Digital Twins or "Asset Management: Industry 4.0"

A Nextspace Whitepaper

Author: Mark Thomas, CEO and Founder; June 2021

Introduction

Digital Twins are the topic of heated and excited discussion everywhere, with good reason. They have been in various forms within science fiction movies for years, and now they are becoming a reality. Offering the potential to join multiple disparate IT systems and data together, to visualise and optimise the world in ways never before explored.

This white paper explores Digital Twin definitions and design principles, outlining the systematic approach in the development of the "Bruce Digital Twin Operating System", a system to create a common framework for the rapid development and interaction between multiple Digital Twin data environments.

In this paper we talk broadly of the concept of "things". This is an encompassing term for assets, entities and datapoints, conceptual or actual. We will discuss the relationships between "things" and their data, which ultimately will allow a true Digital Twin Operating System to create a conceptual model of the world which is relative to the real world. This combines two concepts that originate from the ancient Greeks, the "Method of Loci" or "Memory Palace" (https://en.wikipedia.org/wiki/Method_of_loci) a conceptual memory enhancement tool that allocates things to places within a familiar place to aid in recall, and the idea of categorising the world of things using the "ontological" approach.
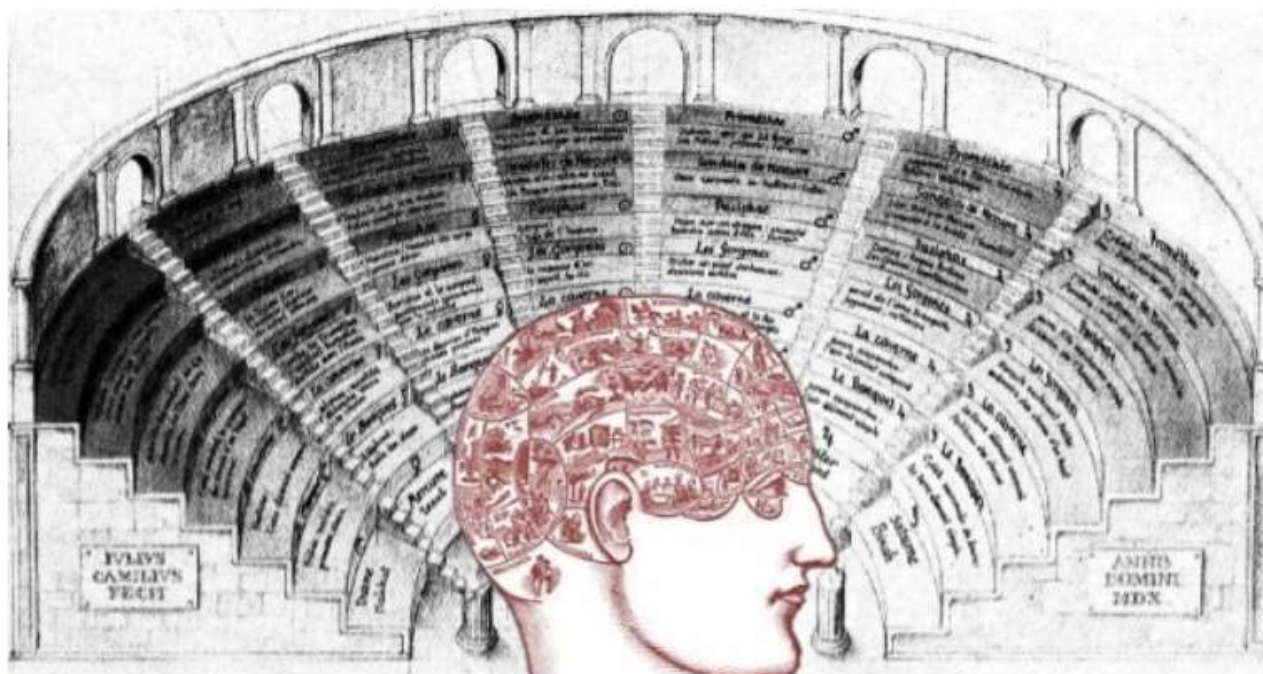
The loci method is about building a "memory palace."

Reference: https://www.inverse.com/article/28879-memory-improve-loci-training-new-study

Digital Twins are widely discussed and debated, with the "why" becoming increasingly clear. With an estimated 75% of a construction's lifecycle costs attributable to its post construction management, optimising the maintenance costs over time can amount to millions of dollars per, year even in a medium sized industrial facility. However, the exact "what" and "how" aspects of Digital Twins are without clear agreement or definition between various discussions, publications and promotional material.

This whitepaper attempts to present an inclusive and encompassing view of what a Digital Twin is, and some important considerations when considering an actual implementation.

Artificial Intelligence (AI) and machine learning will play an enormous part in this Digital Twin triad by finding and defining relationships between *things*. For example, identifying water reservoirs within satellite imagery; finding road signs in video footage; or individual pipes, valves and pumps within a point cloud scan of a process plant. Exactly *how* we can achieve this is now maturing, to the point where it will impact shareholders, boards and IT strategic decisions for the future operational state of their businesses.

"I didn't know" will no longer be an excuse for falling stock prices or more dramatically, keep directors out of jail. Therefore, it is now important to consider issues such as *openness* and *universal access* to those who need to interact, extend or join with a Digital Twin, now and in the future. The utilisation of a Digital Twin will and should span years and potentially centuries. So, it is important to consider how it will incorporate new types of data, new data standards, visualisation tools, services companies and data collection technologies, as they emerge.

By using a well-constructed *Digital Twin ontology strategy,* it is in fact possible to anticipate the future of data, data visualisation and AI technology and to accommodate the changing actors involved in the creation, use and management of Digital Twins. Existing and future stakeholders; public crowd-sourced data; State and Federal amalgamations of data; holographic augmented reality visualisation, etc. These future use cases can all interact throughout the lifecycle of a Digital Twin using the right ontological operating system as the foundational starting point.

Changing Definitions and Scope of the Digital Twin Description

There are many and varied perspectives on what a Digital Twin is. Digital Twins have been specifically associated with visual simulation and analytics applications which seek to model the behaviour of systems with inputs and outputs that attempt to mimic the real-world behaviour of systems. These types of Digital Twins have certainly been valuable for the specific optimisation of assets, processes and systems against goals such as production efficiency, traffic flow, crowd behaviour and so forth. However, these specific types of Digital Twins have not necessarily taken advantage of the power of visual communication back to humans. More recently, it seems any 3D graphic representation or flythrough of cities, factories or other systems, with associated data, are often referred to as Digital Twins. But it is important not to confuse a visualisation with a simulation . Visualisations are *not simulations* per-se, but rather the graphical presentation of pre-calculated data from actual simulation systems with analytical engines running separately. These may be connected real-time or manually.

Rather than suggest that different more specific descriptions are wrong or right we propose a broader, more encompassing vision, of what a Digital Twin is. We suggest it should be considered as a collection of 3 key components, which will enable Digital Twins to create better futures; a way to create value to us as humans. Or *Asset Management for Industry 4.0*.

1. ### A State engine – A way to represent the world as it existed, or exists, or may exist

   This requires a pure data ontology management system that transcends any particular IT or data source system and is able to arrange the world into "things", their attributes of interest, the values of those attributes (past present and future), their relationships to other things, and lastly all the possible valid visual representations of that thing; past, present and predicted. An important ancillary capability the state engine must have is the ability to read and de-construct multiple data sources into this model. And also, to be able to write back to data repositories in the required formats as needed.

   Important Note - while a Digital Twin system may include a data repository and/or data caching function, this is not a required component. It need not be a data repository itself, rather it is more importantly a map between existing data sources and the recording of

relationships within and between these data sources. Therefore, the importance of Global Unique ID (GUID) management, between connected systems, cannot be understated.

2. A Visualisation engine – A Digital "Memory Palace" anyone can use

Visualisation is optimal for human cognition of the state engine (Past present or future). This is why the Greeks invented the Method of Loci or Memory Palace concept. We have the highest data bandwidth through our eyes and so efficient and effective visual communication is critical to understanding the Digital Twin state engine and future possibilities. It takes human understanding, decision making and real-world actions to create better futures. This is an area of rapid development with many new emerging technologies in Virtual Reality, Augmented and mixed reality. Each new development technology potentially requires new data formats and connections. Major developments in this area are the only certainty and so it is important to anticipate and support these new developing technologies with an ontological Digital Twin operating system. New Visualisation systems tend to demand new graphical representations, and structures, but retain standard data association processes. By managing data and its visual representation, as potentially one to many relationships, it becomes relatively trivial to have multiple data synchronised visualisations, of the same things on different visual platforms. Even ones yet to be invented.

3. Simulation and Optimisation – A way to make a better future

Beyond being a useful tool to see some current state of the real world there is of course the need to optimise that world for better performance, quality, ROI, 'greenness' etc. Tools to simulate and then optimise a Digital Twin are a natural next step.

Digital Twins are a symbiotic way to describe not only a current state but also a dynamic system; where behaviour can be embedded in the relationship between *things*. For example, a valve *thing* is connected by the relationship "*connected upstream from*" a pipe *thing, S*o if I change the attribute state of the valve to "*off*", water should stop flowing – virtually in the Digital Twin. This is of course a simple example and relationships may in fact be very complex multi-input, multi output processes, but it highlights how a Digital Twin creates not just a unified model but in fact a connected system. When connected with AI, analytics and simulation the system can be used to optimise the future against set goals such as carbon footprint reduction; economic efficiency; stakeholder satisfaction; optimal maintenance programs; traffic management etc.

Simulation and optimisation are the domain of many different experts in their respective fields. There are 2 significant challenges in applying simulation and optimisation tools

   a) The availability of current and consistently formatted data to feed a simulation engine – This is supported by the State Engine as described in 1. above
   b) The ability to visualise and communicate large and complex data sets of both the inputs and outputs – This is supported by the Visualisation Engine as described in 2. above

If the three core components of a Digital Twin are able to seamlessly and dynamically communicate as a system, then the value of a Digital Twin is greatly increased.

A Data Framework - Why data ontology is so important

A pure ontological approach to data integration is central to Digital Twin advancement - it provides a useful framework to model the world and connect data about the *things* of interest to us. An ontology approach is both respectful but independent of existing IT systems and data schemas, and seeks to define the true nature of *things* in a way that is equally comfortable for a computer as for a human.

Applying this approach creates data models and data flows that can be easily joined together like Lego blocks to form larger and more comprehensive Digital Twins - from a machine level, to a city level, to a

state level, to global scale. This approach also makes it possible to search and find anything based on attributes, position, relationships and so forth without returning to source systems for separate searches.

Using data ontology provides a system-independent way to structure and join data. A good use of ontology for data integration simply asks the following questions of any incoming data set, (static or dynamic).

For example;

- "What are the things that are referenced in this data?" (e.g., walls, valves, doors, pipes, supports etc)

- "What are the attributes that are referenced to these things?" (e.g., Size, manufacture, temperature, pressure, date last maintained etc)

- "What are the implied relationships between things?" (e.g., parent, child, connected to, supplied by, links to other data objects: media, sound, video, documents etc)

- "When and where are these things and their data?" (Lat, long, elevation, relationship to another coordinate system, time of creation, destruction date, time to build etc); and

- "What are the valid visual representations of this thing?" (Point, line, polygon, 3D model, photo, LIDAR scan, point cloud etc; varying resolutions, varying file formats)

It is pleasantly surprising how using this ontology framework to systematically categorise the world into components like this when connecting data from different sources, adds a layer of simplicity and calm to data integration problems.

The reason for this pleasant surprise is perhaps that the universal ontology concept is well tried and tested philosophically independently of the computing world and is, as previously mentioned, as old as the Greeks. The idea of ontology attempts to describe the world in a systematic way that is purely logical, providing that critical bridge between machines and humans.
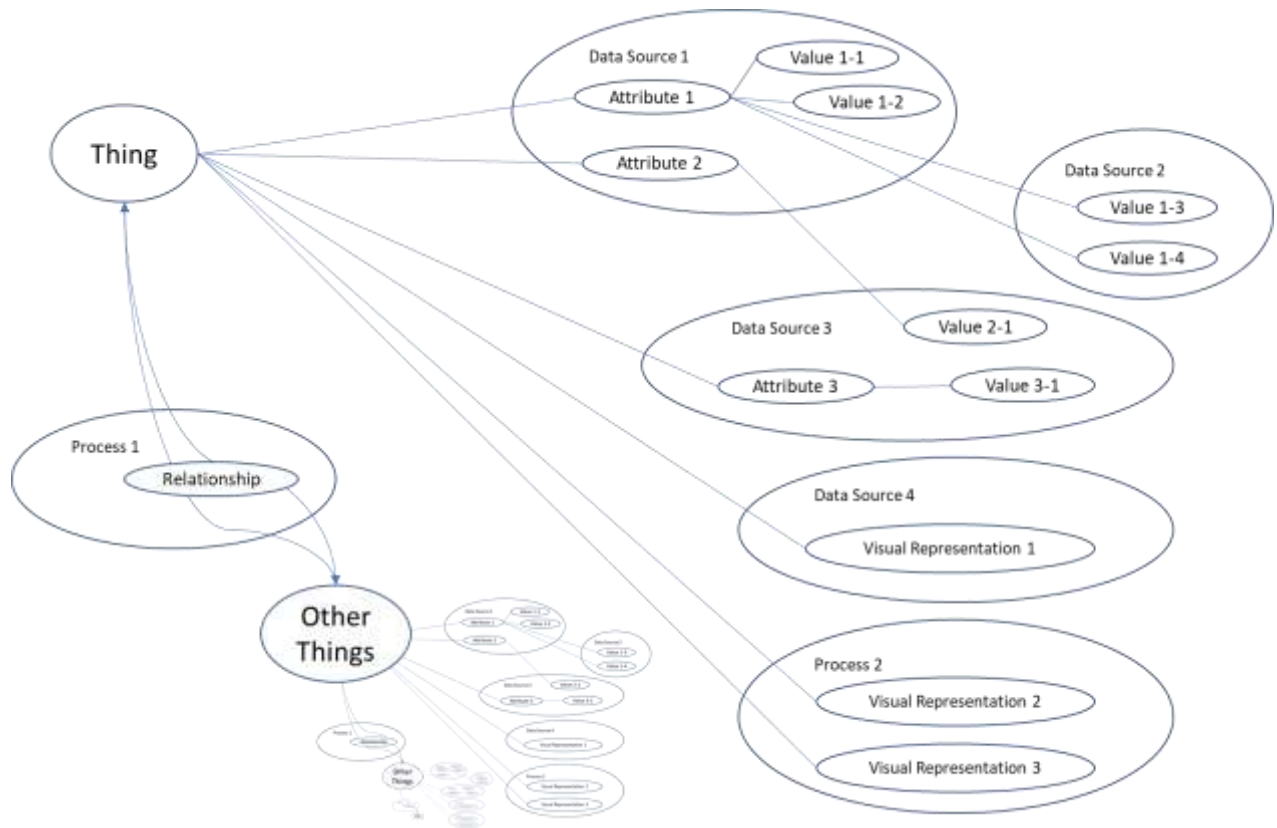
A simple diagrammatic example of a Digital Twin ontology.
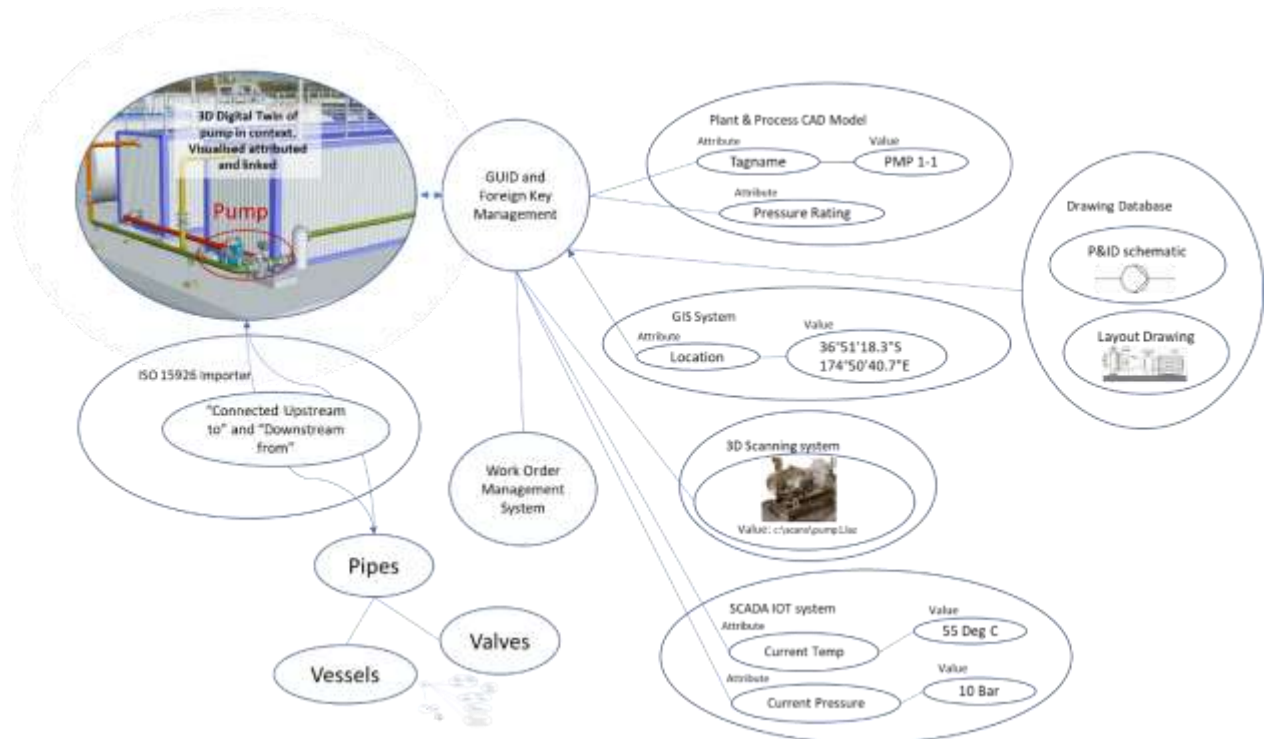
A simple specific diagrammatic example of a Digital Twin ontology in use.

It is important to note here however that an ontology management approach to data integration does not seek to replace standards but rather to provide **a flexible framework within which to define, join and evolve standards**. A well-executed data ontology management approach allows standards to evolve

and change without the sometimes-crippling impact that changes, or even suggested changes, to standards have when embedded in discreet disconnected data repositories and models.

New Technologies and Digital Twin developments

Having expanded the definition of Digital Twins it is useful to consider the impact of recent advances in technology driving the development of more solid foundations for Digital Twins, and other aspects of the useability and scalability of this technology.

In fact, implementing a Digital Twin is no longer a technology problem. The technology is here today and its implementation is imperative for organisations that intend to stay competitive and relevant, now and in the future.

The last 5 years has seen a dramatic step-change in several key technology advancements, crucial to the advancement of Digital Twin technology:

- Cost effective cloud-based technologies - data storage and processing scale is no longer an issue

- The proliferation of IOT applications and devices – easy access to critical, usable, up to date data

- Open Data initiatives demanded by large customers and the provision of Open API's (Application Programming Interface) by most forward-looking IT vendors. These APIs are the tools to connect systems – but a framework for connection management is required.

- Artificial Intelligence and machine learning is making more sense out of the mass of data available in digital form about our world.

- Web-based rendering; advanced data rich visualisation can be delivered anywhere to any device

- Virtual and augmented reality provide new ways to see, experience and absorb richer data for humans

- Rapid advances in computing power especially those using massive arrays of GPU's and quantum computing make simulations of very complex systems feasible and realistic for many more applications.

These critical technology advances enable us to re-imagine the creation, use and users of Digital Twins - what they should now be capable of and where the future of the industry is now headed.

So how can this new technology be applied to Digital Twin enablement and strategy? Below we list out a few of the implications and impacts related to specific new technological developments.

Federated Digital Twins

Comprehensive and inclusive Digital Twins will not be created from the top down, but from the bottom up. They will be sourced from many smaller implementations to create larger and more comprehensive representations of our world, spanning more data sets. The application of universal ontology concepts to data modelling allows separate Digital Twins to be created for widely different purposes. From machine level through to an organisation level, through to city, state and national level, to be joined together to create new insights into our world and businesses and the way they interact with each other. Cloud technology allows for the integration and delivery of federated systems to be realistically achieved.

Integration of dynamic Real-world data

GPS tracking and IoT (or wide area SCADA as some describe it as) is enabling the collection and viewing of massive amounts of data that needs to be seen, analysed and optimised and re-envisaged by many different users for many different use cases. This is all naturally achieved with a good Digital Twin operating system strategy.

IT Agnostic Integration capability

Application Programming Interfaces (API's) are now widespread for any major IT or data acquisition systems. This allows the creation of Digital Twin implementations which are truly open, structured and consistent; future proofing for new technology applications, new data sources, new service providers, and new partners.

Artificial Intelligence

AI offers incredible new ways to create wisdom from data; discovering new relationships, new "things" and better possible, optimised futures. AI requires the structured and consistent data, with hierarchy, states and relationships that an ontological Digital Twin operating system can provide.

New visualisation technology

3D, virtual, augmented and mixed reality, HoloLens and Magic Leap systems; all offer amazing new ways to *see*. An ontological Digital Twin operating system provides for structured data and visuals to be seen and understood, in these evolving computer/human interfaces; making them a truly valuable addition to serious business.

Analytics and simulation systems

The rapid rise in computer power, linked to simulation and optimisation applications, offers solutions to many of the world problems, as they relate to inefficiency or unpredictability. However, these systems need reliable consistent real world data feeds and a way for humans to understand the inputs and outputs. This is a natural connection with Digital Twin implementations.

Some common mistakes in DT system development and delivery

In the author's view and from experience, there have been a few false starts and misrepresentations of Digital Twin technology. Especially by many of the existing CAD, GIS and games developers. Below are some of these perceived missteps and misunderstandings which are frequently seen.

Limitations of CAD, BIM and GIS for Digital Twin frameworks

In the past the compelling nature of animated interactive 3D graphics has created a breakthrough in how we visualise the actual world in all its complexity and possibilities. However, it is important to remember the place which visualisation has within the Digital Twin technology space. It should never form the primary data asset management strategy for a Digital Twin.

It is very tempting to take CAD, GIS, BIM and even games models as a starting framework for Digital Twins. While it is vital that these formats are able to be ingested and understood, as they form a good basis for the visual aspect and even some of the data aspects of a truly useful Digital Twin. However, the "visual first approach" or layered map-based approaches has many pitfalls.

It is vital to be able to manage the attribute structures and data values of *things* outside of the originating authoring system. The field names, values and structure of those things need to be manipulated independently, and amalgamated and extended well beyond the information and structures that are defined early in the lifecycle of *things*.

CAD files, or any graphics containers, are not good long-term data containers for the lifecycle of *things*. These file types change, and continue to evolve over time. Can you find the right software to read a file which is now 20 years old? That is why open data standards are critical. Additionally, typical CAD and other graphic file formats, open or proprietary, are not designed for managing data or data changes over time. Let alone changes to the very structure of the data. That is what database technology is for.

GIS systems typically do use GUID and database oriented and are great source of data for a Digital Twin. However GIS systems approach commonly represent the world from a 2D cartographic map approach – attributed points, lines, polygons and layers – an incomplete model of the real world, and while some allow importing data such as 3D BIM models at visual level and even providing some attribute access, a complete editable scalable Digital Twin with multidimensional relationships between things and a single unified, extendible data model independent of viewing systems is not typically possible in a pure GIS system.

Where and when to control attributes structures and relationships

Reordering *things,* depending on use case or context, is near impossible using traditional file-based representations of complex models and data combined. Users get stuck with the engineering hierarchy, which is different from the manufacturing hierarchy, which is different from the parts ordering structure, which is different from the valid configuration's hierarchy.

For example, most data that is required as part of design, or quote, or build, of a building (or some other structure) is not critical to the ongoing maintenance repair or financial management of that asset. New attribute types such as "last maintained", "current book value" or "owner contact details" are unlikely to exist as values or even attribute types coming fresh from a CAD or BIM system.

Different hierarchical structures of *things* are required to be presented depending on the use case in design, manufacturing, ordering financial management or maintenance (on say an aircraft) and these will differ significantly from the data and hierarchy prescribed in any specific CAD or GIS data system. And any new relationships not created in ERP, CAD or GIS or other applications will need to be discovered, described and attributed, to understand how a collection of assets behaves and can be optimised.

The importance of uniqueness GUID handling

While visualisation - both 2D and 3D - is an important component of Digital Twin technology, it is optional and as suggested already, certainly a mistake using it as the sole basis to establish the uniqueness of a *thing*. This "visual first approach", of taking graphics systems and attempting to add attributes and structural data and even GUID's and store in a single proprietary file, will ultimately create impassable integration problems and inconsistent and out of date Digital Twin models that are of little value.

A modelled state of a visual representation, *is not the essence* of a *thing*. It is only as useful as the specific use case, at a specific time. The essence of a thing should contain a unique identifier GUID. A graphic or 3D model is not a useful consistent Global Unique ID (GUID), yet is often treated as one.

GUID management is an *absolutely essential* part of Digital Twin data change management allowing changes and updates to be processed. Identifying what is a revision, version or replacement at a component level within a partial Digital Twin model update becomes an important aspect of a reliable Digital Twin representation. Extracting GUIDs from existing file or relational database oriented systems is vital but using a graphics representation as the absolute definition of "unique" is simply inadequate for purpose.

Summary

Digital Twins driven by a pure ontology model have the capability to now become truly valuable. They can be seen, communicated and optimized against chosen goals (financial, environmental, social or other), creating meaning, context and value from raw data to a wide range of customers, stakeholders and communities.

The creation of an effective, scalable and lasting Digital Twin relies on some key components and principles as discussed above.

- Things have a life cycle, they are conceived, designed, modified, built, maintained and retired. Data attribute, types, structures of data, visual representations and ways to visualise the data will inevitably change over the lifecycle of things or assets to suit different users and use cases. This needs to be anticipated and accommodated in an ontological Digital Twin operating system strategy

- It is vital to disassemble and categorise these data types into open data sets independent of any particular data authoring system. This is a lesson learned the hard way by leaders in Digital Twin thinking from the aerospace industry.

- The data about *things* and the many possible visualisations of a *thing* and its relationships, need to exist independently and contiguously outside of any specific proprietary file formats. Such combined data structures where relationships between *things*, their data, and each other are effectively hidden within the proprietary file structures become impossible to manage, synchronise, translate and communicate over time.

Without a doubt, this new approach to Digital Twins, with data ontology at its core, is central to the growth and success of the industries in which it serves. Without it, key and important technologies, which will exponentially impact and serve the industry, simply won't be accessible or leveraged.

_____

More Reading

Definitions:

- https://www.merriam-webster.com/dictionary/ontology
- https://en.wikipedia.org/wiki/Ontology_(information_science)
- https://medium.com/sciforce/ontologies-and-semantic-annotation-part-1-what-is-an-ontology-1de10caf2c77

Papers

- Ontology-based modelling of part digital twin oriented to assembly https://journals.sagepub.com/doi/abs/10.1177/0954405420941160
- BIM Ontology – A Visual Knowledge Model https://bimexcellence.org/resources/200series/291in/
- Data management for developing digital twin ontology model
- https://journals.sagepub.com/doi/full/10.1177/0954405420978117
- The ontology in description of production processes in the Industry 4.0 item designing company
- https://iopscience.iop.org/article/10.1088/1742-6596/1059/1/012010/meta
- The projection of cyber and physical systems digital twins in the stage of production technological preparation of the industry 4.0 smart factories
- https://iopscience.iop.org/article/10.1088/1757-899X/537/3/032009/meta

Commercial Articles

- https://www.challenge.org/insights/digital-twin-ontology/
- https://docs.microsoft.com/en-us/azure/digital-twins/concepts-ontologies

Ontology Tools

- https://www.nextspace.nz/brucemainpage
- https://www.w3.org/OWL/
- https://www.w3.org/2001/sw/wiki/RDF

_____

About the Author

Mark Thomas is a visual computing and design veteran, of 30 years, with multiple patents to his name. His previous company, Right Hemisphere offered Digital Twin Solutions for such customers as Boeing, NASA/JPL, SpaceX, Gulfstream, Caterpillar, Chrysler and many others. The company created the Adobe 3D PDF technology and .U3D format licensing them to Adobe as part of the work with Boeing on their "Model Based Definition" strategy; an early version of Digital Twin technology. Right Hemisphere was funded by Sequoia Capital and ultimately acquired by SAP. Mark now heads up Nextspace, a company spun off from Right Hemisphere in 2007 and dedicated to creating the next generation of Digital Twin technology centred around the latest in cloud and web technology and incorporating all previous experience in this field.

About Nextspace

Nextspace is head-quartered in Auckland, New Zealand as a world-class ontology-data product company; bringing the most advanced *'Digital Twin Operating System'* to the world. **Bruce** is an agnostic **Operating System**, which means it will inter-operate with all and any data deemed important for the goals, as set by the user. Enabling sophisticated data integration, visualisation, analytics and machine learning for the Architecture, Engineering, Construction, Agriculture, Facilities Management Industries and more. To find out more about Bruce and Nextspace. Please visit www.nextspace.nz