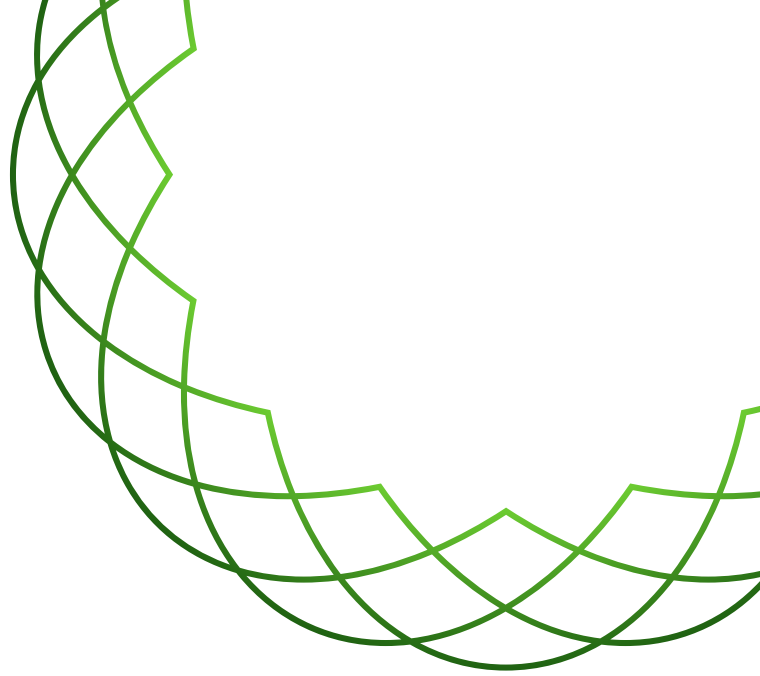


TECHNICAL WHITEPAPER



Predix Architecture and Services

September 2015

Predix



Predix Architecture and Services

Table of Contents

Introduction	3
Architecture Overview	4
Predix Machine	5
Edge Connectivity.....	6
Functional Capabilities.....	7
Deployment Models.....	8
Predix Connectivity.....	9
Predix Cloud	10
Predix Cloud Infrastructure.....	10
Cloud Foundry	11
Predix.io	11
Predix Services	12
Operational Services.....	13
DevOps Services.....	13
BizOps Services.....	14
Industrial Services	15
Asset Services	16
Asset Modeling	16
Data Services	18
Ingestion Pipeline	18
Time Series	19
BLOB Store	21
Relational Database (PostgreSQL).....	21
Analytics Services.....	22
Analytic Microservices	23
Analytic Catalog	23
Taxonomy of Analytics	23
Orchestration of Analytics	24
Application Security Services.....	24
User Account and Authentication (UAA) Service	24
Access Control Service.....	25
Predix for Developers	25
Design Principles	26
Predix Design System	27
Predix Mobility.....	28
Conclusion	29

Predix Architecture and Services

Introduction

Software has begun to influence our society directly. Digital consumer companies are disrupting the old guard and changing the way we live and do business in fundamental ways. Companies such as Uber, Airbnb and Zipcar have disrupted the traditional businesses of taxis, hotels and car rental companies by leveraging software capabilities to create new business models. Opportunities in the industrial world are expected to outpace consumer business cases substantially. General Electric is focused on driving new value for industrial organizations by offering them advanced software capabilities.

At the heart of GE's software portfolio is the Predix platform. Predix leapfrogs traditional enterprise IT solutions with a native cloud architecture that augments industrial operational technologies (OT) for both GE and non-GE assets. In essence, the Predix sweet spot is where IT and OT converge.

The Industrial Internet of Things (IIoT) differs greatly from the Internet of Things touted by mass media. The focus of the IIoT is not on connecting coffee pots to alarm clocks, but rather on connecting industrial assets, such as turbines, jet engines, and locomotives, to the cloud and to each other in meaningful ways. As a leading manufacturer of industrial assets, GE is in a unique position to leverage its understanding of asset models and industrial operations to create new value for industrial customers. The Predix platform provides a set of development tools and best practices that rapidly enable those customers to bridge the gap between software and operations to drive incredible value and innovation.

To provide a simple yet concrete example that highlights the capabilities of the the Predix platform, we will look at a wind turbine. The Predix Machine device gateway connects such assets to the Predix Cloud, regardless of vendor or vintage, enabling operational and historical data to be collected and analyzed, thereby improving operational models and potentially unlocking transformative business value.

As a leading manufacturer of industrial assets, GE is in a unique position to leverage its understanding of asset models to create new value for industrial customers.

For example, with respect to predictive maintenance, Predix allows customers to leverage all of the components of the Predix platform to predict potential problems, conduct preventative maintenance, and reduce unplanned downtimes. The Predix Machine component can monitor data collected from sensors and, using physics-based analytics, detect potential error conditions based on the asset model, and then gracefully shut down the asset. In addition to these edge applications, Predix Machine can also pass the sensor data to the Predix Cloud, where the operational data for all similar machines under management can be stored and analyzed. Over time, data scientists can discover new patterns and create new and improved physics-based analytical models. The new analytic can then be pushed back to all of the assets, effectively improving the performance of all assets simultaneously.

In the following section, this example is discussed in detail as the Predix architecture is examined.

Architecture Overview

Predix is a comprehensive, modern platform with components that span from the machine to the cloud to enable industrial use cases. The primary components are:

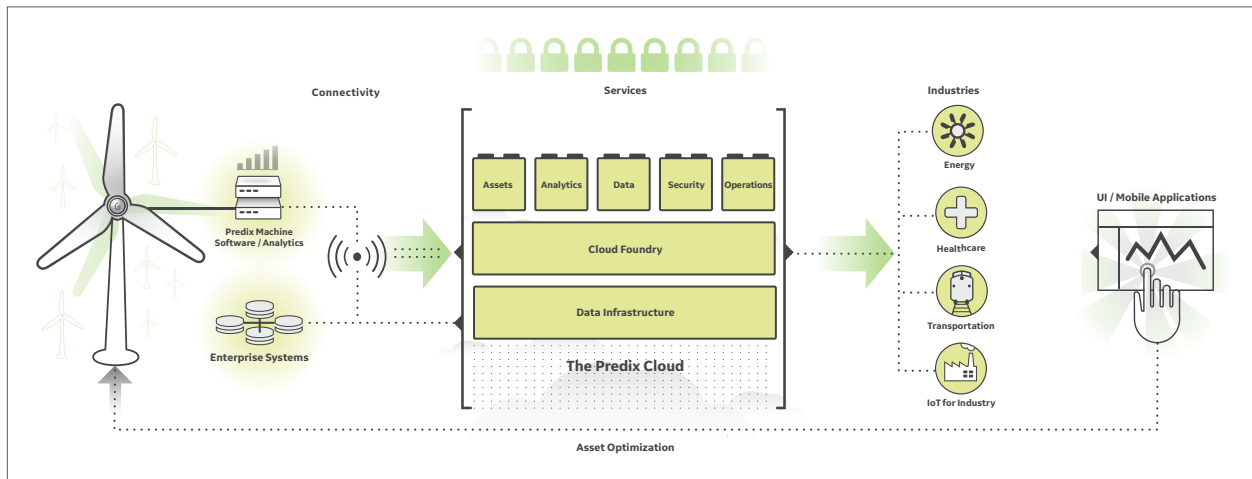
Predix Machine: Predix Machine is the software layer responsible for communicating with the industrial asset and the Predix Cloud, as well as running local applications, like edge analytics. This component can be installed on gateways, industrial controllers and sensors.

Predix Connectivity: Predix Connectivity is for scenarios where a direct Internet connection is not readily available. The service enables machines to talk to the Predix Cloud via a virtual network comprised of cellular, fixed line, and satellite technologies.

Predix Cloud: The Predix Cloud is a global secure cloud infrastructure that is optimized for industrial workloads and meeting regulatory needs.

Predix Services: Predix provides industrial services that developers can use to build, test, and run Industrial Internet applications. It also provides a microservices marketplace where developers can publish their own services as well as consume services from third parties.

Predix for Developers: Predix provides developers with a framework to communicate with services. Its modular design offers a consistent look and feel, as well as a contextual user experience in both web and mobile applications.



Predix Machine

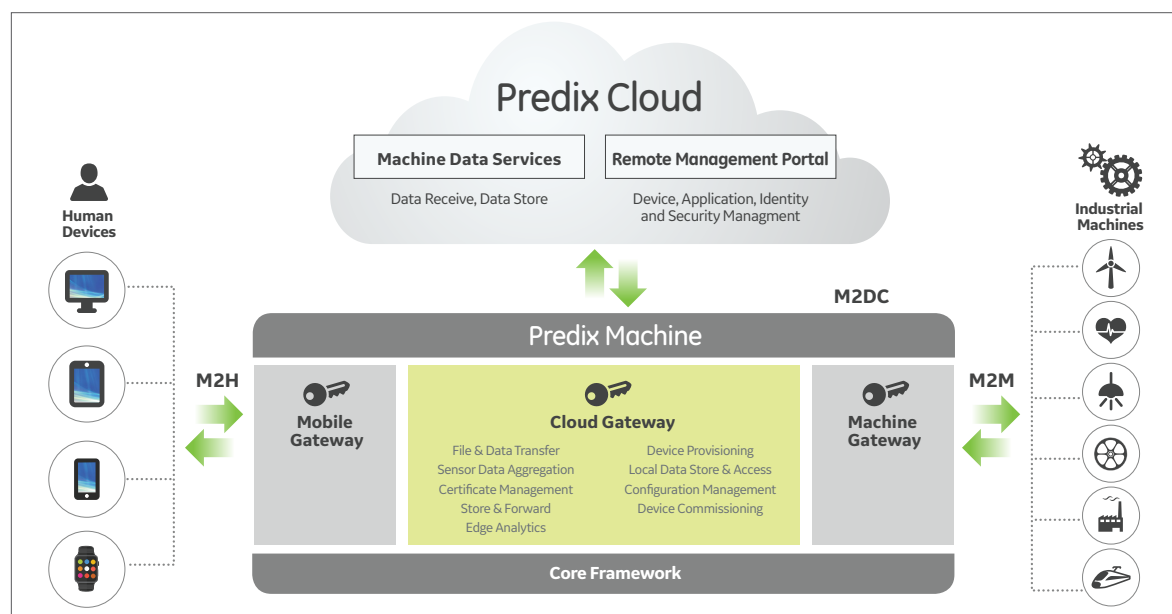
Predix Machine is a software stack that can be embedded into devices such as industrial control systems or network gateways. The software stack is available in multiple form factors and shipped with its own SDK. The SDK includes functions that enable developers to leverage the core features described below.

The primary responsibility of Predix Machine is to provide secure, bi-directional cloud connectivity to — and management of — industrial assets, while also enabling applications (analytical and operational services) at the edge of the Industrial Internet. The latter is particularly important to delivering near-real-time processing in controlled environments.

Predix Machine also provides security, authentication, and governance services for endpoint devices. This allows security profiles to be audited and managed centrally across devices, ensuring that assets are connected, controlled, and managed in a safe and secure manner, and that critical data is protected.

Edge Connectivity

To meet requirements for industrial connectivity, Predix Machine supports gateway solutions that connect multiple edge components via various industry standard protocols. There are three types of edge connectivity options that Predix Machine provides.



1. **Machine gateway (M2M)** — Many assets already support connectivity through industrial protocols such as OPC-UA or ModBus. The Machine gateway component is an extensible plug-in framework that enables out-of-the-box connectivity to assets based on the most common industrial protocols.
2. **Cloud gateway (M2DC)** — The cloud gateway component connects Predix Machine to the Predix Cloud. There are several protocols that are supported, most commonly HTTPS or WebSockets.
3. **Mobile gateway (M2H)** — In addition to connecting to the machines and to the cloud, the mobile gateway component enables people (humans) to bypass the cloud and establish a direct connection to an asset. This capability is especially important for maintenance scenarios. When service technicians are deployed to maintain or repair machines, they can connect directly to the machine to understand its operating conditions or perform troubleshooting. In certain industrial environments, where connectivity can be challenging, the ability to bypass the cloud and create this direct connection to the machine is key.

Functional Capabilities

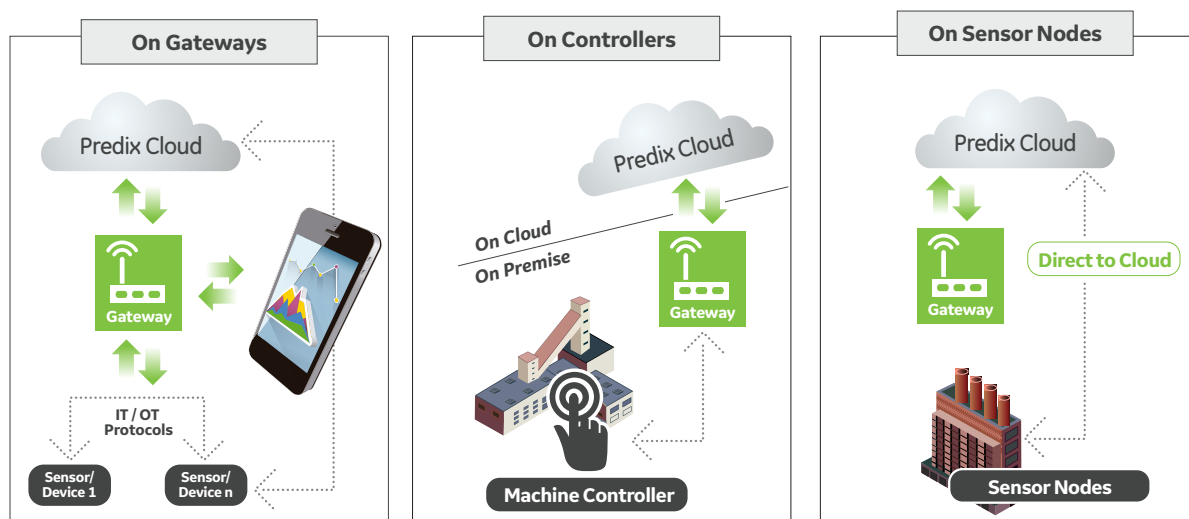
Predix Machine provides a number of core capabilities in industrial scenarios, including edge analytics. Industrial scale data – which can be massive and is often generated continuously – cannot always be efficiently transferred to the cloud for processing. Edge analytics provide a way to pre-process the data so that only the pertinent information is sent to the cloud.

- **File and data transfer:** File and data transfer allows data to be pushed to the cloud by streaming batching or by uploading a file.
- **Store and forward:** Store and forward provides support for intermittent connectivity loss (for example, when a locomotive travels through a tunnel). Data must be collected locally, and then forwarded to the cloud once connectivity is reestablished.
- **Local data store and access:** This capability allows data about machines to be stored on the device so that, for example, a service technician can access the data directly.
- **Sensor data aggregation:** Predix Machine can connect to multiple sensors and then push an aggregated 'fingerprint' to the cloud, which reflects the data gathered from all of the sensors.
- **Edge analytics:** The edge analytics capability enables computational algorithms to be run directly on the data that is streaming off the machine.
- **Certificate management:** In order to provide end-to-end security, Predix Machine supports certificate management to provide SSL-based connections to the Predix Cloud.
- **Device provisioning:** When Predix Machine is installed on an edge device, it can 'phone home' to the Predix Cloud to register itself for further management and software upgrades.
- **Device decommissioning:** When Predix Machine is taken off-line, it can notify the Predix Cloud that it no longer needs to be managed.
- **Configuration management:** Configuration management allows remote configuration of the Predix Machine and the tracking of configuration changes over the lifetime of the machine.

Configuration management allows remote configuration of the Predix Machine and the tracking of configuration changes over the lifetime of the machine.

Deployment Models

The Predix Machine software can be deployed in three ways.



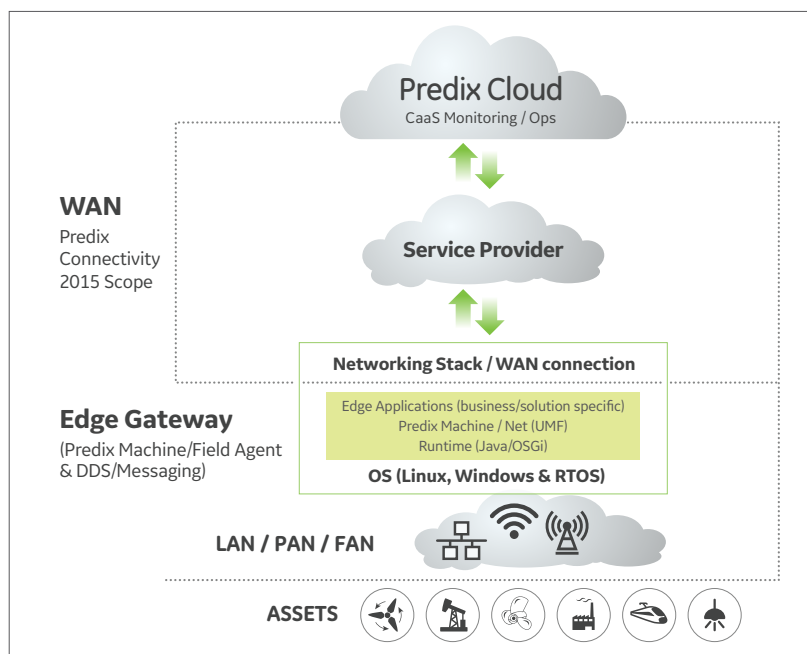
1. **On the gateway:** The gateway acts as the smart conduit between the cloud and the machines. The Predix Machine software is deployed on the gateway device to provide connectivity to machines and other assets via a variety of IT or OT protocols, including HMI devices running IOS, Android or other mobile operating systems.
2. **On controllers:** The Predix Machine software is deployed directly on the machine controller units. This decouples the machine software from the machine hardware, allowing connectivity, upgradability, cross-compatibility, remote access, and remote control. It also enables industrial and commercial assets that have traditionally operated standalone or in very isolated networks to be connected directly to the cloud for data collection and live analytics.
3. **On sensor nodes:** In this scenario, the intelligence lives in the cloud, and simple, low cost sensors are deployed on or near the assets. The sensors collect machine and environment data, and then backhaul it to the cloud (directly or through an IoT gateway), where it is stored, analyzed, and visualized.

While many Industrial Internet applications reside in the cloud, they need to connect to the machines and even process data there. Predix Machine provides the capabilities required to gather sensor data, process it locally, and then push it to the Predix Cloud.

Predix Connectivity

Predix Connectivity provides fast, secure cloud connectivity from Predix Machine to the Predix Cloud. The service eliminates the long lead times and the expensive proposition of designing and operating a custom and potentially sub-optimal connectivity infrastructure.

Predix Connectivity also provides a secure and readily available global virtual network that fulfills Internet requirements in a repeatable fashion and is transparent to the enterprise. The virtual network includes cellular, fixed line, and satellite technologies.



Predix Connectivity is available today through several partners and is offered across the globe. Same day activation and provisioning, combined with continual proactive monitoring support, troubleshooting, and automatic business alerts, are available.

Together with Predix Machine, Predix Connectivity provides plug-and-play, secure, and reliable connections to the Predix Cloud.

Predix Cloud

The Predix Cloud is central to enabling the Industrial Internet. It consists of a scalable cloud infrastructure that serves as a basis for Platform-as-a-Service (PaaS), which is what developers use to create Industrial Internet applications. It also provides an entry point for industrial enterprises to take advantage of new software technology without having to make massive hardware and software commitments.

Cloud technology is now recognized as the way to achieve secure and scalable computing. In order to process the vast amounts of data required by Industrial IoT scenarios, cloud computing is the preferred choice. Traditional enterprise application developers will remember the pain associated with building scalable applications. Typically, they were forced to provision (or, rather, over-provision) high-end, expensive hardware capable of scaling-up to handle the highest envisioned load. That was not only expensive, but also difficult, as developers had to add capacity when more scale was required. Cloud computing solves the problem by leveraging commercial off-the-shelf hardware, making it easy to scale out (add additional compute) instead of scaling up (adding more storage, memory, and CPU to an existing server).

Cloud computing solved many challenges, but developers were still required to perform a significant amount of work to manage scalability at the application layer. PaaS solutions provide developers the tools they need to create cloud-native scalable applications.

Predix manages the complexity of scale so that developers can focus on creating applications that drive industrial value, and brings cloud computing into the currently under-connected industrial world reliably and inexpensively.

Cloud technology is now recognized as the way to achieve secure and scalable computing.

Predix Cloud Infrastructure

Public cloud infrastructure doesn't support the unique and demanding requirements of industry. It is for that reason GE has decided to work with partners around the world to build out a custom cloud infrastructure able to optimize for industrial workloads, especially processing industrial scale data. As a leader in the industrial world, GE provides a robust cloud platform focused on the security and regulatory compliance required by emergent industrial applications.

GE has carefully specified enhanced security controls and optimized the hardware build-out for Predix workloads, including world-class data

processing and networking capabilities. The result is a GE data center blueprint for meeting industry and regulatory requirements around the globe. GE has also created a Software-Defined Infrastructure (SDI) that serves as an abstraction layer above the specified hardware, so that the data center can evolve over time with minimal disruption to the applications. The SDI enables GE to create a shared infrastructure with policy-based provisioning to facilitate dynamic automation and to apply SLA mappings to the underlying infrastructure. This is especially useful when an application requires an underlying hardware configuration. The provisioning management and pooling of resources can be done at a granular level, allowing for optimal resource allocation and ultimately driving costs down and value up.

Cloud Foundry

The Predix platform is based on Cloud Foundry (CF), an open source PaaS that supports multiple developer frameworks and an ecosystem of application services. Cloud Foundry makes it faster and easier for application developers to build, test, deploy, and, perhaps most importantly, scale applications.

Because Predix leverages Cloud Foundry, application developers gain access to the vibrant Cloud Foundry ecosystem and an ever-growing library of CF services. Additionally, because it is open source, CF can be customized for Predix workloads. For example, GE is working with the CF community to build support for industrial protocols.

Predix.io

Predix.io, the self-service portal where developers can access specialized services intended for use in Industrial Internet applications, is the starting point for developers that want to build on Predix. GE has a deep understanding of the capabilities needed to build Industrial Internet applications and has built this platform to meet both its needs and the needs of other industrial companies.

Predix.io also provides a catalog and marketplace of services, so that the ecosystem can contribute additional sub-categories. The goal is to create a robust platform that meets the needs of additional use cases.

GE Industrial Dojo

In collaboration with the Cloud Foundry Foundation, GE Digital has launched the GE Industrial Dojo program to accelerate the ability for developers to contribute code that enables the Industrial Internet.

Participants will be paired with experienced colleagues to immerse themselves in open source projects and quickly learn the core technology in an agile development environment.

Predix Services

Key Predix cloud services can be broken down into two categories, each with its own sub-categories.



Operational Services

Operational Services enable application developers to manage the lifecycle and commercialization of their applications:

- **DevOps Services:** Services to develop and deploy Industrial Internet applications in the cloud.
- **BizOps Services:** Services that enable transparency into the usage of Industrial Internet applications so that developers can ensure profitability.



Industrial Services

Industrial Services provide the core capabilities required by Industrial Internet applications:

- **Asset Services:** Services to create, import, and organize asset models and their associated business rules.
- **Data Services:** Services to ingest, clean, merge, and ultimately store data in the appropriate storage technology so that it can be made available to applications in the manner most suitable to their use case.
- **Analytics Services:** Services to create, catalog, and orchestrate analytics that will serve as the basis for applications to create insights about industrial assets.
- **Application Security Services:** Services to meet end-to-end security requirements, including those related to authentication and authorization.

Operational Services

DevOps Services

One fundamental change brought about by cloud computing is the way software is delivered. Traditionally, software has been shipped to the customer for installation behind a firewall. With cloud computing, the software is no longer shipped, but rather deployed and operated on behalf of the customer, who accesses it over the Internet. This change in delivery model results in changes to the software development lifecycle model, ultimately combining development and operations (DevOps).

DevOps is required to be flexible, focusing on both development and operations. DevOps must do the following:

- Continually integrate and deliver new features
- Build and test often through automation
- Always be ready to deploy to production
- Always place emphasis on speed, efficiency, and stability

As noted above, one key aspect of successfully implementing DevOps is automation. Tools for source control management, agile planning, automated build and deploy, and testing become imperative for efficiently moving code through the pipeline from development to production. By leveraging these types of tools, the time-to-value for a customer is significantly shorter. Predix DevOps offers developers the same collection of tools used by GE to build the platform so they can easily create their own applications.

DevOps Services provide the following capabilities:

- Source control management (SCM)
- Automated software builds and application deployment through the Continuous Delivery (CD) Pipeline service.

Source Control Management

A hosted SCM system is available for storing application source code. Application developers can either move existing project code or create a new project by forking an example Predix application. The CD pipeline can then be configured to automate all software builds and application deployments.

Continuous Delivery Pipeline

The Continuous Delivery (CD) Pipeline automatically builds, tests, and deploys applications to the Predix Cloud. The pipeline is designed to ensure predictability from the beginning of the development cycle through production delivery. It checks that build processes are always working, and enables the rapid and repeatable provisioning of new environments.

With these CD tools, application developers can easily create a pipeline with automated build, test, and deploy capabilities to shorten the time-to-value for Predix customers.

BizOps Services

Along with changes in how software is delivered, cloud computing also requires companies to rethink their business models — the services that they provide to customers and the ways customers buy and pay for those services. Customers expect to pay for services in a variety of ways:

- **Subscription** — The customer pays a fixed amount for the product over a period of time, usually monthly, quarterly, or annually.
- **Utility** — The customer pays as it consumes the product.
- **Freemium** — The customer enjoys the basic product for free and only pays for add-on or premium services.

In addition to shortening the time-to-value, cloud computing offers transparency into the usage of software — that is, visibility into the value of the software. Traditional software vendors and their customers can suffer due to ‘shelfware’ (software that is purchased, but never used). The shelfware cost to customers is obvious — they pay for something that provides no value. But the cost to vendors is real, too, as customer churn rates increase; customers can sign up, upgrade, downgrade, or cancel at any time, impacting profitability:

- Businesses have to invest upfront to acquire customers, without any guarantee of how long a customer will use the products or services, or how much the customer will pay.
- Gross margins are impacted by the costs of serving each customer — technology infrastructure costs, tech/DevOps labor costs, and customer on-boarding and support costs — and fluctuate monthly.

Without understanding the true usage and value of their software, businesses can't determine where to best allocate their development

resources for future product enhancements. But with cloud computing, the software provider is operating the software and can easily instrument its code to better understand the usage of the software. The Continuous Delivery Pipeline of Predix Cloud not only allows businesses to automatically build, test, and deploy applications, but also creates opportunities for product and marketing managers to innovate on business models that create long-term and profitable relationships with customers.

BizOps is a web-based monetization solution that enables continuous business model innovation for industrial services.

BizOps is a web-based monetization solution that enables continuous business model innovation for industrial services. By giving product and marketing managers granular control and visibility over how they roll out, test, and scale key components of their recurring revenue business model, BizOps provides the tools necessary (product catalog, packaging, pricing and policies) to ensure profitability. Product and marketing teams can iteratively build, split-test, and launch each component to one or more customers and distribution channels. As they roll out and scale their business models, they get complete visibility into their revenue and cost metrics so they can measure the impact of the changes. Additionally, they gain insight into the best and worst performing products, packages, customer segments, and distribution channels.

Industrial Services

Predix offers a set of industrial services that work in concert. Asset services provide a way to model industrial assets and associate them with data sources. In our example, an asset model of an industrial turbine would define the hierarchy of parts that make up a turbine, as well as its sensors, including even the acceptable temperature ranges for those sensors. Data services are available to ingest the data from the turbine, cleanse it in the ingestion pipeline, and store it in the appropriate data store. For example, a time series data store would be the most appropriate place in which to store the operational and historical data for the temperature sensor. Finally, analytics services enable applications to pour over the massive amounts of asset data and detect trends or new insights about the assets under management. Analytics services could be used, for example, to determine that the turbine's temperature range could be optimized, and then update the operational analytics and push them back down to Predix Machine to ensure that graceful shutdowns of the turbine can occur instead of a major failure.

Asset Services

The asset model is at the center of all Industrial Internet applications. While assets are the instantiations of asset types (types of industrial equipment, such as turbines), the asset model is the digital representation of its structure.

The Asset service provides REST APIs that enable application developers to create and store asset models that define asset properties, as well as relationships between assets and other modeling elements. Application developers can then leverage the service to store asset-instance data. For example, an application developer can create an asset model that describes the logical component structure of all turbines in a wind farm, and then create instances of that model to represent each individual turbine. Developers can also create custom modeling objects that meet their own unique domain needs.

The Asset service consists of an API layer, a query engine and a graph database:

- **REST API layer** — Applications can access the domain object modeling layer using REST endpoints that provide a JSON interface to describe all of their objects. The service translates data from JSON to RDF triples for storage and query in the graph database, and back to JSON again.
- **Query engine** — The query engine enables developers to use Graph Expression Language (GEL) to retrieve data about any object or property of any object in the asset service data store.
- **Graph database** — The Asset service data store is a graph database that stores data as RDF triples.

Asset Modeling

An asset model represents the information that application developers store about assets, how assets are organized, and how they are related. Application developers use the Asset service APIs to define a consistent asset model and a hierarchical structure for the data. Each piece of physical equipment is represented by an *asset instance*. Assets are organized by a *classification* and by any number of custom modeling objects.

For example, an organization can use a Location object to store data about where its pumps are manufactured, and then use a Manufacturer object to store data about specific pump suppliers.

Asset services provide a way to model industrial assets and associate them with data sources.



It can also use several classifications of Pumps to define pump types; assign multiple attributes, such as Brass or Steel, to each classification; and associate multiple meters, such as Flow or Pressure, to a classification.

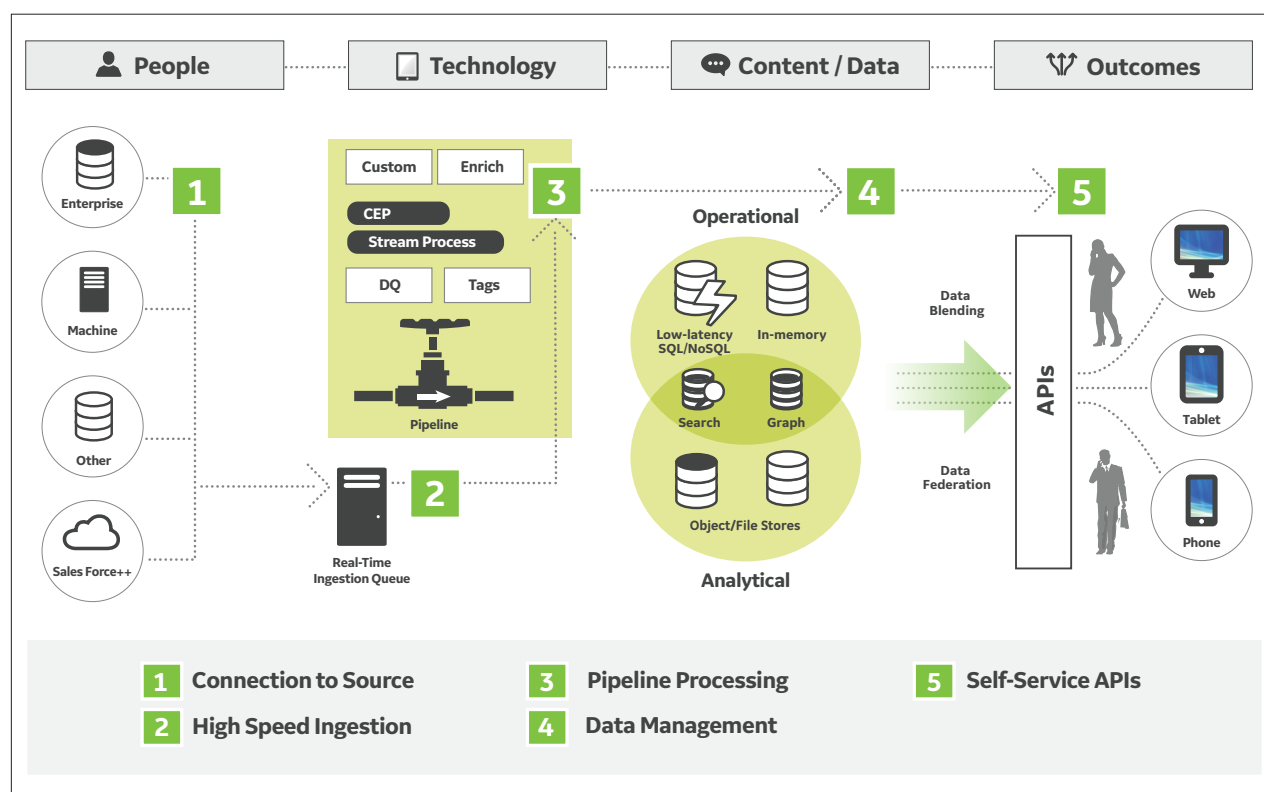
The Asset service APIs support assets, classifications, and custom domain modeling objects:

API Category	Description
Assets	<p>Typically, assets are defined in a hierarchical structure composed of a parent asset and one or many peers and children. Assets can be associated with a classification or any number of custom modeling objects, and can have any number of customer-defined attributes.</p> <p>An asset can also “stand alone” in the system, meaning it is not associated with any other modeling elements.</p>
Classifications	<p>Classifications are arranged in a tree structure and provide a means to group similar assets and to track attributes that are common to those assets. Classifications can be associated with an asset, a meter, or both.</p> <p>A classification can point to multiple assets. Attributes can be assigned at any level in the classification hierarchy.</p>
Custom Modeling Objects	<p>Custom modeling objects are hierarchies that are used to provide more information about assets. For example, custom objects could be created for asset location, manufacturer, and service contract. A location can be associated with multiple assets. Likewise, an individual asset can be associated with multiple locations.</p>

Asset models are central to Predix, tying together all of the industrial services of the platform. The asset hierarchy and the relationships to sensors are what enable application developers to bring in the data for analysis and understand the context for that data.

Data Services

Data services enable Industrial Internet application developers to bring data into Predix and make it available for their applications. The data is brought in through an ingestion pipeline that provides an opportunity to cleanse the data, merge the data with other data sources, and ultimately store the data in the appropriate type of data store, whether it be a time series data store for sensor data, a Binary Large Object (BLOB) store for MRI images, or a RDBMS.



Ingestion Pipeline

The ingestion pipeline is the entry point for all data. It enables data to be received into Predix in a number of ways, including HTTP streaming for real- or near-real-time data ('fast' data) or FTP for more batch-style processing.

The ingestion pipeline allows for the data to be processed before it is stored, which enables developers to do a number of things, including map sensor data to tags, combine asset data with ERP data to assess the financial depreciation of an asset against its performance, and process complex events (look for a combination of certain types of events to create a higher-level business event).

After processing has occurred, the developer can then choose the type of data store that best suits the application's needs. Currently Predix supports a time series data store for sensor data, a BLOB store for image data, and a relational database for everything else.

Time Series

Since Predix is focused on Industrial Internet applications, much of the data that will be brought into the platform for analysis is sensor data from industrial assets. The Time Series service provides a query-efficient columnar storage format optimized for time series data. As the continuous stream of information flows from sensors and needs to be analyzed based on the time aspect, the arrival time of each stream can be maintained and indexed in this storage format for faster queries

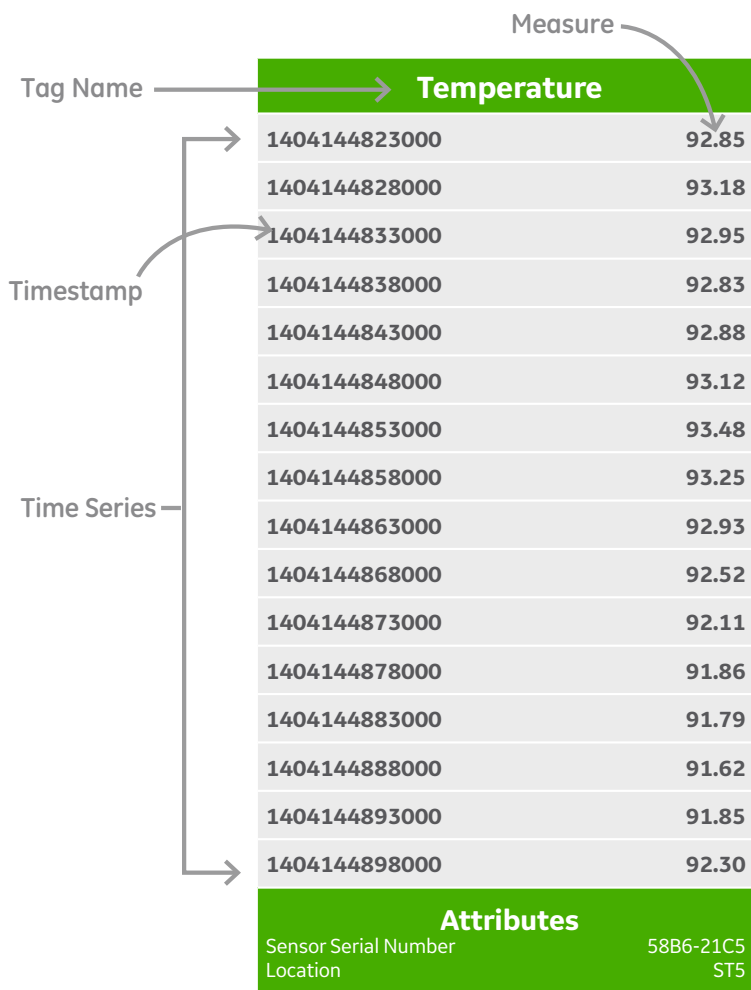
The Time Series service also provides the ability to efficiently ingest massive amounts of data based on extensible data models. The Time Series service capabilities address operational challenges posed by the volume, velocity, and variety of big data:

- Efficient storage of time series data
- Indexing the data for quick retrieval
- High availability
- Horizontal scalability
- Millisecond data point precision

Component	Description
Time Series Database	Fast, distributed, and scalable database
Columnar Storage	Distributed, scalable, and highly available database to manage massive amounts of data.
Data Ingestion	The data-ingestion layer provides real-time data streaming and support for industrial data formats such as Historian and OSI.

Time Series Data Ingestion

A time series dataset uses tags, which are often utilized to represent sensors (for example, a temperature sensor). A tag consists of one Tag Name (Sensor), a TimeStamp (Time), a Measure (Value) and, optionally, one or more attributes (key/value pairs). The ingestion pipeline enables a developer to take the raw data stream and map it to these data structures before being persisted in the Time Series data store.



Term	Definition
Tag name	(Required) Name of the tag (for example, "temperature").
Measure	(Required) Reading (Value) Numeric (for example: 98°C).
Timestamp	(Required) Time Epoch (UNIX epoch time in milliseconds).
Attribute	(Optional) Any data associated with a tag (for example: AircraftID=230, Manufacturer = GE). This is useful for filtering data.

Time Series Data Consumption

The Time Series service provides REST APIs for querying and aggregating time series data. The API allows you to:

- Query time-series data specifying tags (sensors, etc.) and a time window.
- Filter by attribute values.
- Retrieve tags, attribute values and attribute keys from the time-series database.

BLOB Store

The binary large object (BLOB) storage service provides a highly available and horizontally scalable storage service that allows secure storage of byte arrays up to 10 GB, indefinitely. It also provides mechanisms for efficient retrieval to various application containers and tiers. You can store and retrieve both small and large amounts of data in any file type using the BLOB Store.

Relational Database (PostgreSQL)

PostgreSQL is an open source relational database management system (RDBMS). Although there is a proliferation of so-called NoSQL databases for special purposes like time series data management, SQL databases are still the most commonly used.

Time series, BLOB store and Relational Databases are the data stores currently supported and available to application developers; more will be added over time. The data stores, plus the ingestion pipeline, are the data services that are needed by developers to bring data in for analysis.

Analytics Services

The real power of the Predix platform is its ability to analyze data to create insights that can then be turned into actions that provide real business outcomes. Primarily, the platform performs two types of data analyses: operational and historical.

Operational analytics — Data is analyzed in real time at the source — an aircraft engine, wind turbine, MRI machine, etc. — to detect problems so that split-second changes can be made in the operation of the asset to prevent damage and optimize performance.

Historical analytics — The collection and analysis of petabytes of historical operational data. From this analysis, it is possible to build predictive models that can be used to more efficiently operate entire manufacturing plants or fleets of equipment.

Predix offers data science capabilities that enable a feedback loop between operational and historical analytics. Operational analytics at the edge ensure the efficient operation of assets, but those analytics can be improved over time based on historical analysis. For example, an operational analytic might continually monitor the temperature of an asset and initiate a shutdown in the event of unacceptable operating temperatures. However, after collecting years of data across many assets, it could be determined that the threshold should vary by the physical elevation of each asset. In that case, a new analytic that takes elevation into account could be created and then be deployed to the Predix Machines associated with assets of that type around the world.

Predix Analytic services provide a framework for developing and embedding advanced analyses in business operations. This framework can be used to manage the execution of analytics through configuration, abstraction and extensible modules. It can also serve as a sandbox for analytic testing.

An *analytic* is a function or small program (often a physics-based algorithm) that processes machine data. It can be used directly by an application or at a step in the orchestration of multiple analytics, where the output of one analytic can be the input for another. The inputs and outputs for each analytic are typically expressed as parameters to enable an analytic to be reused for different use cases. Analytics and orchestrations are organized under a taxonomical (category with sub-category) structure and stored in a catalog.

Predix Analytics services currently support analytics written in Java, Matlab and Python, all of which can be uploaded to the Analytic Catalog and executed by the Runtime service.

Analytic Microservices

The following table provides a summary of the Analytics services.

Microservice	Description
Analytic Catalog	Repository to host and manage analytics that enable the execution of analytics as a microservice.
Orchestration Configuration	Orchestration defines the flow between groups of analytics to be run as a single unit. The orchestration configuration is defined using the BPMN 2.0 standard.
Orchestration Runtime	A microservice that runs the orchestration BPMN configuration. Use this service to run an orchestration on demand.

The real power of the Predix platform is its ability to analyze data to create insights that can then be turned into actions that provide real business outcomes.

Analytic Catalog

Analytics developed in Java, Matlab, or Python languages can be published to the Analytic Catalog, and multiple versions of each analytic can be maintained. The author of each analytic can add metadata to the catalog entry to improve retrieval, specifying the location in the analytic classification taxonomy where the analytic is to be placed. Analytics can be browsed or searched according to specific taxonomies.

The primary modes of interaction with the catalog are through REST APIs and the user interface. A key catalog feature is analytic validation, which enables the developer to test the analytic against sample data while preparing it for production. Additionally, the Analytic Catalog service makes it easy to deploy an analytic independently as a microservice.

Taxonomy of Analytics

Analytics hosted in the analytic catalog are categorized by their taxonomy locations. This provides a structured method of grouping related analytics and aids in analytic retrieval. The analytic catalog comes with a default taxonomy for predefined Predix-included analytics. Taxonomy locations can be added through the API, which expects the desired taxonomy in a JSON structure (as either a single node object or an array of node objects).

Orchestration of Analytics

The goal of the Runtime service is to support elastic execution of analytics from the catalog, whether individually or as part of an orchestration. It is a collection of coordinated microservices that handle the dispatch of analytics or orchestrations based on rules, timers and events.

Each analytic is executed as a separate microservice; the orchestration execution microservice coordinates their work. Third-party analytics that are not deployed from the analytics catalog may also be orchestrated, provided the orchestration execution microservice can interact with them. The user can interact with a configuration microservice that allows an orchestration to be defined in BPMN XML including references to analytics in the catalog as well as third-party analytics. The initial execution of an analytic or orchestration is done for the purpose of validation. Following testing, they can be moved into production.

Orchestration is a group of analytics to be run as a single unit. Its analytic workflow is defined within an Orchestration BPMN file (an XML file conforming to the BPMN 2.0 standard).

The ability to create and orchestrate analytics for both operational analysis and historical analysis is a key differentiator for the Predix platform. Through data analysis, insights can be turned into actions that provide real business outcomes.

Application Security Services

The end-to-end security requirements of the Industrial Internet are stringent. Predix offers application security services that are key to building Industrial Internet applications. The two most important are authentication and authorization.¹

User Account and Authentication (UAA) Service

The UAA service provides a way for applications to authenticate users. An application developer can bind to the UAA service in the marketplace and then use the industry standards SCIM and OAuth to handle identity management and authentication, respectively. Together, these two capabilities provide the basic login and logout support that every application needs.

¹ A companion paper outlining all of the security features incorporated into Predix will be released shortly.

Additionally, UAA supports SAML (Security Assertion Markup Language), which enables users to login using third-party identity providers. For example, many of our industrial customers leverage GE Single Sign On (SSO) to access their applications.

The basic UAA features have been extended to include the following:

- **User whitelisting:** Ensures only a qualified subset of authenticated users can login to an application.
- **Client-side token validation:** Eliminates extra network round trips and significantly improves performance.

After users have been authenticated, the application then needs to control access to its resources.

Access Control Service

Predix Access Control service is a policy-driven authorization service that enables applications to create access restrictions to resources based on a number of criteria. The policy language is JSON-based and was developed as an answer to the deficiencies in XACML. The access control service is well integrated with UAA and provides a Spring security extension to make it easy for Spring Boot applications to make access decisions



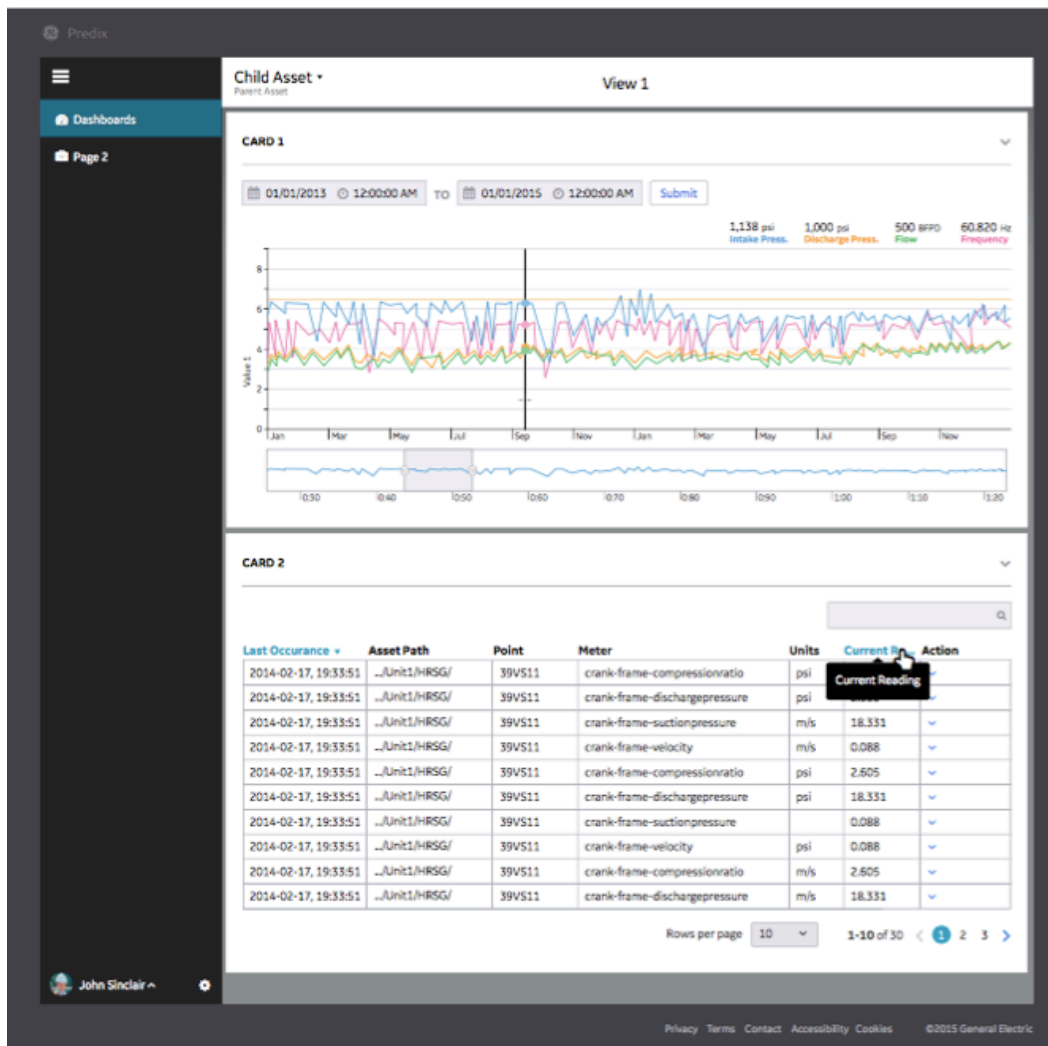
Predix for Developers

Building Industrial Internet applications comes with some unique requirements and needs that separate these applications from traditional IT applications. Context drives any user experience, but it is particularly crucial to provide the right information to the right user at the right time and in the right way when it comes to industrial applications. GE Digital has years of experience building software applications for industry and the design of Predix incorporates best practices to enable the rapid development of high quality Industrial Internet applications.

Design Principles

The central theory of Predix user experience technology is context-based design. This philosophy hinges on the idea that users need the right information presented to them at the right time and in a way that supports efficient and effective interaction. Therefore, it's critical to understand the types of users (personas). By looking at each persona and taking time to understand the tasks that they perform, the application developer can create a data model that reflects real-world business flow. Beyond that, because many users have unique constraints in the field, it's also important to understand the context in which the application will be used.

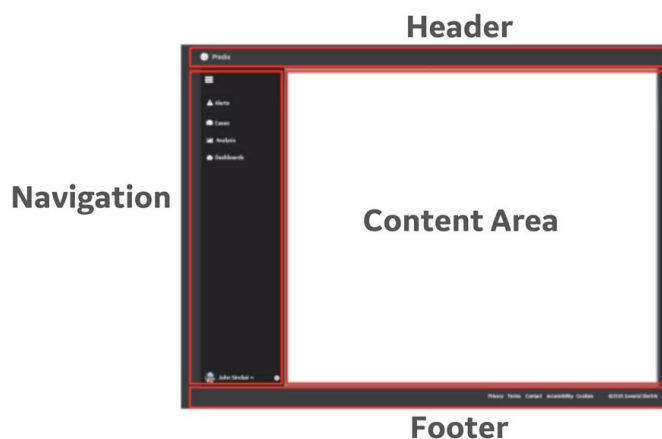
Predix leverages these design principles and offers two means to build applications: The Predix Design System and Predix Mobility.





Predix Design System

The Predix Design System provides a comprehensive set of web components that are instrumental to Industrial Internet applications. The primary unit of application development is a 'page,' which is typically comprised of four primary areas of interaction: header, navigation bar, content area and footer. A 'card' is the fundamental building block for any view; it is created from web components, is self-contained and can be shared and saved.



One of the fundamental principles of the Predix Design System is composable design and creation. A page may have several views, which in turn may have several “decks,” or groupings of cards. Views are context-aware and responsive. They present content in an intelligent, consumable format. The “context” of the page, which is derived from a variety of factors (e.g., the currently selected asset or assets, the step in the process, or the role of the user), is used to enumerate views, decks and cards.

A component is a part of an application that is built to serve a specific function. It is a self-contained feature that uses UI Elements from the Predix Design System. UI Elements are individual pieces of a consistent visual language used to represent items such as buttons, form fields,

drop-downs, and other coded patterns. Together, the components and UI Elements provide a library of useful application widgets that can rapidly be composed in an application and help to maintain a consistent look and feel.

Predix Mobility

Predix Mobility architecture is a framework that simplifies the building of mission-critical, high performance, always-available mobile applications that function even despite network outages. The system provides cross platform, multiform-factor support for major mobile devices, laptops, desktops, and browsers. It does so with a consistent look and feel, and a rich set of responsive web components. Predix Mobility is purpose-built to support Industrial Internet applications that can operate in various environments, from the board room to rugged, industrial environments, such as oil rigs. Predix Mobility is designed to provide the services that enable the broadest use of mobile devices in the most secure manner possible.

Predix Mobility architecture is based on a cross-platform, distributed persistence engine that synchronizes data between remote devices and enterprise data domains, as well as with Predix Machine, using the Predix Mobility backend services. With templates designed to support standard cloud connectivity patterns, integration to enterprise data domains is delivered as a managed microservice.

On the client side, Predix Mobility is a flexible and layered component system, with extensible services that support advanced application behaviors, including remote workflow and analytics. Designed to support standard web components, as well as the Predix Design System, the architecture simplifies the process of building enterprise-level mobile applications. The Offline-First development system supports rich, native user experiences while retaining the simplicity of a web-based application.

Designed to work everywhere and keep data synchronized with backend systems and other devices, Predix Mobility provides all of the capabilities for building great Industrial Internet mobile applications.

The responsive Predix Design System with web components for online scenarios, and the native experience with off-line capabilities of Predix Mobility, enable Industrial Internet application developers to create applications that are intuitive and easy to use.

Predix mobility is a framework that simplifies the building of mission-critical, high performance, always-available mobile applications that function even despite network outages.

Conclusion

GE is pioneering the industrial cloud computing space. Predix represents the starting point of a journey that will yield increasingly important improvements in key industrial and business processes.

While progress on the automation of industrial equipment has been made over decades, and assets have become 'smarter,' the intelligence of any individual piece of equipment pales in comparison to the intelligence that is unleashed when smart devices are connected. Aggregating the data and operational capabilities of intelligent devices enables industrial companies to significantly improve business processes. Whether turning to the industrial cloud for real-time asset optimization in power generation, improved analytics and diagnostics in healthcare delivery, or improved production management in large industrial factories, Predix is designed to support the future of business.

Predix provides the capabilities needed to develop new, valuable industrial solutions more easily than ever before.

For more information on how to get started using Predix, visit <http://www.predix.io>.

About the Author

Steve Winkler is a software industry veteran with almost two decades of experience in enterprise software. He is a principal architect at GE Digital.

predix.com

©2015 General Electric Company – All rights reserved.

GE, the GE Monogram and Predix are trademarks of General Electric Company.

No part of this document may be distributed, reproduced or posted without the express written permission of General Electric Company.

THIS DOCUMENT AND ITS CONTENTS ARE PROVIDED "AS IS," WITH NO REPRESENTATION OR WARRANTIES OF ANY KIND, WHETHER EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO WARRANTIES OF DESIGN, MERCHANTABILITY, OR FITNESS FOR A PARTICULAR PURPOSE. ALL OTHER LIABILITY ARISING FROM RELIANCE UPON ANY INFORMATION CONTAINED HEREIN IS EXPRESSLY DISCLAIMED.