FEATURE

Investigative analytics

Leveraging data for law enforcement insights

Alan Santos, Isaac Jenkins, Joe Mariani, Mike Gelles, and Alex Mirkow

Technological advances can improve law enforcement agencies' investigative methods, but investigators often struggle to decode the massive wall of data. Here's where digital tools and techniques can help.

CAPTAIN JOHNSON HAS a problem. As the head of investigations for her agency, she and her personnel are charged with both drug and death investigations. The past few months have witnessed activity on both fronts: an alarming rise in drug overdose deaths from opioids. Her team faces questions about the source of the drugs, with inconclusive toxicology results from the medical examiner.

Where can she begin? First, Captain Johnson asks her team to determine which criminal actors or organized criminal groups are involved in the diversion, distribution, and trafficking of similar opiates in the region. Her narcotics team finds a number of leads but, in fact, have too much data to analyze—from prescription drug monitoring program (PDMP) data and public health records to financial records and social media. Homing in on the exact threat is like finding a needle in a haystack. Many days and many cups of coffee later, the investigators are overwhelmed and about to give up. But every day the problem grows, and Captain Johnson needs results.

Individuals and organizations involved in criminal and illicit activity are becoming increasingly sophisticated. Bad actors are harnessing the power of new technologies as quickly as they're being invented. The good news is that technological advances can also improve the investigative methods of law enforcement agencies. In the words of one homicide supervisor, "In 1993, there were no social media or surveillance cameras. We just had blood typing that was used to help identify people from a crime scene. Now DNA is used, which can identify an individual using his DNA to one in a billion, quadrillion, or greater."[1]

These benefits bring their own challenges. The digital world presents investigators with a massive, seemingly impenetrable wall of data. For example, in one year alone, a single FBI investigation collected six petabytes of data—the equivalent of more than 120 million filing cabinets filled with paper.[2] When targeting illicit activity, law enforcement and intelligence agencies must not only grapple with a wide variety of new and unfamiliar data sources, but also need to make better use of the data they already collect. Without effective data analysis, law enforcement agencies will struggle to counter the criminal actors they are charged with targeting.

This kind of data analysis is no easy task for law enforcement agencies, with their resources mostly dedicated to core law enforcement functions of keeping communities safe. Most law enforcement personnel don't have training in data science and digital research, which are generally required to perform advanced analytics. To get access to data-driven insights, many police departments have turned to data science volunteers from local universities for help, and even amateur genetics enthusiasts.[3] Resource constraints, however, can limit an agency's ability to bring outside experts in on a consistent basis, possibly resulting in a daily struggle to find relevant information for investigations, even though data is now everywhere.

But new investigative tools and techniques can help law enforcement agencies overcome resource constraints and come to grips with massive digital data. Artificial intelligence (AI), open source data management tools, predictive analytics solutions, and social media exploitation capabilities are helping many investigators and operators make sense of mountains of data. New tools can help

make previously unseen connections between information and identify where key information is lacking, while new sources can provide access to reams of data, and new partnerships can help exploit the data. But most of all, these practices can give investigators more of what they really need: time. The FBI's Counterterrorism Division's efforts to streamline its data systems has enabled a 98 percent reduction in manual work for analysts and a 70 percent cost reduction.[4] Leveraging these tools and approaches can help investigators return to where they belong: not writing code or searching through databases, but rather tracking criminals and keeping our communities safe.

The problem is not the data but the approach

Captain Johnson and her team's data problem is a common one. Historically, investigative analysis and targeting have often been guided by what information an agency has access to. Just as with a hammer every problem looks like a nail, often every investigation begins with whatever data is most readily available. But there may be critical data sources that are missing, too hard to access, or too complicated to analyze, leading to blind spots in an investigation.

To simplify the investigator's data problem, there are three fundamental questions every organization should answer: 1) What information do we have? 2) How do we make sense of it all? And 3) What key pieces of information are we missing?

Ironically, the first step in the solution to being overwhelmed with data is often to create an even larger pool of data. That is because many organizations have huge volumes of data but are unable to use it effectively, due to computing and integration challenges. Outdated and insufficient computing power and platforms hinder advanced analysis. Silos, both within and outside the organization, inhibit meaningful access to integrated data that

may help an investigation. Breaking down these walls is typically a key first step. For one police department in the United Kingdom, that meant bringing together intelligence data sets, command and control systems, and operational data streams, to give officers a unified picture on a single screen.[5]

Then this larger, fuller pool of data can be sorted and managed by a small team of professionals. By bringing data sources together, many investigators can be supported by a team of data scientists and analysts that can help them make sense of the data that they do have through "exploratory data analysis," a standard process by which data sets are measured for breadth, depth, and quality (explored later in this article). Important data sources include in-house data maintained by law enforcement and intelligence agencies; commercial data sources; and open sources such as social media activity, property records, criminal histories, professional licenses, medical databases, and a host of other sources.

Once an organization fully understands the breadth of data it currently has access to, it can begin pondering the next question: What key pieces of information are we missing? Starting with a target- or problem-centric approach can help investigators create a holistic picture of persons, places, and objects of interest from a large pool of data. That holistic picture, in turn, can begin to show where key holes in the analysis lie. Perhaps there is no information on how two known associates communicate, or where a suspect receives his mail. Knowing these gaps can help guide future collection and surveillance to build the right case much faster.

Five steps to harnessing data more effectively

Knowing there is a more efficient way to look at data, what should agencies do to achieve results? A few key steps can help agencies of all sizes come to grips with the data challenges presented by complex investigations.

1. **Take stock of your current data sources and how they are utilized and managed.**[6] Often critical to this effort is performing a detailed and honest data audit, including who controls access to which data sources and which can be used together. One data analysis company stated that their law enforcement clients often find that they already have more data than they think. Simply making this data organized and available to investigators can jump start cases.[7] Achieving greater data integrity upfront can enable greater insight downstream. For Captain Johnson's department, that process might include examining internal databases, such as the records management system, and external databases, such as the relevant PDMP, and ensuring that data can be accessed and integrated.

2. **Find new data sources.** As investigators apply the target-centric approach to data the organization has now pooled and organized, they can begin to see where there are common gaps in knowledge. Everyone leaves a digital trail, including criminal actors and organizations, and this can allow investigators to home in on what's missing. One new source of data can connect dots and break open cases. For example, one detective for the Orlando Police Department stated that simply beginning to treat opioid overdoses as crime scenes rather than accident sites led them to look through the phones of overdose victims. Cross-referencing the contact lists from different overdoses quickly led to the identification of a handful of dealers.[8]

   New sources of data can be anywhere from the databases of other agencies to lying in the street next to victims. Identifying data sources of value from other agencies, commercial or open sources, or even publicly available social media data, can fill critical gaps in an agency's investigations.

3. **Determine how to capture and store collected data.** A massive pool of all available data is not helpful to investigators if it cannot easily be used to find the exact information needed. This requires that the data be structured, organized, and stored in specific ways. The specific set of data management tools needed will depend on the exact volume and nature of the data an agency is dealing with. For example, powerful graph and novel data structures can allow organizations to create data stores that fuse different types of data into a single descriptive visualization of how specific persons, places, and objects relate to one another.[9] The appropriate data structure can also allow new tools and capabilities to be added on as needed. For example, with a well-defined data structure, an agency could use predictive modeling to predict where certain types of crime are likely to take place, allowing more efficient use of resources. And agencies are seeking creative ways to store their data: In Utah and Oregon, some police departments are finding that they can store data together in a consortium, leading to reduced costs and increased insight.[10]

4. **Address the need for analytics skill sets in personnel, more capable computing systems, and better integration across data stores.** Investigators are not data scientists, nor should they be. But agencies also often lack dedicated training or career tracks for analytics. Combine that with the fact that siloed legacy systems are often inadequate for processing and analyzing the larger, more complex data sets produced by modern investigations, and you likely get a recipe for poor use of data. This problem is not unique to law enforcement. Many commercial organizations face exactly the same problem when trying to identify potential fraud within their organizations. As a recent Deloitte report found in a forensic financial investigation, "When data volumes are small, basic analytical skills and spreadsheet programs might be adequate to handle preliminary analysis of structured data from enterprise systems and other software applications, as well as unstructured data such as emails, texts, and voice

recordings. But when that volume grows into the millions, analysis can require technology, advanced analytics, and forensic skills that aren't readily available within many organizations."[11]

5. **Think about the type of reporting and outputs that are best suited to help investigators and analysts in the most efficient way possible**. The days of handing someone a ream of paper or a complex set of spreadsheets—and leaving it at that—are over. While the specifics of data reporting and visualization will necessarily vary from case to case, departments should put some thought into how officers typically use data and what the most seamless way to consume it would be.

How data analysis drives investigations in the real world

Technologies are not just new, interesting tools that improve existing processes; rather, they can open up entirely new ways of doing investigations. Where today, many investigators spend large amounts of time trying to find the right data, potentially leaving only a short time for analyzing it and piecing it all together, in the future, investigators should be able to rely on new data tools to more rapidly find data, allowing them to spend the majority of their time analyzing data. In 1970, a study of one medium-sized police department found that 50 percent of a patrol officer's time was spent on tasks directly related to crimes (with the remaining half spent on administrative tasks).[12] Some four decades later, a 2012 look at six different police forces found that this percentage has risen to 80 percent—a 60 percent improvement.[13]

The massive impact from using these seemingly simple tools can be seen by revisiting the story from the introduction: Faced with the rash of overdose deaths in her community, Captain Johnson convinces her agency to develop a robust analytics capability. First, her agency assesses its existing data sources, performing an audit that includes storage, management, access, and utilization. Data analysts perform an exploratory analysis to understand their data quality and comprehensiveness. Realizing that they lack consistent access to critical data, including communications, prescriptions, and geotagged business data, the team compiles publicly available sources, data subscriptions, and social media data from nearby locations. Agency leaders then agree on a plan to ingest, process, and store data in the cloud, enabling instant access to the latest computing systems and eliminating data silos. They hire a couple of data scientists and provide training for interested officers to develop their own data skills.

Captain Johnson's team was collecting information on pharmacies believed to be illegally supplying local addicts with prescription opioids. Equipped with new capabilities around the ingestion, processing, and storing of bulk data sets, along with new personnel resources with advanced analytics skill sets, Captain Johnson's team is able to establish a risk ranking of the practitioner prescribers and patients found in the pharmacy data and can quickly cross-reference it against available drug overdose information. Upon doing this, they discover that a number of the overdose victims had received prescriptions and/or had opioids dispensed from some of the pharmacies in question.

Next, through the bulk ingestion and exploitation of open-source property record data, a capability the team did not have in the past, it realizes one of the high-volume prescribing practitioners appears to have an excessively large number of financial assets. In addition, analytics professionals on the team determine that the husband of the practitioner in question is linked via social media to a known dealer involved in heroin and fentanyl trafficking, both illicit opioids.

Investigators then track available social media profiles of recent overdose victims and determine that many of the victims had connections with both the known drug trafficker and the practitioner's husband. By taking the property addresses of both the practitioner's husband and the drug trafficker and by cross-referencing them against commercially

available foreign-based shipping records, investigators determine that a number of packages have been shipped to the properties from a known industrial park in China.[14]

Captain Johnson and her team then work with their representative on a local federal drug task force to connect to a limited access database and ingest sharable information, which shows that the industrial park is home to an unregulated pharmaceutical facility. Another agency is working on spotting the shipments before they enter the United States, and a number of other local agencies are fighting a similar problem. Now, investigators have enough information to pursue warrants for enforcement activity.

With these new tools and methods, investigators are no longer overwhelmed and hoping for a break. They are now actively targeting a criminal network and getting to key targets faster. Better data sources, coupled with targeting expertise, have allowed Captain Johnson and her investigators to quickly obtain better results and more significant investigative outcomes.

Setting up systems for faster, easier data collection

We've seen how investigators across law enforcement and intelligence organizations use sophisticated computer programs to document linkages between people, places, and things. While this represents a leap forward from the analog era, it still requires painstaking data collection before an investigator can plot, analyze, and understand a network of actors.

What if this process could be partially automated, to enable analysis from day one? As with Captain Johnson's department, setting systems up for an analytics capability starts with understanding the landscape of available data sources, from in-house data sources to social media activity. Then, the data sources have to be able to speak to each other. That means designing a system that organizes, formats,

and stores data accessibly. For example, by formatting records so that names, dates, and locations are easily searchable, an investigator can reach across dozens of data sources to compile histories, profiles, and activities automatically.
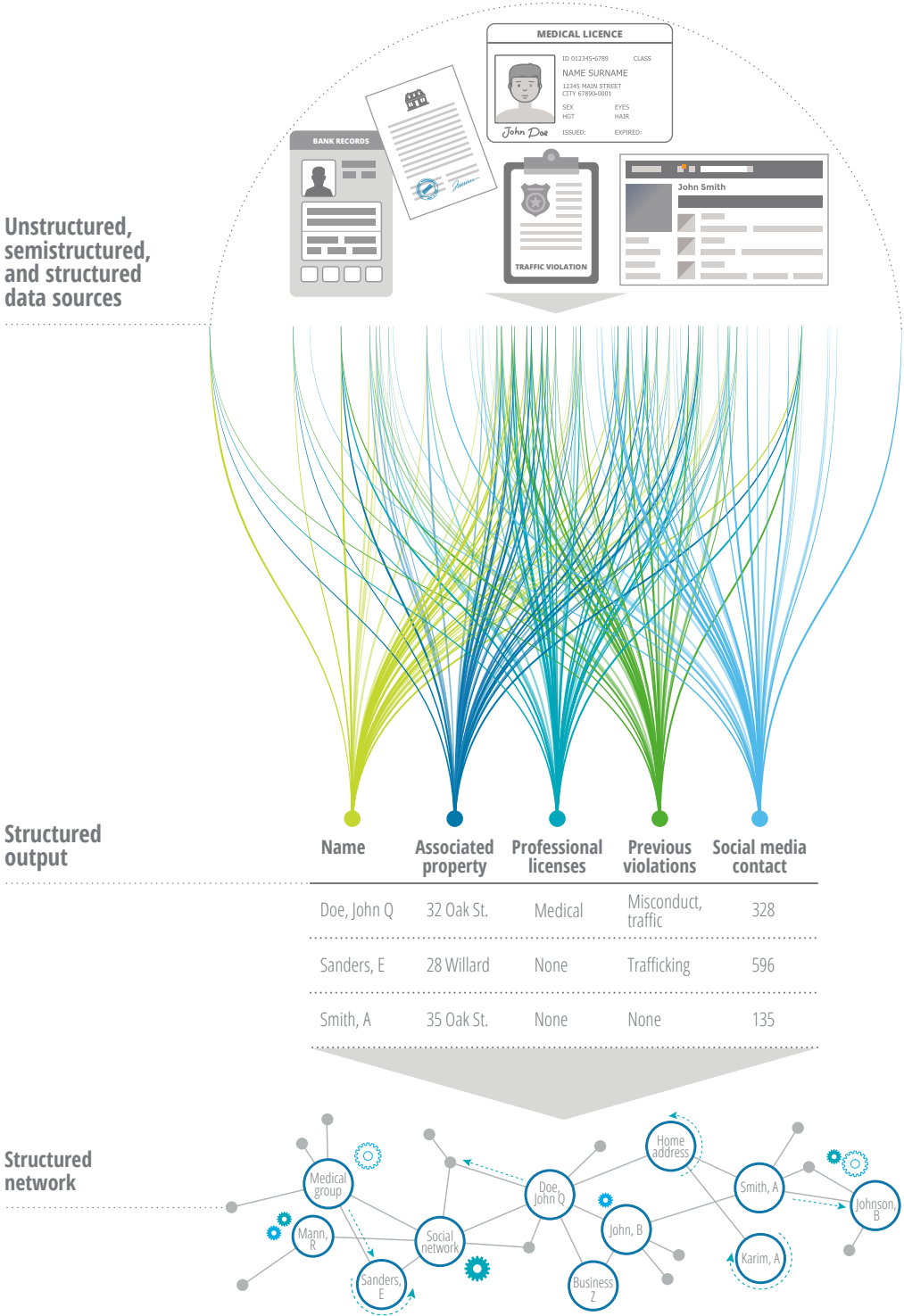
In figure 1, various data sources are compiled into a single, structured output. Inputs include unstructured data such as scans from documents; semistructured data such as websites and social media; and structured data including property records, professional license records, and criminal records. The output is a single, multidimensional profile of an individual of interest, just as Captain Johnson might be interested in finding information about an individual suspected of overprescribing opioids.

But investigators rarely need just one person's profile; instead, they often need to understand multiple forms of connections, from known relationships, to presence in the same geographies or businesses, to shared connections on social media. Using automated data analysis to find links among multiple structured output profiles, investigators can develop networks of actors and understand activities and behaviors across a group (see figure 1). For example, natural language processing can perform what is called "named entity recognition." In layman's terms, this means that AI can use contextual clues to tell the difference between the State of Georgia, the country, and the first name, for example.[15] For investigators, this can mean differentiating between a suspect and an innocent citizen with the same name.

These tools all serve to increase the accuracy of investigations and helping police find the right conclusions faster. Instead of spending weeks or months to develop a detailed link chart, they can leverage both stored data and live data streams to develop an advanced starting point for their analysis, allowing them to more quickly get results. The increased speed and accuracy of investigations can be attested by the Durham, North Carolina police department. That department used natural language processing to study incident reports and

FIGURE 1

Semistructured data sources can become structured profiles of people or places, which, in turn, can offer networks for investigators to analyze

Unstructured, semistructured, and structured data sources

Structured output

Name	Associated property	Professional licenses	Previous violations	Social media contact
Doe, John Q	32 Oak St.	Medical	Misconduct, traffic	328
Sanders, E	28 Willard	None	Trafficking	596
Smith, A	35 Oak St.	None	None	135

Structured network

Source: Deloitte analysis.

find patterns in criminal activity. With that information, they were able to identify areas with a high incidence of specific types of crime, allowing them to deploy the right assets to the right areas ahead of time. The end result was a 39 percent drop in violent crime in Durham from 2007 to 2014.[16]

Aligning the organization with new capabilities

It's not always easy to change the way an agency goes about conducting its investigations. But as criminals become increasingly sophisticated, law enforcement and intelligence professionals have to harness all available data or risk being left in the dark. To be sure, new solutions, tools, and methods will not take root overnight. They need continued support and may require new technical and targeting capabilities for investigators, operators, and analysts alike.

Organizations have a natural inertia to keep doing things in the same ways, with the same tools. The more changes that better technologies force upon an organization, the more likely they are to be resisted.[17] However, a few concrete steps can help

any agency get started on the journey to investigative transformation:

- **Leaders championing the change**. Having an active and involved champion can increase the likelihood of success for technology implementation.[18]

- **Bringing technology and investigative staff together**. Technology should not be developed in a bubble. By working together in cross-functional teams, technology staff can develop the right tools faster, and investigators can be sure they are getting tools that they can actually use.[19]

- **Communicating with partners**. No organization, however large, has all the data it needs. Therefore, interagency partners can be crucial to long-term success. Continuous and open communication with these partners regarding agency goals and data needs can increase the likelihood that they share needed data when you need it most.[20]

While these steps are just the beginning, they can offer a path to a new era of investigations, one where law enforcement is not overwhelmed by data, but can harness it for good.

Endnotes

1. Christopher Barling, "I'm a homicide detective in the LAPD. What do you want to know?," *Guardian*, July 25, 2013.

2. For FBI statistic, see Zach Noble, "Why law enforcement agencies need to share data," FCW, May 6, 2016. For Petabyte analogy, see Jesus Diaz, "How large is a petabyte," Gizmodo, July 8, 2009.

3. Drake Bennett and Kristen V. Brown, "Your DNA is out there. Do you want law enforcement using it?," Bloomberg Businessweek, October 27, 2018.

4. Billy Mitchell, "The FBI is looking to the cloud to stop the next terrorist attack," FedScoop, November 28, 2018.

5. Sam Trendall, "How the police wants to use AI and analytics to 'adopt a public-health approach to crime,'" PublicTechnology, January 15, 2019.

6. John Ferraioli and Rick Burke, *Drowning in data, but starving for insights: Starting the digital supply network journey with legacy systems*, Deloitte Insights, April 11, 2018.

7. Author conversation with a leading data organization and analysis company, January 9, 2019.

8. John Kennedy and Ed Michael, "Opioid crisis in America: From digital clues to a murder conviction," Cellebrite training webinar, November 7, 2018.

9. International Association of Crime Analysts, "Social network analysis for law enforcement," February 2018.

10. Andi Mucklow, "*Law enforcement consortiums*: The new face of records management," *Central Square*, August 15, 2018.

11. Deloitte, *Overcoming data challenges in forensic investigations*, 2018.

12. John A. Webster, "Police task and time study," *Journal of Criminal Law Criminology and Police Science* 61, no. 1 (1970): DOI: 10.2307/1142102.

13. Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services, "Taking time for crime: A study of how police officers prevent crime in the field," September 27, 2012.

14. Matt Sledge, "New Orleans, Chinese officials announce international investigation into fentanyl ring," *New Orleans Advocate*, August 29, 2018.

15. Christopher Dozier et al., "Named entity recognition and resolution in legal text," chapter from *Semantic Processing of Legal Texts* (Berlin: Springer, 2010).

16. Kevin Ebi, "How Durham, N.C. fights crime with data—and wins," Smart Cities Council, September 17, 2014.

17. Frank Geels, "Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and case study," *Research Policy* 31, no. 8 (2002): pp. 1,257–74.

18. E. W. T. Ngai, C. C. H. Law, and F. K. T. Wat, "Examining the critical success factors in the adoption of enterprise resource planning," *Computers in Industry* 59, no. 6 (2008): pp. 548–64.

19. Gerald C. Kane et al., *Achieving digital maturity: Adapting your company to a changing world*, Deloitte Insights, July 13, 2017.

20. Joe Mariani, "Leading to chaos: A conversation with General Stanley McChrystal," *Deloitte Review* 19, July 25, 2016.

About the authors

ALAN G. SANTOS is a specialist leader for Deloitte & Touche LLP working with both the Forensics & Investigations and the Regulatory & Compliance Service practices. Santos served as a member of the US government's Senior Executive Service (SES) at the US Drug Enforcement Administration (DEA) in a regulatory and law enforcement leadership position, and also has considerable private sector experience working with commercial pharmaceutical and life science clients as the CEO of two companies in the health care consulting and services space.

ISAAC JENKINS is a research manager with Deloitte's Center for Government Insights. Jenkins' research focuses on technology and strategy in commercial, defense, national security, and development contexts. His experience includes defense industrial policy development, work as a defense consultant, conducting research with leading think tanks, education, and international development.

JOE MARIANI of Deloitte Services LP is a research manager with Deloitte's Center for Government Insights. Mariani's research focuses on innovation and technology adoption for both commercial businesses and national security organizations. His previous experience includes work as a consultant to the defense and intelligence industries, high school science teacher, and Marine Corps intelligence officer.

DR. MICHAEL GELLES is a director with Deloitte Consulting LLP's Federal practice, consulting in the areas of law enforcement, intelligence, and security. Gelles is a thought leader and widely published author on critical national security issues such as insider threat, security processing, secure workforce, asset loss, exploitation, sabotage, and workplace violence. Previously, Gelles served as a naval officer and the chief psychologist for the Naval Criminal Investigative Service.

ALEX MIRKOW leads Deloitte Consulting LLP's relationships with the US Department of Justice and the US courts. He specializes in advising federal, international, and private sector clients in the area of strategy and operations. In particular, Mirkow brings more than 20 years of experience in strategic planning, financial modeling, performance measurement, and evaluating operating results. Previously, Mirkow led Deloitte's relationship with the Transportation Security Administration (TSA) and served as the national sponsor for Deloitte's Hispanic Network (HNet) business resource group (BRG).

Acknowledgments

About the Deloitte Center for Government Insights

The Deloitte Center for Government Insights shares inspiring stories of government innovation, looking at what's behind the adoption of new technologies and management practices. We produce cutting-edge research that guides public officials without burying them in jargon and minutiae, crystalizing essential insights in an easy-to-absorb format. Through research, forums, and immersive workshops, our goal is to provide public officials, policy professionals, and members of the media with fresh insights that advance an understanding of what is possible in government transformation.

Contacts

Alex Mirkow
Principal
Deloitte Consulting LLP
+1 571 814 7610
amirkow@deloitte.com

Alan Santos
Specialist leader
Deloitte & Touche LLP
+1 571 858 1963
alsantos@deloitte.com

William Eggers
Executive director
Center for Government Insights
Deloitte Services LP
+1 571 882 6586
weggers@deloitte.com

Deloitte's Analytics and Information Management practice helps clients design, build, and run insight-driven organizations by helping to maximize the potential value of analytics and information to deliver operational excellence, new products and services, competitive agility, and growth. Analytics and Information Management services offer an integrated approach with our broad range of solution offerings that help shape strategy, manage data, deliver information, improve performance, optimize insights, amplify intelligence, build capabilities, and manage environments. Please contact the authors or read about our **AIM services** to learn more.

Deloitte.
Insights

Sign up for Deloitte Insights updates at www.deloitte.com/insights.

Follow @DeloitteInsight

About Deloitte Insights

Deloitte Insights publishes original articles, reports and periodicals that provide insights for businesses, the public sector and NGOs. Our goal is to draw upon research and experience from throughout our professional services organization, and that of coauthors in academia and business, to advance the conversation on a broad spectrum of topics of interest to executives and government leaders.

Deloitte Insights is an imprint of Deloitte Development LLC.

About this publication

This publication contains general information only, and none of Deloitte Touche Tohmatsu Limited, its member firms, or its and their affiliates are, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your finances or your business. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser.

None of Deloitte Touche Tohmatsu Limited, its member firms, or its and their respective affiliates shall be responsible for any loss whatsoever sustained by any person who relies on this publication.

About Deloitte

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as "Deloitte Global") does not provide services to clients. In the United States, Deloitte refers to one or more of the US member firms of DTTL, their related entities that operate using the "Deloitte" name in the United States and their respective affiliates. Certain services may not be available to attest clients under the rules and regulations of public accounting. Please see www.deloitte.com/about to learn more about our global network of member firms.