Digital continuity of heterogeneous data using a digital twin for infrastructure's asset management

Moussa Issa, Guillaume Ducellier, Bruno Landes, Jean-Christophe Michellin, Sébastien Remy

Digital continuity of heterogeneous data using a digital twin for infrastructure's asset management.

Moussa ISSA[1], Guillaume DUCELLIER[2], Bruno LANDES[3], Jean-Christophe MICHELLIN[4], Sébastien REMY[5]

[1,3,4]SNCF Réseau, Paris, France
[1,2,5]UTT, Troyes, France
Corresponding Author: Moussa ISSA (moussa.issa@reseau.sncf.fr)

Abstract

Digital data continuity is a real issue for companies that are looking to make a successful digital transformation. At SNCF Reseau we have the same concerns which relate to factors which hinder progress towards digitization. In addition, the fact that in today's world, with the emergence of new technologies, we have a huge amount of data that comes from different sources and in different formats. We should better use these data to satisfy the requirements of end users and better manage our assets. In the context of a railway network, the design, construction, operation, and maintenance phases are often Ensuring a continuous flow of information between the four phases is often difficult to achieve. In this article, we propose a new methodology for managing digital data continuity using the new concept of a digital twin. This new approach will offer business-oriented services. Once set up, the asset manager will be equipped with a decision support tool, and it will also allow them to better plan interventions on construction sites.

After working with different specialists in the rail system, with different end users, we will outline how a data model can be constructed and then easily implemented within the SNCF Computer System. Then this data model will consider the heterogeneity of the data sources to update the data bases.

Keywords: Digital twin, Asset management, Heterogeneous data, Digital information's continuity

1. Introduction

The industrial world is currently experiencing a real transformation to the digital era. Digital transformation is a continuous and complex undertaking that can substantially shape a company and its operations (Christian Matt, Thomas Hess, et Alexander Benlian s. d.). The concept of the digital twin was first described in the context of manufacturing processes (Grieves 2015) which aimed to create a digital representation of a physical product.(Tao et al. 2018) define the digital twin as an integrated multi-physics, multi-scale, probabilistic simulation of a complex product and uses the best available physical models, sensor updates, etc. to reflect the life of its corresponding twin. We have seen the use of digital twin models for asset management in the aerospace sector (Cai, Zhang, et Zhu 2019), healthcare care and manufacturing domain (Zhu, Liu, et Xu 2019) and other fields (Kritzinger et al. 2018). The digital twin approach is new for the management of a linear infrastructure which is geographically widely distributed. Representing an infrastructure as-built digitally, will allow stakeholders of every part of the infrastructure to access to an update database in real-time. The digital continuity of data between business lines is essential to exploit the full potential of digital twins and enhance the performance of the railway system (Al-Douri, Tretten, et Karim 2016). One well known example of using a digital twin is in the manufacturing sector, with a concrete example of a test bench for the bending of beams (Haag et Anderl 2018). French railway company, SNCF Reseau, is one of the first railway companies in the world who initiated the transformation to the digital era. Historically, the SNCF worked with solely with paper documents and old processes to manage their assets. In this paper, we intend to move into the digital era and manage all our assets by using a virtual representation of our infrastructure. In the section 2, we will explain the methodology which has been used, then in section 3, we will give more details about the use of digital twin models for asset management. We will focus on the concept of a digital twin and in particular the continuity of digital data during the life cycle of the physical assets of our railway system. The infrastructure is more than 30.000 km of line with assets of different ages and complexity to manage (crossings and switch, signal, catenary …etc). Initially, we made the choice to work with the track and catenary assets but intend to explore their use to all other assets. Before concluding, we will present the first results obtained from this approach.

2. Methodology

2.1 Definition of a railway system digital twin

The digital twin of a railway system for asset management purposes integrates the digital representation of the physical infrastructure with three service levels. The first level offers a descriptive view of the railway system network, its functionalities, its state, its behaviour, and the history of interventions and/or works? carried out on the physical system. The Analytics layer aims to complete level one through the integration of simulation services and data analytics algorithms. The third level consists of projecting the impacts of a functional change or making decision (Figure 1 : Digital twin service Levels).

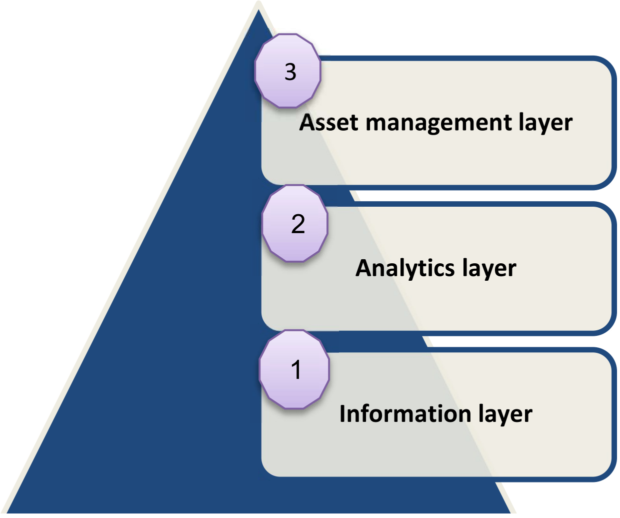
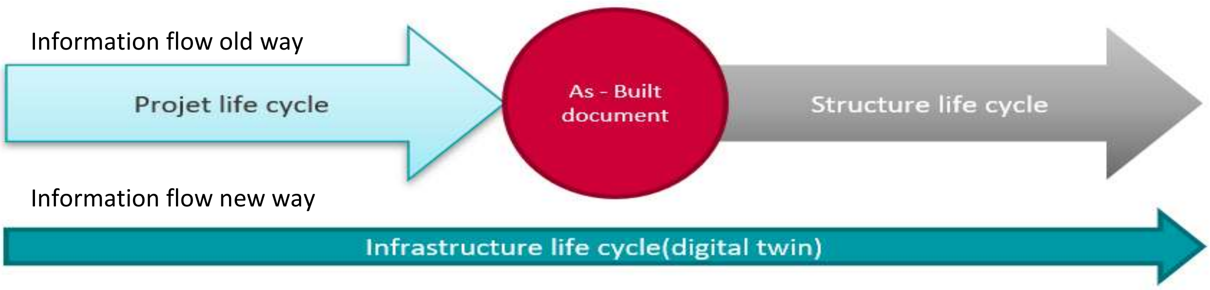


Figure 1 : Digital twin service Levels

The methodology aims to fulfil each layer in turn. The first step was to get all the information needed by the infrastructure manager to control their assets. To accomplish this task, many interviews were made with all the relevant actors and specialists. The aim of these interviews was to understand how the operational processes of our maintainers and asset managers. This information is then used to make a digital data skeleton of the future digital twin. This skeleton is initially empty has a representation of all data needed by maintainers and assets managers.

As mentioned previously, we have initially chosen to work on the track and catenary data. These data are gathered after the work on a site after analysis in the high output machines (Track and catenary renewal). After working with the specialists, we have identified what data are requested and used by them. The transition from the phase of a track and catenary regeneration project to the phase of infrastructure management by the project owner is undertaken through an as-built document, which is essentially a set of paper files. These files are in most cases incomplete or even unusable (see Figure 2 : Information discontinuity source). therefore, this generates a significant loss of information on the completed works. Based on these documents of completed works, a list of the data was made (see Figure 3 : Data collected from As-buit document).

Figure 2 : Information discontinuity source



Domain	Object	Data	Output format
Catenary	Contact wire	Contact wire position	Enum(X/Y/Z)
Catenary	Tensioning equipment	Types	Enum(25kV (with contact wire), 25kV (with contact wire and catenary wire), 1500V)
Catenary	Tensioning equipment	Localization	Enum(Line id/Track number/Kp)
Catenary	Tensioning equipment	Id(Devise number)	String
Catenary	Tensioning equipment	anti-tracking	
Track	switch and crossing	switch and crossing id	string
Track	switch and crossing	Commitionning date of switch and crossing	Date
Track	switch and crossing	End of activity date of switch and crossing	date
Track	switch and crossing	Installation date of switch and crossing	date
Track	switch and crossing	CWR(Continuous Welded rail)/Welded/Jointed switch and crossing	Enum(CWE, welded,jointed)
Track	switch and crossing	Type of switch and crossing	String
Track	switch and crossing	Tangent of type of switch and crossing	Real
Track	switch and crossing	Type of tangent (Example: C, L or NaN (short, long))	Enum

Figure 3 : Data collected from As-buit document

2.2 Digital data continuity

Digital continuity of information is today one of the biggest issues in various industries. In this era of digital transformation, companies like SNCF Reseau are looking for solutions to get information about the as-built asset in a completely digital way. This company used to work mainly with paper documents and some practices that are not standardized. Also, the structure of the company heavily compartmentalized and sharing documents and data are still a real issue. All these challenges make the digital continuity of data a difficult goal to achieve.

To attain level two of the digital twin's service, our digital twin must be based on data that are up to date and representative of the physical facilities. It is the importance of having digital data continuity. (Lee et al. 2021develop and test an integrated digital twin and blockchain framework that can selectively store, trace and share important project-related data. The framework proposed in previous articles shows how project data are collected on site and updates the As Built information in the digital twin. In our case at SNCF Reseau, we will know the data needed and how to collect them. We therefore focused on how to set up a workflow of data lifecycle and allow our asset managers to share the updated data.

We initially started with the empty skeleton of a digital twin, then a functional representation of physical objects was made. This functional representation was undertaken with the collaborations of all track and catenary specialists to ensure no data was forgotten.

Based on this functional representation, a database dedicated to the different sources of information was structured in the cloud. The asset managers will have access to the virtual representation of his physical installations without necessarily going to the site or consulting a multitude of databases.

This database must contain all the information on the objects that were represented in the data model, before, during and after the works, quite simply during the life cycle of the installations. This database is updated thanks to the different data acquisitions, which is explained in the next section: Data collection.

2.3 Data collection

The digital twin is updated in real time thanks to the integration of information from different data sources. The sources of data are multiple: data from specific High Output renewal trains during renewal works, data from internal asset description databases, data from specific LiDAR and photogrammetry surveys, IoT, such as illustrated on the picture below (Figure 4 : different sources of data). The data model is implemented in the digital twin's database and end users are simply required to login to the database to access the information they need.

The objective of the data collection phase was to determine the right amount and quality of data to provide the most convenient response to operational needs and avoid including unnecessary data to the demonstration use case.

At different times to initiate and develop the "digital twin"

- Before the renewal phase.

- During the renewal phase to update the initial "augmented ABR (As-buit Records)" and generate the initial release of the digital twin.
- At the end of the renewal phase, to have a final release at the end of this phase.
- At a regular time interval, to update the digital twin.

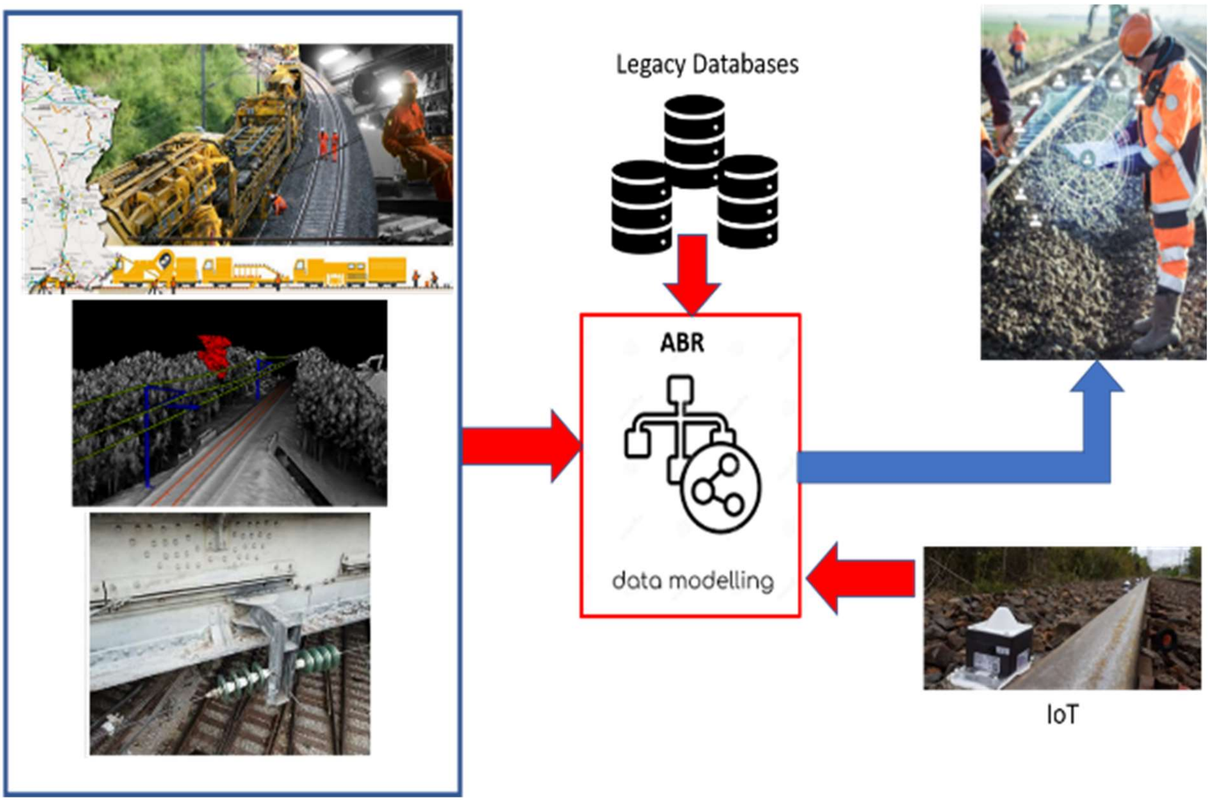


Figure 4 : different sources of data

3. Asset management

Asset Management covers all the activities of an organization which contribute to optimize the risks, the costs, and the performances of physical assets, regarding its Strategic and Business objectives. Rail asset managers have a variety of requirements, depending on their domain (catenary or track). For our test use case "the establishment of a digital twin for track and catenary sections (The High Output Machines replaces existing rail and catenary components with high-performing new ones), we consulted the representative key people who are potential future users of the product to identify their needs and work processes.

The keys actors that we met are those who best understand their assets. We have identified with them what information they needed for their works and know the description and state of their work sectors.

We observed that the information's that they have defined as necessary are paper document files which contain a lot of data on the assets (for example catenary arming diagram, the site description for the track .... etc) or local Excel files.

Means of acquisition	Methodo-logy	Precision accuracy
LiDAR on tripod	"Static" surveying	< 1 cm
Trolley LiDAR	Dynamic surveying	1 cm
Train LiDAR	Dynamic surveying	1.5 cm
UAV LiDAR	Dynamic surveying	2 cm to 3 cm
Photogrammetric UAV	Dynamic survey	> 10 cm

Table 1: Precision of acquisition tools

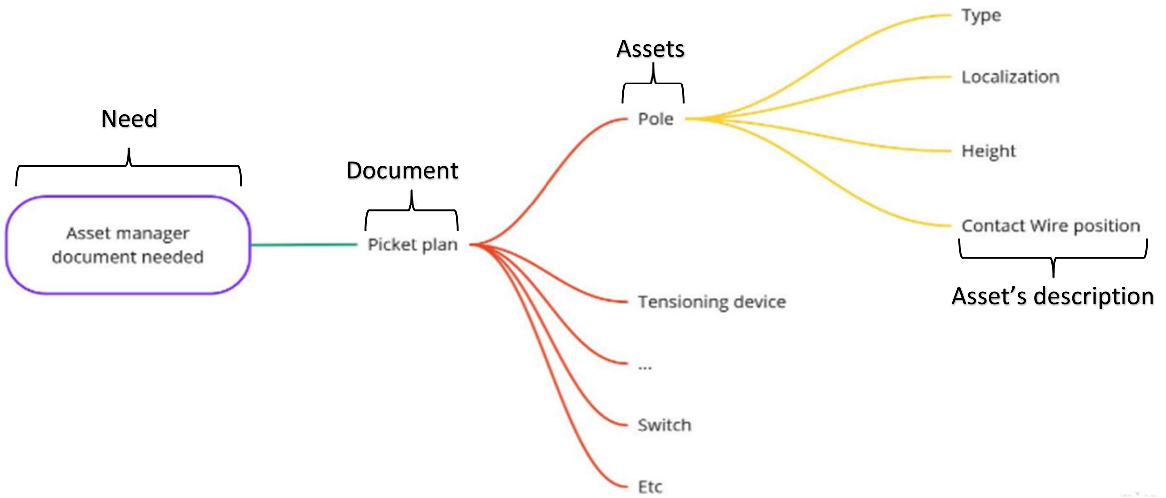


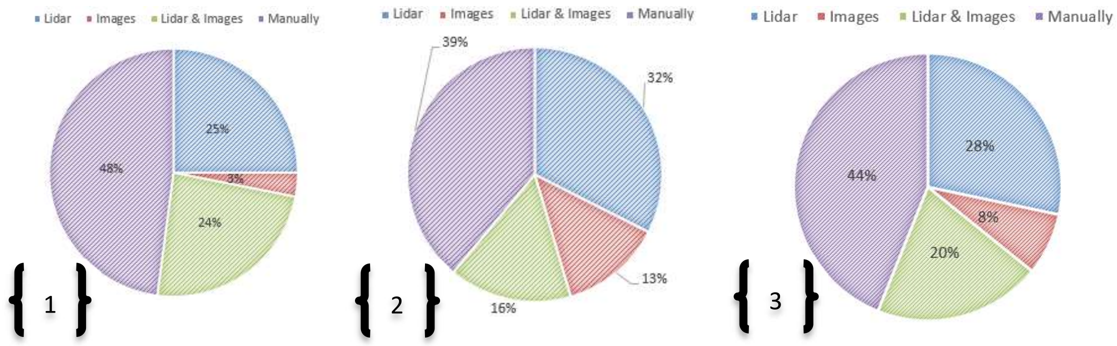
Figure 5 : Object definition methodology

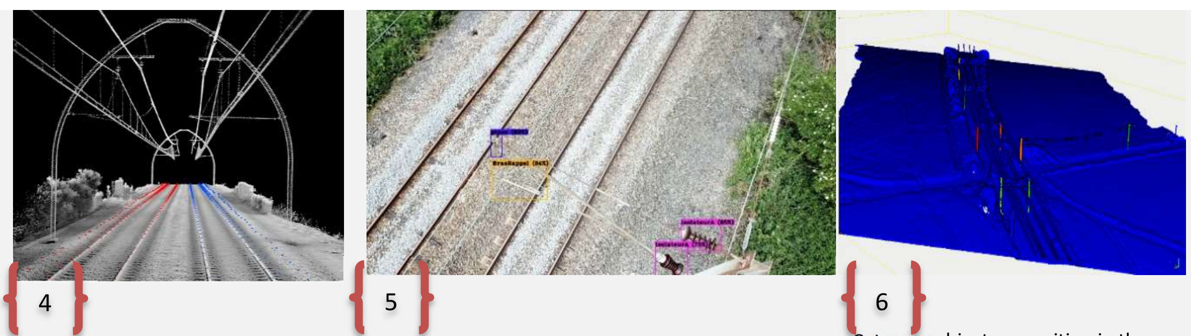
The asset manager, based on the digital twin updated using the various heterogeneous data acquisitions, may have several possible services depending on its use case (see figure 6).

4. Results

To obtain a relevant Digital Twin for asset management, we need to build it around exhaustive and up to date information (from maintenance experts). The digital twin structure enables evolutions and accurate technical requirements which should be easily to maintain and update the DT. The first results based on interviews with experts.

1. The amount of detectable data in the Point clouds, images or manually for the catenary section.
2. The amount of detectable data in the Point clouds, images or manually for the track section.
3. The amount of detectable data in the Point clouds, images or manually for Track & Catenary.





Example of data automatically extracted from Track lidar data

Catenary object extraction based on drone imagery

Catenary object recognition in the Lidar data

5. Conclusion

In this article, the methodology that has been put in place to ensure the digital continuity of heterogeneous data of form and format has started to bear fruit. We have succeeded in modelling the data and setting up a database rich in information. This database is updated using different techniques, such as object recognition in point and image clouds, in addition to operators in the field who will have access to it according to their different use cases. The real added value of this method is to ensure a better representation of the rail system and to avoid making applications by use case which end up cluttering the information system. Due to the misinterpretation or insufficiently detailed specifications, this approach becomes a necessary step in the functional representation of physical assets. This digital twin of the track and catenary sections will improve their operations and maintenance. thus, it will attract other use cases, for example from the communication sector, civil engineering, engineering structures and so on.
A user interface and access rights will be made available to the various actors in the track and catenary sections depending on the data they wish to handle. In the end, this heterogeneity of data will not block the digital continuity of information if it is properly considered at the start of the process of updating digital data.

Acknowledgment

This research is made with collaboration with other EUROPEEN infrastructure stakeholders project called In2smart-2. This project represents the continuation of the work initially carried out through "the In2Rail lighthouse project" and mainly through the work of the IN2SMART project. These projects together form part of the framework of research and innovation that will deliver the vision and strategy of Innovation Program shif2rail.

References

Al-Douri, Y.K., P. Tretten, et R. Karim. 2016. « Improvement of railway performance: a study of Swedish railway infrastructure ». *Journal of Modern Transportation* 24(1): 22-37.

Cai, Hongxia, Wei Zhang, et Zheng Zhu. 2019. « Quality Management and Analysis of Aircraft Final Assembly Based on Digital Twin ». In *2019 11th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, Hangzhou, China: IEEE, 202-5. https://ieeexplore.ieee.org/document/8941382/ (29 avril 2020).

Christian Matt, Thomas Hess, et Alexander Benlian. « Digital Transformation Strategies ».

Grieves, Michael. 2015. « Digital Twin: Manufacturing Excellence through Virtual Factory Replication ».

Haag, Sebastian, et Reiner Anderl. 2018. « Digital Twin – Proof of Concept ». *Manufacturing Letters* 15: 64-66.

Kritzinger, Werner et al. 2018. « Digital Twin in Manufacturing: A Categorical Literature Review and Classification ». *IFAC-PapersOnLine* 51(11): 1016-22.

Lee, Dongmin et al. 2021. « Integrated Digital Twin and Blockchain Framework to Support Accountable Information Sharing in Construction Projects ». *Automation in Construction* 127: 103688.

Tao, Fei, He Zhang, Ang Liu, et Andrew YC Nee. 2018. « Digital twin in industry: State-of-the-art ». *IEEE Transactions on Industrial Informatics* 15(4): 2405-15.

Zhu, Z., C. Liu, et X. Xu. 2019. « Visualisation of the digital twin data in manufacturing by using augmented reality ». In , 898-903.