Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

DATA REQUIREMENTS FOR A DIGITAL TWIN OF A CNC MACHINE TOOL

Deogratias Kibira, and Guodong Shao

Engineering Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, USA

ABSTRACT

Digital twins can enable the intelligent operation of computer numerical control (CNC) machine tools. The data for the digital twin are collected from the machine controller, sensors, or other Internet of Things (IoT) devices. Creating a valid digital twin for a specific purpose requires identifying and specifying the right types and quality of data. However, challenges exist, such as unlabeled data and a lack of clarity on the sufficiency of data required to build a digital twin for a specific purpose. This paper discusses the data types, sources, and acquisition methods for creating digital twins for machine tools with different capabilities. Depending on the purpose, any digital twin can be categorized as descriptive, diagnostic, predictive, prescriptive, or autonomous. Data requirements for each of these categories are discussed. This paper can be used as a guide for developing and validating different types of digital twins for CNC machine tools.

1 INTRODUCTION

This section provides a background and description of the role of data in the development of digital twins of CNC machine tools. It also describes the challenges of data selection, scoping, and preparation for different types of digital twins of machine tools.

1.1 Background

The introduction of digital twins to CNC machine tools is transforming how machines are managed and maintained. In addition, digital twins are helping enhance the quality of machined products, reduce waste, and improve the sustainability and productivity of machining operations. As a result, CNC machine tool digital twins are becoming more prevalent. Major machine tool manufacturers are developing digital twins of these machines to support decision-making during product development, design, production, and service.

A digital twin is a representation of the structure, functions, and behavior of a physical system, product, phenomenon, or process in the digital world in real-time. The architecture of the digital twin of a machining system has multi-dimensional viewpoints, including the machine tool, the machined part, the material removal process, and the cutting tool. However, most machine tool digital twins focus on one dimension.

Data are critical in constructing virtual models, building cyber-physical connections, and executing intelligent operations for any of the above dimensions. For example, the linear axes of a machine tool are composed of a linear motor, sliding system, guideline, cooling system, and other components. However, the information about each element is visible only through data, making data acquisition the key to studying the equipment. Data requirements for a digital twin are the specifications of what data must be collected, accessed, fused, stored, manipulated, analyzed, and used to build the digital twin models. Our previous research developed a digital twin of a Pocket NC machine tool with emphasis on methods, tools, and standards (Kibira et al. 2024). This paper explores the data requirements for developing a digital twin with a focus on machine tool data. Ren et al. (2024) observed that current digital twin research focuses on algorithm design and development driven by the digital twin's purpose but pays little attention to the data requirements. Some studies have addressed technical aspects of data and its acquisition for various applications. Examples include Zhang et al. (2022), who explore the basic principles and methods for digital twin data gathering, interaction, storage, association, fusion, and evolution. Tong et al. (2020) address methods to collect data from a machine tool, including sensors and measurement devices. Data sources for a digital twin of a machine tool have been described as internal data, external data, control system data, and production data (Sicard et al. 2023). There is limited literature discussing the data requirements for developing machine tool digital twins.

1.2 Objectives and Motivation

The challenges to identifying and using data to build a digital twin include dealing with unlabeled data, a lack of clarity on the sufficiency of data, and a lack of knowledge on preparing validation data sets (Weiss and Brundage 2021; Kibira and Weiss 2022). Dihan et al. (2024) observe that the abundance and variety of the data enhance the accuracy and authenticity of the digital twin. However, large data sets require identifying and scoping because the needed data should not only be complete for the digital twin purpose but also relevant, focused, and obtained on time (Zhang et al. 2022).

Completeness of data refers to the coverage of all relevant data for the specific objective of the digital twin. The motivation of this paper is to provide a guide to the identification of data for various applications of the digital twin for a machine tool. For example, machine tool manufacturers often specify a subset of available data through a standard such as MTConnect. If some data items cannot be directly obtained through the standardized interface, they could be obtained using virtual sensors. Figure 1 illustrates an example of data acquisition from a machine tool and its usage in the digital twin.

1.3 Contribution, Scope, and Organization of the Paper

Few publications focus on the data required for building a digital twin of a machine tool. Emphasis has been placed on real-time data collection, communication, and analysis without first categorizing and identifying the range of data necessary for a digital twin with a specific purpose (Tong et al. 2020). Performing a data requirements analysis explores complete and valid reference data sets. Although this paper does not provide an exhaustive coverage of all possible digital twin data, it offers an analyst a guide on what data they require to build a digital twin of their machine tool for a given set of services. Digital twin services are defined at a high level, and a developer will first determine where their intended services fall to select the most relevant data needed. Other factors relevant to developing effective digital twins, such as data quality, sensor synchronization, preprocessing, edge computing, and low-latency requirements, are out of scope.

The rest of the paper is organized as follows. Section 2 presents related work; Section 3 overviews data sources from machine tools and associated standards; Section 4 discusses data required for different digital twin categories; Section 5 provides an example of data for a descriptive digital twin; Section 6 is the discussion, conclusion, and future work.

2 RELATED WORK

This section reviews the relevant literature. Tong et al. (2020) describe data acquisition methods for processing and analysis. They describe data acquisition methods where signals are related to machine tool kinematics, dynamics, acoustics, thermodynamics, optics, energy consumption, and vision. Data are classified as attribute, real-time control, and condition data. Cai et al. (2017) discuss techniques for deploying sensors to capture machine-specific features for modeling and developing digital twins of machine tools. The focus is on integrating manufacturing information from controllers and sensory data into virtual machine tools to improve their accuracy and capabilities. Three categories of data and information are recorded and passed on for information extraction: the machine data stream, the sensor data stream, and the machine health status. Hänel et al. (2021) identify basic information and data models necessary for a digital twin of machining operations. Data categorizations are workpiece data, process data, technology data, machine tool data, and tool data.

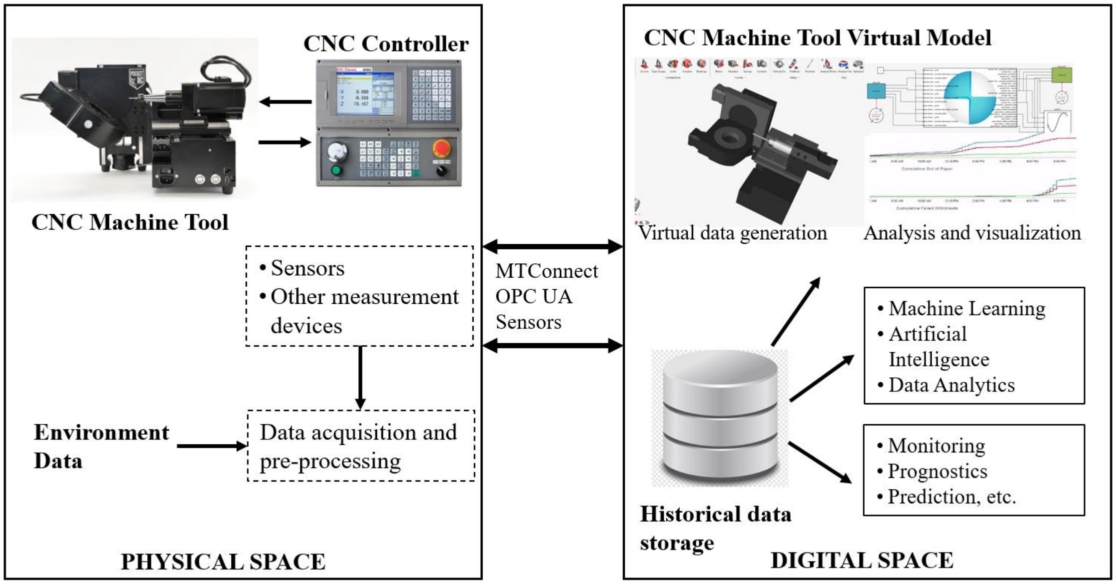

Figure 1: Data acquisition, usage, and exchange between a CNC machine tool and its digital twin.

Lai et al. (2021) provide a classification of machining data for developing a digital twin. There are four categories: property data, process planning data, sensory data, and control system data. Pantelidakis et al. (2024) demonstrate data acquisition, processing, and distribution, and virtual representation for a digital twin of a Pocket NC machine tool. Both online and offline digital twins are investigated, and the offline digital twin is demonstrated for collision detection during machining operations.

The review shows no standard or organized structure for identifying data for building digital twins of CNC machine tools. Each case identifies its own data. Most discussions on data requirements have been done in the context of digital twins for general application, not a machine tool digital twin. Examples are the categorization of data for the digital twin and a methodology to support the identification of data from outside of the physical system (Hildebrandt et al. 2022) and Yan et al. (2023), who investigated data requirements for a digital twin through sourcing from multiple recipients.

3 RELATED STANDARDS AND DATA SOURCES

This section overviews standards that support digital twin development for manufacturing applications and the different sources of data for machine tool digital twins. It also discusses typical machine tool components from which data are collected.

3.1 Related Standards

Literature includes case studies of digital twins for machine tools. However, each case study follows a different approach to developing the digital twin because there is no unified approach, including terminology and related guidelines. There are also no technical standards related to digital twins of machine tools leading to the adoption of high-level manufacturing digital twin standards (Cabral et al. 2023; Kibira et al. 2024). Some of the applicable manufacturing digital twin standards are:

- ISO 23247: This series of standards provides a generic development framework that can be instantiated for case-specific implementations of digital twins in manufacturing (ISO 23247, 2024).
- ISO 14306: This standard defines the syntax and semantics of a file format for 3-dimensional (3D) visualization and interrogation of geometry and product manufacturing information derived from Computer-aided design (CAD) systems using visualization software tools (ISO 14306, 2024).

- ISO 14649: This series of standards supports data modeling and data transfer between CAD/Computer-aided manufacturing (CAM) systems and CNC machines. It is also called the Standard for the Exchange of Product Model Data for Numerical Control (STEP-NC). Part 1 provides an overview and fundamentals (ISO 14649-1, 2003).
- IEC 63278-1:2023 defines the structure of a standardized digital representation of an asset, called Asset Administration Shell (AAS), which enables uniform access to information and services (IEC 63278-1, 2023).
- Data collection standards that can support the development of machine tool digital twins include MTConnect, Message Queuing Telemetry Transport (MQTT), OPC Unified Architecture (OPC UA), blockchain, Open API, and Modbus TCP (Hsiao et al. 2021).

Further development for the digital twin data standard should focus on (1) data structure and communication, (2) cybersecurity, and (3) open data. The interoperability of digital twins, with clear definitions of inputs and outputs, connections, communication, and data exchange, is a significant factor in a heterogeneous environment.

3.2 Data Sources

Data for machine tool digital twins can be obtained from different sources, as elaborated below:

- Physical sensors: Sensors provide data that is internal to the system. These sensors include servo-control systems, position/angle encoders, torque, current, and voltage sensors. Many of the sensors are integrated into the machine design. Others are IoT devices, mobile devices, and wearable devices.
- Controller: Machine controllers are the primary source of data. Data collected includes axis positions, spindle speed, tool offsets, current machine status (running, paused, alarm), alarms and error codes, cutting tool data, and program execution status.
- Virtual sensors: Virtual sensors are needed when it is impossible to measure the data directly. Virtual data must be inferred from directly collected data through simulations and computations. Doerrer et al. (2022) modeled a virtual sensor to detect vibrations at the tool center point based on internal machine data.
- External data: These are data from outside the machine tool that can influence its operation. They are obtained by measuring the working conditions: thermometers, vibration displacement sensors, accelerometers, vibration displacement sensors, a microphone, acoustic emission sensors, an RFID reader for code recognition, energy consumption sockets, and industrial cameras (Tong et al. 2020).
- Production data includes information such as part history, quality, and machine reliability. It also contains data such as the number of cuts performed by a specific machine tool, the number of defects or rejects, surface roughness, or the processing time for each operation.
- Data from production systems: Existing applications in an industrial environment are Computer-aided Process Planning (CAPP), Enterprise Resources Planning (ERP), Supervisory Control and Data Acquisition (SCADA), and Manufacturing execution system (MES).

3.3 Machine Tool Components

Most research efforts for digital twins of machine tools focus on the moving parts since they consume the most energy and are more liable to wear and tear. Considering a machining center and milling machine, the major components include the servo motor, spindle motor, Automatic Tool Changer (ATC), coolant pump, carousel rotation, fan, unloaded motor, jog axis, spindle, tool magazine, tool changer, coolant pump, and machine table (Kordonowy et al. 2002; Avram et al. 2011). Figure 2 shows major CNC machine tool components.

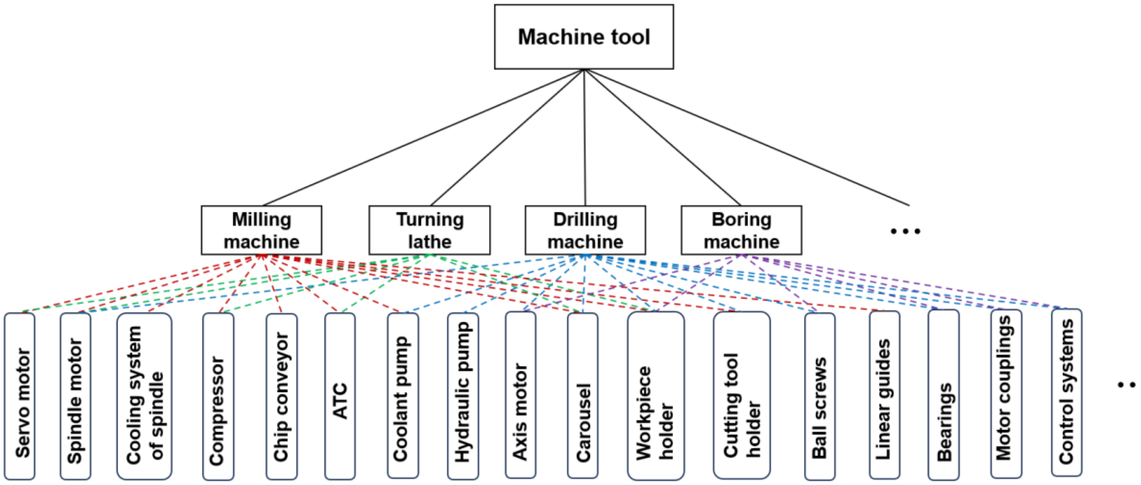

Figure 2: CNC Machine tools and some of their components from which data are collected.

4 DATA REQUIREMENTS FOR DIFFERENT DIGITAL TWIN TYPES

A digital twin can be developed to fulfill multiple objectives such as monitoring, aggregating, analyzing data, visualizing data, obtaining insights from data, predicting future state and behavior, and providing decision support (Stadtmann and Rasheed 2024). Digital twins are context-dependent, so the capabilities and data integrated into a digital twin depend on the intended purpose (Shao and Helu 2020). However, digital twins play larger roles as their capabilities and the set of services that they deliver increase. Sundby et al. (2021), and Altamiranda et al. (2019) use the concept of ‘‘capability level’’ to categorize and analyze the levels of functionality of digital twins and the required data sets. The levels used in this paper are 1-descriptive, 2-diagnostic, 3-predictive, 4-prescriptive, and 5-autonomous. These digital twin types are connected to their physical counterparts.

However, a digital twin can also be developed to support testing, optimizing, and identifying potential design issues before the system is built. The required data are obtained from CAD models of the machine tool, kinematic simulations, dynamic modeling, G-codes, or finite element analysis. These digital twins, also called digital twin prototypes, not only provide for design validation and performance analysis but also virtual commissioning (Norberger et al. 2020). This section discusses data requirements for digital twins at different capability levels. The identification of data includes description, sources, the purpose of collecting this data, and the data format.

4.1 Descriptive Digital Twins

Descriptive digital twins mirror and describe the physical system and enhance understanding of the prevailing state of a system. The data are obtained from the controller, sensors, and other sources. While not providing or recommending a course of action, the visualization of the status provided by descriptive digital twins helps to detect an anomaly or problem so that appropriate actions can be taken. The data required for building descriptive digital twins includes that which helps model the physical system and provides visualization capabilities. The data should also be in a suitable format for developing the digital twin. A standard such as ISO 14306 helps define the syntax and semantics of collected data for 3D visualization (ISO 14306, 2024). The relevant data are (1) workspace data, (2) data of the machine, tool, and workpiece, (3) machine tool dynamic data, (4) CNC program data, and (5) production plan data. These data are described next.

4.1.1 Workspace Data

Description: Workspace data refers to the information that defines the three-dimensional space within which the machine tool operates. This includes the reach of its axes and limitations based on its geometry and the machine footprint within the workcell or factory layout. These data are provided as CAD models.
Purpose: Workspace data provides boundaries of the machine tool's workspace in relation to other equipment, such as material handling. This data also ensures that the designed part can be machined within the machine's capabilities, avoiding collision and maximizing efficiency.
Source: The source of these data is the original designs of the workcell or the factory. Direct measurements of the existing or proposed workspace expansion plans can also provide this data.
Format: Various formats are available, including those that are standardized, such as ISO 10303 (ISO 10303, 2024), also called STEP (Standard for the Exchange of Product Model data). Others are IGES (Initial Graphics Exchange Specification), STL(Stereolithography), and DXF (Drawing Exchange Format), which are considered "neutral" formats allowing data exchange between various CAD applications. Others include JSON (JavaScript Object Notation), XML (Extensible Markup Language), and spreadsheets.

4.1.2 CAD Data of the Machine Tool, Cutting Tool, and Workpiece

Description: These are the data for the machine tool 3D CAD model. Each machine tool can be described by its kinematic model. These data are used to model, visualize, analyze, and program the machine tool's movements during CNC machining to simulate toolpaths and optimize manufacturing processes.
Purpose: The purpose of this data is to allow engineering designers to accurately represent the machine tool's geometry, work area, tooling positions, and other features. These data enable virtual testing and optimization before the machine tool is built.
Source: The source of these data is the machine tool vendors, who often provide the machine CAD data. Usually, the data is also available through manufacturers' websites or a dedicated download service. There are also websites such as SketchUp, Free3D, and GrabCAD, which are cloud-based platforms for sharing CAD models. Workpiece data are available in the product or part design with embedded Product and Manufacturing information.
Format: The models and embedded data are often available in STEP or a similar format.

4.1.3 Machine Tool Dynamic Data

Description: These are the machine tool data that change with time. This category also refers to data related to forces and torques that cause the motions during the machining process. These data include cutting force, motor torque, energy consumption, vibration, and noise levels.
Purpose: In the context of descriptive digital twins, these data are monitored and compared with expected values to understand the health state and performance of the machine tool.
Source: The main source of these data is the machine tool, often from the interior, using sensors.
Format: Machine data are often provided in XML or a spreadsheet.

4.1.4 Computer Numerical Control Program Data

Description: The CNC program data provides commands to control the operation of the machine tool. It translates part design specifications into coded instructions (G-code) that direct the movements of the cutting tool to manufacture a part according to the part's design. Each "G" command signifies a specific geometric operation such as linear movement, circular interpolation, or rapid positioning.
Purpose: The purpose of this data is to direct the position of the tooltip on the part within the working envelope or workspace.
Source: These data are generated from the CNC control program or G-code. Although this program can be generated by direct programming, engineers more often use CAM software for this purpose.
Format: The format of CNC programs and data is ASCII code in a plain text file.

4.1.5 Production Plan Data

Description: The production data provides information about the status of production orders, which is compared with the original production plan. Data are needed from all stages of the production cycle, such as the number of parts completed, rejected, and those that need rework. Key performance indicators, such as schedule attainment, availability, quality rate, and overall equipment effectiveness, monitor and assess the production plan.
Purpose: The purpose is to track the efficiency of the operation of equipment as well as the frequency of nonscheduled equipment stops during production.
Source: The source of this data is the historical production information and production plans
Format: Various data formats are available for structuring this data, including JSON (JavaScript Object Notation), XML (Extensible Markup Language), databases (e.g., SQL), or spreadsheets (e.g., EXCEL). Specialized manufacturing software (e.g., ERP) often has its specialized formats.

4.2 Diagnostic Digital Twins

Diagnostic digital twins help identify why an unexpected event happened or why it is happening to the asset. The digital twins created for prognostics and health management monitor degradations, detect anomalies, and diagnose failures and their cause(s). Diagnosis requires data from at least two levels of sensing: machine-level sensing and individual component sensing. Machine-level sensing data establishes that something is wrong with the machine tool. Component-level data, on the other hand, determines the source and cause of the observed machine degradations. These levels of sensing are illustrated in Figure 3.

4.2.1 Upper-level Sensing Data for CNC machine tools.

Description: The data to sense machine tool performance include tool position accuracy, velocity accuracy, repeatability, cutting forces, spindle speed, energy consumption, vibration, and noise levels.
Purpose: The purpose of upper-level sensing data is to indicate the machine's state regarding behavior or performance.
Source: These data are obtained from the machine controller, CNC program, or direct measurements using sensors. Energy consumption is obtained through metering at the CNC's electrical power input port. Machine vibration and temperature data are obtained through accelerometers and temperature gauges.
Format: The format is the same as that for the descriptive digital twin.

4.2.2 Lower-level Sensing

Description: The lower level of sensing is where data is analyzed to determine a relationship between observed machine tool behavior and the machine tool component(s) responsible for that behavior. The data includes vibration peak amplitude, acoustic emission frequency spectrum, and ultrasonic emission. Temperature data include motor temperature, spindle bearing temperature, and coolant temperature. Power consumption data, such as motor power draw and fluctuations in power usage. The cutting tool data such as tool status, tool wear, and surface finish. Other relevant data are overheating, chatter, mechanical wear in the machine components (bearings, guideways), spindle runout, and gear backlash.
Source: These data are derived from the interior components of machines, especially those that are prone to failure. See Figure 2 for CNC machine tool components.
Format: The format is the same as that for the descriptive digital twin.

4.3 Predictive Digital Twins

Description: The predictive digital twin's aim is to forecast the system's future state, which supports decision-making. Cai et al. (2017) identified three main types of data for machine tool prediction: machine control data, sensor data, and machine status data. The operator specifies machine control data, including coolant level, tool number, cutter location, and spindle speed. Sensor data consists of vibration, current

acoustic emission, and power. Machine status includes on/off, busy, and idle. Most of these data are the same as that for diagnostic digital twins. However, larger volumes of this data are required to build the needed predictive models.
Purpose: The purpose is to predict future states of the machine tool to facilitate scheduling downtime for testing or maintenance. Predictions especially help guide production and maintenance scheduling by recommending replacing components just before their lifetimes expire to prevent wasting valuable productive time.
Source: These data are obtained from the machine controller, numerical control program, or sensors.
Format: The format is the same as that for the descriptive digital twin.

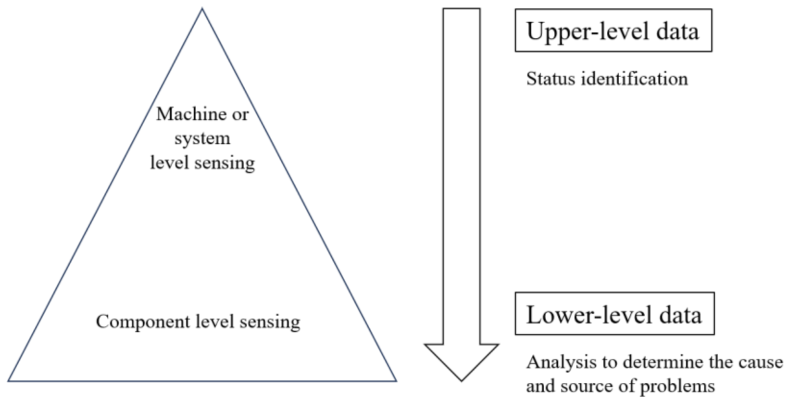

Figure 3: Major Levels of sensing to support machine tool diagnosis.

4.4 Prescriptive Digital Twins

Description: “Prescription” is the recommendation of the optimal course of action or strategy and the reason for selecting that course of action. Prescriptive digital twins are mainly built on what-if analysis through simulation and uncertainty quantification (Stadtmann et al. 2023). Most of the required data are the same as that for predictive digital twins. It includes the data to assess the current and anticipated health states of equipment, the status of production orders, and machine tool loading. In addition, historical labeled data is required on machine control data, sensor data, and machine tool behavior and performance to train prescriptive models.
Purpose: The purpose of prescriptive digital twins is to recommend the course(s) of action to realize or achieve the intended purpose(s) and goal(s) of the digital twin.
Source: The sources of equipment data for prescriptive digital twins are maintenance logs, manufacturer recommendations, and similar equipment. Production data are obtained from the production equipment workload. Degradation data, which provides anticipated health states of equipment, are obtained from published data of similar mechanisms that are installed in machine tools such as linkages, brakes, clutches, preselection mechanisms, sliding mechanisms including cams, screws, racks, bearings, cams, gears, and motors.
Format: The format of the data is the same as that for descriptive digital twins, but differences may exist depending on the application and specific case. JSON is one of the most standard data formats for digital twins. Other formats include XML, Digital Twin Definition Language (DTDL), and Apache Parquet.

4.5 Autonomous Digital Twins

Description: An autonomous digital twin operates independently with minimal human intervention, monitoring, analyzing, and making decisions based on new data streams to improve accuracy and adapt to changing conditions (Latifah et al. 2024). Further, autonomous digital twins support a machine tool to adapt to changes in conditions, such as the part’s design, level of production, or the cutting tool (Zhang et al.

2024). Prediction and self-optimization are particularly essential for autonomous digital twins. The predictions are compared with the real values in real-time to determine prediction accuracy and update the prediction models. The data required includes what is needed for prescriptive digital twins but needs to be integrated in real-time for analysis and the application of decision-making algorithms and data-driven models. In addition to the real-time production data, historical data and surrounding conditions are particularly important.

Purpose: The purpose of autonomous digital twins is to speed up the decision-making process and improve efficiency by learning from incoming data to refine the predictions and decision-making capabilities. Autonomous digital twins reduce costs and enhance product quality.

Source: The data comes from the machines. These data include the machine data stream, the sensor data stream, and the machine health status.

Formats: The formats include spreadsheets, CSV, PDF-ECG, and XML-based formats.

5 DATA FOR DESCRIPTIVE DIGITAL TWIN OF A POCKET NC MACHINE TOOL

A simulation model provides visualization capabilities, which enhance verification/validation, system understanding, and decision-making for a descriptive digital twin. This section provides a case study of streaming data from a machine tool to create a descriptive digital twin of a machine tool.

5.1 Machine Tool Description and Features

The machine tool is a Pocket NC milling machine, which executes tasks through three translational and two rotational movements. Linear motions are those made in a straight line, i.e., X, Y, and Z axes. Rotary motions take place about a fixed point or axis. Figure 4 shows the major moving components of the Pocket NC machine tool (left) and the digital representation (right).

5.2 Identification of Digital Twin Data

- Workspace data: The workspace consists of two robot arms for material handling, the Pocket NC machine tool for producing parts, and a Coordinate Measuring Machine (CMM) for inspection and metrology. The machine tool is mounted on an optical table to provide a rigid and stable support. The dimensions of the table are obtained. The distance between mounting holes on the table allows for easy and precise placement and data on the machine location relative to the table and the other equipment in the lab. These workspace data are obtained in STEP or .x_b formats.
- CAD data of the machine tool, cutting tool, and workpiece: The vendor provides CAD models for the machine tool, cutting tool data, and fixture in the STEP format. The models are "grouped" into the major components, namely, base, carriage, vertical saddle, spindle, A Table, and B Table. The part design has pockets, slots, holes, and edge chamfering.
- CNC program data: This program is obtained from the part design. The CAD/CAM software allows you to visualize how the machine produces each geometric feature and validate the machining process. The result was a tool path to clear the bulk material and produce the final geometry. The program is obtained, loaded as a .ngc file, and copied to the machine's computer. Other program data include cutting parameters such as feed rate, cutting depth, width of cut, and spindle speed.
- Machine tool dynamic data: These data change with time during machine operation. If the machine operates normally, these data are expected to take predictable values. However, deviation from the predicted levels may indicate a fault or failure. These data include tool point position, cutting force, spindle speed, energy consumption, vibration, and noise levels.

5.3 Visualization of position data

The positions commanded by the CNC Machine tool are Absolute X position, Absolute Y position, Absolute Z position, Absolute A position, and Absolute B position. Figure 5 is a plot of the X-axis positions and position errors. The plot shows that significant errors coincide with changes in the direction of motion or rotation about the axes.

5.4 Observation

The data visualized for this case study was only the kinematic motion data for the descriptive digital twin. However, additional data can be streamed, analyzed, and the results visualized. These data include torque, force, vibration, temperature, tool condition, and environmental conditions. The selection would be determined by the purpose of the digital twin.

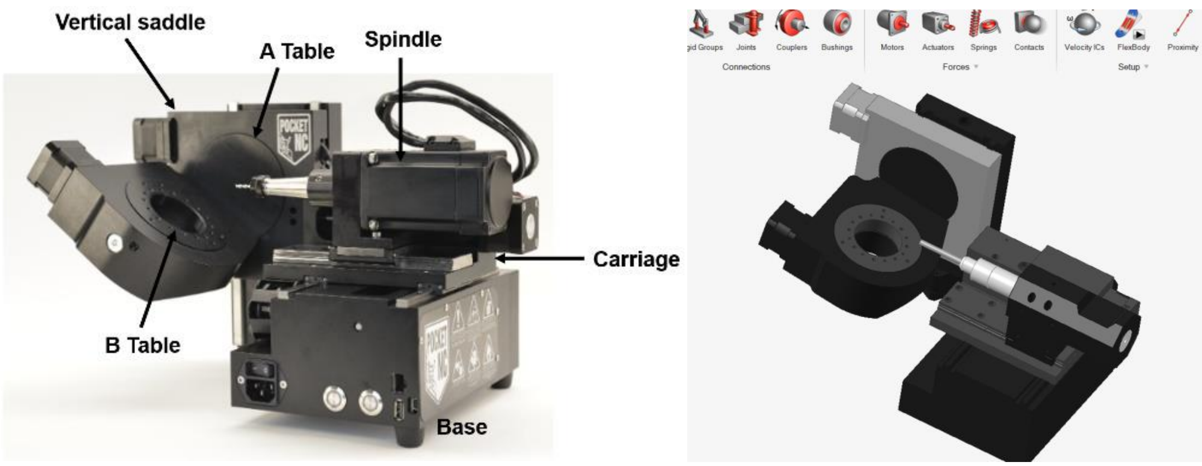

Figure 4: Main parts of a Pocket NC machine tool (left) and its digital twin (right) (Kibira et al. 2024).

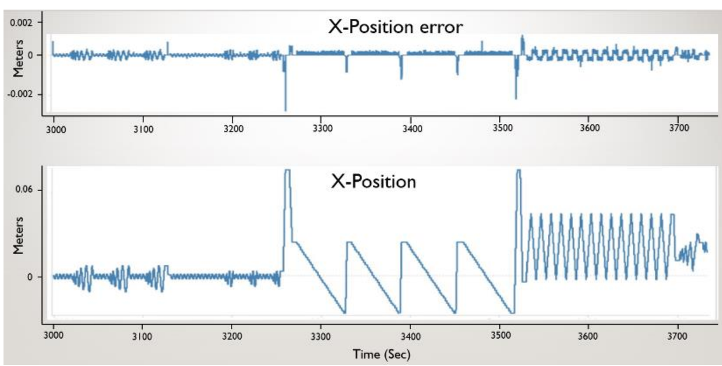

Figure 5: Plot of the Absolute X position and position errors (Kibira et al. 2024).

6 DISCUSSION AND CONCLUSION

The challenges for building a digital twin stem from a lack of standards specific to digital twins for CNC machine tools, including, among others, the specification of the required data sets. This lack results in an inability to specify a consistent approach to data identification, collection, communication, validation, and use. Yet specifying, collecting, and analyzing the correct data required is pivotal for realizing a valid digital twin. This paper has used a digital twin capability scale (from 1 to 5) to identify data requirements for a digital twin for a given purpose, application, or project. We also discussed the different data types, sources, and acquisition methods for building digital twins for machine tools.

The ongoing digital twin research efforts at NIST have been bolstered by installing a manufacturing workcell comprising collaborative robots, a machine tool, and a coordinate measuring machine. The objective is to develop scenarios representing different digital twin objectives and identify and specify the relevant data sets. Digital twin verification and validation are among the major concerns for further development and adoption of digital twin technology. Just as a digital twin can be used to validate a program, data from a physical machine tool can validate a digital twin before deployment.

In furthering this work, efforts will be directed towards data requirements focusing on the cutting tool, production process, and the workpiece. Research will also focus on enhancing approaches for integrating digital models with real-time streaming data. Many digital twin concepts in literature lack full integration, especially the feedback loop. This is partly due to the challenges of data identification, gathering, filtering, and processing in real-time. Future work should also focus on data specification, data processing, interfaces, digital twin configuration, twin model construction, and integrations. The results of this work will address interoperability, adoption barriers, and protocol limitations particularly in heterogeneous CNC environments.

DISCLAIMER

Certain commercial products and systems are identified in this paper to facilitate understanding. Such identification does not imply that these software systems are necessarily the best available for the purpose. No approval or endorsement of any commercial product by NIST is intended or implied.

REFERENCES

Altamiranda, E., and E. Colina. 2019. "A System of Systems Digital Twin to Support Lifetime Management and Life Extension of Subsea Production Systems". In *OCEANS 2019*, June 17th-20th, Marseille, France, 1-9.

Avram, O., I. Stroud, P. Xirouchakis. 2011. "A Multi-criteria Decision Method for Sustainability Assessment of the use Phase of CNC machine tool Systems". *International Journal of Advanced Manufacturing Technology* 53: 811-828.

Cabral, J. V. A., E. A. R. Gasca, and A. J. Alvares. 2023. "Digital Twin Implementation for Machining Center Based on ISO 23247 Standard". *IEEE Latin America Transactions*, 21(5):628-635.

Cai Y., B. Starly, P. Cohen, and Yuan-Shin Lee. 2017. "Sensor Data and Information Fusion to Construct Digital-twins Virtual CNC machine tools for Cyber-physical Manufacturing". *Procedia Manufacturing,* 10:1031-1042.

Dihan, M. S., A. I. Akash, Z. Tasneem, P. Das, S. K. Das, M. R. Islam, M. M, Islam, F. R. Badal, M. F. Ali, M. H. Ahmed, and S. H. Abhi. 2024. "Digital Twin: Data Exploration, Architecture, Implementation and Future". *Heliyon* 10(5):1-27.

Doerrer, F., A. Otto, M. Kolouch, and S. Ihlenfeldt. 2022. "Virtual Sensor for Accuracy Monitoring in CNC Machines". *Journal of Manufacturing and Materials Processing*, 6(137):1-11.

Hänel, A., A. Seidel, U. Frieß, U. Teicher, H. Wiemer, D. Wang *et al.* 2021. "Digital Twins for High-tech Machining Applications - A Model-based Analytics-Ready Approach". *Journal of Manufacturing and Materials Processing*, 5(3):1-18.

Hildebrandt, G., P. Habiger, D. Dittler, M. Barth, R. Drath, and M. Weyrich. 2022. "A Methodology for Classifying Data Relevance to Utilize External Data Sources in the Digital Twin". In *2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA)*, September 6th - 9th, Stuttgart, Germany, 1-4.

Hsiao, C.H., and W. P. Lee. 2021. "OPIIoT: Design and Implementation of an Open Communication Protocol Platform for Industrial Internet of Things". *Internet of Things*, 16(1):100441.

IEC 63278-1:2023: "Asset Administration Shell (AAS) for Industrial Applications - Part 1: Administration Shell Structure". Document IEC 63278-1 ED1. https://webstore.iec.ch/en/publication/65628, accessed 5th April 2025.

ISO 10303: 2024. "Industrial Automation Systems and Integration - Product Data Representation and Exchange". https://www.steptools.com/stds/step/, accessed 8th April 2025.

ISO 14306: 2024. "Industrial Automation Systems and Integration - JT File Format specification for 3D Visualization". https://www.iso.org/standard/86063.html, accessed 11th April 2025.

ISO 14649-1: 2003. "Industrial Automation Systems and Integration - Physical Device Control – Data Model for Computerized Numerical Controllers". https://www.iso.org/standard/34743.html, accessed 11th April 2025.

ISO 23247: 2024. "Digital Twin Framework for Manufacturing". https://ap238.org/iso23247/, accessed 11th April 2025.

Kellens, K., W. Dewulf, M. Overcash, M. Hauschild, and J. R. Duflou. 2012. "Methodology for Systematic Analysis and Improvement of Manufacturing Unit Process Lifecycle Inventory (UPLCI). Part 2: Case Studies". *International Journal of Life Cycle Assessment* 17(2): 242-251

Kibira, D., G. Shao, R. Venketesh, and M. J. Triebe. 2024. "Building a Digital Twin of a CNC machine tool". In *2024 Winter Simulation Conference (WSC),* 2915-2926. https://doi.org/10.1109/wsc63780.2024.10838819.

Kibira, D., and B. A. Weiss. 2022. "Towards a Digital Twin of a Robot Workcell to Support Prognostics and Health Management". In *2022 Winter Simulation Conference (WSC),* 2968-2979. https://doi.org/10.1109/wsc57314.2022.10015371.

Kordonowy, D. N. 2002. “A Power Assessment of Machining Tools”. *B.S. Thesis*, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Lai, X., Y. Zhou, L. Jiang, and G. Ding. 2021. “A Review: CNC Machine Tools Digital Twin Modeling and Application”. *2021 26th International Conference on Automation and Computing*, September 2nd - 4th, Portsmouth, United Kingdom, 1-6.

Latifah, A., S. H. Supangkat, E. Leksono, and A. Indraprastha. 2024. “Navigating the Future of Building Management: A Deep Dive into Prescriptive Digital Twins.” *https://www.researchgate.net/publication/381928884_Navigating_the_Future_of_Building_Management_A_Deep_Dive_into_Prescriptive_Digital_Twins*, accessed 5th April 2025.

Liu, J., D. Yu, Y. Hu, H. Yu, W. He, and L. Zhang. 2021. “CNC machine tool Fault Diagnosis Integrated Rescheduling Approach Supported by Digital Twin-driven Interaction and Cooperation Framework”. *IEEE Access*, 9:118801-118814.

Norberger, M., R. Apitzsch, A. Sewohl, H. Schlegel, and M. Putz. 2020. “A Holistic Approach for the Development of a Digital Twin Focused on Commissioning and Control of Electromechanical Feed Axes”. In *Proceedings of the 17th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2020)*, July 7th-9th, Lieusaint, France, 769-774.

Pantelidakis, M., and K. Mykoniatis. 2024. Extending the Digital Twin Ecosystem: A Real-time Digital Twin of a LinuxCNC-Controlled Subtractive Manufacturing Machine”. *Journal of Manufacturing Systems*, 74:1057-1066.

Ren, Z., K. Wan, R. Zhang, and P. Qiao. 2024. “Digital Twin Portrait: A Fusion and Application Method of Multisource Twin Data for Flexible Manufacturing Line,” in *IEEE Journal of Emerging and Selected Topics in Industrial Electronics* 5(2):753-762.

Shao, G., and M. Helu, M. 2020. “Framework for a Digital Twin in Manufacturing: Scope and Requirements”. *Manufacturing Letters* 24:105-107.

Sicard, B., Q. Butler, P. Kosierb, Y. Wu, Y., Ziada, and S. A. Gadsden. 2023. “Design Considerations for Building an IoT Enabled Digital Twin CNC Machine Tool Sub-System”. In *2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings)*, September 16th – 17th, Mount Pleasant, MI, USA, 1-5.

Stadtmann, F., and A. Rasheed. 2024. “Diagnostic Digital Twin for Anomaly Detection in Floating Offshore Wind Energy”. In *International Conference on Offshore Mechanics and Arctic Engineering*, June 9th – 14th, Singapore, Singapore, 1-10.

Sundby, T., J. M. Graham, A. Rasheed, M. Tabib, and O. San. 2021. “Geometric Change Detection in Digital Twins”. *Digital* 1(2):111-129.

Tong, X., Q. Liu, S. Pi, and Y. Xiao. 2020. “Real-Time Machining Data Application and Service Based on IMT Digital Twin”. *J Intell Manuf* 31:1113–1132.

Ward, R., C. Sun, J. Dominguez-Caballero, S. Ojo, S. Ayvar-Soberanis, D. Curtis *et al.* 2021. “Machining Digital Twin Using Real-time Model-based Simulations and Lookahead Function for Closed Loop Machining Control”. *Int J Adv Manuf Technol* 117: 3615–3629.

Weiss, B. A., and M. P. Brundage. 2021. “Measurement and Evaluation for Prognostics and Health Management (PHM) for Manufacturing Operations–Summary of an Interactive Workshop Highlighting PHM Trends”. *International Journal of Prognostics and Health Management* 12(1):1-34.

Yan, J., Q. Lu, N. Li, and M. Pitt. 2023. “Stakeholder Identification-based Data Requirements Specification Approach for City-level Dynamic Digital Twin”. In *30th EG-ICE: International Conference on Intelligent Computing in Engineering*, July 4th – 7th, London, United Kingdom, 1-10.

Zhang, L., X. Wang, S. He, X. Mao, B. Li, and H. Liu. 2024. “Multi-models Associated with Process Information-driven Process Autonomous digital twin for Multi-variety Production of Intelligent Machines”. *Science China Technological Sciences* 67:3825-3842.

Zhang M, F. Tao, B. Huang, A. Liu, L. Wang, N. Answer *et al.* 2024. “Digital Twin Data: Methods and Key Technologies, *Digital Twin*, 1:1, 2.

AUTHOR BIOGRAPHIES

DEOGRATIAS KIBIRA is a Professional Research Experience Program (PREP) fellow in the Engineering Laboratory at the National Institute of Standards and Technology. His research interests are in sustainable manufacturing, digital twin development, virtual manufacturing environments, and PHM for smart manufacturing. He has a PhD in Manufacturing Engineering. His e-mail address is deogratias.kibira@nist.gov.

GUODONG SHAO is a Computer Scientist in the Systems Integration Division (SID) of the Engineering Laboratory (EL) at the National Institute of Standards and Technology (NIST). He manages the Digital Twins for Advanced Manufacturing project and focuses on generic guidelines and methodologies for implementing digital twins and relevant standards development and testing. He is a technical expert on relevant standards committees and a co-chair of the Digital Twin Verification and Validation Working Group at the Digital Twin Consortium (DTC). His email address is gshao@nist.gov.