Directive Driven System of Systems Approach to Visualise Data Chasms

Gash Bhullar[1] and Rachel Davies[2]

[1] Control 2K Limited, Waterton technology Centre, Bridgend, CF313WT, UK
[2] Digital Manufacturing Innovation Hub Wales, Waterton Technology Centre, Bridgend, CF31 3WT, UK

Abstract

With the continuous growth of automated data systems to control major tasks, it is clear that a new approach is needed to manage and decipher information across a range of interconnected, proprietary systems. Increasingly, process and production companies are progressing toward low carbon Net Zero and decarbonisation models to be deployed across the entire enterprise. This necessitates a system of systems approach, and this paper discusses how a "Directive Driven" model incorporating a vendor neutral approach can create interoperable systems that can visualise entire ecosystems beyond the standard Digital Twin concepts and bridge the chasms between operating domains of supply chains that provide products and services from raw materials to consumer goods.

Keywords

Industry 4.0, system of systems, interoperability, circular system development, directive-driven, vendor-neutrality, digital twin

1. Introduction

Building upon the concept of 'Big Data,' where the information extracted and shared between platforms and systems is often 'too large or too complex' to be handled by traditional data processing methods, it stands to reason that a new approach to visualising information will be required to meet the ever-increasing demands of business and society, both in terms of interoperability and interconnectivity between data services. The Variety and Velocity of information - two of the 'five V's' of Big Data [1] is increasing with each evolution of technology at each stage of development; from the processing of raw materials, to the development of embedded devices for use in low energy systems, and the end users of those systems who will demand a minimum impact on the environment.

The data created at each stage of a product's lifecycle is typically vendor-specific and inaccessible to other systems, creating vast 'Knowledge Chasms' between proprietary systems, as described in Figure 1.

Proceedings of the Workshop of I-ESA'22, March 23–24, 2022, Valencia, Spain
EMAIL: gbhullar@control2k.co.uk (G. Bhullar); rae@dmiw.co.uk (R. Davies)

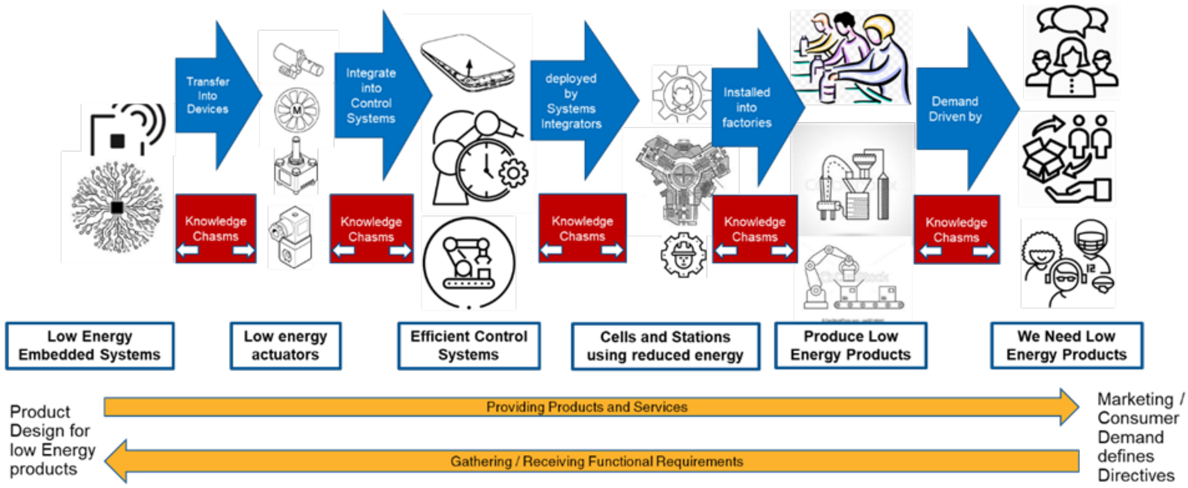

Figure 1: Complexity of data exchange throughout the supply chain

Fully understanding the impact of a product requires system data held within these knowledge chasms to be extracted and presented in a meaningful way so that the user can make sense of the data provided and take appropriate action.

1.1. Key challenges

New modeling tools are needed to support the visualization of the data, in particular the exchange of information in the chasms between domains. The challenge is to adopt a 'neutral format' for all stakeholders and create a vendor neutral platform to allow the flow of data across multiple platforms, using multiple protocols, using a System of Systems (SoS) approach [2].

The paradigm of system data has changed rapidly over the last 20 years, from manual distributed systems (local) to virtual centralisation (Cloud) and with it the need for advanced security and authentication methods. System developers now need to ensure data integrity across a very wide range of data sources as the risks of data breaches increase by 50% each year as volumes of data continue to increase [3]. Traditional 'Relational Database Management Systems' (RDBMS) are unable to guarantee data security and integrity across multiple domains [4], even when enhanced authorisation architecture such as 'Single Sign On' (SSO) is deployed [5].

Advanced Industrial Digital Technology (IDTs) such as Digital Twins were originally developed to help support the visualisation of system data by simulating a physical and functional representation of that system, to support design tasks or validate system properties [6]. However, in most cases these Digital Twins are based on a single data system or closed vendor ecosystem and are unable to interpret data from external data sources.

Providing access to multiple systems or linking data from a variety of sources poses a serious challenge for industry and there is a clear requirement to pull common data from these systems and extract the data held between systems in knowledge chasms, to develop an approach to deliver directives for key issues, such as decarburization, reduction in the use of rare minerals and materials and the circular economy as a whole, rather than continually trying to develop new systems to determine the overall impact of disparate components.

2. Bridging knowledge chasms

Whilst the majority of platform development begins with the definition of specifications for the target application domain, platforms within a technology genre are usually developed in parallel with other vendors rather than jointly and most only represent a partial solution to a problem rather than an end-to-end solution. It is up to the end user to connect the domains and glean meaningful insights from the information presented.

Research into this area has highlighted a number of trial industrial platforms and projects, aimed at linking together sources of data using either an Application Programming Interface (API) between proprietary services; or have focused on the use of Open Development Platforms. In most cases the Volume and Variability of the data has proven too complex for visualisation of the end-to-end lifecycle in one place [7], platforms are hampered by software engineering challenges such as the availability or security of data; and most fail to address the issue of why the data is needed in the first place.

The concept of 'blueprinting' or defining a directive or platform purpose has been explored with huge successes [8] where platforms are designed for a specific purpose, for example a product marketplace or a federated platform, such as a Data Spine model [9]. However, there is little evidence at the present time to demonstrate how this concept has been transferred to delivering on key industrial directives such as the circular economy, decarburization, or NetZero production across multiple vendor systems.

2.1. Managing multi-purpose platforms

Based upon our understanding of information architecture, the key to unlocking the data between these domains is to create and deploy software 'Agents' [10] that are "capable of autonomous, reactive and proactive operation in response to changes in their local environment". Agents, operating at the network edge, enable leveraging cloud resources into the proximity of the user devices. to interact with the core data in these systems and 'pull' information in a vendor-neutral format that can be accessed via most standard communication protocols, such as MQTT.

These Agents are essentially a code-driven protocol link between the edge devices of a system and the Cloud services that support that ecosystem, using the primary directives of the platform to interrogate the knowledge held in each system. Using this model, the market or consumer demand provides the directive for a platform, much like a user journey template which forms the basis of Agile Software Development, i.e. "As a (WHO), I want to (WHAT), so that (WHY)" such as is described in Figure 2.

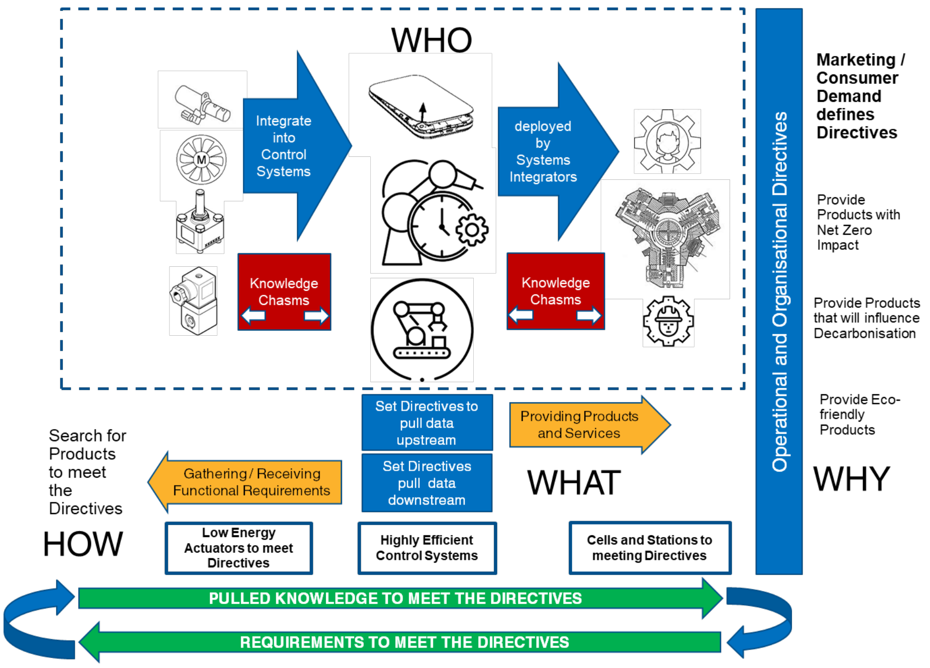

Figure 2: Creating Directives to fulfil Data Transfer

The agents are deployed downstream to stipulate the required data to meet the specific Directives and the corresponding knowledge is gained and pulled upstream to fulfill the Directives and presented via a visualisation system such as a Digital Twin or dashboard. Linking, passing through or sharing data minimises the traffic, optimises and rationalises the data (Smart Data); these event-driven Agents can provide a level of control over the flow of this data.

For example: In a data model that seeks to understand the NetZero impact of a water treatment facility, the user needs to understand data from a range of different systems. A common data value would be current consumption for a given motorised pump. An agent needs to span across all Industrial Control System (ICS) ecosystems to pull the current consumption parameters from multiple vendors that will be in a variety of different formats. So, the agent is given the directive:

"Please provide energy consumption of all pumps operating in your water treatment area stipulating a KWh value"

The Directive needs to be specific enough to minimise the amount of data returned so that the 'search' algorithm can be specifically channeled to the most appropriate system. The sheer volume of data that could end up being interrogated could potentially run into Petabytes and this warrants the use of AI Technologies incorporating XAI (explainable AI) in order to have a chance of understanding the required data, its format and its specific properties. In addition, audit trails need to be maintained to ensure track and trace functionality to prove the validity of the data. XAI is more suited to handling unpredictable events and provide diagnostic data to trace the potential breakpoints within the knowledge retrieving process.

2.2. Modeling the Interface of data

Economically, the advantages of crossing the chasms between knowledge ecosystems are held within a shared value proposition, or mutual complementarity where shared value can offer competitive advantage [9]. We need to understand the key issues and possible future solutions in order to fully exploit the potential value of digital twins, whether that be towards improved product development, more efficient production systems.

As an example of the types of interfaces to data systems, Figure 3 shows an Agent-based management of support systems for distributed brainstorming [11] where Usability-aware Service Orchestration System (USOS) and a Flexible Support System (FSS) for distributed brainstorming can have multiple connections to edge or cloud services. Agent-Based Computing (ABC) is suitable for implementing robust scalable systems and interoperable and virtualisable 'things'. ABC is suitable for supporting the design and implementation of autonomous IoT systems [12].

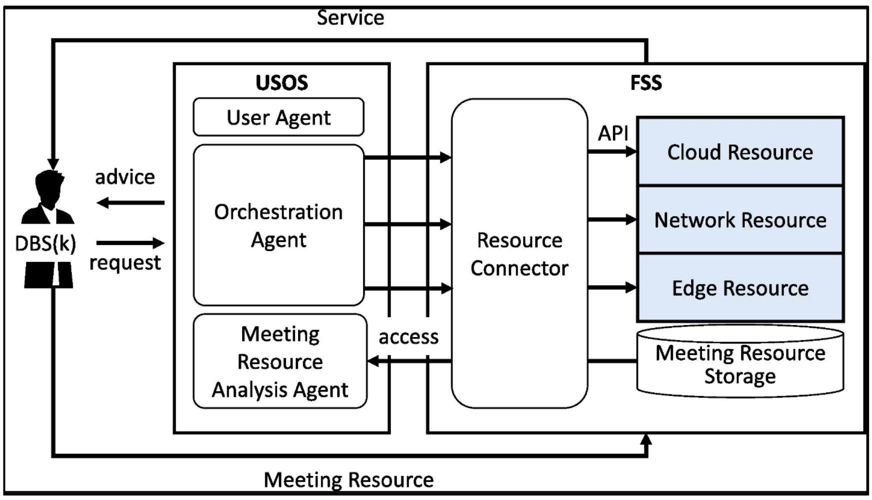

Figure 3: Agent-based and resource-oriented FUSSDB architecture

2.3. Visualisation systems

Graphical representation of data has always been the easiest way to present data to the users / decision makers as a majority of people respond to graphical representation rather than steams of digits or lists but the challenge is finding the right representation. Modern methods to present data now include holographic imaging [13] and the more recent developments to present data via the Metaverse [14] which incorporates an substantive range of characteristics as shown in Figure 4.

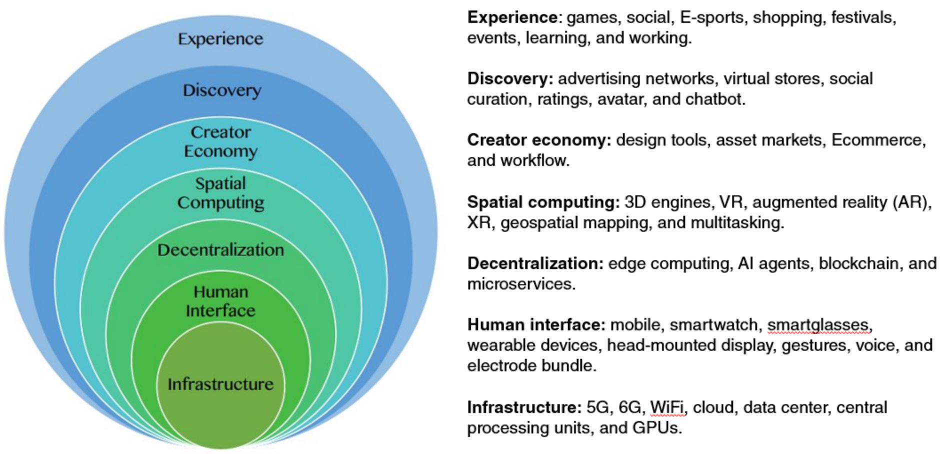

Figure 4: Metaverse – Seven Layers

Whilst these visualisation systems are clearly going to dominate the marketplace in the coming years, most people still prefer a more conventional 3D image of the world without additional layers of complexity. This is the normal cultural shift delay whilst systems become normalised and accepted way to visualise the real world in the same way as digital twins became accepted a way to represent the real world.

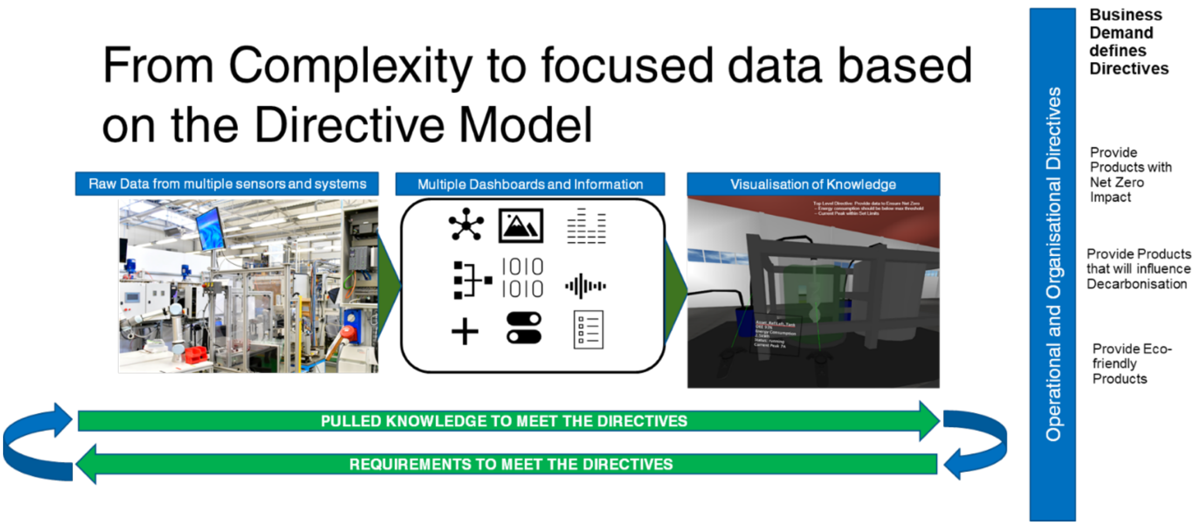

Figure 5: Pulling data into the digital twin and drilling down

The methods to visualise the knowledge coming from our Directive driven approach needs to concisely represent the original Directive and a funnel approach pulls complex and raw data from multiple ecosystems and initially presents it at a dashboard level to condition the data before finally presenting to the Directive owner in a manor that suits the knowledge being presented.

Figure 5 shows the way that data from a typical manufacturing environment can be conditioned via the agents to pull smart data that in this case be presented through the Industreweb Vactory™ environment [16] to visualise data from multiple shopfloor systems and presenting in a way that engineers and managers can understand.

2.4. Vendor neutrality and neutral formats

Vendor neutrality is the result of creating an open dialogue within an ecosystem allowing multiple systems to intercommunicate using agreed protocols. This compatibility and interchangeability / interoperability works at a systems level. Vendor-neutral specifications encourage the development of competing yet compatible implementations, freeing the purchaser to choose from a multitude of vendors without suffering a loss of functionality" (Agnostic solutions) [15]. The adoption of cloud technologies and digital platforms has been restricted in the SME sector due to potential 'lock-in' issues with vendors, resulting in limited interoperability and fixed formats for cloud data. Neutral formats (agreed formats to exchange data and protocols) have been a subject of many EU projects such as STASIS [17] and conceptually appear to be a good way to semantically map information from one system to another but with the volume of data now present globally, the storage of such a repository would cause the same problems as relational databases being stored globally.

3. Summary

A Directive Driven based Approach to pull data from complex systems including data from multiple digital twin environments serves multiple purposes to streamline operations, reduce waste, improve efficiency but most importantly provided the opportunity to maximise returns on investments through evidence provided by Visualise KPIs which are essentially what are behind the Directives.

Data Chasms are reduced by all systems working towards the same directives and thereby the types of data and information is normalised across several sectors when considering common goals for key drivers such as carbon reduction and recycling of goods and materials. The approach outlined in this paper also acts as an Enabler of Circular Economy by ensuring a closed loop approach to the data requested from the ecosystem and the alignment of that data to the key Directives required by stakeholders and decision makers. By being very specific with the Directives, it is the most effective way of interrogating multiple data systems.

4. References

[1] A. Jain, The 5 V's of big data, 2016. URL: https://www.ibm.com/blogs/watson-health/the-5-vs-of-big-data/

[2] J. Boardman, B. Sauser, System of Systems - the meaning of of, in: 2006 IEEE/SMC International Conference on System of Systems Engineering, Los Angeles, 2006, 6. doi: 10.1109/SYSOSE.2006.1652284.

[3] S. Wheatley, T. Maillart, D. Sornette, The extreme risk of personal data breaches and the erosion of privacy, The European Physical Journal B: Condensed Matter and Complex Systems 89 (2016) 1-12. doi: 10.1140/epjb/e2015-60754-4.

[4] Codeacademy, What is a Relational Database Management System, 2022. URL: https://www.codecademy.com/article/what-is-rdbms-sql.

[5] J. De Clercq, Single Sign-On Architectures, in: G. Davida, Y. Frankel, O. Rees (Eds.), Infrastructure Security, InfraSec 2002. Lecture Notes in Computer Science, vol 2437. Springer, Berlin, 2002. doi: 10.1007/3-540-45831-X_4.

[6] S. Boschert, R. Rosen, Digital Twin - The Simulation Aspect, in: P. Hehenberger, D. Bradley (Eds.), Mechatronic Futures, Springer, Cham, 2016, pp. 59-74. doi: 10.1007/978-3-319-32156-1_5

[7] A. B. M. Moniruzzaman, S. A. Hossain, NoSQL Database: New Era of Databases for Big data Analytics - Classification, Characteristics and Comparison, International Journal of Database Theory and Application 6 (2013) 1-13.

[8] S. Garcia-Gomez, M. Jimenez-Ganan, Y. Taher, C. Momm, F. Junker, J. Biro, A. Menychtas, V. Andrikopoulos, S. Strauch, Challenges for the Comprehensive Management of Cloud Services in a PaaS Framework, Scalable Computing: Practice and Experience 13 (2012) 201-213.

[9] R. A. Deshmukh, D. Jayakody, A. Schneider, V. Damjanovic-Behrendt, Data Spine: A Federated Interoperability Enabler for Heterogeneous IoT Platform Ecosystems, Sensors 21 (2021) 4010. doi: 10.3390/s21124010.

[10] T. Leppänen, C. Savaglio, L. Lovén, W. Russo, G. Di Fatta, J. Riekki, G. Fortino, Developing Agent-Based Smart Objects for IoT Edge Computing: Mobile Crowdsensing Use Case, in: 11th International Conference on Internet and Distributed Computing Systems, Tokyo, 2018, pp. 235-247. doi: 10.1007/978-3-030-02738-4_20.

[11] Y. Kaeri, K. Sugawara, C. Moulin, T. Gidel, Agent-based management of support systems for distributed brainstorming, Advanced Engineering Informatics 44 (2020) 101050. doi: 10.1016/j.aei.2020.101050.

[12] G. Fortino, R. Gravina, W. Russo, C. Savaglio, Modeling and Simulating Internet-of-Things Systems: A Hybrid Agent-Oriented Approach, Computing in Science & Engineering 19 (2017) 68-76. doi: 10.1109/MCSE.2017.3421541.

[13] H. Jung, Y. Jung, M. Fulham, J. Kim, Mixed reality hologram slicer (mxdR-HS): a marker-less tangible user interface for interactive holographic volume visualization, arXiv – Human-Computer Interaction 2022. doi: arxiv-2201.10704

[14] T. Huynh-The, Q.-V. Pham, X.-Q. Pham, T. T. Nguyen, Z. Han, D.-S. Kim, Artificial Intelligence for the Metaverse: A Survey, 2022. URL: https://arxiv.org/pdf/2202.10336.pdf.

[15] J. Opara-Martins, R. Sahandi, F. Tian, Critical review of vendor lock-in and its impact on adoption of cloud computing, in: Proceedings of the International Conference on Information Society, IEEE, New York, 2014, pp. 92-97, doi: 10.1109/i-Society.2014.7009018.

[16] G. Bhullar, S. Osborne, M. J. Núñez Ariño, J. Del Agua Navarro, F. Gigante Valencia, Vision System Experimentation in Furniture Industrial Environment, Future Internet 13 (2021) 189. doi: 10.3390/fi13080189.

[17] European Commission, Software for ambient semantic interoperable services, 2009. URL: https://cordis.europa.eu/project/id/034980.