



HAL
open science

Aide à la conception de l'architecture d'un jumeau numérique par la modélisation des données

Doha Nefti, Mohammed Adel Hamzaoui, Nathalie Julien, Fayçal Balkaid

► To cite this version:

Doha Nefti, Mohammed Adel Hamzaoui, Nathalie Julien, Fayçal Balkaid. Aide à la conception de l'architecture d'un jumeau numérique par la modélisation des données. CIGI QUALITA MOSIM 2023 Propulser la performance, Jun 2023, Trois-Rivières, Canada. hal-04121075

HAL Id: hal-04121075

<https://hal.science/hal-04121075>

Submitted on 7 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CIGI QUALITA MOSIM 2023

Aide à la conception de l'architecture d'un jumeau numérique par la modélisation des données

DOHA NEFTI¹, MOHAMMED ADEL HAMZAOU², NATHALIE JULIEN³, FAYÇAL BELKAÏD⁴,

¹ Université Abou Bekr Belkaïd
Tlemcen 13000, Tlemcen
nefti.doha@gmail.com

³ Lab-STICC, Université Bretagne Sud
Lorient 56100, France
nathalie.julien@univ-ubs.fr

² Lab-STICC, Université Bretagne Sud
Lorient 56100, France
mohammed.adel@univ-ubs.fr

⁴ MELT, Université Abou Bekr Belkaïd
Tlemcen 13000, Tlemcen
f_belkaid@yahoo.fr

Résumé –

La gestion des données dans le contexte du jumeau numérique est un enjeu majeur en raison du manque de formalisation, de modèles et de conceptualisation du cycle de vie des données, malgré l'intérêt croissant suscité par cette technologie émergente dans divers domaines. Pour répondre à ce défi, l'étude présentée propose un modèle de cycle de vie des données ainsi qu'une typologie des données pour caractériser exhaustivement les différentes étapes de la gestion des données. La formalisation de la gestion des données ainsi obtenue permettra d'assurer la qualité, la cohérence et la fiabilité des données utilisées dans les différentes applications du jumeau numérique, et représente une contribution significative pour son développement et son déploiement à grande échelle. Ceci permet également de dimensionner l'architecture du jumeau numérique du point de vue des ressources (stockage, puissance de calcul, fréquences de transmission, etc.), d'un point de vue financier (cout d'investissement et de fonctionnement), ainsi que celui de la consommation énergétique.

Abstract –

Data management in the context of digital twin technology is a major challenge due to the lack of formalization, models, and conceptualization of the data lifecycle, despite the growing interest in this emerging technology in various fields. To address this challenge, the presented study proposes a data lifecycle model and a data typology to comprehensively characterize the different stages of data management. The formalization of data management thus obtained will ensure the quality, consistency, and reliability of data used in various digital twin applications and represents a significant contribution to its development and large-scale deployment. This also makes it possible to dimension the architecture of the digital twin from the point of view of resources (storage, computing power, transmission frequencies, etc.), from a financial point of view (investment and operating costs), as well as that of energy consumption.

Mots clés – Jumeau Numérique, Modèles de données, Modèles et architectures, Méthodologie de déploiement.

Keywords – Digital Twin, Data models, models and architectures, Deployment methodology.

1 INTRODUCTION

Aujourd'hui, le monde connaît une avancée considérable dans le domaine de la technologie numérique, notamment avec le développement de l'intelligence artificielle, l'internet des objets, la réalité virtuelle, mais aussi l'émergence du jumeau numérique (JN). Ce dernier a connu un véritable boom depuis 2017, cela s'illustrant notamment dans la multitude d'applications dans différents domaines tels que la santé, l'industrie, le transport, et la logistique, pour ne citer que ceux-là.

Ceci a inévitablement suscité un intérêt accru, d'une part, des industriels dont les finalités sont beaucoup plus opérationnelles et dépendantes des applications finales, et d'autre part, des chercheurs dont les objectifs s'orientent plus vers une quête de conceptualisation et formalisation de cette technologie. Néanmoins, la « hype » autour de cette thématique a un peu

dilué et dispersé les efforts dans ce sens, puisque nous pouvons compter pas moins d'une cinquantaine de définitions différentes dans la littérature [Julien & Martin, 2021b].

Nous observons donc des développements et déploiements de technologies jumeaux numériques dans différents domaines, faisant appel souvent à des technologies différentes et hétérogènes, cela sans recours à des standards ou des modèles communs. Pour remédier à cela, il est nécessaire de mettre au point des formalismes rigoureux pour définir le concept à travers des normes, des modèles et des architectures communes. En plus de méthodologies génériques de développement et de déploiement du jumeau numérique, permettant ainsi une grande capacité d'adaptation de cette technologie à différentes problématiques issues de domaines variés, tout en gardant toute la robustesse et le rigorisme nécessaire pour ce type de technologie dont le potentiel de rupture est clairement avéré.

Par ailleurs, ce fort potentiel promet à cette technologie une présence de plus en plus forte dans les prochaines années dans l'ingénierie, l'industrie et les services, ainsi que dans la recherche (le nombre de publications traitant du jumeau numérique triple chaque année depuis 2016 [Tao et al., 2019]). Des travaux s'intéressent même aux futurs réseaux de jumeaux qui pourraient voir le jour prochainement [Hamzaoui & Julien, 2022], ainsi qu'à des modes de structuration de ces derniers, ainsi que leur interfaçage avec les différents environnements qui les entourent. De ce fait, il convient de souligner que la gestion des données est un élément clé dans ce contexte, notamment en ce qui concerne la génération, l'utilisation, le traitement et le cycle de vie de ces données. Ici également, le manque de formalisation, de modèles et de conceptualisation du cycle de vie des données dans le contexte du jumeau numérique représente un défi majeur pour le développement et le déploiement de cette technologie à grande échelle, ainsi que pour son interopérabilité et sa scalabilité. Il est donc essentiel de développer des normes et des modèles communs pour la gestion des données dans le jumeau numérique, afin d'assurer la qualité, la cohérence et la fiabilité des données utilisées dans les différentes applications du jumeau numérique.

Dans l'optique d'identifier les verrous scientifiques relatifs au cycle de vie de la donnée dans le contexte du jumeau numérique, le présent article ambitionne de proposer des solutions concrètes. Ainsi, une nouvelle modélisation du cycle de vie de la donnée est présentée, mettant l'accent sur une étape clé de la méthodologie de déploiement du jumeau numérique proposée par [Julien & Martin, 2021b], décomposée en cinq étapes principales qui feront l'objet d'une présentation concise dans la section suivante. Par ailleurs, l'article s'inscrit dans la continuité du développement de la méthodologie (5C), dont la pertinence a été justifiée par le travail préalable réalisé. Pour assurer l'applicabilité de la méthode au-delà du jumeau numérique, une analyse des différents modèles de cycle de vie des données dans la littérature a été entreprise. Ainsi, un modèle concis et facile à suivre a été élaboré, fédérant les modèles existants et incluant toutes les étapes clés. L'objectif consiste à faciliter le développement et le déploiement d'applications et de produits numériques, tout en mettant l'accent sur le jumeau numérique et en garantissant l'applicabilité de la méthode à toute entité numérique.

Le présent article poursuit son exposé en procédant à une analyse critique de la littérature existante portant sur la gestion des données dans le contexte des jumeaux numériques. Suivie d'une présentation de la méthodologie générique pour le déploiement des jumeaux numériques susmentionnée (5C). Par la suite il sera question de présenter la contribution principale de ce papier, le modèle du cycle de vie des données dans le jumeau numérique (5A), ainsi qu'une proposition de typologie de la donnée adaptée au jumeau numérique. Enfin, une synthèse des résultats obtenus et des recommandations pour les recherches futures dans ce domaine concluent l'article.

2 REVUE DE LA LITTÉRATURE

La nécessité de disposer de formalismes rigoureux et robustes permettant le développement et le déploiement de la technologie jumeau numérique à grande échelle n'a pas échappé à l'attention de la communauté scientifique. Des efforts ont été faits dans ce sens, apportant à chaque fois des contributions visant à adresser des verrous scientifiques et techniques différents. À travers cette section, nous présentons

les principaux travaux qui participent à cette finalité tout en étant complémentaires les uns avec les autres, et sur lesquels nous nous sommes basés pour alimenter notre réflexion.

L'un des modèles les plus robustes formalisant le jumeau numérique est celui proposé par [Tao et al., 2019], et communément appelé le modèle 5D. Ceci, en référence au fait de décrire le jumeau numérique en le décomposant en 5 dimensions différentes, ajoutant donc deux éléments supplémentaires aux trois proposés par [Grieves, 2014] (élément physique, élément virtuel, données et informations). Dans le modèle 5D, on retrouve donc les éléments « Services » et « Connexions » comme dimensions supplémentaires, comme illustré dans la Figure 1.

Le modèle 5D est donc composé de ces cinq dimensions Eq. (1), et chacune d'elles est décrite en détails dans (Tao et al., 2019). Néanmoins, nous proposons une brève description pour une meilleure compréhension du modèle.

$$M_{DT} = (PE, VE, S_s, DD, CN) \quad (1)$$

PE : une entité physique qui existe dans le monde réel.

VE : une entité virtuelle constituée d'une collection de modèles.

SS : un service fourni à la fois par PE et VE. DD : les données qui constituent le JN.

CN : les liaisons entre les différentes parties du JN.

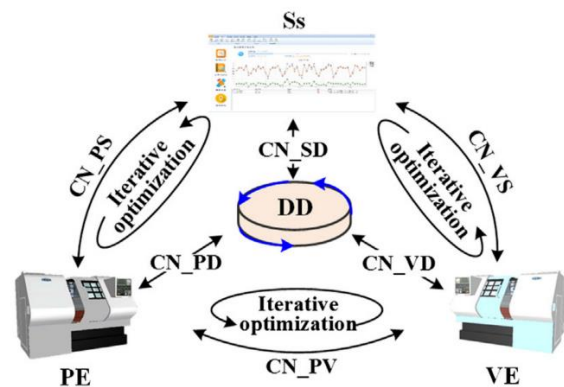


Figure 1. Le modèle 5D du jumeau numérique [Tao et al., 2019]

Dans [Tao et al., 2019], le jumeau numérique est comparé au corps humain, de la sorte, le PE agit comme un "squelette" qui prend en charge le JN. En tant que "cœur" du système, le VE agit comme une "pompe" qui alimente les résultats et les techniques de simulation vers d'autres composants. La dimension SS représente quant à elle un "organe sensoriel" qui parle directement à l'utilisateur. L'élément DD est le "sang" qui alimente en permanence le JN en informations vitales. En conséquence, les connexions CN agissent comme un "conduit sanguin" transmettant des informations.

Parmi les efforts fournis dans la formalisation et la standardisation du concept de jumeau numérique, nous pouvons citer ceux de l'Organisation Internationale de Normalisation. Cette dernière a récemment mis au point la norme ISO 23247 (International Organization for Standardization, 2021), qui fournit un cadre permettant le développement et l'utilisation efficace des jumeaux numériques dans un environnement industriel. En plus de fournir une nomenclature et une terminologie précise, elle propose une architecture cohérente et structurée. En plus de fournir des lignes directrices et des bonnes pratiques pour la

mise en œuvre du JN, et de favoriser la réutilisation des données communes et des composants du modèle.

La norme se compose de quatre parties qui fournissent des lignes directrices pour la création de jumeaux numériques pour des cas d'utilisation spécifiques. Cela inclut l'analyse des besoins en matière de modélisation, la définition de la portée et de l'objectif, ainsi que la promotion de l'utilisation d'une terminologie commune.

Dans un registre plus bas niveau, en d'autres termes, qui s'intéresse plus aux technologies mises en jeu pour le développement du JN qu'aux définitions et formalismes ou architectures et modèles. Nous trouvons le tableau périodique des capacités du jumeau numérique [Pieter, 2022], proposé par le Digital Twin Consortium. Cet outil vise à aider les organisations à concevoir, développer, déployer et exploiter des jumeaux numériques en fonction des exigences de capacités de cas d'utilisation spécifiques, plutôt que de s'appuyer sur les fonctionnalités des solutions technologiques.

Le CPT (Capabilities Periodic Table) utilise une architecture et un cadre de définition des exigences indépendant de la technologie pour définir les exigences des cas d'utilisation individuels, qui peuvent ensuite être agrégées pour déterminer les exigences globales en matière de capacité, d'architecture du jumeau numérique, ainsi que des solutions technologiques nécessaires pour répondre aux besoins spécifiques de l'entreprise. Les travaux mentionnés ci-dessus traitent des définitions des concepts et formalismes autour du jumeau numérique suivant des approches et angles de perception différents. Le modèle 5D définit les composants élémentaires du JN ainsi que leurs interconnexions, la norme ISO 23247 offre un cadre de base pour la construction de JN dans un environnement industriel, alors que le CPT vient apporter une dimension plus opérationnelle avec une approche par fonctionnalité et capacités du JN. Cependant, bien que nécessaires, voire cruciaux, ces formalismes ne sont pas suffisants s'ils ne sont pas intégrés dans une méthodologie générique de développement et de déploiement du jumeau numérique.

Comme évoqué plus haut, la mise en place d'une méthodologie générique et technologiquement indépendante afin de développer et déployer efficacement le JN est une nécessité. Dans ce sens, [Julien & Martin, 2021b] ont proposé une méthodologie se déclinant en cinq étapes (Figure 2) basée sur l'architecture CPS 5C établie par [Lee et al., 2015]. Cette dernière se présentant selon une approche top-down donne une grande importance à l'identification des besoins dictés par les usages pendant les premières étapes du développement du JN [Julien & Martin, 2021a].

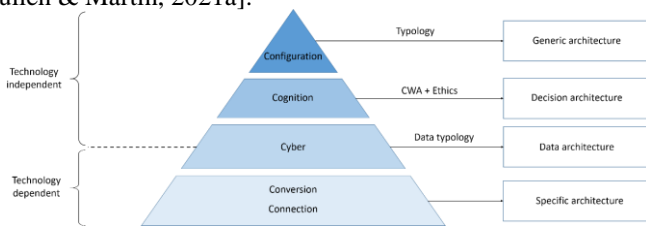


Figure 2. Méthodologie en 5C pour le déploiement du jumeau numérique [Julien & Martin, 2021b]

Cette méthodologie se présente donc en cinq étapes distinctes, les deux premières d'entre elles étant indépendantes de la

technologie, alors que les deux dernières en sont dépendantes. La troisième étape étant particulière, car servant d'une étape charnière, est justement l'étape qui a monopolisé notre intérêt dans ce papier. Nous décrivons néanmoins, de manière brève, en quoi consiste chacune des étapes :

Configuration : Afin d'aider les concepteurs du JN à développer une architecture initiale générique, des propriétés clés doivent être définies en suivant concisément une « Typologie » constituée en 11 critères [Julien & Martin, 2021b] [Blanchet et al., 2022]. Ici, il s'agira d'esquisser le profil général du JN à mettre en place.

Cognition : L'interaction entre le jumeau numérique et l'agent humain est primordiale, principalement en ce qui concerne l'affectation des tâches de prise de décision. À ce niveau-là, les auteurs avaient proposé la méthode « Cognitive Work Analysis » (CWA). Un travail préliminaire appliqué à la maintenance prédictive a été présenté par [Kamoise et al., 2022].

Cyber : La phase "cyber" représente une étape clé dans la conception et l'implémentation du JN. Elle consiste à prendre des décisions technologiques initiales, telles que la décentralisation, la cybersécurité et la réactivité du système. Dans cet article, cette étape est analysée du point de vue de la donnée, de la typologie de la donnée vivant dans le JN et de son cycle de vie.

Conversion et connexion : Arrivant à ces niveaux-là, l'architecture précédemment définie sera progressivement raffinée en spécifiant les interconnexions et les protocoles de communication, ainsi que la sélection des capteurs et des méthodes de collecte de données en fonction de l'environnement de l'application. L'objectif est de produire des métriques pour sélectionner l'architecture la mieux adaptée en termes de coût, de performance, d'agilité et de sécurité.

Il convient de souligner que l'étape « Cyber » représente une étape charnière entre la partie indépendante de la technologie et la partie dépendante de la technologie dans le développement du jumeau numérique. Pour saisir correctement le cycle de vie de la donnée dans le jumeau, une architecture de données bien définie est nécessaire à cette étape. Nous allons passer en revue les différents modèles de données disponibles et souligner leurs points forts et leurs faiblesses.

[Zhang et al., 2021] ont posé les fondements théoriques de la DTD (Digital Twin Data) qui sont indispensables à la promotion et à l'application du JN. Ils soulignent l'importance de la DTD, qui est la base du JN et qui permet la création et la maintenance du jumeau numérique d'un objet physique ou d'un système en temps réel. En fournissant des orientations pour la collecte, le stockage, l'analyse et la sécurité des données du jumeau numérique, ces orientations sont essentielles à la promotion et à l'application du JN. Ainsi, la mise en place de la DTD peut aider les professionnels à concevoir et à mettre en œuvre des jumeaux numériques plus fiables et plus efficaces. Néanmoins, l'article fournit une vue d'ensemble complète des méthodes et des technologies clés, mais il manque d'approfondissement dans l'analyse de la structuration des données pour les jumeaux numériques.

Par ailleurs, [Gallala et al., 2022] se penchent sur l'utilisation des technologies de l'Industrie 4.0 afin d'améliorer les interactions entre humains et robots au moyen de jumeaux

numériques. Une analyse complète des approches de structuration des données est présentée pour faciliter le déploiement de ces jumeaux numériques. Cependant, le cycle de vie des données proposé, bien que linéaire, ne peut représenter adéquatement la complexité et la dynamique de la gestion de données dans les projets de type jumeau numérique. En effet, cette gestion peut souvent comporter des itérations et des boucles plutôt qu'une progression linéaire, rendant ainsi la tâche plus difficile. Il est donc crucial de mettre en place une gestion de données flexible, accompagnée d'outils et de méthodologies adaptés, pour assurer une gestion efficace des jumeaux numériques au niveau des données.

3 LE MODÈLE 5A, UN MODÈLE POUR LE CYCLE DE VIE DE LA DONNÉE DANS LE JUMENTAUX NUMÉRIQUE

La phase « cyber » comprend plusieurs aspects clés du développement du jumeau numérique, notamment la gestion des données, l'organisation architecturale et les décisions technologiques. La gestion des données est essentielle pour répliquer avec précision les systèmes physiques, de plus, la modélisation des données permet d'organiser et de structurer les informations. L'organisation de l'architecture est également importante, car elle influe sur le stockage et le traitement des données, ainsi que sur la construction du système, ce qui peut avoir un impact significatif sur la qualité, la précision, la performance et la consommation d'énergie du jumeau numérique [Julien & Hamzaoui, 2022].

Enfin, les décisions technologiques prises durant cette phase peuvent influencer l'efficacité, les coûts et l'impact sociétal du jumeau numérique, d'où la nécessité de les considérer attentivement dans sa conception. En somme, la phase "cyber" dans la création d'un jumeau numérique implique une planification minutieuse et une exécution rigoureuse de la gestion des données, de l'organisation de l'architecture et des choix technologiques pour créer un jumeau numérique précis, efficace, économe en énergie et prenant en compte les implications sociétales des technologies employées.

Dans cette étude, notre objectif a été d'harmoniser le modèle visuel proposé par [Zhang et al., 2021], basé sur une représentation en blocs du flux de données, avec le tableau périodique du Digital Twin Consortium [Pieter, 2022]. Pour y parvenir, nous avons élaboré un modèle composé de cinq étapes (ou états) par lesquelles passe une donnée durant sa vie dans le jumeau numérique. Ceci permet de représenter de manière globale le cycle de vie de la donnée, en prenant en compte toutes les étapes, de l'acquisition à la destruction des données (Figure 3). L'objectif était de mettre en place un modèle qui soit, d'une part, assez précis et discriminant dans l'identification et l'énumération des étapes clés, et d'autre part, ne pas se noyer dans la « sur-minutie » et la perte de généralité. Nous sommes donc arrivés à identifier les cinq étapes clés du cycle de vie de la donnée dans le jumeau numérique de la collecte à l'application, toutes commençant par la lettre « A », d'où le nom de modèle 5A, celles-ci étant :

Acquisition : cette étape consiste à collecter les données à partir de différentes sources en se posant des questions sur leur nature, leur origine et leurs propriétés (interne ou externe, humaine ou non). Les données sont donc recueillies et stockées pour une utilisation ultérieure.

Agrégation : une fois que les données ont été collectées, elles peuvent être agrégées avec d'autres données pour former des modèles dynamiques afin d'alimenter la phase suivante.

Analyse : l'étape d'analyse permet de tirer des informations des données agrégées en modèles en utilisant des méthodes telles que l'intelligence artificielle et les méthodes d'apprentissage basées sur les données, pour comprendre les tendances, les comportements et les modèles dans les données.

Application : une fois que les données ont été analysées, elles peuvent être appliquées à diverses fins, en fonction des besoins des utilisateurs, qu'il s'agisse de prédire des tendances, de prendre des décisions éclairées ou de développer de nouveaux produits et services.

Access : cette étape concerne les aspects liés à la gestion de l'accès aux données, la sécurité et la confidentialité. Il s'agit de stocker, archiver, effacer ou détruire les données, ainsi que de gérer les aspects juridiques et de cybersécurité associés à la gestion des données.

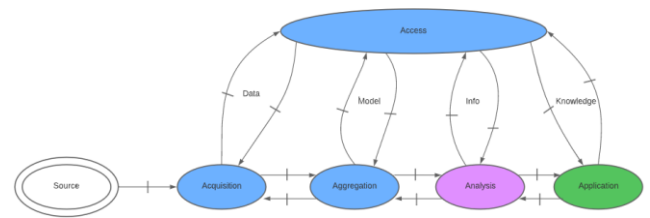


Figure 3. Modèle 5A du cycle de vie de la donnée dans le jumeau numérique

Nous pouvons remarquer que qu'il existe implicitement une notion d'unicité d'unicité de la donnée dans ce cycle de vie, et qui est un trait commun à la très grande majorité des cycles de vie de la donnée trouvés dans la littérature.

Toutefois, il convient de noter que l'étape « Access », en tant que dimension transversale par rapport aux autres étapes, peut conduire à une duplication de la donnée. En effet, au niveau de « Access », il existe des processus de stockage, d'archivage et de suppression qui peuvent affecter la singularité des données. Cela étant dit, il est important de souligner que l'origine des données et leur parcours demeurent identiques à chaque cycle de vie, commençant toujours par la source avant de se diriger vers les autres étapes. Bien que des copies subséquentes de ces mêmes données puissent être créées et dupliquées, le cycle de vie initial de la donnée demeure inchangé.

Comme indiqué plus haut, ce modèle permet de conceptualiser et de formaliser de processus de vie de la donnée au sein du jumeau numérique. Cela, avec une approche par « états » de l'unité de donnée ainsi que son rôle dans le jumeau numérique dans sa globalité à chaque état. Cependant, il est également nécessaire d'avoir connaissance de certaines caractéristiques spécifiques des données manipulées. Il est crucial de connaître les attributs clés de ces dernières, ceci afin de -là aussi-correctement dimensionner les besoins en ressources technologiques. De ce fait, nous proposons ici également une typologie de la donnée, qui permettra elle aussi de dresser les profils des données à manipuler au sein du JN en identifiant des attributs clés.

4 TYPOLOGIE DE LA DONNÉE

Comme mentionné plus haut, établir des profils des données à manipuler dans le jumeau numérique est crucial. La typologie des données proposée permettra d'esquisser les premiers traits de ces dernières en identifiant et décrivant les critères qui permettent de caractériser une donnée :

La source : c'est l'origine de la donnée. Il est important de savoir d'où provient la donnée pour en évaluer la qualité et la pertinence.

Le type : il s'agit de la nature de la donnée, par exemple si elle est environnementale, géométrique, fonctionnelle, temporelle, etc. Cette information est importante pour la manipulation et l'analyse de la donnée.

La valeur : c'est la mesure ou l'observation que représente la donnée. On doit connaître la valeur de la donnée pour pouvoir l'utiliser correctement.

L'accès : cela concerne les modalités d'accès à la donnée, telles que l'accès libre ou sous condition, l'accès en lecture, en écriture, ou en modification, etc.

La criticité : il s'agit de l'importance de la donnée pour l'entreprise ou l'organisation. Cette information est essentielle pour la gestion de la cybersécurité et la protection de la donnée.

Le stockage : il est important de déterminer où la donnée sera stockée et comment elle sera sauvegardée pour éviter toute perte de donnée. La synchronisation : cela concerne la fréquence à laquelle la donnée sera collectée et mise à jour. Il est important de synchroniser la donnée selon les besoins pour garantir la pertinence des analyses.

La précision : cela concerne la qualité et la précision de la donnée. Il est important de connaître la précision de la donnée pour garantir des analyses précises.

La variabilité : cela concerne la variation de la donnée, elle fait référence à la mesure de la fluctuation ou du changement de la valeur d'une donnée au fil du temps. Plus précisément, la variabilité est utilisée pour déterminer la rapidité à laquelle une donnée change, ce qui peut avoir un impact sur la synchronisation avec d'autres données ou événements.

Le modèle : cela concerne le modèle ou l'algorithme qui sera utilisé pour analyser la donnée. Il est important de connaître le modèle pour interpréter correctement les résultats.

L'information : cela concerne les informations que l'on peut tirer de la donnée. Il est important de connaître les informations que l'on peut tirer de la donnée pour en évaluer la pertinence.

La connaissance : cela concerne les connaissances que l'on peut acquérir grâce à la donnée. Il est important de connaître les connaissances que l'on peut acquérir grâce à la donnée pour en évaluer la pertinence.

L'utilisateur : cela concerne les utilisateurs de la donnée. Il est important de connaître les utilisateurs de la donnée pour en évaluer l'accessibilité et la pertinence.

La communication : cela concerne les modalités de communication de la donnée, par exemple si elle sera présentée sous forme de tableau de bord ou de courbe d'évolution. Il est important de communiquer la donnée de manière adaptée aux besoins des utilisateurs.

5 CAS APPLICATIF

Le modèle 5A et la typologie présentée ci-dessus ont été appliqués aux données d'une entreprise agroalimentaire spécialisée dans la fabrication de gâteaux, afin de mettre en évidence leurs avantages. Le Tableau 1 présente un résumé de la typologie des données obtenue.

a) Le modèle 5A :

Au sein d'un environnement d'entreprise agroalimentaire, la gestion de l'architecture de données est une tâche complexe qui exige une attention particulière pour garantir une gestion efficace des données. La première étape de ce processus consiste en l'acquisition des données, qui est généralement réalisée par le biais de capteurs. Ces capteurs collectent les informations pertinentes provenant de différentes sources, qui sont ensuite stockées dans une base de données (celle du MES dans notre cas) hébergée sur le serveur sur site de l'entreprise.

La deuxième étape consiste à agréger les données collectées avec d'autres sources de données pour créer des modèles dynamiques. Cette tâche est généralement exécutée en utilisant des logiciels d'agrégation de données. Les modèles ainsi créés sont soumis à une analyse détaillée à l'aide de différentes méthodes adaptées, (issues de l'intelligence artificielle ou autre) pour détecter des tendances ou des comportements spécifiques.

Dans le contexte de notre cas d'étude (entreprise agroalimentaire), les modèles prédictifs de contaminations et les modèles temporels sont des exemples des principaux modèles utilisés pour l'analyse des données remontées de la ligne afin d'identifier des situations critiques et d'identifier l'emplacement des produits potentiellement contaminés. Cependant, la gestion de l'accès aux données doit être réalisée avec prudence pour garantir la sécurité et la confidentialité des données stockées. Cette tâche nécessite également une attention particulière pour gérer les aspects juridiques et de cybersécurité liés à la gestion des données.

b) Typologie de la donnée :

Pour la première donnée, il s'agit de la température mesurée par un capteur environnemental. Elle est de type quantitatif et son accès est libre, car elle n'est pas critique. Elle est stockée dans le MES de manière synchrone avec une précision de 0,1 °C et une faible variabilité, ce qui signifie que les valeurs mesurées sont relativement proches les unes des autres. Elle alimente le modèle de contamination pour aider à évaluer le risque de contamination et à prendre des décisions en conséquence. Les utilisateurs internes, humains et logiciels, pourront accéder à cette donnée et la visualiser sous forme de courbe.

Pour la deuxième donnée, il s'agit de l'ordre de fabrication qui provient également du MES. Cette donnée est de type temporel et son accès est conditionnel, car elle concerne des informations sensibles sur la production de gâteaux. Elle est

stockée dans le MES de manière asynchrone avec une précision de l'ordre de la seconde. Cette donnée alimente le modèle temporel pour aider à identifier le produit et à prendre des décisions en cas d'incident. Les utilisateurs internes, humains et logiciels, pourront accéder à cette donnée et la visualiser sous forme de tableau.

Dans les deux cas, la typologie de la donnée permet de mieux comprendre les caractéristiques de la donnée, son importance pour l'entreprise, son stockage et son utilisation, pour aider à une prise de décision éclairée à un niveau précoce dans le développement du JN et en amont du déploiement de ce dernier, cela avant d'engager des ressources matérielles ou financières.

Tableau 1. Typologie de la donnée appliquée à un cas d'étude

Critères	Température	Date d'ordre de fabrication
Source	Capteur	MES
Type	Environnementale	Temporel
Valeur	Quantitative (mesurable)	Quantitative (mesurable)
Accès	Oui (Ouvert)	Conditionnel
Criticité	Faible	Moyenne
Stockage	MES	MES
Synchronisation	Synchrone	Asynchrone
Précision	Faible (0.1° C)	Faible (1s)
Variabilité	Faible	Faible
Modèle	Modèle de contamination	Modèle temporel
Information	Risque de contamination	Identification de produit
Connaissance	Aide à la décision	Aide à la décision
Utilisateurs	Interne : humains – logiciel (MES)	Interne : humains – logiciel (MES)
Communication	Oui – Courbe (passé, présent)	Oui – Tableau (passé, présent)

6 CONCLUSION

Dans le présent article, nous avons présenté un modèle de cycle de vie des données en soulignant l'importance cruciale de la phase "cyber" dans le développement d'un jumeau numérique. Cette phase englobe divers éléments critiques, tels que la gestion des données, l'architecture organisationnelle et les décisions technologiques, qui peuvent exercer une influence substantielle sur la qualité, la précision et la performance du jumeau numérique. Nous avons proposé un modèle de données nommé 5A qui permet de représenter de manière exhaustive le cycle de vie des données, qui se décline en 5 étapes : acquisition, agrégation, analyse, application et accès. Le formalisme en graphe du 5A permet une représentation sous forme de réseaux de Petri qui permet de représenter, fusionner, modifier les graphes de façon formelle et optimiser l'architecture du jumeau. De plus, nous avons présenté la typologie des données qui facilite l'identification et

la description des critères permettant de caractériser une donnée.

Les travaux prospectifs se concentreront sur l'obtention d'une vision microscopique de chaque étape clé, en se penchant sur les détails spécifiques relatifs aux données. Cela implique l'identification des formalismes pertinents pour chaque étape, tels que les normes, modèles et conventions, ainsi que l'identification des outils techniques appropriés pour mettre en œuvre ces formalismes.

7 RÉFÉRENCES

- Blanchet, A., Julien, N., & Hamzaoui, M. A. (2022). Typology as a Deployment Tool for Digital Twins: Application to Maintenance in Industry. 8.
- Gallala, A., Kumar, A. A., Hichri, B., & Plapper, P. (2022). Digital Twin for Human–Robot Interactions by Means of Industry 4.0 Enabling Technologies. *Sensors*, 22(13), 4950. <https://doi.org/10.3390/s22134950>
- Grieves, M. (2014). Digital Twin: Manufacturing Excellence through Virtual Factory Replication.
- Hamzaoui, M. A., & Julien, N. (2022). Social Cyber-Physical Systems and Digital Twins Networks: A perspective about the future digital twin ecosystems. 6. <https://doi.org/10.1016/j.ifacol.2022.08.006>
- International Organization for Standardization. (2021). Automation systems and integration—Digital twin framework for manufacturing (ISO No. 23247). <https://www.iso.org/obp/ui/#iso:std:iso:23247:-1:ed-1:v1:en>
- Julien, N., & Hamzaoui, M. A. (2022). Integrating Lean Data and Digital Sobriety in Digital Twins through Dynamic Accuracy Management. 12.
- Julien, N., & Martin, E. (2021a). How to characterize a Digital Twin: A Usage-Driven Classification. *IFAC-PapersOnLine*, 54(1), 894–899. <https://doi.org/10.1016/j.ifacol.2021.08.106>
- Julien, N., & Martin, E. (2021b). Typology of Manufacturing Digital Twins: A First Step Towards A Deployment Methodology. 12.
- Kamoise, N., Guerin, C., Hamzaoui, M., & Julien, N. (2022). Using Cognitive Work Analysis to deploy collaborative digital twin. 6.
- Lee, J., Bagheri, B., & Kao, H.-A. (2015). A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3, 18–23. <https://doi.org/10.1016/j.mfglet.2014.12.001>
- Pieter, van S. (2022). Digital Twin Capabilities Periodic Table: A Digital Twin Consortium User Guide. Digital Twin Consortium.
- Tao, F., Zhang, M., & Nee, A. Y. C. (2019). Digital twin driven smart manufacturing. Academic press.
- Zhang, M., Tao, F., Huang, B., Liu, A., Wang, L., Anwer, N., & Nee, A. Y. C. (2021). Digital twin data: Methods and key technologies. *Digital Twin*, 1, 2. <https://doi.org/10.12688/digitaltwin.17467.1>