Information Systems Engineering with Digital Shadows: Concept and Case Studies

An Exploratory Paper

Martin Liebenberg[(⊠)] and Matthias Jarke

Computer Science Department, RWTH Aachen University, Aachen, Germany
`liebenberg@kbsg.rwth-aachen.de`, `jarke@dbis.rwth-aachen.de`

Abstract. The production sector has faced many difficulties in taking full advantage of opportunities found in other web application domains. Production research has focused on sophisticated mathematical models ranging from molecular materials modeling to efficient production control to inter-company supply network logistics. Often, these models have no closed-form solutions; this led to intense simulation research for individual modeling viewpoints, often labeled "Digital Twins".

However, the complexity of the overall system precludes Digital Twins covering more than just a few system perspectives, especially if near-realtime performance is required. Moreover, the wide variety of individual situations and behaviors is usually captured only as statistical uncertainty. In order to achieve better performance and more context adaptation, the interdisciplinary research cluster "Internet of Production" at RWTH Aachen University is exploring the concept of "Digital Shadows". Digital Shadows can be understood as compact views on dynamic processes, usually combining condensed measurement data with highly efficient simplified mathematical models. In this exploratory paper, we argue based on a couple of initial case studies that Digital Shadows are not just valuable carriers of deep engineering knowledge but due to their small size also help in reducing network congestion and enabling edge computing. These properties could make Digital Shadows an interesting solution to address resilience in other information-intensive dynamic systems.

Keywords: Digital shadows · Internet of Production (IoP) · Database views · World Wide Lab (WWL) · Collaborative information systems

1 Introduction

In recent years, Digital Twins became increasingly important in the research fields of production planning and controlling [25]. Digital Twins are digital

---

artefacts which represent in a production environments, for instance, a work piece or product in a production scenario. The representation is done with mathematical models or simulations which can also be regarded as models. These models are used, for example, to develop new products or improve processes [14]. Digital Twins are thought of as a digital counterpart to physical production artefacts. So they have a high resolution to be useful for every purpose in its environment. But, Digital Twins are often too slow to use them in real-time scenarios. In terms of simulations, Finite Element Method (FEM) [32] simulations are widely used in the field of production [4,18,20]. These simulations are rather expensive to be calculated because production parts are divided into hundreds to millions of discrete elements for which certain physical properties are calculated in relation of its neighbors. So that typically, a duration of one simulation run ranges from several minutes to hours or days. That is not feasible for fast reaction to defects in a production process.

Therefore, we explore the concept of Digital Shadows. The term "Digital Shadow" was first introduced about ten years ago [8] in the privacy debate on the dangers of your "Digital Footprint" where such a footprint consists of a large number of shadows or traces left by all your digital actions; since that time, there exist even companies that help you or your company against misuse of your digital shadow[1].

During the early Digital Twin debate in production engineering around 2015, Digital Shadows played only a marginal role as documented traces of the "real" production processes or its digital twin. In contrast, we want to study and treat digital shadows as "first-class citizens". In this setting, Digital Shadows are dynamic digital views or traces on a physical process or a simulation, where only those aspects are represented which are necessary for a specified purpose. This can be subsets of data from production processes, or functions mapping certain data onto subset of other data. These can then be used in scenarios where fast reaction is crucial. From a physical data management perspective, a Digital Shadow is condensed data of small size generated for a particular task in a production process so that it can be transferred in networks reducing congestion.

In the interdisciplinary research cluster "Internet of Production" at RWTH Aachen University, more than 20 institutes from mechanical engineering, material science, humanities and computer science are working on the vision of an Internet of Production (IoP) for a new level of cross-domain collaboration. In this cluster we work on systems realising the generation and usage of Digital Shadows in a global network of production facilities.

In this paper we explore Digital Shadows as a main aspect of the IoP in more detail in terms of requirements for information systems. In the next section we sketch in brief the vision of the IoP. In Sect. 3 we explore the role of Digital Shadows from two perspectives. First, we discuss in Sect. 3.1 Digital Shadows from the perspective of Database Views. Second, in Sects. 3.2 and 3.3 we present

---

[1] This is not to be confused with another stream of political economy literature concerning the "digital shadow economy" [9] related to the concept of (usually illegal) shadow economy.

two interdisciplinary use cases for very useful Digital Shadows in Plastics and Steel Engineering and demonstrate their Information Systems Engineering in a prototypical implementation. We end with a discussion on further research questions in making Digital Shadows and the IoP a reality.

2   IoP: Beyond Production Monitoring and Control

One core vision of the IoP is called the World Wide Lab (WWL). Here, we assume that the models can only be improved if the data is as diverse as possible. For instance, if we train a Neural Network with data coming from only one type of metal in hot rolling, it is likely that for another material the model does not work properly. So, we claim that we need to share data and models so that Digital Shadows can be improved and furthermore be shared for use in other production scenarios. We regard each experiment or even process step as part of one huge experiment in the WWL. Therefore, we want to use lightweight purpose-driven Digital Shadows to avoid, for instance, expensive simulations and network congestion [22,23]. That is, Digital Shadows are meant to be reduced models or subsets of data for certain purposes. These models should be used to calculate much faster than simulations relevant aspects of a production step so that they are enablers of important aspects of the fourth industrial revolution as mass customization of products.

With exchanging data and Digital Shadows in the IoP we go further than just monitoring processes or controlling them. We collect data from the monitoring and aggregate it to build Digital Shadows but not only from our own processes but also from others in the WWL. That means cross-domain exchange of data which can not only improve the processes in ones own company. So we can make data in its data-silos in different companies more valuable as, for example, present traffic information makes the routes in Google Maps more valuable.

3   Exploring the Role of Digital Shadows

Above we presented a rough idea of what we understand about a Digital Shadow in the WWL within an IoP. In this section we want to elaborate more on the meaning of Digital Shadows and how we can generate and use them.

First, we will take a look on Digital Shadows from the perspective of data base systems. Then we present two use cases, where we show the usage of Digital Shadows. Finally, we present a concept for a first prototype of a node of the IoP dealing with the generation and usage of Digital Shadows.

3.1   From Database Views to Digital Shadows

Our approach to formalizing the Digital Shadow concept is inspired by the highly successful 50 year old concept of database views. A view is defined as a named query on a database which can be reused in other queries or applications in the

same way as a stored relation. In other words, views are Janus-faced objects which can be seen both as a mathematical model (query) and as condensed transformed data according to a specific user interest.

We are particularly interested in adapting the following roles of views to Digital Shadows:

(1) Views can be used for information hiding in conceptual schemas. In other words, data providers can employ views to determine which excerpt of their data consumers can see, and in which granularity. This makes cleverly designed views (Digital Shadows) a potentially valuable trade object and also helps privacy concerns. Research in this area dates back to the 1975 Ph.D. thesis of Mike Stonebraker on view unfolding. In the opposite direction, data mining and data-driven machine learning have been widely interpreted as detecting interesting queries/views from huge data sets, i.e. creating models from data.

(2) Conversely, recent heterogeneous data integration and exchange mechanisms make it possible for data consumers to map different data provider views to an integrated own perspective. Starting from pionieering work in IBM's CLIO project [6], recent research on mapping strategies for heterogeneous data exchange using ontologies [16] or tuple-generating dependencies across different data models [11,13] has enabled both semi-automatic generation of mappings between provider and consumer schemas, and the automated generation of code from such mappings. These recent results also enable completeness, consistency, and other data quality checks in data sharing settings – extremely important aspects that need to be linked to quality management in the production sector through the Digital Shadow concept.

(3) In many such scenarios, the Digital Shadows quickly become independent objects which must be used by the data consumers detached from their sources. The analogous research question how to maintain externally materialized views with minimal data transfer has been intensely studied in the late 1990s [30]. The equally relevant question how to answer questions only on materialized views, without access to sources has also led to useful algorithms such as MiniCon [24].

For the planned knowledge transfer from view theory to a future Digital Shadow theory, important differences must also be considered, most importantly the fact that most digital shadows will be views on processes, and that the models (analogy to the view definitions) are much more complex than simple database queries, reflecting decades on highly specialized engineering and mathematical research. In the IoP research, we have therefore decided to first get an intuition for realistic Digital Shadows relevance and challenges through a number of interdisciplinary case studies in which especially the combination of mathematical engineering theories with deep learning and related data mining techniques in Digital Shadows is explored.

3.2   Two Case Studies

We discuss here two highly relevant production use cases where Digital Shadows can improve processes drastically. The first is an injection moulding use case which considers the production of plastic pieces. The second is one of the most energy-intensive production steps worldwide, the hot rolling task within steel-based production where metal slaps are rolled to thinner plates. These use cases are very different. Not only the material is different but also the production process itself. The data from these processes are very different and so are the Digital Shadows.

In the following we describe the use cases in more detail and address the relation to Digital Shadows.

Injection Moulding. In injection moulding, an elaborate plastic piece is produced in one single complex process step using injection moulding machines. On the one hand, that makes it very efficient to produce a large number of plastic parts. On the other hand, there exists no mathematical model which describes the process as a whole. That complicates the initialization of a new process and also makes it difficult to create a Digital Shadow for the injection moulding process.

This process is executed on an injection moulding machine. A turning screw transports small plastic pellets through a long, horizontal, heated barrel. On its way through the barrel the plastic melts and is compressed in a nozzle towards the end of the barrel. For each plastic piece, the screw moves rapidly forward and forces the molten plastic through pipes into the cavity of the mold in the shape of the desired workpiece. After cooling down and hardening of the plastic, the mold opens and the workpiece is ejected with pins which are pulled out of the mold. After moving the screw back into its initial position, the next part can be produced. So, a substantial number of pieces can be produced in a short time depending on the properties of the material and the size of the cavity. There are many variations of that process, for instance, in some cases multiple pieces can be produced at once [12,27]. There exist mathematical models for some parts of that process but there are no closed-form solutions covering the whole process. Until today, the process parameters are often still determined by hand [7]. Instead of models, costly simulations such as FEM simulations, mentioned above, are used to support the parameterization. The calculation time of such simulations ranges from minutes up to hours, which is far from real-time. So it is not efficient to switch from one part to another as would be needed, for example, for mass customization. Instead, today it is more efficient to produce the same part as often as possible to save the time of reconfigure the parameters for another part.

In Fig. 1, our approach to generate Digital Shadows is depicted for predicting quality parameters of a plastic part for given parameter settings [19]. Given this Digital Shadow, it was possible to find parameters for the injection machine to attain the desired quality of the part within seconds. Because there was no closed-form mathematical models of the process which could be simplified to gain
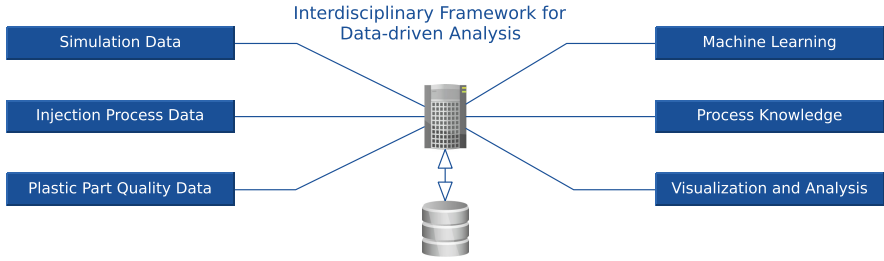
Fig. 1. Setup for building Digital Shadows with machine learning in the injection moulding use case.

functions representing a Digital Shadow, we used machine learning to learn these functions. Getting data from experiments takes time because the dimensions and the weight of a workpiece we needed for the quality parameters had to be measured by hand. So it was not possible to generate the amount of data which is usually needed for machine learning.

Nevertheless, it is possible to obtain Digital Shadows even from less data although data-driven machine learning approaches need a large amount of data. Experiments in our studies showed that it is possible to use the combination of simulation data and data from experiments to train Neural Networks so that data from simulations could compensate the reduced amount of data coming from the process itself. On the left in Fig. 1, is shown that for experiments in our studies two ways to get data were chosen: (a) manual experiments on the injection machine with measurements of the resulting product quality and (b) FEM simulations. The FEM simulation took about 10 min per part so that it was possible to obtain several hundred of data points. In the middle of the picture is the prototypical framework depicted we present later in this paper. It was used to store the data from the machines and the simulations. On the right-hand side in the picture is depicted that this data from the framework was used for the aforementioned machine learning. These networks learned for a range of process parameters which value a certain quality criterion, for instance weight, will be gained during the injection moulding process. Furthermore, on the right is depicted that knowledge of the process could be obtained from the framework and with visualization of the data it was possible for data analysts to select the data they needed for the machine learning approaches. For more details of the findings in this studies we refer to [19].

Hot Rolling. As Fig. 2 illustrates on the example of the production process of the so-called B column (the core stabilizing element in a car), hot rolling is an essential element in all steel-based production processes; moreover, it is one of the most energy-intensive production tasks world-wide (several percent of industrial energy usage), such that improvements here have a significant potential in reducing carbon footprint worldwide.
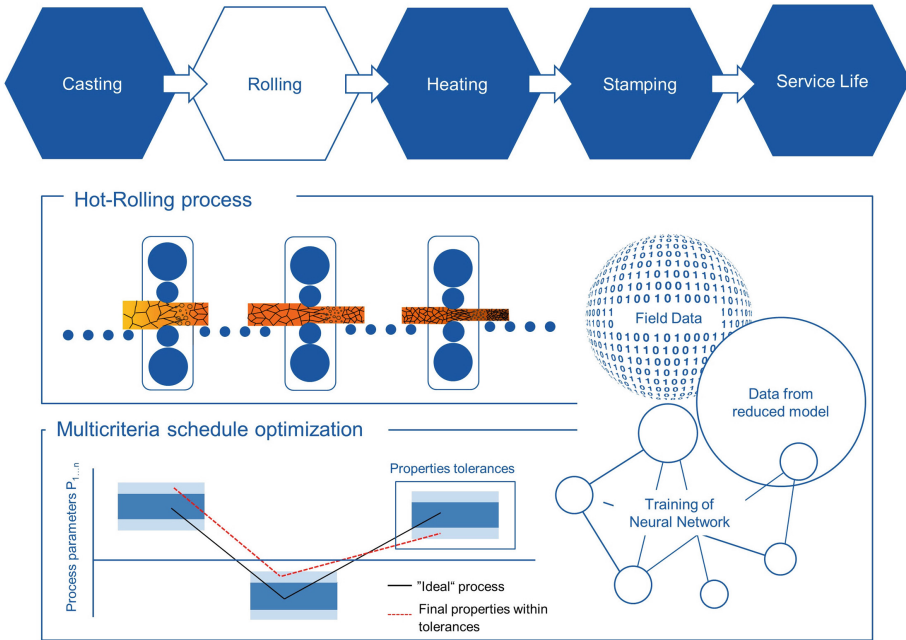
Fig. 2. Hot rolling within the steel production chain: schedule optimization by combining fast reduced models, field data and neural network learning in a two-layer Digital Shadow

In the hot rolling process, a hot slab of metal is pressed between a pair of rolls resulting in a reduction of the height and additional lengthening of the slab; a hot rolling mill usually comprises a series about 30 such pairs (the figure shows 3). Our study considered only reversing hot rolling mills [19], where the slab as a whole is rolled back and forth through a mill with only one pair of rolls. During one pass, with the height reduction also the microstructure within the material is changed. That microstructure influences the properties of the resulting product [15]. That is an intermediate product further processed in following production steps such as fine-blanking [1] or deep-drawing [5]. The latter is called stamping in some cases, as for the B column in Fig. 2. The core issue influencing the quality of the product and the energy consumption of the process is the combined design of the individual steps, and their overall scheduling.

In most current practice, experienced engineers manually design such a schedule, and evaluates it using FEM methods which take 30 min to 4 h in typical production processes; thus, to plan production for the next day, at most a couple of plans can be explored. In prior research, six different engineering groups from production and material science had developed reduced versions of the very complex mathematical models for the many aspects influencing the dimensions of the resulting slabs, its microstructure, the thermal condition, and energy impact

of this process. Taken together, the purely model-based Digital Shadow from the clever combination of these models reduces the evaluation time to about 50 ms with little quality loss over the 30–240 min FEM computations – a speedup over 100.000 [2,29].

Unfortunately, this still does not solve the problem as there is an exponential number of possible schedules to be evaluated. To turn the fast evaluation Shadow into one that can also actively search for fast, high-quality and energy-efficient rolling schedules, given a desired combination of target parameters, a second-level Digital Shadow has been constructed. In this schedule (shown in the lower right of the figure) the reduced model evaluations are embedded in the training of a Deep Neural Network, where the training data stem both from production process and experiment data, and artificially generated schedules gained by other permutations of those schedules. The trained Neural Network can then be used to directly recommend suitable inputs (schedules) for desired outputs [19]. Note that here, the purpose of the NN-based Digital Shadow is quite different from the injection moulding use case. Even withing the hot rolling use case, a Digital Shadow could have been constructed in a similar way as for injection moulding, but without the fast models, the learning process would have been rather meaningless.

In current work, we try to demonstrate that the potential of the multi-layer Digital Shadow for improving hot rolling processes goes even further. The microstructure in the material cannot be detected during the process, but the fast models are able to calculate the inner structure of the material. That information can be used to adapt later production steps of the resulting slab. But even during the process, if it is noticed that a certain process parameter is not as expected in the rolling schedule (e.g. humidity changes due to unexpected rainfall), the Digital Shadow could re-calculate the inner structure and thus avoid regarding the slab as scrap such that the whole energy-intensive process must be repeated [26]. The in-process resilience thus achieved would additionally allow more "courageous" energy-saving schedules than the present ones which have to be extremely careful with tolerances, as they must take all likely failures into account from the start.

3.3   Linking Processes, Models, and Machine Learning: An Experimental Infrastructure for the IoP

Loucopoulos et al. [17] present an early requirements engineering approach to transform existing requirements for traditional production systems to requirements of cyber physical production systems. The requirements mentioned there obviously overlap with our work. However, gaining purpose-driven Digital Shadows from data out of a global multi-site WWL and share them in the multi-disciplinary IoP demands new requirements for information systems. In this regard, our ongoing IoP research also profits from the multi-year requirements engineering effort on alliance-driven data platforms within the international Industrial Data Space initiative [21] which focuses on controlled data sharing in so-called alliance driven data platforms which do not have a keystone
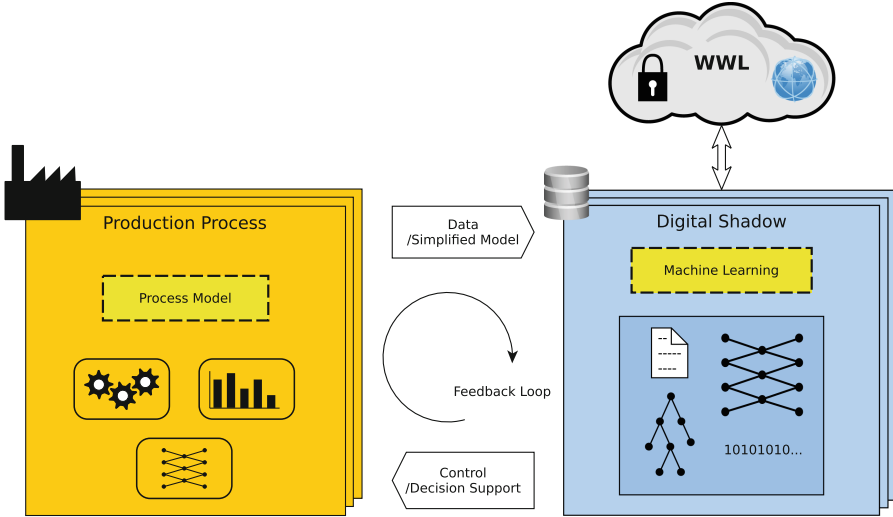
Fig. 3. A first concept for a node in the IoP. Digital Shadows should be exchanged securely in the WWL. The Digital Shadows are generated with data coming from different processes. A feedback loop improves Digital Shadows and processes on both sides. (Color figure online)

player – exactly the World Wide Lab setting we are envisioning for the IoP. We shall therefore not repeat this aspect here but focus on the specific aspects relevant also for the cases above.

In this section we discuss the requirements, possible solutions and a prototypical implementation meeting important requirements of the IoP. To present these diverse requirements of information systems in the IoP, in Fig. 3, we depict the concept for an exemplary node of the IoP in a diagram. On the left-hand side the orange boxes represent the production process side. On the right-hand side the blue boxes represent the storing side for data and Digital Shadows. Above the storing part we have the connection to the WWL.

The multiple orange boxes on the left-hand side, indicate that different production processes come, for example, from one or more production facilities. A production process might have already a model but does not necessarily need to have one, which is indicated by the dashed line around the yellow box. A model can also be gained by using machine learning as shown above in our injection moulding use case. The three rounded boxes in a production process box mean that on the production process side the data should represented in visualizations of the data, the charts on the right-side, and its processes, the gears on the left, for interdisciplinary usage and understanding about the meaning of the data. Especially in a cross-domain collaboration, it is important to understand where the data comes from. The visualization should also help to decide which data can be stored and what can be used to generate purpose-drive Digital Shadows. This is similar to other data discovery approaches used in machine learning or data

mining [10]. In addition, the visualization of Digital Shadows, the rounded box below, can give a better understanding of the process and can also be used for explaining decisions of an optimization system using a certain Digital Shadow.

The multiple blue boxes on the right should represent the different Digital Shadows for specific purposes stored in a database system. But to gain digital Shadows, the systems need to handle a large amount of data. But storing is only one aspect. The amount of data of a big production facility can be so high that, on the one hand, it is physically not possible to store all of it and, on the other hand, there might be no point in storing every single value of the data. However, machine learning methods need large amounts of data. Thus, we need systems which enable the analysis of sample data so that one can chose the appropriate data from a large data set for the machine learning algorithms. In addition, automatic aggregation can help to reduce the amount of data. The storage should also be able to handle data of different types. That is represented by the inner box which depicts documents, raw data and graphical models. The latter can be like Neural Networks or decision trees [28] generated by machine learning to learn models. These models should be stored for inspection or sharing. With reduced mathematical models, we do not need machine learning because we can use these models directly. Therefore, the yellow box for machine learning has a dashed line as border in the picture. Finally, in the IoP we need systems which provide Digital Shadows for application in production processes. So we think that the data and Digital Shadows should meet the FAIR principles [31] which state that data should be findable, accessible, interoperable and reusable. That encompasses, for instance, that data needs to have unique identifier, metadata describing the data and open protocols to get the data. Especially the latter is importing regarding standardized protocols so that not every machine speaks another language.

The arrow, from left to right in the middle, indicates that data and fast, simplified, mathematical models has to be transferred from the production processes to the database system to gain Digital Shadows. The lower arrow, from right to left, indicates that with Digital Shadows it should be possible to control production processes or give decision support. The latter can also incorporate other techniques from the field of artificial intelligence, for example knowledge-based approaches [3, 28], which is beyond the scope of this paper.

The lock in the WWL cloud at the top indicates that we need a secure exchange of Digital Shadows and data in the WWL. Especially for data from private companies it is crucial that only trusted partners should see the data. The globe on the right in the cloud indicates that we want to use common internet technology like HTTP in the WWL. The feedback loop in the middle indicates that Digital Shadows should be improved either with more data from different production processes or with data from other facilities in the WWL. Additionally, the production processes are improved by the Digital Shadows.
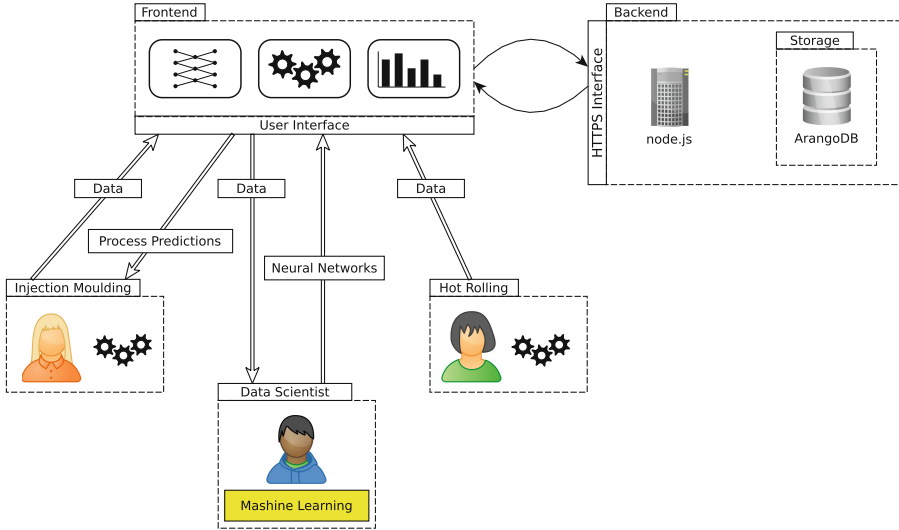
Fig. 4. Architecture of an early prototype of a node in the IoP. Users from different domains can upload their data and data scientists can apply machine learning on it. These models would be used to predict quality parameters of the injection moulding process, for instance.

In conclusion, we summarise the requirements for information systems in the IoP we discussed above:

1. We need data storage for storing a great number of multifarious data and models.
2. We need secure connections for transferring data from machines to a storage.
3. We need FAIR data [31].
4. We need techniques for automatic aggregating and reducing the data.
5. We need analysation methods for the data and the models.
6. The system needs to give feedback for decision support and controlling.
7. We need secure and fast protocols for the WWL.

A Prototypical Implementation. In an interdisciplinary subproject of our research cluster, together with data scientists, material scientists, mechanical engineers and experts in artificial intelligence, we investigated building and usage of Digital Shadows [19]. For that endeavor, we developed a prototype for a node in the IoP with a web interface working with the above described use cases of hot rolling and injection moulding. With this implementation we concentrated on meeting Requirements 1, 5, 6 from the list above. Due to the lack of a WWL we implemented only the parts below the cloud in Fig. 3 without a connection to the WWL. But, this prototype can be seen as a node in the IoP which could be connected to the WWL. For the security part of Requirement 7 we used HTTPS for the web UI.

In Fig. 4, the architecture of our prototype is depicted. It consists of a user-friendly UI in the frontend and a database and small applications on a node.js web-server in the back-end. In our project, the scientists of both use-cases could upload their data from experiments and simulation runs onto the platform. The data could be downloaded from other scientists for exploring Digital Shadows with machine learning techniques. With the results of the machine learning, we developed some applications on the platform.

In the upper right corner in Fig. 4, the backend of our prototype is depicted, where we used the schema-less database ArangoDB for storing. We chose ArangoDB to address Requirement 1 because it provides document, key/value and graph storing models and thus allows to store various sorts of data in the JSON format and is able to store huge amount of data very fast. The graph storing model was interesting for us because we wanted to store connective models such as Neural Networks representing the Digital Shadow. It is worth noting that although, we did not need to have FAIR data and, therefore, address Requirement 3 in our project it would be easy to realise it with ArangoDB and JSON. The database itself has unique identifiers and with JSON it is unproblematic to add metadata needed for FAIR.

In Fig. 4, in the upper left corner the frontend of our prototype is depicted which runs on the user's browser. As mentioned above it communicates via HTTPS with the backend to address security issues. Via the UI in the frontend, it is possible to upload data from machines and experiments which meets Requirement 2. The focus in our project was not on uploading the data directly from the machines into the database, but in principle it is possible. For the data upload, we used hand-written parsers for the different data formats of the raw data so that the users of the two use cases could upload their process data easily. In the future here one needs standardised formats so that no manual step is needed anymore.

To meet Requirement 5, one can filter, visualize and select the data for download in different formats. Because we expected confidential data from industrial partners, we restricted the view of particular data to certain registered persons only. Users can inspect the data, the models and the results of the models on the platform, which is depicted by the rounded boxes in the frontend box in Fig. 4. Among others, in the IoP, the Digital Shadow should be used to improve processes. Furthermore, the Digital Shadows should be improved in further experiments. Therefore, we visualized the learned Neural Networks so that the nodes and the weights of the connections can be inspected. That can give one interesting knowledge about the relevance of certain parameters if some connections in a Neural Network are less used then others. In the IoP a crucial aspect is to utilize the compiled Digital Shadows. Therefore, for the injection moulding use-case, we built an application to see the results of the learned Neural Networks. The user can choose two or three dimensional visualizations of the results of the Neural Network. It is possible to set the input parameters of the network with sliders. That way one can see by a range of input parameters which quality of the product is predicted. For utilizing the learned Neural Networks, we realized

a micro service in a Python program. In the future we want to setup autonomous agents as micro services which should act in the WWL to find answers similar the traffic information in Google Maps but for production scenarios.

4   Discussion and Conclusion

In this exploratory paper, we explore the potential of establishing collections of Digital Shadows as a complement to the much-discussed Digital Twin approach in Industry 4.0 and many other settings. From a theoretical point, we argue that the enormously successful view concept in databases can serve as the starting point for a theoretical foundation of Digital Shadows, as it constituted an early combination of model-based and data-based methods which is also the core of the Digital Shadow concept.

From an empirical point of view, we explore the relevance of the idea through two initial, but rather ambitious use cases in the important industrial fields of plastics and steel-based production; the preliminary results of both case studies which combined latest advances from mathematical model-based approaches with AI-based data-driven methods illustrate the enormous potential we can expect from further pursuing this avenue.

In both the theoretical and the use case discussion, we have pointed out the strong need for intensive further research in this area not just in the production engineering, but also in the IS engineering field.

References

1. Altan, T., Tekkaya, A.E.: Sheet Metal Forming: Processes and Applications. ASM International (2012)
2. Bambach, M., Seuren, S.: On instabilities of force and grain size predictions in the simulation of multi-pass hot rolling processes. J. Mater. Process. Technol. **216**, 95–113 (2015)
3. Brachman, R.J., Levesque, H.J.: Knowledge Representation and Reasoning. Morgan Kaufmann, Burlington (2004)
4. Ceretti, E., Lucchi, M., Altan, T.: FEM simulation of orthogonal cutting: serrated chip formation. J. Mater. Process. Technol. **95**(1–3), 17–26 (1999)
5. Colgan, M., Monaghan, J.: Deep drawing process: analysis and experiment. J. Mater. Process. Technol. **132**(1–3), 35–41 (2003)
6. Fagin, R., Kolaitis, P.G., Miller, R.J., Popa, L.: Data exchange: semantics and query answering. Theoret. Comput. Sci. **336**(1), 89–124 (2005)
7. Fei, N.C., Mehat, N.M., Kamaruddin, S.: Practical applications of Taguchi method for optimization of processing parameters for plastic injection moulding: a retrospective review. ISRN Ind. Eng. **2013**(2013)
8. Gantz, J., Reinsel, D.: The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. IDC Analyze the Future (2013)

9. Gasparenienene, L., Remeikiene, R.: Digital shadow economy: a critical review of the literature. Mediterr. J. Soc. Sci. **6**(6), 402–409 (2015)

10. Grinstein, G.G., Fayyad, U.M., Wierse, A.: Information Visualization in Data Mining and Knowledge Discovery. Morgan Kaufmann, Burlington (2002)

11. Hai, R., Quix, C., Wang, D.: Relaxed functional dependency discovery in heterogeneous data lakes. In: Laender, A.H.F., Pernici, B., Lim, E.-P., de Oliveira, J.P.M. (eds.) ER 2019. LNCS, vol. 11788, pp. 225–239. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33223-5_19

12. Kashyap, S., Datta, D.: Process parameter optimization of plastic injection molding: a review. Int. J. Plast. Technol. **19**(1), 1–18 (2015). https://doi.org/10.1007/s12588-015-9115-2

13. Kensche, D., Quix, C., Chatti, M.A., Jarke, M.: *GeRoMe*: a generic role based metamodel for model management. In: Spaccapietra, S., et al. (eds.) Journal on Data Semantics VIII. LNCS, vol. 4380, pp. 82–117. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-70664-9_4

14. Kritzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W.: Digital twin in manufacturing: a categorical literature review and classification. In: 16th IFAC Symposium on Information Control Problems in Manufacturing (2018). 51(11), 1016–1022

15. Lenard, J.G.: Primer on Flat Rolling. Newnes, Burlington (2013)

16. Lenzerini, M.: Direct and reverse rewriting in data interoperability. In: Giorgini, P., Weber, B. (eds.) CAiSE 2019. LNCS, vol. 11483, pp. 3–13. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-21290-2_1

17. Loucopoulos, P., Kavakli, E., Chechina, N.: Requirements engineering for cyber physical production systems. In: Giorgini, P., Weber, B. (eds.) CAiSE 2019. LNCS, vol. 11483, pp. 276–291. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-21290-2_18

18. Makinouchi, A., Teodosiu, C., Nakagawa, T.: Advance in FEM simulation and its related technologies in sheet metal forming. CIRP Ann. **47**(2), 641–649 (1998)

19. Meyes, R., et al.: Interdisciplinary data driven production process analysis for the internet of production. Procedia Manuf. **26**, 1065–1076 (2018)

20. Oezkaya, E., Biermann, D.: Segmented and mathematical model for 3D FEM tapping simulation to predict the relative torque before tool production. Int. J. Mech. Sci. **128**, 695–708 (2017)

21. Otto, B., Jarke, M.: Designing a multi-sided data platform: findings from the International Data Spaces case. Electron. Markets **29**(4), 561–580 (2019). https://doi.org/10.1007/s12525-019-00362-x

22. Pennekamp, J., et al.: Towards an infrastructure enabling the internet of production. In: 2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS), pp. 31–37. IEEE (2019)

23. Pennekamp, J., et al.: Dataflow challenges in an internet of production: a security & privacy perspective. In: Proceedings of the ACM Workshop on Cyber-Physical Systems Security & Privacy, pp. 27–38. ACM (2019)

24. Pottinger, R.A.: Processing queries and merging schemas in support of data integration. Ph.D. thesis, Citeseer (2004)

25. Qi, Q., Tao, F.: Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison. IEEE Access **6**, 3585–3593 (2018)

26. Reck, B.K., Graedel, T.E.: Challenges in metal recycling. Science **337**(6095), 690–695 (2012)

27. Rosato, D.V., Rosato, M.G.: Injection Molding Handbook. Springer Science & Business Media, New York (2012)

28. Russell, S.J., Norvig, P.: Artificial Intelligence: A Modern Approach. Pearson Education Limited, London (2010)
29. Seuren, S., Willkomm, J., Bücker, H., Bambach, M., Hirt, G.: Sensitivity analysis of a force and microstructure model for plate rolling. In: Metal Forming 2012, pp. 91–94. Wiley-VCH (2012)
30. Staudt, M., Jarke, M.: Incremental maintenance of externally materialized views. Citeseer (1996)
31. Wilkinson, M.D., et al.: The FAIR guiding principles for scientific data management and stewardship. Sci. Data **3** (2016)
32. Zienkiewicz, O.C., Taylor, R.L., Zhu, J.Z.: The Finite Element Method: Its Basis and Fundamentals. Elsevier, Amsterdam (2005)