Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 277 (2026) 3521–3530

Procedia Computer Science

www.elsevier.com/locate/procedia

7th International Conference on Industry of the Future and Smart Manufacturing
(former International Conference on Industry 4.0 and Smart Manufacturing)

# Data Platforms in Biointelligent Value Creation Systems - Towards a Biointelligence Metaverse

Arber Shoshi [a,b,c]*, Peter Reimann [a,d], Daniel Schel [b], Thomas Bauernhansl [b,c], and Robert Miehe [b,c]

[a]*Graduate School of Excellence advanced Manufacturing Engineering (GSaME), University of Stuttgart, Nobelstraße 12, 70569 Stuttgart, Germany*
[b]*Fraunhofer Institute for Manufacturing Engineering and Automation (IPA), Nobelstraße 12, 70569 Stuttgart, Germany*
[c]*Institute of Industrial Manufacturing and Management (IFF), University of Stuttgart, Allmandring 35, 70569 Stuttgart, Germany*
[d]*Institute for Program Structures and Data Organization (IPD), Karlsruhe Institute of Technology (KIT), Am Fasanengarten 5, 76131 Karlsruhe, Germany*

## Abstract

The biological transformation of industry introduces new demands for data management in production systems. Biointelligent value creation (BVC) systems integrate biological, technical, and digital components within decentralized, modular, and adaptive production networks. In these systems, data platforms are essential for enabling integration, coordination, and data-driven control across heterogeneous actors and domains. However, a systematic understanding of which data platforms exist, how they support biointelligent value creation, to what extent they fulfil their requirements, and what is currently lacking. This paper addresses this gap by conducting a systematic literature review of 40 peer-reviewed publications, identifying and categorizing key technological approaches, infrastructure concepts, and application domains. A functional framework is developed to assess platform capabilities across core, additional, and future-oriented requirements—from data integration and automation to semantic processing, predictive modeling, and real-time analytics. The findings show that while current platforms address selected requirements, no existing system fulfils the comprehensive demands of BVC. Major challenges include limited interoperability across biological and technical domains, insufficient support for non-expert users, and the lack of standardized, scalable architectures. The paper concludes with design implications and proposes the vision of a Biointelligence Metaverse – a modular, semantically enriched ecosystem combining data lake architectures, AI-based analytics, and IoT infrastructures to enable sustainable and collaborative value creation.

---

* Corresponding author. Tel.: +49 711 970 16950; fax: +49 711 970 1220.
*E-mail address:* arber.shoshi@gsame.uni-stuttgart.de

*Keywords:* Biointelligent Value Creation Systems, Data Platforms, Biointelligence Metaverse, sustainable production

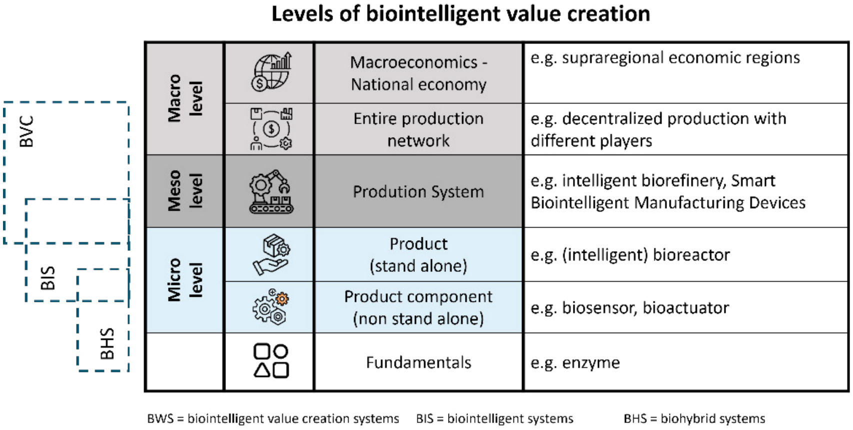
---

## 1. Introduction

The biological transformation of industry marks a fundamental shift in production and value creation by embedding biological principles into industrial systems for more sustainable and efficient value creation [1]. Within this context, biointelligence has emerged as a key approach. It is the result of the convergence of biology, hardware and software [2]. Biointelligent Systems (BIS) integrate biological components (e.g., cells, enzymes) with technical infrastructure (e.g., reactors, sensors) and digital controls to form adaptive, self-optimizing systems. Applications include cell-based manufacturing, 3D-printed food, and decentralized therapeutic production [3, 4].

When scaled, BIS form Biointelligent Value Creation (BVC) Systems – modular, decentralized production or service networks that utilize living biological resources while involving both industrial and non-industrial actors. BVC aims to reduce waste, minimize transport, and support circular processes [5, 6]. A defining feature is the inclusion of diverse stakeholders – companies, small enterprises, communities, and individuals – broadening access to production capabilities.

One of the most fundamental issues is the unpredictability and variability of biological inputs such as cells, bacteria, and biomass. Cells or microorganisms react non-deterministically, and real-time monitoring is often limited due to inadequate sensor technologies. These uncertainties complicate process control and modeling. Moreover, BVC is socio-technically heterogeneous: participants differ in expertise, infrastructure, and digital maturity. While some require high-end data management systems, others rely on intuitive interfaces. This results in dynamic, trust-dependent environments with increased coordination demands. Figure 1 illustrates how BVC span micro to macro levels of production, embedding biointelligent capabilities across layers.



**Levels of biointelligent value creation**

| | | |
|---|---|---|
| Macro level | Macroeconomics - National economy | e.g. supraregional economic regions |
| | Entire production network | e.g. decentralized production with different players |
| Meso level | Prodution System | e.g. intelligent biorefinery, Smart Biointelligent Manufacturing Devices |
| Micro level | Product (stand alone) | e.g. (intelligent) bioreactor |
| | Product component (non stand alone) | e.g. biosensor, bioactuator |
| | Fundamentals | e.g. enzyme |

BWS = biointelligent value creation systems  BIS = biointelligent systems  BHS = biohybrid systems

Fig. 1. Levels of biointelligent value creation and categorization of biohybrid system (BHS), BIS and BVC.

A representative example is the vision of decentralized CAR T cell therapy mini-factories [7], where patient-derived cells are processed on-site. Here, biological materials, digital tools, and technical infrastructure converge across clinical staff, technicians, and suppliers [8]. Such adaptive, localized systems require robust information logistics (IL) ensuring that the right information reaches the right stakeholders at the right time, in the right format and quality [9, 10].

Information logistics systems (ILS) operate alongside physical supply chains, handling data collection, storage, processing, and dissemination – often in the opposite direction of material flows. In BVC, the decentralized and interdisciplinary structure increases IL complexity. Seamless data exchange is essential for coordinating biological, technical, and human processes, ensuring traceability and shared interpretation [11]. Despite increasing interest in biointelligent production, there is little research on the specific data platforms used and to support IL in BVC. This paper contributes to the emerging discourse on biointelligent value creation by systematically reviewing and evaluating data platforms that aim to support BVC or comparable production systems. It provides three key contributions: (1) a synthesis of 40 peer-reviewed publications to map technological and conceptual developments; (2) a functional framework that classifies platform requirements into core, additional, and future-oriented capabilities specific to BVC; and (3) an assessment of how well current platforms meet these requirements, identifying critical gaps-particularly in interoperability, semantic processing, and inclusivity. To address those gaps, the paper introduces the concept of a Biointelligence Metaverse – a novel, AI-supported data and automation infrastructure designed to accelerate the implementation of BVC by enabling inclusive, modular, and real-time coordination across domains.

The remainder of the paper is structured as follows: Section 2 presents the methodology of the systematic literature review. Section 3 summarizes the results, including a functional analysis of existing data platforms. Section 4 discusses the findings and proposes the Biointelligence Metaverse as a future-oriented platform concept. Section 5 concludes the paper and outlines directions for future research.

## 2. Methodology

This study conducts a systematic literature review to provide a structured overview of current research on data platforms in biointelligent value creation. The 40 selected studies were systematically examined regarding methods, results, and discussions. Extracted information was consolidated to answer the research question and synthesize the current state of knowledge. The review follows the PRISMA protocol to ensure transparency, standardization, and reproducibility [12]. The central research question guiding the study was: *Which data platforms are currently discussed in the context of biointelligent value creation, and to what extent do they fulfill the requirements of biointelligent value creation systems?* In the context of this paper, data platforms are integrated digital infrastructures that enable the ingestion, storage, management, analysis and exploitation of heterogeneous data from biological, technical, and organizational domains.

Table 1. Publication databases and search string.

| Publication data bases | Search String |
|---|---|
| Scopus, Web of Knowledge | TITLE-ABS-KEY ("data storage architecture" OR "data management" OR "data platform" OR "data lake" OR "data warehouse" OR "data repository " OR "information archive" OR "data organization") AND TITLE-ABS-KEY ("bio") AND TITLEABS-KEY ("production" OR "manufacturing" OR "production AND process" OR "industrial AND production" OR "value AND creation" OR "production AND systems" OR "manufacturing AND systems" OR "industrial AND systems" OR "production AND networks" OR "value AND creation AND systems" OR "manufacturing" OR "non-expert" OR "on-demand" OR "modular production") |

Figure 2 illustrates the selection and review process. A keyword-based search was conducted in Scopus and Web of Science, targeting publications that explicitly address data platforms in the context of biointelligent production or value creation (see Table 1).

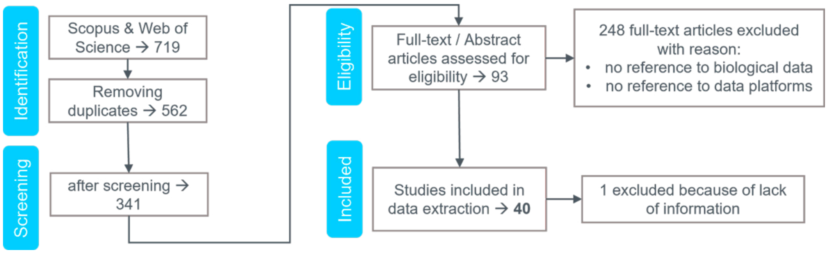


Fig. 2. Flowchart of the systematic literature review using the PRISMA protocol for studies on data platforms in BVC.

The initial search yielded 719 results; after removing duplicates, 562 records remained. Title and abstract screening reduced this number to 341 potentially relevant studies, since the other removed studies had no reference to BVCs. A detailed eligibility assessment based on full texts and abstracts led to 40 publications meeting the inclusion criteria, i.e., discussing concrete aspects of biointelligent value creation (see Section 1). Of these, 40 were included in the final analysis; 52 were excluded due to insufficient relevance, primarily lacking a biological data link and one for inadequate methodological clarity.

## 3. Review Results: Data Platforms for Biointelligent Value Creation

In this review, 40 publications on data platforms in BVC were analyzed, all dealing with integrating biological and technical components. Figure 3 shows the platforms analyzed by infrastructure type (left) and application context (right). Efficient value creation" refers to platforms focusing on improving resource efficiency and circular economy through optimized data management. Half of the studies focus on circular models; 65% examine decentralized or modular systems that coordinate distributed actors. On-demand production is only marginally addressed (5%), and no study explicitly addresses non-expert users, defined here as those without advanced BIS or data management skills. A major gap remains in platforms integrating biological, technical, and organizational data into intuitive, user-friendly interfaces. This is key to enabling adaptive, comprehensive production models.

The results are divided into two parts: (1) functional requirements and platform capabilities and (2) use case-specific applications in different BVC configurations.

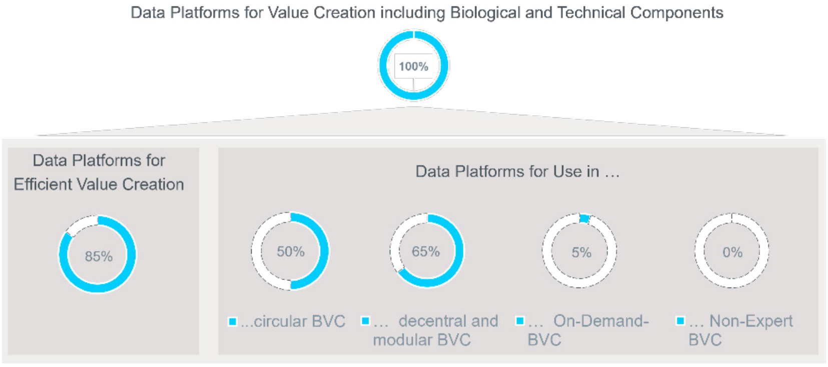


Fig.3 Distribution of reviewed publications by BVC application focus. The chart shows how many studies address efficiency, circularity, modularity, on-demand use, and non-expert accessibility.

### *3.1. Functional Roles of Data Platforms in BVC*

BVC systems tightly integrate biological, technical, and organizational domains, demanding data platforms that surpass conventional industrial systems in functionality, flexibility, and adaptability. Central to this is managing heterogeneous data across various fields such as biomedical research [13–17], plant pathology [18], biofuel production [15, 19], and microalgae cultivation [20, 21]. Three core requirements are consistently emphasized in the literature: Interoperability and standardization, semantic modeling and adaptive analytics.

Interoperability remains a primary challenge due to the heterogeneity of biological and technical data. Numerous studies propose domain-specific platforms [13, 18–20, 22, 23], often using cloud infrastructures for scalability [13, 14, 17, 24]. Standardized data formats and metadata schemes are essential [14–16, 18, 25, 26], with ontologies like LOINC supporting consistent representation [16]. Provenance tracking and data quality assurance enhance transparency [14, 16, 25].

To address data heterogeneity, architectural solutions such as lakehouses [14], data lakelands [24, 27] and data lakes [28] are employed. These support raw data storage and flexible querying. The FAIR principles-findability, accessibility, interoperability, and reusability-are frequently cited as foundational guidelines [18, 28, 29]. These principles serve as a foundational framework for developing robust data infrastructures.

Integrated platforms are applied across various domains, from omics data analysis [30] and short read mapping [31] to bioprocess optimization [20]. Managing unstructured data (e.g., lab notes, clinical documentation) remains a challenge [18, 24, 28, 32]. Platforms like Kibio and KibioR facilitate large-scale biological data integration [29] , and targeted applications such as bioenergy field data collection are discussed in [27].

Some studies demonstrate real-world integration of biological and technical data. For instance, [33] combines biobank and wearable sensor data with electronic health records using HL7 FHIR and NLP. Lean Digital Twin architectures are applied to biorefineries [34] and regenerative bone therapy [35], incorporating additive manufacturing and in silico modeling.

Sector-specific platforms include biosensor-based systems in precision agriculture [36], phenotyping tools for crop breeding [37], livestock monitoring [38], and mobile diagnostics for uric acid detection [39]. Further applications span biogas tracking [40], biomass logistics [41], and pharmaceutical manufacturing [45]. Despite domain diversity, common challenges persist: lack of interoperability [22, 42], insufficient standardization [22, 33–35], and limited scalability. Proposed solutions include integrated platforms [35, 36], semantic models [22, 43], and machine learning for adaptive control [36]. Cloud computing remains a key enabler [37, 39, 40, 44], while the FAIR principles continue to serve as a normative reference [44].

## *3.2. Platforms for Data Utilization in Different BVC Aspects*

BVCs extend BIS by enabling interaction between multiple systems in modular, decentralized networks (see Figure 1). They follow principles like circularity, modularity, decentralization, demand-responsiveness, and inclusivity for non-expert users. Data platforms are essential to realize these principles by coordinating diverse stakeholders, managing bio-technical processes, and optimizing resource use through adaptive control.

### *3.2.1. Data Platforms for Circular Value Creation Systems*

In BVC systems designed around circular economy principles, data platforms are critical in optimizing resource loops and enhancing traceability. However, the reviewed literature reveals a significant gap in platforms explicitly tailored to circular value creation. While some systems address relevant industrial processes, they often do not contextualize their functionality within a broader circular framework [15, 21]. For instance, SABANA [21] is an IoT-based platform for algae cultivation, and Super-O [15] supports biofuel optimization, yet neither system explicitly connects with circularity principles.

In the bioenergy sector, spatial data for biomass sourcing [27] and GIS-based logistics modeling [41] are explored, whereas Buyel et al. [16] highlight integration challenges in biopharma. Technologies such as blockchain are used to enhance transparency and accountability, as seen in biofuel tracking systems [45] and cold chain monitoring [46]. In agriculture, cloud-based databases improve resilience [47], and participatory monitoring approaches support decentralized governance [48]. Nevertheless, only a few contributions, such as [43], address circularity at the platform level, signifying a clear research gap.

### *3.2.2. Platforms in Decentralized and Modular Production Networks*

The literature reflects growing attention to decentralized and modular production as central tenets of BVC. These systems require data platforms capable of supporting distributed operations, real-time coordination, and modular dataflows. Examples include bioinformatics frameworks with modular designs [49], IoT-based systems for algae production [21], and SERS sensor data platforms for automated quality control [23]. Kadi4Mat [50] serves as a rare example of a platform explicitly designed for decentralized bioprinting.

The challenges in managing data across transnational networks are illustrated in case studies such as IMPC and IKMC [26]. Key concerns include standardization, funding sustainability, and governance. Some systems, like ProBioRefine [15], manage diverse feedstocks and processing routes, hinting at latent modularity. Tools like BETY-db [24] and Kibio [29] exhibit potential for decentralized applications but lack explicit integration frameworks.

In general, the literature emphasizes the need for interoperable and sustainable data infrastructures [22]. Cloud-based and containerized workflows [44], Lean Digital Twin models [34], and RDF databases such as SBOL Stack [25] are examples of supporting technologies.

### *3.2.3. Platforms for Data Utilization in Dynamic, Demand-Oriented Value Creation*

Dynamic and demand-oriented BVC systems require high-performance platforms capable of real-time data processing and adaptive decision-making. Several contributions propose scalable architectures such as lakehouses [18], distributed systems [21, 32], and HPC-enabled platforms [23, 31]. Applications range from cohort discovery in biomedical research [51] to bioprocess optimization [15] and bioprinting [50]. Data quality and metadata management emerge as critical success factors. Robust curation, ontological metadata, and privacy-preserving data sharing methods are emphasized [24, 26, 49]. Platforms like Kadi4Mat support sensitive data handling through techniques such as double-blind testing [50]. Sustainability concerns are also addressed, with emphasis on funding models and public infrastructure [26]. Cloud computing is widely employed to support scalability, as seen in platforms operating on Azure [30], GCP [44], and AWS [40]. Domain-specific applications include DSS tools in agriculture [36], digital twins in biorefineries [34], and microbial databases [42]. Challenges such as data governance, standard inconsistencies, and platform interoperability are recurrent themes [22, 33].

### *3.2.4. Platforms for Data Utilization in Non-Expert Production Models*

Although user-friendly platforms are essential to make biointelligent production accessible to non-experts, the literature reveals a clear research gap: no study explicitly addresses platform design for non-specialist users in BVC contexts. Insights from related domains suggest that inclusive, low-threshold platforms could significantly broaden participation and enhance the accessibility of BVC systems.

## 4. Discussion

### *4.1. Alignment of Platform Capabilities with BVC Characteristics*

The literature reveals a fragmented platform landscape with varying degrees of coverage, architectural coherence, and alignment to BVC-specific needs. These findings are synthesized across three dimensions (Figure 4): (1) types of decentralized BVC systems, (2) application domains for data platforms, and (3) functional platform requirements. These requirements are categorized into core, additional, and future-oriented functions.

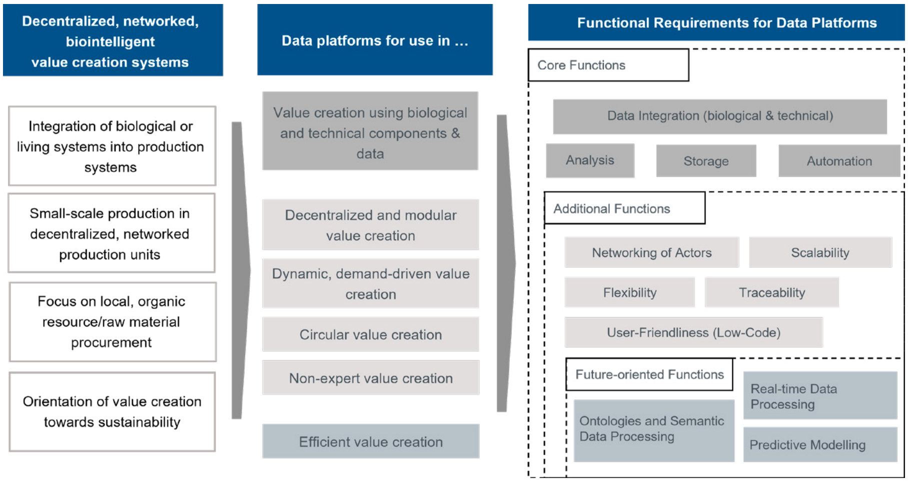


Fig. 4. Categorization of functional requirements for data platforms in BVC, grouped into core functions, additional capabilities, and future-oriented functionalities.

Core functions represent the baseline: integrating data, storing it, analyzing it and automating it. Additional functions (e.g. scalability, traceability) are critical in distributed BVC networks. Future-oriented functions (e.g. real-time processing, semantic modelling) enhance system adaptability and intelligence once operational stability is achieved. Importantly, these functional categories are not exclusive to BVC but reflect general requirements for robust platform design. Their relevance intensifies in BVC due to increased heterogeneity and complexity. Table 2 synthesizes the capabilities of reviewed platform types by mapping them to three categories of requirements and assigning coverage levels based on treatment in the literature. The ratings reflect an evaluative assessment focusing on functionality or conceptual relevance. Technologies such as data lakes or semantic models may support multiple purposes; however, Table 2 only lists technologies where there is clear evidence of specific application to ensure methodological transparency. While core functions such as data integration and automation are well supported, additional and future functions remain underdeveloped. User-friendliness, flexibility, inclusivity and modularity are rarely implemented, and semantic models and predictive capabilities are absent, limiting customization.

Table 2. Evaluation of existing data platforms based on the functional requirements of BVC to determine the extent to which a requirement is met by existing data platforms. Functional requirements were rated from 0 (not addressed) to 5 (fully addressed) based on the information provided in the publications, considering reported features and application examples.

| | Requirement / Function | Rating (0–5) | Applied Technologies – Data Platforms | Sources |
|---|---|---|---|---|
| Core functions | Data Integration (biological, technical data) | 4 | FIWARE, LIMS LabWare, SABANA Cloud, Web of Microbes (WoM), KibioR | [21, 29, 42, 49] |
| | Data storage | 4 | Azure Data Lake, SciDB, U-SQL, Kadi4Mat, Google Cloud, Kibio, AagingBase, @Web, Bio2RDF, Data Lakelands, WoM, de.NBI, GIS-based databases, MongoDB | [17, 18, 24, 28–30, 33, 35, 37, 41, 42, 47–51] |
| | Data analysis | 3 | PostgreSQL, SciDB, R Shiny, CogStack platform | [13, 19, 23, 24, 33, 44, 47, 49] |
| | Automation | 4 | FIWARE Orion Context Broker, Apache Spark | [13, 20, 22] |
| Additional Functions | Flexibility | 2 | Genome and proteome databases, blockchain systems, SQL databases, Powersim, OLAP, data warehouses, biosensor databases | [16, 22, 36] |
| | Traceability | 2 | Hyperledger Fabric, IoT gateways, blockchain systems | [45, 46] |
| | Scalability | 2 | Data lakehouse, data lakes, ELIXIR, DataCafe | [24, 26, 32, 42] |
| | Actor Networking | 3 | Data lakehouse, Data Lakelands, ELIXIR, DataCafe | [24, 26, 31, 32, 42] |
| | User-Friendliness (Low-Code) | 0 | - | - |
| Future-Oriented Functions | Ontologies and Semantic Data Processing | 2 | BioPAX, EBI RDF, ELIXIR, Natural Language Processing Platform, @Web, ProBioRefine | [24, 28, 33, 42, 43] |
| | Real-Time Data Processing | 2 | MinIO, Apache Kafka, Node-Red | [46] |
| | Predictive Modelling | 2 | Amazon Web Services (AWS), DynamoDB | [25] |

This shows the difference between technology focus and what's needed at the system level. Platforms mostly focus on efficiency and infrastructure but ignore user-orientation, integration of meaning and being adaptive, which is vital for scalable and strong BVC systems. Functions for coordinating people in a decentralized way, such as interoperability, tracking and modular control, are not dealt with equally.

### *4.2. Implications for Platform Design in Biointelligent Systems*

For core requirements such as data integration and storage, many platforms-especially data lakes [18, 28], lakehouse architectures [18], and RDF-based systems [25, 26]-receive high coverage ratings (4). These technologies have been implemented in industrial and biomedical applications, enabling scalable ingestion and cross-format data consolidation. Automation is also widely supported through machine learning and workflow orchestration components, as discussed in platforms such as Kibio and Kadi4Mat [29, 50]. Additional requirements such as user-friendliness and modularity are less mature. Few contributions explicitly address non-expert usability, such as low-code environments, and most platforms lack modular plug-in architectures. The functional gap is due to a lack of design focus on inclusive and flexible platform use, not technical infeasibility [50].

Future-oriented requirements, including semantic reasoning, real-time coordination and predictive analytics, are the least covered. While some studies refer to ontologies or digital twins [34], the integration of semantic models into operational workflows remains conceptual. Predictive modeling is mentioned in the context of bioprocess optimization [20, 31], but is rarely embedded in the platform architecture in a generalizable way.

### *4.3. Biointelligence Metaverse*

Current data platforms do not meet the diverse and complex requirements of BVC. A novel digital infrastructure that supports existing industrial processes is needed to address these shortcomings. The biointelligence metaverse (BM) is a forward-looking framework for developing such systems. BM is an AI-supported, semantically structured data and automation environment that makes biological systems digitally accessible, modifiable and controllable down to the cellular level. It builds on functional requirements, including modularity, interoperability, traceability and support for dynamic biological variability. The metaverse combines real-time data integration, transfer learning, virtual design environments and digital process modelling.

A core objective is to democratize access to biological system design. This includes enabling non-expert users to explore, configure, and adapt biological components for specific needs-without requiring deep domain knowledge. Through intuitive interfaces, embedded simulation, and guided decision-support tools, the BM could make complex bio-design tasks operable by actors from healthcare, agriculture, municipalities, and SMEs. Use cases range from AI-supported process optimization and adaptive automation in cultivation systems to simulation-guided scaling of biotechnological production. The concept supports modular digital twins and domain-specific extensions, ensuring technical depth and sectoral relevance. Unlike existing solutions, the BM is not a closed system, but an open, interoperable infrastructure designed to evolve. Ultimately, the goal is to create a robust foundation for the next generation of biointelligent platforms-capable of enabling inclusive, adaptive, and sustainable value creation. As such, the BM outlines a research and innovation trajectory rooted in the challenges identified throughout this study.

In this context, the FabOS reference architecture is a model for building modular, open, AI-integrated platform ecosystems. Its principles offer key design impulses for the envisioned BM and underline the importance of cross-domain platform standards in industrial applications [52].

## 5. Conclusion

The present study shows that although existing data platforms address selected functional aspects of BVC, none of the solutions examined fully meet the identified requirements. Basic functionalities such as data integration, storage and automation are partially implemented, but there are significant deficits in of interoperability, semantic data structuring, real-time processing and usability for non-specialized user groups. The requirements identified can be divided into core, additional and future functionalities. While basic capabilities can be found in several systems, more advanced functions, such as traceability, actor coordination and adaptive control, are still underdeveloped. In particular, the integration of semantic models, predictive analytics and dynamic system control is only partially realized. One limitation of this survey is that it focuses exclusively on peer-reviewed academic publications, which may result in industrial developments and proprietary solutions not being considered. In addition, the functional analysis is based on reported system characteristics. It does not include empirical testing or practical validation, which limits the assessment of performance and user experience in operational environments.

Further research should develop scalable architectures that enable semantically integrated, cross-domain dataflows—combining data lakes, AI, digital twins, and IoT for adaptive information logistics. In addition, the environmental and social impact of decentralized BVC architectures, especially regarding their potential contribution to sustainable production, must be empirically investigated. Finally, future work should explore BM as a unifying framework that integrates platform technologies, intelligent analytics and user-centered interaction to enable integrative and collaborative biointelligent value creation.

## Acknowledgements

This work was supported by the Landesministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg (Ministry of Science, Research and the Arts of the State of Baden-Wurttemberg) within the Nachhaltigkeitsförderung (sustainability support) of the projects of the Exzellenzinitiative II.

## References

[1] Miehe, R., Bauernhansl, T., Schwarz, O., Traube, A. et al. The biological transformation of the manufacturing industry – envisioning biointelligent value adding 72, p. 739.

[2] Miehe, R., Horbelt, J., Baumgarten, Y., Bauernhansl, T. Basic considerations for a digital twin of biointelligent systems: Applying technical design patterns to biological systems 31, p. 548.

[3] Miehe, R., Full, J., Scholz, P., Demmer, A. et al. The Biological Transformation of Industrial Manufacturing-Future Fields of Action in Bioinspired and Bio-based Production Technologies and Organization 39, p. 737.

[4] Shoshi, A., Xia, Y., Fieschi, A., Baumgarten, Y. et al., 2025. An Analysis of Monitoring Solutions for CAR T Cell Production. Healthc Technol Lett 12, e70012.

[5] Full, J., Shoshi, A., Gamero, E., Baumgarten, Y. et al. Biointelligent Waste-to-X systems: A novel concept for sustainable, decentralized and interconnected value creation 116, p. 576.

[6] Miehe, R., Waltersmann, L., Sauer, A., Bauernhansl, T. Sustainable production and the role of digital twins–Basic reflections and perspectives 3, e10078.

[7] Fraunhofer-Institut für Produktionsanlagen und Konstruktionstechnik. Solid CAR-T: Modular mini-factories for autonomous CAR-T cell production. https://www.ipa.fraunhofer.de/en/expertise/laboratory-automation-and-biomanufacturing-engineering/adapt/solid-cart.html. Accessed 4 May 2025.

[8] Shoshi, A., Xia, Y., Fieschi, A., Ackermann, T. et al., 2024. A Flexible Digital Twin Framework for ATMP Production – Towards an efficient CAR T Cell Manufacturing 125, p. 124.

[9] Hartleif, S., Bauernhansl, T., Erlach, K. Schlanke Informationslogistik: Konzept für ein nachfragegesteuertes Informationslogistiksystem 33.

[10] Krcmar, H. Einführung in das Informationsmanagement, 2nd edn. Springer Gabler.

[11] Miehe, R., Buckreus, L., Kiemel, S., Sauer, A. et al. A Conceptual Framework for Biointelligent Production—Calling for Systemic Life Cycle Thinking in Cellular Units 3, p. 844.

[12] Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G. et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement 339, b2535-b2535.

[13] Xu, L., Qian, L., Chang, Z., Wu, Z. The Bidirectional Data Flow Based On The Data-Lake, in 2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), IEEE, p. 971.

[14] Batko, K., Ślęzak, A., 2022. The use of Big Data Analytics in healthcare. J Big Data 9, p. 3.

[15] Bertran, M.-O., Orsi, A., Manenti, F., Woodley, J.M. et al. Synthesis of Sustainable Biofuel Production Processes: A Generic Methodology for Superstructure Optimization and Data Management, in Advances in Energy Systems Engineering, Springer International Publishing, p. 651.

[16] Buyel, J.F. Towards a seamless product and process development workflow for recombinant proteins produced by plant molecular farming 75, p. 108403.

[17] Beavis, W.D. Architectures for Integration of Data and Applications: Lessons from Integration Projects, in Genome Exploitation, Springer US, p. 31.

[18] Begoli, E., Goethert, I., Knight, K. A Lakehouse Architecture for the Management and Analysis of Heterogeneous Data for Biomedical Research and Mega-biobanks, in 2021 IEEE International Conference on Big Data (Big Data), IEEE, p. 4643.

[19] Charaniya, S., Le, H., Rangwala, H., Mills, K. et al. Mining manufacturing data for discovery of high productivity process characteristics 147, p. 186.

[20] Janzen, N.H., Striedner, G., Jarmer, J., Voigtmann, M. et al. Implementation of a Fully Automated Microbial Cultivation Platform for Strain and Process Screening 14, p. 1800625.

[21] Munoz, M., Guzman, J.L., Torres, M., Acien, F.G. An IoT Platform for Data Management in an Industrial-Scale Microalgae Cultivation Plant 10, p. 127128.

[22] Rodríguez-García, E.E., Soto-Mendoza, A., Guajardo, D. Serialization and Data Management of Bioreactors through Digitization: Bioprocessing 4.0 a Systematic Review, in 2023 Portland International Conference on Management of Engineering and Technology (PICMET), IEEE, p. 1.

[23] Zuo, Y., Vernica, R., Lei, Y., Barcelo, S. et al. A Big Data Platform for Surface Enhanced Raman Spectroscopy Data with an Application on Image-Based Sensor Quality Control, in 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), IEEE, p. 463.

[24] Eder, J., Shekhovtsov, V.A. Data quality for federated medical data lakes 17, p. 407.

[25] Skelton, D.J., Hallinan, J.S., Park, S., Wipat, A. Computational intelligence for metabolic pathway design: Application to the pentose phosphate pathway, in 2016 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), IEEE, p. 1.

[26] Mishra, A., Schofield, P.N., Bubela, T.M. Sustaining large-scale infrastructure to promote pre-competitive biomedical research: lessons from mouse genomics 33, p. 280.

[27] Surendran Nair, S., Kang, S., Zhang, X., Miguez, F.E. et al. Bioenergy crop models: descriptions, data requirements, and future challenges 4, p. 620.

[28] Che, H., Duan, Y. On the Logical Design of a Prototypical Data Lake System for Biological Resources 8, p. 553904.

[29] Ongaro-Carcy, R., Scott-Boyer, M.-P., Dessemond, A., Belleau, F. et al. KibioR & Kibio: a new architecture for next-generation data querying and sharing in big biology 37, p. 2706.

[30] Mrozek, D., Dąbek, T., Małysiak-Mrozek, B. Scalable Extraction of Big Macromolecular Data in Azure Data Lake Environment 24, p. 179.

[31] Natarajan, S., KrishnaKumar, N., Pal, D., Nandy, S.K. ReneGENE-GI: Empowering Precision Genomics with FPGAs on HPCs, in Applied Reconfigurable Computing. Architectures, Tools, and Applications, Springer International Publishing, p. 178.

[32] Kathiravelu, P., Sharma, A. A Dynamic Data Warehousing Platform for Creating and Accessing Biomedical Data Lakes, in Data Management and Analytics for Medicine and Healthcare, Springer International Publishing, p. 101.

[33] Zhang, J., Symons, J., Agapow, P., Teo, J.T. et al. Best practices in the real-world data life cycle 1, e0000003.

[34] Gamero, E., Shoshi, A., Full, J., Sauer, A. et al. Data Management in Biorefineries: Conceptual Thoughts on Lean Digital Twinning 125, p. 48.

[35] Poh, P.S., Lingner, T., Kalkhof, S., Märdian, S. et al. Enabling technologies towards personalization of scaffolds for large bone defect regeneration 74, p. 263.

[36] Golub, B., Hudz, A., Dudnyk, A., Bushma, A. Production of Biotechnological Objects using Business Intelligence, in 2019 9th International Conference on Advanced Computer Information Technologies (ACIT), IEEE, p. 200.

[37] Kim, J.Y. Roadmap to High Throughput Phenotyping for Plant Breeding 45, p. 43.

[38] Mofakkarul Islam, M., Renwick, A., Lamprinopoulou, C., Klerkx, L. Innovation in Livestock Genetic Improvement 12, p. 42.

[39] Li, N.-S., Chen, Y.-T., Hsu, Y.-P., Pang, H.-H. et al. Mobile healthcare system based on the combination of a lateral flow pad and smartphone for rapid detection of uric acid in whole blood 164, p. 112309.

[40] Kianijaya, M.R., Hasanuddin, M.O., 2022. Implementation of Data Storage for the Monitoring System for Biogas Production Optimization, in 2022 8th International Conference on Wireless and Telematics (ICWT), IEEE, p. 1.

[41] Cafferty, K.G., Muth, D.J., Jacobson, J.J., Bryden, K.M. Model Based Biomass System Design of Feedstock Supply Systems for Bioenergy Production, in Volume 2B: 33rd Computers and Information in Engineering Conference, American Society of Mechanical Engineers, V02BT02A023.

[42] Kosina, S.M., Greiner, A.M., Lau, R.K., Jenkins, S. et al. Web of microbes (WoM): a curated microbial exometabolomics database for linking chemistry and microbes 18, p. 115.

[43] Fabre, C., Buche, P., Rouau, X., Mayer-Laigle, C. Milling itineraries dataset for a collection of crop and wood by-products and granulometric properties of the resulting powders 33, p. 106430.

[44] Gnimpieba, E.Z., Hartman, T.W., Do, T., Zylla, J. et al. Biofilm marker discovery with cloud-based dockerized metagenomics analysis of microbial communities 25, bbae429.

[45] Yi, H. A traceability method of biofuel production and utilization based on blockchain 310, p. 122350.

[46] Zhang, Y., Liu, Y., Jiong, Z., Zhang, X. et al. Development and assessment of blockchain-IoT-based traceability system for frozen aquatic product 44, e13669.

[47] Gojon, A., Nussaume, L., Luu, D.T., Murchie, E.H. et al. Approaches and determinants to sustainably improve crop production 12, e369.

[48] Danielsen, F., Eicken, H., Funder, M., Johnson, N. et al. Community Monitoring of Natural Resource Systems and the Environment 47, p. 637.

[49] Bellgard, M.I., Bellgard, S.E. A Bioinformatics Framework for plant pathologists to deliver global food security outcomes: Keynote address from the 18th Australasian Plant Pathology Society Conference 2011 41, p. 113.

[50] Schmieg, B., Brandt, N., Schnepp, V.J., Radosevic, L. et al. Structured Data Storage for Data-Driven Process Optimisation in Bioprinting 12, p. 7728.

[51] Dobbins, N.J., Spital, C.H., Black, R.A., Morrison, J.M. et al. Leaf: an open-source, model-agnostic, data-driven web application for cohort discovery and translational biomedical research 27, p. 109.

[52] Stock, D., Schel, D., Schneider, M., Götz, B., Rauh, L., Schnicke, F., Schulz-Zander, J., 2022. FabOS - Towards an open, distributed, real-time capable and secure operating system for AI-assisted manufacturing. Fraunhofer-Gesellschaft.