sectors' risk management frameworks to apply to consequential AI (particularly, machine learning) systems. These tools and techniques can help organizations address concerns about AI system properties such as fairness, operational risk, safety, and credibility. Chuck has over 30 years of experience working in High Assurance Systems Engineering and AI. He previously held roles at Mitretek, Sun Microsystems, Reliable Software Technologies, Verdix, and Computer Sciences Corporation. He was a member of the Institute of Electrical and Electronics Engineers (IEEE) Software Engineering Body of Knowledge Industrial Advisory Board. Chuck chaired the First Annual Assurance Case Workshop in Florence, Italy, and co-chaired the Fall 2015, 2016, 2017, and 2018 Association for the Advancement of Artificial Intelligence (AAAI) national workshops on Cognitive Assistance in Government and Public Sector Applications. He is co-author of the book *Solid Software* (Prentice Hall, 2001). Chuck is a Senior Member of the IEEE and a member of the AAAI and the Association for Computing Machinery.

About the Day One Project

The Day One Project is dedicated to democratizing the policymaking process by working with new and expert voices across the science and technology community, helping to develop actionable policies that can improve the lives of all Americans, and readying them for Day One of a future presidential term. For more about the Day One Project, visit dayoneproject.org.