

# Unlock the potential of data with a modern data management platform

Segmented data stored in disparate systems prevents proactive incident response and effective law enforcement.

Red Hat's open and trusted data platform provides the flexibility and cross-agency access to achieve mission success.

## Introduction

The amount of data that is being collected, produced, processed, and used worldwide is growing at an exponential rate. The World Economic Forum<sup>1</sup> predicts the number will reach 44 zettabytes<sup>2</sup> by the end of 2020. Although this is a global outlook, a similar data footprint story is unfolding within the enterprise. Collecting data has become easier. Storing data has become cheaper, and computing power has become more readily available. The challenge of managing data volume, velocity, veracity, variety, and value is compounded further when you start to bring together structured, semi-structured, and unstructured data.

Not unlike other commercial organizations and government agencies, U.S. federal law enforcement agencies (U.S. law enforcement) face similar data management challenges. U.S. law enforcement collects troves of data (including evidentiary, case, cyber, and physical threats data) to make critical decisions with speed, accuracy, and trust. The complexity of this kind of decision making requires sound data management practices, which in turn, can support the sophisticated mission analytics needed by U.S. law enforcement.

## How Red Hat can support law enforcement data needs

Common data needs of U.S. law enforcement agencies include:

- ▶ Access to relevant, accurate, and trustworthy data – with search capabilities.
- ▶ Transformation of different types of data.
- ▶ The ability to provide data that is discoverable, secure, extensible, and scalable.

Red Hat is uniquely positioned to address these challenges based on the following areas:

- 1. Data management, governance, and access.** Red Hat's automated management tools help agencies create a firm, Data Management Body of Knowledge (DMBOK<sup>3</sup>)-based, data governance foundation. This governance enables a streamlined process to manage systems and content across on-premise, as well as all public and private cloud environments in alignment with mission timelines.
- 2. Open, integrated, cloud-native architecture.** Red Hat provides practical and strategic expertise in helping organizations operationalize containers, connect to various sources to model, and develop modern hybrid cloud artificial intelligence and machine learning (AI/ML) solution architectures. This expertise is key in establishing an ecosystem of, not only Red Hat® solutions, but a



facebook.com/redhatinc  
@RedHat

linkedin.com/company/red-hat

<sup>1</sup> Desjardins, Jeff. "How much data is generated each day?" World Economic Forum, April 17, 2019.

<sup>2</sup> A zettabyte is a measure of storage capacity and is expressed as 1,000,000,000,000,000,000 bytes. One zettabyte is approximately equal to a thousand exabytes, a billion terabytes, or a trillion gigabytes. Definition found in: Fitzgibbons, Laura. "Zettabyte definition." TechTarget, accessed Feb. 2021.

<sup>3</sup> DAMA international. "Body of Knowledge," accessed Feb. 2021.

network of strategic partnerships with vendors like [Cloudera](#), [SAS](#), and IBM, as well as enriching communities and leading projects, like [OpenDataHub](#), that allow customers to build their own AI platform.

- Open source technology.** Red Hat plays an active role in upstream open source communities, leading and advancing data technologies and roadmapping innovations.

### A holistic view of enterprise data management

Organizations are making big gains in establishing modern data architectures, including analytics at the edge and AI/ML workloads. However, the success of those efforts is dependent on data management: knowing what data you have, where it is located, who owns it, who has access to it, and how it is accessed. Holistically, these data management components allow organizations to provide the right information to the right people at the right time with the right level of controls to gain analytical insights and make informed decisions.

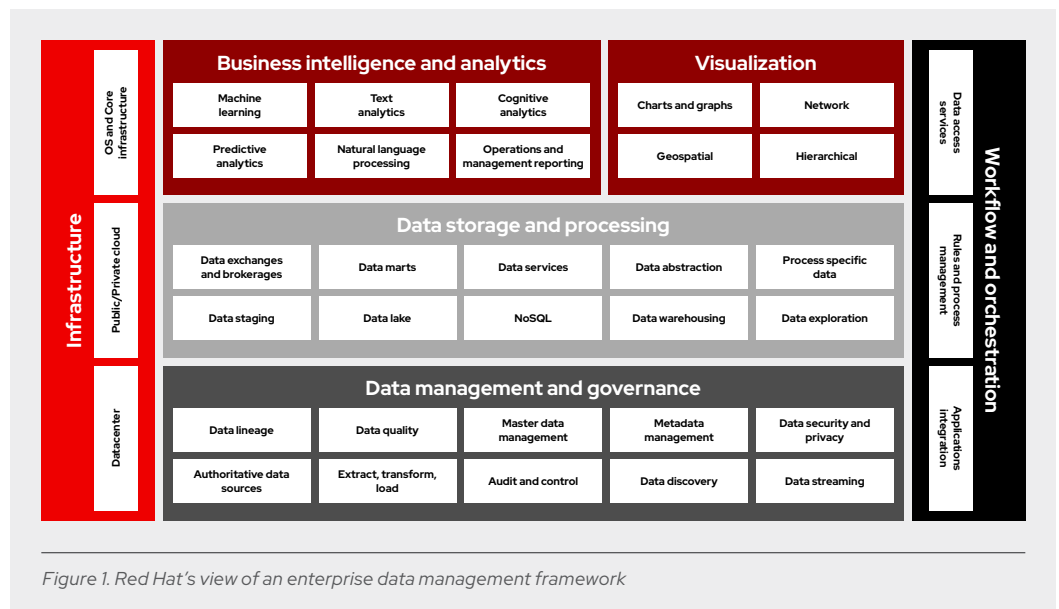


Figure 1 depicts a framework for how Red Hat breaks down the data management architectural components that organizations need to consider as they address their enterprise data needs. This framework illustrates that to move up the stack, you must first implement the foundational layers. While we do not own all aspects of data management, Red Hat can help organizations manage their data and use it as a strategic enterprise asset.

### Red Hat's trusted data platform

U.S. law enforcement requires a federated data approach at the headquarters, field offices, and in the field that ensures data is timely, accurate, and accessible at the point of need with the required security – and is available at the edge.

## What is a trusted data platform?

You may have heard of a trusted software supply chain (TSSC), which is a software development process that uses a set of guardrails to accelerate and enforce security, compliance, privacy, and transparency. A trusted data platform is an intersection of data, AI/ML, and TSSC.

Red Hat’s approach to data management makes information actionable and operational and extends the data’s reach through various channels and partners, while enforcing standards and compliance. Red Hat’s process provides managed collaboration by shifting from a reactive, batch-based, reporting approach, to one where U.S. law enforcement can access actionable information that relies on a flexible data fabric that is extensible and augments decisions. This process allows for a proactive approach to operationalizing the data.

Red Hat’s data platform operates across a hybrid, multicloud ecosystem with myriad data sources. It does not lock users into a single computing, processing, and analytics solution provided by a single vendor. Red Hat’s approach focuses on choice, transparency, and flexibility – including capabilities, vendors, and solutions that run on a standardized platform to meet U.S law enforcement’s data management, analytics, storage, integration and AI/ML requirements today and into the future.

Red Hat understands that innovation is happening in both the open source community and by other product vendors, and believes in the importance of investing in a platform that is extensible to offer a blend of sanctioned and vetted solutions. This is offered via [Red Hat Marketplace](#) and the [OperatorHub](#), including Cloudera for advanced reporting, SAS for scientific modeling, ServiceNow for Case Management, Elastic and Kibana for search and reporting, Couchbase and CrunchyDB for alternative database sources, and IBM Watson for AI/ML specific semantic analysis.

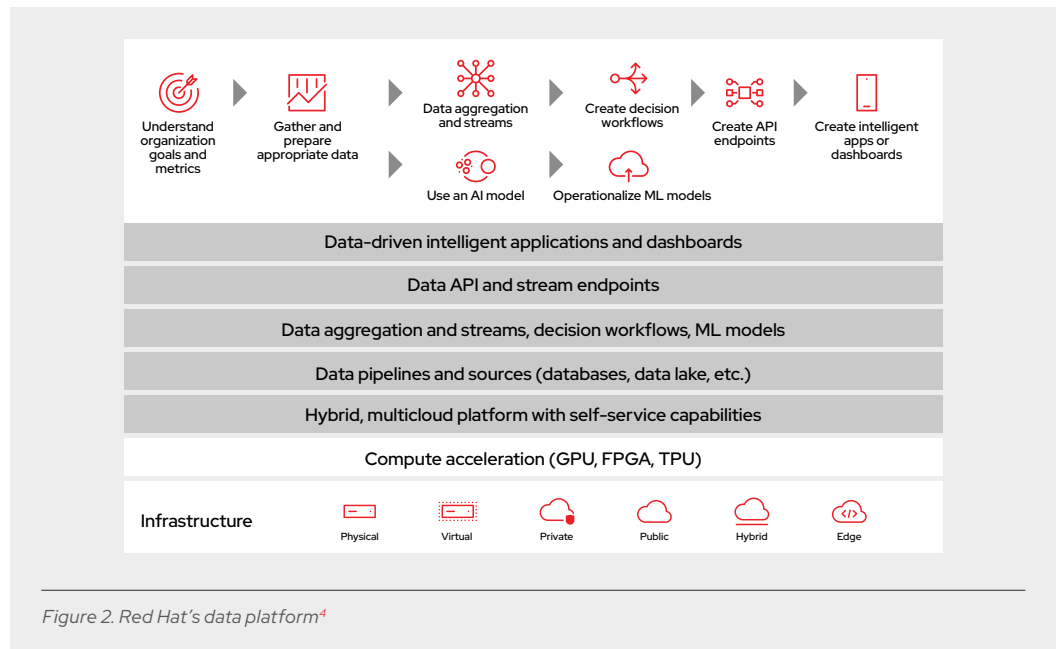


Figure 2. Red Hat’s data platform<sup>4</sup>

The data platform, as depicted in Figure 2, provides better API access, higher trust in data, improved data quality, and serves as the basis for Red Hat’s [Open Data Hub](#) initiative, providing customers with the components to build an AI/ML platform using open source technologies. Using a Red Hat technology stack gives U.S. law enforcement a choice in infrastructure and product, all supported by

<sup>4</sup> GPU: graphics processing unit, FPGA: field programmable gate array, TPU: tensor processing unit

a scalable, consistent, and performant underlying platform. By using the Open Data Hub reference architecture, data scientists can create models using [Jupyter](#) notebooks and select from popular tools, such as [TensorFlow](#) and [Apache Spark](#), to develop proactive data solution models. Open Data Hub becomes the initial solution to engage other training- and event-based architectures to extend authoritative data across a mixture of sources. It is backed by powerful streaming solutions, based on [Apache Kafka \(Red Hat AMQ Streams\)](#), to harness that power and an autonomous workflow solution that can be integrated into applications to allow more intelligent choices. Applied AI models can be used to more effectively implement the rigorous data protection and compliance requirements that are needed by U.S. law enforcement. The core workflow solution for this platform is provided by [Red Hat Decision Manager](#) for complex event processing, and the data feeds are harnessed and transformed using [Red Hat Fuse](#).

### Supporting the data scientist

Modern solutions require modern platforms and capabilities. However, we often find platforms that are not connected back to the development and operations teams, lacking an enterprise, or standardized, approach to how data science is conducted. In addition, disparate vendor capabilities make scalability and portability cost prohibitive. To bring the data science capability to the heart of establishing data as a strategic enterprise asset, we believe that a number of personas need to come together: data scientists to prepare or harness the data, data engineers to structure or connect various data sources, application developers to create the intelligent applications, IT operations teams to scale the operational models, and business owners to define the workflow, metrics, and success criteria for model activation and operation.

Figure 3 depicts how these personas come together as part of a data-enabled DevSecOps software factory, termed the [trusted intelligent application supply chain](#).

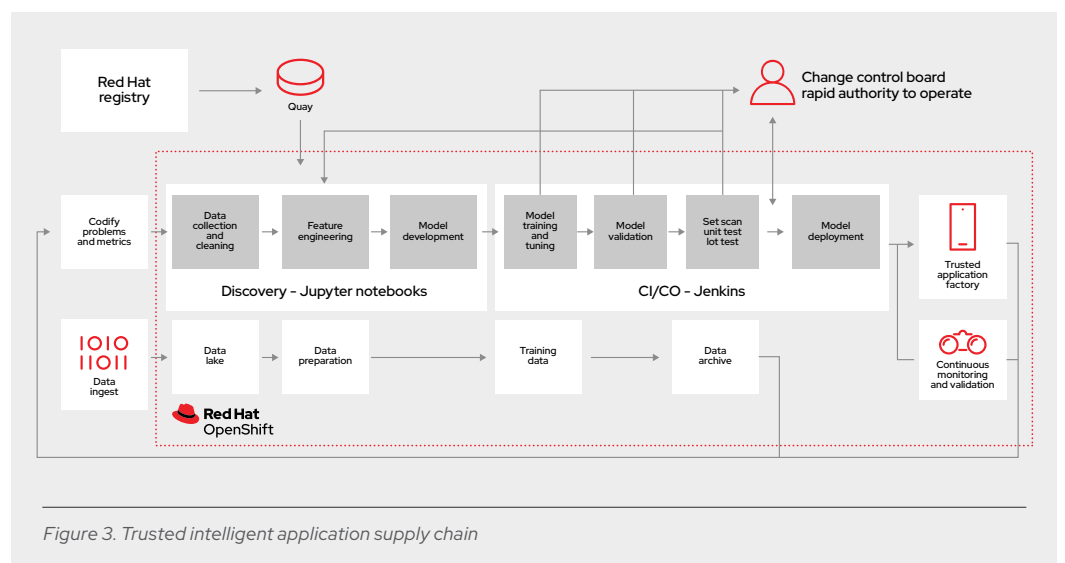


Figure 3. Trusted intelligent application supply chain

### **Implementing zero trust policy**

In an effort to prevent data violations, U.S. law enforcement requires a zero trust policy in enterprise network architecture. Implementing this policy requires the right data, the right environment, and the right capabilities at scale. As such, architecting a zero-trust solution in a hybrid cloud world, requires an open hybrid cloud architecture, powered by [Red Hat OpenShift®](#). When you operate and launch a model or need access to data services and data endpoints, you need high standards of operational compliance and security. Red Hat's solution allows zero-trust and event-driven architectures at scale that will power current investments and future models, workflows, and applications.

### **Storage considerations**

The simplicity of public cloud storage should be the model for all storage resources. Many organizations use public cloud storage due to its scalability, efficiency, and early return on investment (ROI). However, it can tie you to the specific protocols, standards, and tools of the cloud vendor, and affect migration costs as you try to pull data out. The same huge dataset will be used for multiple projects, with limited on-premise computational resources and considerable cost when stored in the cloud.

Using a Red Hat stack, including our [storage offerings](#), U.S. law enforcement is able to eliminate vendor lock-in and use virtualization to combine multiple storage systems into a single, scalable storage fabric. This approach allows smarter data choices and the ability to use cloud storage services with private resources for scaling, blending, migration, overflow capability, and disaster recovery. Red Hat offers application independence by running the specific application on the specific platform it requires, without unnecessary worry about the data location. For on-premise storage and cloud computing, we can minimize the volume of data you need to push to the cloud by automatically splitting out relevant data and only moving it to the cloud. Our strategy helps you easily create multiple data repositories, isolate or share data, and locate the data anywhere.

### **Customer examples: Red Hat data services experience**

**ExxonMobil** conducts research in machine learning, statistics, signal processing, and optimization. It applies the research findings to large-scale problems in physics, chemical, and engineering data sets and models. ExxonMobil is accelerating and performing machine learning for production optimization. It improved operations collaboratively across multiple business units using Red Hat OpenShift and removed technical hurdles for hundreds of data scientists. [Learn more.](#)

When **Boston Children's Hospital** wanted to provide a central collaboration platform for the global medical imaging community, it used Red Hat technologies to develop a web-based medical imaging platform. This platform improved image processing times and provided the ability to share critical data through real-time collaboration. It also simplified the building, deployment, and scaling of applications for imaging, analytics, and diagnosis. [Learn more.](#)

**The General Services Administration (GSA)** generates massive amounts of data. To provide easier big data analysis and make more intelligent decisions, the GSA teamed with Acuity Systems to build an open source platform using Red Hat technology. As a result, the GSA gained unified, easier data access and analysis, as well as flexibility and scalability to meet growing data needs. [Learn more.](#)

**A large grocery chain** needed to better understand customer data. Its data was incomplete and inconsistent, and the company lacked a central view and data integration across systems. Red Hat implemented a unified platform with full metadata and data lineage, providing quicker access to data and an enterprise data catalog of all data assets.

### Learn more

To learn more about how Red Hat can help government IT innovate, visit [redhat.com/government](https://redhat.com/government).



### About Red Hat

Red Hat is the world's leading provider of enterprise open source software solutions, using a community-powered approach to deliver reliable and high-performing Linux, hybrid cloud, container, and Kubernetes technologies. Red Hat helps customers integrate new and existing IT applications, develop cloud-native applications, standardize on our industry-leading operating system, and automate, secure, and manage complex environments. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500. As a strategic partner to cloud providers, system integrators, application vendors, customers, and open source communities, Red Hat can help organizations prepare for the digital future.



facebook.com/redhatinc  
@RedHat  
linkedin.com/company/red-hat

**North America**  
1 888 REDHAT1  
www.redhat.com

**Europe, Middle East,  
and Africa**  
00800 7334 2835  
europe@redhat.com

**Asia Pacific**  
+65 6490 4200  
apac@redhat.com

**Latin America**  
+54 11 4329 7300  
info-latam@redhat.com