Informatica

Solution Brief

Data Management for Crisis Response and Recovery

Key Benefits

- Validate and enrich address, email, and phone data
- Automatically catalog the entire data landscape
- Intelligently profile data to find the best-quality data at scale
- Leverage more than 65 pre-built data quality accelerators, in the cloud or on-premises
- Rapidly ingest petabytes of data from files, databases, or streaming sources
- Ensure compliance with CCPA/GDPR
- Detect and protect sensitive PII data

Establish Capabilities to Ensure Future Resilience

The success of an organization's response to a crisis hinges to a large degree on its ability to manage the accompanying deluge of data. Whether you're responding to a natural disaster, an economic collapse, or a disease outbreak, the capability to leverage data as an asset is critical. For example, in a crisis such as the COVID-19 worldwide pandemic, government agencies, healthcare organizations, and higher education institutions must contend with data on topics ranging from stimulus fund disbursement to contact tracing to helping students safely return to campus.

Having the appropriate tools in place to productively manage an influx of data from a wide variety of sources can help organizations coordinate a timely, integrated, and wide-reaching response for the crisis at hand, smooth the path to recovery, and prepare for future incidents. Organizations must establish a way to collect and store data, as well as ensure that everyone from data scientists to policy makers are conducting analyses and making decisions based on high-quality data.

Ensuring That Data Is Fit for Use

Using spreadsheets and non-digital, manual mechanisms for reporting purposes increases the risk that you will feed delayed or inaccurate data into the prediction models, artificial intelligence, advanced analytics tools, and dashboards that are used to make policy. As a result, these policies may be based on data that is out of date or incomplete, leading to consequences such as delayed response times, improper allocation of resources, or an erosion of public trust.

Putting in place methods to ensure data is fit for use helps counter such risks. Instituting best practices for discovering, collecting, cleansing, governing, and cataloging information and standardizing processes across organizations results in cleaner data sets. This in turn helps assure organizations that the quality, completeness, and accuracy of data is high enough to deliver insights that can be trusted for use in decision-making.

In times of crisis, such standardization may not be possible at the ground level. For example, during the COVID-19 pandemic, a diverse array of organizations needed to perform contact tracing. These included government agencies whose efforts may involve highly manual and labor-intensive methods; places of employment and higher-education institutions, who may deploy Bluetooth-enabled applications; and healthcare providers, who may deploy their own custom tracing solutions. Modern data management solutions enable organizations to reconcile this disparity of collection methods, data sources, and formats, automatically recognizing and eliminating outliers and duplicate data.

In addition, data that is collected may contain corrupt or poor quality information. It's important to correct inaccuracies and fill in the blanks if some entries are incomplete. Modern data management systems also include data cleansing tools that leverage statistical analysis to review data sets based on predefined parameters. Any data that falls outside those parameters is set aside for human review. Some systems utilize machine learning and artificial intelligence (AI) for data cleansing, speeding up the process and automatically compensating for incomplete data sets using statistical processes such as imputation.

By automating data quality, data governance, and metadata management, organizations can get a better handle on the vast amounts of structured and unstructured data flowing in—including IoT data from devices and sensors—and can more effectively develop and execute a comprehensive plan for response, recovery, and future preparedness.

Key Capabilities

Depending on your current state, a thorough data management crisis response strategy could require an extensive digital transformation, especially if you're relying on siloed, legacy applications and systems. That said, there are steps that any organization can take to quickly get a data management program off the ground. For example, providing access to real-time information is crucial for informed decision-making. In crises such as a natural disaster or a pandemic, conditions can change rapidly. The ability to assess the allocation and distribution of resources to meet critical needs such as shelter or appropriate medical care depend on having a 360-degree view of a situation as quickly as possible.

Additional key data management capabilities include the following:

Automatically Discover and Prepare Data

Validate and enrich contact data, including citizen, employee, and business addresses and phone numbers, and instantly verify that email addresses are not only accurate but also safe to send to, with autocorrelation and corrections. Use automated black list and spam trap detection features to help ensure email addresses are legitimate. Leverage global geocoding to accurately pinpoint latitude and longitude to develop an understanding of trends, hot spots, and other issues. Informatica Data as a Service (DaaS) allows you to easily identify whether phone records correspond to mobile or landlines, which is critical for enrolling people to receive SMS alerts. API-based high-volume SMS notification services with support for both inbound and outbound messages and a built-in opt-in or opt-out flow enable you to send SMS messages and alerts in a timely manner.

Standardize and catalog all data—including structured data such as information contained in spreadsheets and unstructured data such as written documents—with context and glossaries. Streamline search and retrieval efforts by automating the discovery and understanding of your data across cloud and on-premises systems using artificial intelligence and machine learning capabilities with an Enterprise Data Catalog. Use intelligent profiling to find your best-quality data at scale with Google-like search capabilities.

Ingest and Cleanse Data at Scale

Rapidly ingest petabytes of data and integrate data to and from files, databases, or streaming sources, including on-premises and multi-cloud environments such as Microsoft Azure, Amazon Web Services, Google Cloud Platform, and more. Easily support new and complex integration patterns and achieve high performance, reliability, and near-universal connectivity for mission-critical processes.

Leverage the Informatica Data Quality Accelerator for Crisis Response. With more than 65 pre-built, re-usable data quality rules that work with both Informatica Data Quality and Cloud Data Quality, the accelerator includes support for rapid data cleansing to speed data handling both on-premises and in the cloud. In cases where data quality rules don't cover your requirements (such as missing values) or sensitive data merges, sophisticated exception management will distribute tasks throughout your team to ensure an issue is solved.

Data Quality Accelerator for Crisis Response	
Healthcare	Parse comorbidities from diagnosis strings
	Parse ICD codes from strings
	Standardize healthcare facility names
	Parse healthcare facility type from facility name
Geospatial	Calculate distance from two geocoordinates
	Identify country
	Standardize country name
	Check to see if a point is in the polygon
Address	Validate USA county
	Validate USA state
	Validate USA zip code
Email	Parse email from string and validate
	Validate email address format
Phone	Parse USA phone number
	Standardize USA phone number
	Validate area code of USA phone number
	Parse standardize validate USA phone number

Name Handling	Identify suspect names
	Convert nicknames to formal name
	Parse multiple names from string
	Parse name FML / LMF format
Contact Validation and Dedupe	Validate USA drivers license
	Household USA contact data
	Dedupe USA contact data
	Standardize race
	Validate gender
Dates and Times	Check date completeness
	Parse date from string
	Validate and standardize date
	Days between dates
Numerical	Validate positive or negative number
	Validate number
	Parse strings from numbers
String Parsing	Parse number at end or beginning of string
	Validate positive number
	Remove non-printable characters and punctuation
	Convert diacritic characters to normal

Figure 1: The Informatica Data Quality Accelerator for Crisis Response enables organizations to dramatically reduce the amount of time they spend executing basic data handling tasks.

Master and Relate Data to Create a Single Source of Truth

Combine transaction, interaction, and other citizen data types and manage billions of records across all data sources in existing data lakes or Hadoop environments. For unstructured data, use natural language processing to infer citizen, employee, student, or patient attributes like health risk, prognoses, lifecycle events, and more. With Informatica Master Data Management solutions, you can link citizen, patient, or claimant data, as well as social data, and create relationships using machine learning, right out of the box. Visualize relationships, contact points, families or households, social networks, and organizational hierarchies at scale using a graph data store.

Organizations can also ensure compliance with CCPA/GDPR and detect and protect sensitive personal identification information (PII). With data security solutions such as Informatica Data Masking, you can de-identify, de-sensitize, and anonymize sensitive data from inadvertent release or unauthorized access for application users, mission intelligence, application testing, and outsourcing.

Key Benefits

Provide a Trusted Foundation for Decision-Making

Ensuring the quality and accuracy of data gives organizations confidence that their decisions are based on the best information available. Well-intended but misguided decisions based on low-quality data can erode public trust and result in the misallocation of resources, such as life-saving medicine or protective and testing equipment being improperly prioritized and distributed. It's also paramount to ensure that data is complete so that no gaps exist between data that is supposed to be collected and what actually is collected. Using automated means to eliminate duplicated data sets can also help to make the best use of potentially limited bandwidth and data storage capacities.

Optimize the Execution of Vital Tasks

When carrying out crucial tasks such as contact tracing during a pandemic, organizations can leverage automated verification and enrichment of address data collected for the infected individual—as well as the individuals and locations potentially exposed to the infected person—to help ensure that they can successfully contact and check in with persons of interest. Similarly, email verification and phone validation also help reinforce trust that important messages and details are going to reach persons of interest. By identifying phone types, organizations can pinpoint which individuals could be enrolled in an SMS alerts-and-notifications program for efficient messaging of relevant details.

Organizations can also create a real-time comprehensive risk assessment. For example, they can generate heat maps of affected areas and track data points, such as whether persons of interest have been tested for a disease, whether they've traveled, and what their quarantine status is, as well as notify other citizens of exposure risk. They can also manage at-risk individuals across services by providing physicians, insurance companies, higher education institutions, and other organizations with a minimum shared profile and deliver better support based on all needs.

Guard Against Fraud and Protect Citizen Privacy

In a crisis, governments typically disperse emergency funds as part of their response, setting up reporting and monitoring requirements to guard against fraud, abuse, and waste. Recipients include citizens, health organizations, educational institutions, and businesses to provide crisis response resources, replace lost income, or help maintain solvency during the crisis, as well as to spur post-crisis recovery. The ability to search and retrieve, in real time, all records associated with a recipient despite any errors or variations that might exist—within either the search record or the data that the agency already has on file—enables users to see a complete profile of the recipient that includes data such as address, geographic location, phone numbers, past benefit claims, audits, or investigations. This 360-degree view helps reduce the risk of making poor decisions based on incomplete data. By reducing waste and fraud, government agencies can also bolster public confidence that tax dollars are being disbursed appropriately to eligible beneficiaries.

About Informatica

Digital transformation changes expectations: better service, faster delivery, with less cost. Businesses must transform to stay relevant and data holds the answers.

As the world's leader in Enterprise Cloud Data Management, we're prepared to help you intelligently lead—in any sector, category, or niche. Informatica provides you with the foresight to become more agile, realize new growth opportunities, or create new inventions. With 100% focus on everything data, we offer the versatility needed to succeed.

We invite you to explore all that Informatica has to offer—and unleash the power of data to drive your next intelligent disruption.

Having a 360-degree view of a citizen's data also makes it easier to respond to critical requests—such as Freedom of Information Act (FOI) or court-ordered—and inform other departments of status changes. It enhances the ability to notify citizens of benefits changes and new programs, as well. And by executing security measures to shield sensitive data, government agencies, healthcare organizations, and higher education institutions can reinforce public trust and encourage cooperation with public safety measures.

Next Steps

The capabilities outlined here can serve as a starting point for organizations to design and execute an initial crisis response, recovery, and preparedness data management strategy. To help ensure future resiliency, organizations should pursue a comprehensive digital transformation that incorporates intelligent, automated data management. Find additional guidance here.

Worldwide Headquarters 2100 Seaport Blvd., Redwood City, CA 94063, USA Phone: 650.385.5000, Toll-free in the US: 1.800.653.3871

IN17_1120_03911