



INFOR OS

## Infor Data Lake

Raw data is a valuable resource. Capturing data in an unaltered, pre-cleansed, pre-standardized form allows it to be reused for many new, different, or unanticipated applications. Instead of altering data ahead of time, if you apply structure and context to data on-the-fly (known as "schema-on-read"), you can drive iterative answers to different business questions without losing track of the original dataset. Infor Data Lake stores raw data from across your enterprise—whether generated by disparate applications, people, or IoT devices—and provides all the infrastructure and analytics for you to unlock your data's true potential.

### **A comprehensive approach to data**

With Infor Data Lake, structured, semi-structured, and unstructured data from across your enterprise feed into a secure, encrypted data repository, while advanced metadata intelligence prevents your data lake from becoming unmanageable. From this advanced data ingestion, to cataloging, to policy-based data management, and advanced analytics, Infor Data Lake is the end-to-end solution for your big data needs.

## Data is everything

Data is the single most valuable asset in today's world of digital business. There's so much potential inherent in data that you may have heard the catchphrase, "data is the new oil," bandied about. You may also have heard that big data is nothing without "big insights," and analytics are the key to big insights. This is all true. But there's a bigger story here as well. Data, in its raw form, is incredibly versatile and powerful.

To understand this, let's take a look at the different types of data out there:

- Structured data is highly organized, in relational databases, columnar databases, and mainframes.
- Semi-structured data contains tags and markers to denote hierarchies, like in XML, JSON, spreadsheets, retail point-of-sale (POS), and Internet of Things (IoT).
- Unstructured data doesn't have a pre-defined data model or form of organization and often includes context-heavy information such as social media, logs, analytics, clickstreams, audio, video, and geographical data.

To be analyzed, these different types of data need to undergo various levels of preparation. Preparing data correctly requires a deep understanding of business process—you need to know what sort of insight you're looking for so that you can transform data into a format that works for a particular type of analysis. 80% of a traditional analytics pipeline is consumed by this preparing/transforming process, which involves lots of data cleansing and aggregation tooling.

Of course, once you prepare data in a certain way, if you store it as such, then you reduce the potential that data has. Data prepared for one type of analysis might not easily allow for another type of analysis. And this is why raw data is valuable.

Raw data offers versatility and flexibility. Preparing data on-demand, as-needed, provides extraordinary benefits when storing raw data that can be reused at any time. You may have one business question that you're trying to answer now, but in the future, you could use the same raw dataset to answer completely different business questions.

On-demand analysis can take care of data preparation on-the-fly, including: metadata extraction, format conversions, blending, relational metagraphing, aggregation, de-normalization, and indexing.

This new paradigm, marshaling raw data for as-needed analysis, can create tremendous business value when data becomes flexible to address many needs. After all, data is the foundation of new, innovative business models, and the more ways you can harness it, the greater the potential reward. Data lakes provide the foundation for organizing data in a unified storage repository and exposing access to that data with a variety of new tools.

## Data lakes, not data swamps

A data lake is a central repository for all data types and formats, structured, semi-structured, or unstructured. Data is stored as-is, without any transformation. Since data is stored raw, there's no need for it to conform to pre-defined schema. This also means data lake storage is low-cost and highly scalable.

One of the caveats with data lakes, though, is that an unmanaged data lake can become a "data swamp" if the data lake is used as a dumping ground with poor integrity, poor quality, stewardship, governance, and data protection. To avoid this, you need a data lake infrastructure that comes with the right tools to manage your data.

Infor Data Lake delivers many tools for schema-on-read intelligence along with a fast, flexible data consumption framework to enable new ways of making key decisions. Leveraged access to your entire Infor ecosystem lets you start capturing and delivering big data to power your next generation analytics and machine-learning strategies.

## Marshaling data with Infor Data Lake

Infor Data Lake takes the variety and scale of your enterprise's information network to drive better decisions, better insights, and faster ROI on your technology investments.

Let's take a closer look at the various elements of Infor Data Lake:

### Data ingestion

Being part of the Infor OS platform helps make use of your existing data ingestion and acquisition technologies to maximum benefit. This starts with leveraging your loosely-coupled integrations, powered by Infor ION®, to begin ingesting all your ecosystem's system-to-system transactions. The Infor ION API Gateway expands that acquisition arsenal by providing fully-documented APIs to ingest content streams directly from any application registered in the gateway.

### Metadata intelligence

The Infor Data Catalog ensures there's always a semantic definition of the content stored in the data lake. Schema versioning provides a baseline so that your organization knows and understands exactly what the data looked like the day it was captured—whether it was yesterday or five years ago. New visualization and user experiences help guide and interact with your enterprise metadata so that you're able to identify the systems and experts associated with your metadata, fields of interest and security, and a suite of APIs to interface directly with the catalog for real-time reporting needs.

### Front-and-center security

Governing your data begins with Infor Data Lake's integration with Infor Federation Services. Entitle and restrict access to content landed in the data lake using authorizations designed at the user, group, and accounting entity levels. Your enterprise data is always secured and encrypted at rest and, when you need it most, content is secured in-transit.

### Data-lake-as-a-service

Data lakes are built on strong foundations centered around metadata. We expose that metadata-driven approach to storing and consuming information as a suite of APIs registered in the ION API Gateway so you're able to search, catalog, and marshal data to help deliver on heterogeneous integration requirements, ad-hoc reporting and referencing needs, and networking collections of data based on key metadata attributes.

### Traverse the data lake with metagraphs

Data Lake Metagraphs provide a simple, intuitive designer that guides you in modeling the relationships between your data—regardless of data format and content. Targeted metagraphs help composite collections of data and raw datasets so you can begin deriving intelligence and value from the data you're already storing.

### Trace, troubleshoot, and resolve

Track and trace the lineage of each message entered into the data lake. Any messages ingested directly via Infor ION are automatically registered in a searchable timeline, logging any process and integration impressions made along the way, before ultimately landing in the data lake. The ION API gateway, likewise, provides a robust lineage service to identify and catalog the sequence of hops and statistics your enterprise data takes before finally arriving in the data lake.

## Establish your digital business with Infor Data Lake

Every modern business has the opportunity to turn data into an advantage, whether that's improving operations, reaching more customers, or creating innovative products and services. Embarking on such a digital transformation can be challenging, because it requires a strong data foundation, along with infrastructure designed to meet your needs.

With Infor Data Lake, raw data from across your Infor OS applications becomes a key ingredient for enormous business potential.

[Learn more >](#)



Infor builds business software for specific industries in the cloud. With 16,500 employees and over 90,000 customers in more than 170 countries, Infor software is designed for progress. To learn more, please visit [www.infor.com](http://www.infor.com).

Follow us: [in](#) [f](#) [t](#)