SkySpark's Role in Achieving the Promise of Digital Twins for the Built Environment

Comprehensive Data Platform Providing Live and Historic IoT Data, Analytic Results, KPI's and Data Visualization Tools

SkyFoundry

SkySpark – The Data Platform for Operational Data

Providing normalized data from sensors, equipment systems, databases and web-services in a comprehensive, open, data platform

As a platform for acquiring, managing, visualizing and performing analytics on facility and equipment data, SkySpark has a key role in achieving the promise of comprehensive digital twins. By combining SkySpark's industry leading capabilities for data acquisition from real world equipment systems, the highly efficient storage of its Folio database, its high-speed industrial historian, analytics processing engine, and comprehensive suite of data visualization Apps, SkySpark provides essential capabilities to achieve the vision of digital twins for the built environment.

Contents

Introduction – Digital Twins – Clarifying the Term

One of the new areas of technology gaining traction in the built environment is “digital twins”. The term conveys the concept of a digital representation of a facility implemented via a software application. The representations can extend from the physical structure of the building to the actual data produced by the equipment systems that maintain the occupant environment, and theenergy and resource flows associated with those systems.

A familiar example of a digital twin is a 3D model of a building which allows a user to drill down into each and every space showing the size and shape of the space, and even the finish materials and furnishings. Today’s state of the art takes those examples further to include sensor, equipment and environmental data such as space temperature, air quality, occupancy status, and energy use. Bringing all of these sensor and equipment data sources together into a normalized data store is a key function SkySpark brings to digital twins.

More Than Pretty Pictures

At first look, it’s easy to conclude that digital twins are nothing more than fancier 3D “color graphics” of buildings and equipment systems that have been the mainstay of automation systems for more than 30 years. While this may appear to be the case, the difference lies below the surface with range and types of the data embraced by the digital twin, and the techniques provided to generate value from that data.

Where basic automation system graphics focus primarily on showing live equipment data on a graphical representation of the equipment, a digital twin may include a software model of the building that can respond to inputs to show expected responses and performance of the building. For example, a digital twin might show how the building responds to external environmental conditions of temperature or solar load, and model energy and resource flows in response to these factors along with occupancy patterns and space utilization.

Digital Twins – a Definition – or Two

As noted, the term Digital Twin is often associated with a 3D model of a facility. These physical representations are typically derived from a BIM data model of the physical asset (BIM – Building Information Model – https://en.wikipedia.org/wiki/Building_information_modeling). However, Digital Twins are not limited to this type of implementation. Some definitions will help us in our discussion.

The Industrial Internet Consortium (IIC) defines a digital twin as follows:

A digital twin is a formal digital representation of some asset, process or system that captures attributes and behaviors of that entity suitable for communication, storage, interpretation or processing within a certain context. The digital twin information includes, but is not limited to, combinations of the following categories:

- physics-based model and data
- analytical models and data
- time-series data and historians
- transactional data
- master data
- visual models
- computations

https://www.iiconsortium.org/news/joi-articles/2019-November-JoI-A-Short-Introduction-to-Digital-Twins.pdf

The Digital Twin Consortium™ offers this definition:

A digital twin is a virtual representation of real-world entities and processes, synchronized at a specified frequency and fidelity.

https://www.digitaltwinconsortium.org/

At SkyFoundry, we introduce an additional term – “digital data replicas”. We find this term is more representative of the reality involved, especially considering that the term “twin” implies there are only two instances of the digital data representation. As you will see, there can be many digital replicas of facility and equipment data to serve a wide range of applications that provide the digital twin experience for users.

Combining Physical Data and Operational Data

Combining digital data that represents both the physical model of a building and the operational data associated with equipment systems, sensors, meters, and devices that operate the facility is an essential element of achieving digital twins. Effective normalization of that data provides the basis for a range of digital twin applications. SkySpark plays a key role in acquiring, processing, storing, normalizing, and visualizing operational data and efficiently conveying it to other applications.

Getting There: Combining Data from Diverse Systems and Sources

Even in modern buildings with "smart" devices, equipment systems are often siloed due to a variety of factors. Some examples:

- different systems use different protocols and data formats creating barriers to integration
- building automation system data is often not combined with utility rate data, which is essential to calculate actual energy costs
- in many cases, data comes from external websites via "web-services" requiring integration via API's (Application Programming Interfaces) –this involves some level of software development work, which can range from simple to complex
- software applications often utilize proprietary databases that do not provide documented "schemas" to describe the meaning of the source data in a standardized format, this creates a barrier to easily query and utilize data in other applications
- even the newest IoT devices and sensors are often "islands" with their own separate networks, user interfaces, communication protocols, and "closed" data repositories which are often available only via API's that require "pay to play" to access the data

The message – no one system or data source has it all. A fundamental requirement to achieve a digital twin is the ability to combine data from diverse systems and normalize it for consumption by the applications that form the digital twin experience. We consider that "Job 1" of an IoT data platform.

Why It Matters

Facility managers cannot have effective situational awareness and the ability to ensure optimal performance of their facilities without a platform that effectively brings together the data coming from diverse equipment systems. Unifying this operational data is a key element in the "stack" of technology to accomplish a digital twin and is an essential capability provided by SkySpark.

Key Requirements in Achieving Comprehensive Digital Replicas of Operational Data

Effective digitization of the built environment involves more than just onsite sensors, meters, equipment, and automation systems. The ability to connect to diverse data sources "beyond sensors" is essential. Some examples include utility feeds of consumption, demand, pricing data and tariff models, asset data including facility characteristics, and production metrics. All of these need to be brought together to create a comprehensive digital representation of a facility.

Data Integration – Key Requirements:

- Communication connectors for diverse systems and data sources
- Ability to normalize data by applying metadata using an accepted industry standard

Combining diverse data heightens the need to add descriptive information to the data (metadata). This needs to be done using a standard metadata methodology that provides consistent, high-fidelity semantic information to represent the meaning of the data.

This means that the data platform must provide more than an aggregated storage repository that data is simply "poured" into, i.e., a "data lake". It is essential that the data platform combine descriptive "metadata" with the source data to make it useful to other applications. Applying metadata is an essential element of implementing an effective digital twin. For modeling of the physical attributes of a building, BIM is the industry standard. For modeling of the operational data associated with building equipment systems, Haystack is the most widely deployed standard and is the core data modeling standard utilized by SkySpark (http://project-haystack.org).

Is It Live or….

Digital twins require more than the presentation of "live" sensor and equipment data. Many users needto be able to work with a complete functional "replica" of their data – even when offline from the data producing systems. This is a key distinction *–providing features for analysis of historical data is an essential requirement.*

Ideally, we want to provide users with access to a complete replica of all data up to the last moment the data source was available. This is especially important in the case of interruptible/unreliable network connections, and in applications such as energy and performance analysis, benchmarking and M&V that are often performed "offline" and require extensive stores of historical data. SkySpark accomplishes this through a feature we call "replication" which we will highlight later in this paper.

SkySpark's Role – The Data Platform for Operational Data

Bringing Together Sensor and Equipment Data, Web-Services Data, Historical Data, Analytic Results, Fault Patterns, KPI's, Energy Cost Analysis and More in a Single Platform

As a platform for acquiring, managing, visualizing and performing analytics on facility and equipment data, SkySpark has a key role in achieving the promise of comprehensive digital twins.

By combining the platform's industry leading capabilities for acquisition of data from real world equipment systems (see sidebar), the highly efficient storage of data in its Folio database, its high-speed industrial historian and analytics processing engine, and comprehensive suite of data visualization apps, SkySpark provides essential capabilities to achieve the vision of the digital twin.

To see how buildings perform over time, we need to collect and efficiently manage large volumes of sensor and equipment data, aswell as asset data and other data. For example, to determine howa building responds in relation to design expectations we need to be able to compare actual results with expected performance. Theoutput of a building energy model can be brought into SkySpark as a data set for comparison to actual operating data.

Going further, mathematical models can be processed directly in SkySpark using its scientific math package and machine learning tools to forecast and predict outcomes such as energy use, and trends leading to loss of performance and failures.

SkySpark supports a variety of data acquisition connectors including:

- BACnet IP
- Modbus TCP
- Obix
- Haystack
- SNMP
- Sedona
- OPC UA
- MQTT
- SQL
- CSV import (manual batch or automated)
- REST API

SkySpark also includes a connector development toolkit for custom data connectors

AND all data in SkySpark is easily available to external applications via a fully supported REST API which provides data in a number of standard formats

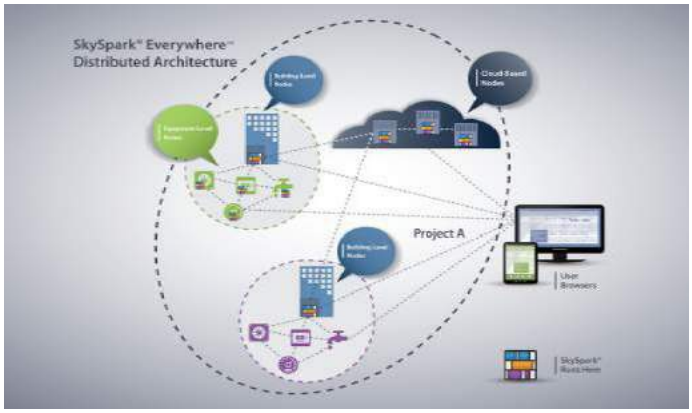
SkySpark is open to get data in and data out!

Delivering Data to External Applications

Because all of the data contained in SkySpark, including all analytic results, KPIs and other calculations, are directly available to other applications, SkySpark provides the data foundation to serve the wide range of applications that provide users with their "digital twin experience".

Digital Data Replicas – More Than Just Twins

SkySpark is unique in the world of IoT software because it can be applied from the "edge" to the "cloud", providing a fully distributed data and compute platform. Each node performs data collection, storage, analytics processing and data visualization whether at "the edge", the mid-tier, or the cloud. Applications that present data and results to users work as a unified, seamless system. Key benefits of this distributed architecture capability include:



- Scaling for Very Large Enterprise Systems
  - Multiple nodes of SkySpark can be deployed to support extremely large enterprise systems that process analytics on billions of data samples collected from equipment systems and IoT devices. SkySpark has numerous deployments operating at enterprise scale: https://skyfoundry.com/file/437/Case-Study-Proven-at-Scale---Enterprise-scale-Deployments.pdf
- Greater fault tolerance
  - By collecting data, processing analytics and creating visualizations as close to the data source as possible, SkySpark provides greater fault tolerance
  - An example is enabling in-building personnel full access to their data and analytic results – *evenwhen they cannot communicate to an external cloud-based application or data source*
- Low latency
  - Computing at the edge provides near real-time data acquisition, processing of analytics and control actions that cannot be accomplished with the latency of sending data to a cloud
- Support applications with "constrained networks"
  - IoT devices and equipment are often connected to slower, bandwidth limited or intermittent networks, or use cellular connections with high data transfer costs. By processing analytics at the edge, network traffic can be reduced by as much as 1000:1 !!!
- Security
  - Many projects require data be kept on premise. SkySpark meets requirements for projects that cannot send data to an external cloud or are subject to regulations for data storage location
  - Isolating in-building systems from the Internet – SkySpark acts as a security barrier for connected equipment with its Arcbeam, WebSocket-based protocol
- Save Engineering Costs
  - Using SkySpark from the edge to the cloud means you engineer yourapplication once with one uniform set of tools.
- Reduce hardware costs
  - SkySpark at the edge can eliminate the need for additional gateway andsecurity appliances in many applications.

Managing Distributed Systems – Replication, Clustering and Provisioning

When you have systems that consist of multiple computing nodes distributed across a facility or the world, you have another challenge – the need to manage the data they collect, the computational results they produce, as well as the software they contain. Distributed IoT applications also need to support the ability to connect "clusters" of nodes into a seamless system.

The term clustering refers to the ability of multiple, distributed SkySpark nodes to be connected into a system using SkyFoundry's highly efficient and secure Arcbeam protocol. Once connected into a cluster, users interact with their data, analytic results, reports and views as if they were interacting with a single computer and single database. The result is a seamless user experience even as data and processing are distributed across many computing nodes.

Replication

- SkySpark provides automated replication of nodes in distributed systems
- Replication provides a fully operational replica of each node that users can interact with – EVEN WHEN the actual node is OFFLINE
- User queries can work with last available data in a replica – saving significant data transfer costs
- Replication also provides a full automated backup of individual nodes
- Saves time, engineering and IT support effort

Going beyond simple backup of data. Replication is a SkySpark feature that enables copies of SkySpark databases from distributed nodes to be automatically copied (replicated) to one or more servers. Replicas are not simple data backups. Rather, SkySpark replicas are fully operational copies of individual distributed nodes that provide the full user experience. Even when the original data source(s) are not available (offline), replicas allow users to work with the last available data and analytic results to perform analysis and reporting using the complete suite of SkySpark Apps. And replicas can continue to make all of this data available to other "digital twin" applications, providing those users with continued operations, even when live data feeds are interrupted. This capability is highly unique.

Replication features allow databases from edge nodes to be synched (copied) to one or more servers on-premise, or in the cloud, on a configurable basis. The result are replicas that allow users to fullyinteract with their data from remote systems that may be offline or only intermittently connected.

The User Experience – Providing Operators with Relevant, Actionable Information – The "Last Mile"

SkySpark Comprehensive Suite of Visualizations Serves Different User Roles and Needs

Operators – service technicians, financial managers, energy engineers or customer experience professionals – need the ability to access and view the data they care about using a variety of apps, analysis tools and reporting applications.

Beyond acquiring, normalizing and storing the data from your diverse systems, SkySpark provides visualization tools that enable users to address their individual needs for operational management, analysis and reporting. In some cases, these augment the visualizations provided by digital twin applications. In many cases, they provide the digital twin representations needed by users without further development.

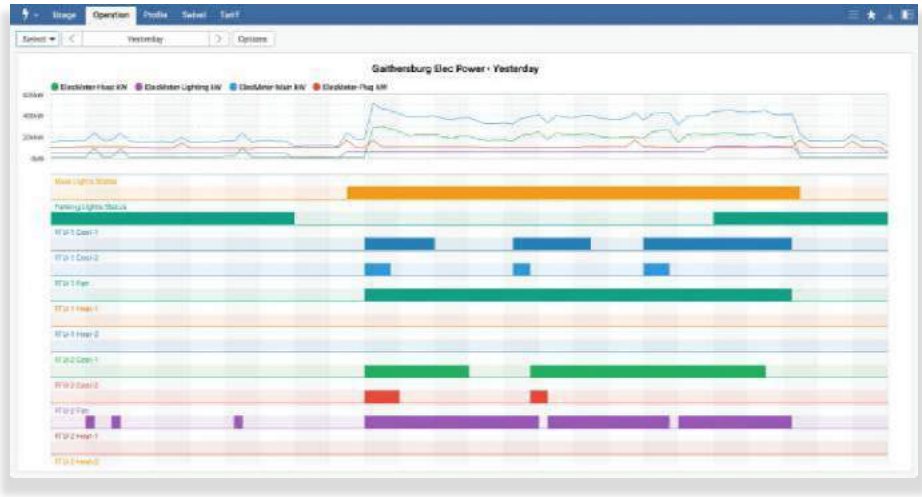
Digital Twin Applications Independent of Physical Models

Many uses of operational data are independent of physical models. One good example is combining energy consumption and demand data with tariff rate charges to show the impact of energy use patterns on actual costs.

In this view, the lines represent consumption (kWh) and demand (kW) and the vertical bars represent energy cost calculated based on actual tariff rate charges.

Another example shows the correlation of energy meter data (top) and equipment operation status (bottom)

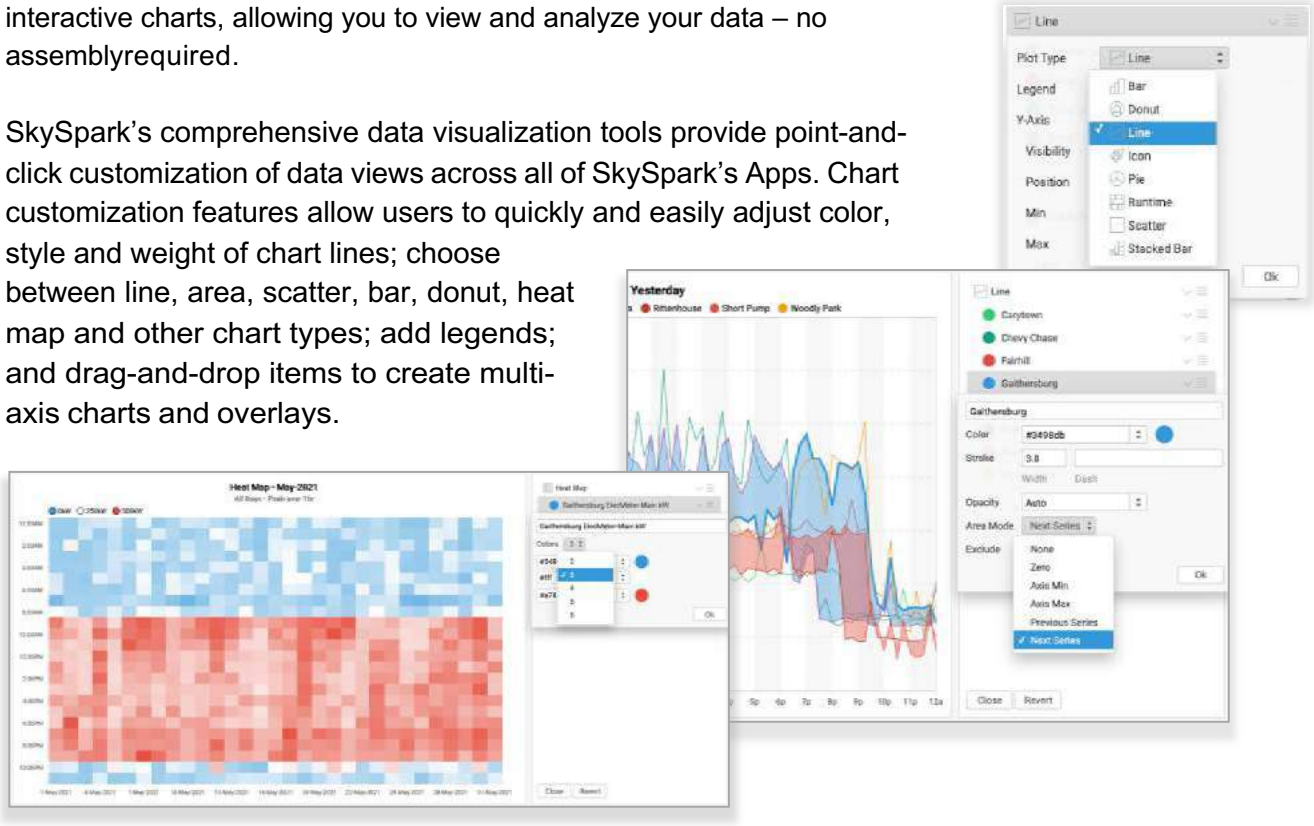
➔



Automatically Chart and Graph All of Your Data

Seeing your data provides insight. SkySpark provides the ability to instantly chart and graph your data using a wide range of data analysis tools and visualization tools. SkySpark's automated charting features mean that as soon as data is in SkySpark it automatically creates interactive charts, allowing you to view and analyze your data – no assemblyrequired.

SkySpark's comprehensive data visualization tools provide point-and-click customization of data views across all of SkySpark's Apps. Chart customization features allow users to quickly and easily adjust color, style and weight of chart lines; choose between line, area, scatter, bar, donut, heat map and other chart types; add legends; and drag-and-drop items to create multi-axis charts and overlays.



Energy Data Analysis – Comprehensive Tools Are Built-In

SkySpark's Energy App provides a suite of tools to view and analyze meter data for electricity, oil, gas, water – or virtually any metered resource – critical data for most digital twin applications. Whether it's data from physical meters or virtual meters based on calculations, the Energy App allows operators to quickly view and analyze meter data across any desired period of time. The App supports meter hierarchies and association of meters with specific equipment. Built in tools allow for easy normalization of energy data for weather, facility size and virtually any other normalization factor desired. Baseline analysis allows actual energy performance to be compared to previous periods and data from design models. And the SkySpark GhG App provides tracking and reporting of Greenhouse Gas Performance.

Calculating and Tracking KPIs. Virtually any metric that can be expressed with a math formula can be defined as a KPI. Once defined, SkySpark continuously calculates and stores KPI values. The KPI App then allows the user to view KPI values and their historical trends over time.

The SkySpark Tariff Engine – Calculate True Energy Costs

By combining energy use data with tariff-based costs, SkySpark provides true insight into the financial impacts of your actual energy use patterns. The rate modeler feature of SkySpark's Tariff Engine allows rates to be defined based on a wide range of billing "charges" including:

- Consumption, Demand, Distribution and Generation charges
- Service and equipment charges (both fixed-rate and percentage-based)
- Minimum contract charges
- Ratchets
- Time of Use including both time of day and monthly use factors
- Ranges (or blocks)
- Currency
- Custom charges, which can be expressed as math functions
- Definition of billing periods (including support for variable billing periods)
- Variable fees based on data retrieved from external systems

SkySpark's Energy App suite provides more features and capabilities than can be effectively presented in this paper. These documents provide further details on the Energy suite:

SkySpark for Energy Analysis, Commissioning and M&V: https://skyfoundry.com/file/411/SkySpark-for-Energy-Analysis-Commissioning-and-MV.pdf
SkySpark's Tariff Engine for Energy Cost Calculation: https://skyfoundry.com/file/402/SkySpark-Tariff-Engine-White-Paper.pdf
Tracking and Reporting Greenhouse Gas Performance: https://skyfoundry.com/file/459/Tracking-and-Reporting-Greenhouse-Gas-Performance---the-SkySpark-GhG-App.pdf

Summary – SkySpark's Role in Achieving the Promise of DigitalTwins for the Built Environment

We hope that this white paper has helped provide an understanding of the role SkyFoundry's SkySpark software plays in achieving the vision of digital twins for the built environment. As a platform for acquiring, managing, visualizing and performing analytics on facility and equipment data, SkySpark provides essential functionality necessary to achieve the promise of comprehensive digital twins.

By combining industry leading capabilities for data acquisition from real world equipment systems, the highly efficient storage of its Folio database, its high-speed industrialhistorian, analytics processing engine, and comprehensive suite of data visualization Apps, SkySpark provides essential capabilities to achieve the vision of "digital twins" for the built environment.

SkySpark is offered through a worldwide network of authorized resellers and is available directly from SkyFoundry for OEM's to combine with or embed in their software offerings. Learn more by contacting us at: info@skyfoundry.com

ABOUT SKYFOUNDRY

SkyFoundry's mission is to provide software solutions for the age of "the Internet of things". Areas of focus include:

- Building automation and facility management
- Energy management, utility data analytics
- Remote device and equipment monitoring
- Asset management

SkyFoundry products help customers derive value from their investments in smart systems. Contact us to learn more.

https://skyfoundry.com/

info@skyfoundry.com