AI-Driven Workforce Training through Industry Partnerships Using Educational Biodigital Twins (BDT) and Programmable Cloud Laboratories (PCL)

Prof. Reem Khojah, University of California, San Diego

Reem Khojah serves as an assistant teaching professor in the Shu Chien-Gene Lay Department of Bioengineering at the University of California, San Diego. With experience in instructing bioengineering at introductory and graduate levels, she actively contributes to enhancing accessibility to research tools for undergraduate research experiences. Her primary focus is on optimizing engineering education through data-driven pre-and post-lecture formative assessments and designing AI-proof assignments. Her educational background includes a B.S. in Medical Technology, a Master's degree in Chemical and Biological Engineering from KAUST, and a Ph.D. in Bioengineering from the University of California, Los Angeles. Reem has also engaged in post-doctoral research at the University of California, Santa Cruz, and the University of California, Irvine.

Dr. Isgard S. Hueck, University of California, San Diego

Dr. Isgard Hueck (Ph.D. in Higher Ed/ Leadership & Policy; M.Sc. in Bioengineering; M. Phil. in Education) Affiliations: UCSD - Dept of Bioengineering, Whitaker Institute of Biomedical Engineering & UCSD School of Medicine; Moores Cancer Center.

Born and raised in Germany, Isgard Hueck studied Biology at the Wilhelms- University in Munster and received her license as Cyto-Pathologist in Cologne, Germany, in 1987. After years of clinical work in hematology, cancer diagnosis and therapy, Isgard Hueck received a Master of Science degree in Bio-Medical Engineering in 1998 from the University of Applied Sciences in Aachen, Germany. In conjunction with the University California in San Diego, Dr. Hueck conducted biomedical research focussing on long-term medical implants, angiogenesis in diabetes, and stem cell research. She has published in scientific journals like Microcirculation and the American Journal of Physiology. She contributed two chapters to the Springer Verlag Books 'Applications in Stem Cell Engineering' and 'Biological, Physical, and Technical Basics of Cell Engineering.' In 2012, Dr. Hueck was recruited as the Program Assessment Specialist for the Dept of Bioeng., UCSD, leading continuous improvement processes in engineering education and for the Accreditation Board of Engineering and Technology. In 2022, Isgard Hueck completed her Ph.D. in leadership in higher education with a focus on engineering education. Dr. Isgard Hueck founded the Office of Industrial Relations (IRO) in the Department of Bioengineering at UCSD to provide engineering students better opportunities to build close industry relationships within the academic education. In addition, Dr. Hueck engages in enrichment programs for "learning beyond the classroom". She is actively assessing and researching opportunities to improve eduction for the modern and holistic engineer of tomorrow.

Dr. Alyssa Catherine Taylor, University of California San Diego

Dr. Alyssa Taylor, Teaching Professor of Bioengineering, joined the UC San Diego Jacobs School of Engineering in 2022. She previously served as an Associate Teaching Professor in the Department of Bioengineering at the University of Washington. There, she was awarded the American Society for Engineering Education Pacific Northwest Section Outstanding Teaching Award in 2022; the University of Washington College of Engineering Faculty Teaching Award in 2021; and the University of Washington Distinguished Teaching Award, the university's top teaching award, in 2020. She was recently voted Best Undergraduate Teacher in Bioengineering at UCSD in both 2024 and 2025. She teaches introductory, lab-based, and grad-level design courses in bioengineering, and ultimately aims to foster the development of inclusive, thoughtful engineering graduates who will integrate their technical and professional skills to positively impact society. Taylor earned a PhD in biomedical engineering from the University of Virginia and a bachelor's degree in biological systems engineering from UC Davis.

Adam M Feist, University of California, San Diego
Benjamin L Smarr, University of California, San Diego

AI-Driven Workforce Training through Industry Partnerships Using Educational Biodigital Twins (BDT) and Programmable Cloud Laboratories (PCL)

Abstract

The rapid adoption of artificial intelligence (AI) and generative AI (genAI) in biomedical and bioengineering industries is reshaping workforce expectations and exposing persistent gaps between academic preparation and industry practice. Effective AI training depends strongly on data context, including how data is generated, curated, validated, and used within operational workflows, rather than on algorithmic techniques alone. These contextual factors directly influence model generalization, which remains a central challenge for responsible AI deployment, particularly when models are applied beyond narrowly defined training conditions. In biomedical engineering, this challenge is further compounded by the use of patient-derived and clinical data, where data provenance, regulatory oversight, ethical constraints, and limited accessibility fundamentally shape AI development and evaluation. As a result, traditional academic training environments often struggle to provide students with exposure to the data practices and workflow constraints that govern real-world AI performance. This study examines industry perspectives on the potential role of educational biodigital twins (BDT) and programmable cloud laboratories (PCL) as scalable infrastructures for AI workforce preparation. The results indicate that industry views BDT and PCL as promising when they preserve workflow structure and decision constraints through the use of synthetic or constrained data, while avoiding direct exposure to sensitive datasets. However, scalability is limited by data governance requirements and internal resource constraints. These findings inform academic–industry collaboration models that emphasize context-aware training environments and support the design of partnerships aligned with industry realities.

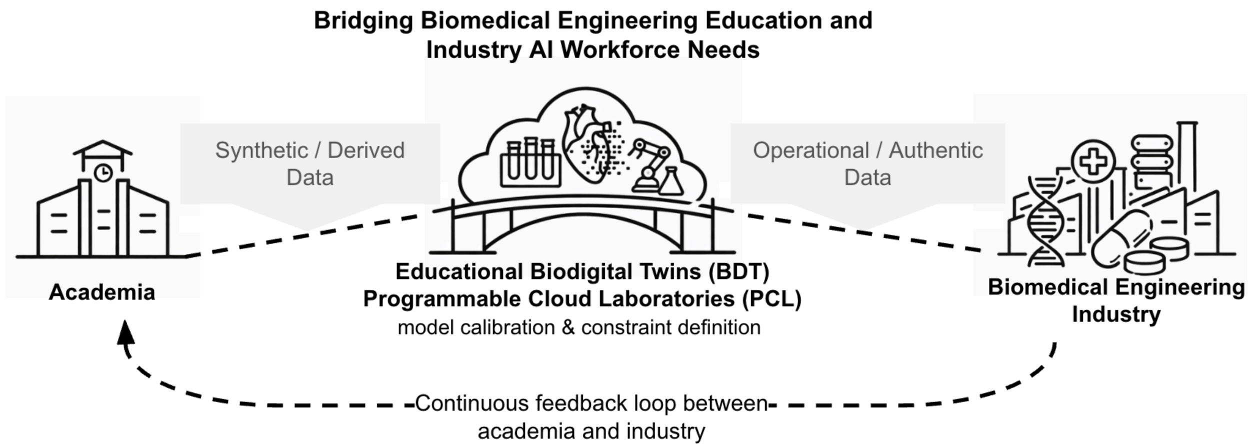


Figure 1. Conceptual framework illustrating a continuous feedback loop between academia and the biomedical and bioengineering industry to support AI-ready workforce preparation. Industry practice defines workforce demands, operational constraints, and data context, which are characterized through industry surveys and inform identification of skill gaps. These insights guide curriculum design and workforce-aligned training implemented through educational biodigital twins (BDT) and programmable cloud laboratories (PCL). BDT and PCL mediate between operational industry data and educational use by enabling model calibration, constraint definition, and generation of synthetic or derived data, providing students with realistic exposure to data-driven, automated, and regulated workflows. Ongoing industry feedback is then used to iteratively refine academic curricula, closing the loop between education and evolving industry practice.

Introduction

Rapid adoption of artificial intelligence (AI) and generative artificial intelligence (genAI) across the biomedical and bioengineering industry is reshaping workforce expectations, motivating increased attention to industry partnerships that can inform educational approaches for preparing future bioengineers [1], [2], [3]. Recent workforce and education studies indicate that employers increasingly prioritize applied data skills, computational literacy, and the ability to operate within established workflows over narrow technical specialization [4], [5]. For undergraduate and master's engineering programs geared towards preparing students to work in the biomedical industry, it is imperative that the curricular experiences and training address and align with modern industry computational tools. However, the integration of AI and genAI introduces challenges that differ from prior computational tool transitions, as AI performance and reliability depend strongly on data provenance, representativeness, and context of use [6], [7]. In biomedical engineering, these factors are closely tied to regulated datasets, ethical constraints, and physical laboratory workflows, limiting the scalability of authentic AI training across large student populations and contributing to persistent gaps between academic instruction and industry practice.

Digital twins (DT), and their biological extension as biodigital twins (BDT), together with programmable cloud laboratories (PCLs), have emerged as potential educational infrastructures capable of supporting realistic, workflow-oriented AI training at scale [7], [8], [9], [10]. In this role, biodigital twins and PCLs function as mediating infrastructures through which industry data context is translated into model calibration and constraint definitions, enabling the use of synthetic or derived data for educational purposes. However, their effectiveness for workforce preparation depends critically on alignment with industry needs, constraints, and expectations. This study therefore examines industry perspectives on AI readiness, workforce gaps, and the perceived role of biodigital twins and programmable cloud laboratories in biomedical engineering education (**Figure 1**).

Background

In accordance with the need to incorporate industry-related experiences into the bioengineering curriculum across many programs [11], informed usage of modern computational tools in industry into educational context is desirable [12]. To continue to understand the need for industry-desired skills of future employees [13], [14], we first need to collect industry feedback on the new computational tools, which this study will address. As educational theoretical frameworks (e.g. STL) established the importance of "context" in developing workforce readiness with realistic industry-used tools, educational biodigital twins have the potential to provide industry-used data and workflow directly in an educational environment. This will better prepare engineering graduates to be ready to engage in industry workflow processes and address the need for frequent stakeholder feedback in an environment of rapid technology advances [15].

Situated Learning Theory (STL)

Situated learning theory provides a framework for understanding why learning in engineering is closely tied to the technical context, tools, and workflows in which knowledge is applied rather than acquired as decontextualized skills. From this perspective, knowledge develops through interaction with authentic tasks, representations, and technological systems, and is shaped by the constraints and affordances of the environments in which it is used [16]. Engineering education research has applied this framework to show that competencies emerge through engagement with realistic problem structures, data representations, and operational tools, and that instruction grounded in authentic technical contexts is associated with improved transfer to professional practice [17], [18]. In practice, the application of situated learning involves immersion in real-world engineering environments through apprenticeships and internships, which is particularly critical in generative AI settings where understanding data provenance, structure, and constraints directly shapes system behavior. However, such experiences are inherently difficult to scale across programs and cohorts.

Educational Biodigital Twins

Digital twins are persistent virtual representations of physical systems that remain linked to real-world data streams, enabling prediction, monitoring, and decision-making across system lifecycles [8], [19]. Biodigital twins extend the digital twin paradigm by integrating mechanistic and data-driven models across molecular, cellular, tissue, and organ scales [9], [20], [21], [22]. Prior studies have described biodigital twins as enabling infrastructures for precision medicine, personalized decision support, and systems-level biomedical research, particularly when coupled with AI-based modeling approaches [23], [24], [25]. In educational settings, biodigital twins refer to digital twin infrastructures applied to biological or biomedical systems, where computational models are grounded in real or realistically structured biological, experimental, or clinical data and organized around engineering workflows.

Educational digital twins have been adopted across multiple instructional contexts to support authentic, data-driven learning aligned with professional practice [26]. In electronics and microprocessor education, digital twin environments enable students to interact with embodied sensors and collaborative virtual laboratories that support circuit design, debugging, and system-level reasoning [27], [28]. In nuclear engineering education, digital twins provide a safe instructional setting for exploring complex and high-risk systems while preserving core operational and analytical workflows [29]. In bioprocess engineering, digital twin models are used to teach process monitoring, control strategy development, and decision-making under industrial constraints [30]. An educational biodigital twin has also been demonstrated in the context of biomanufacturing through the BioDT framework, which integrates biological process models, data streams, and virtual representations to support learning and analysis in intelligent biomanufacturing systems [31]. Immersive and multimodal digital twin learning environments have been explored in engineering education contexts [32] as well as in non-engineering tactical training settings, illustrating the broader pedagogical potential of such systems [33].

In biomedical engineering education, biodigital twins have been examined as instructional environments that support learner engagement with evolving system representations reflecting constraints such as data provenance, uncertainty, validation, and regulatory considerations [34]. Rather than emphasizing visualization alone, these environments enable participation in iterative modeling, evaluation, and documentation aligned with disciplinary practice. However, adoption in industrial settings remains constrained by data sensitivity, regulatory requirements, intellectual property concerns, and validation challenges [25], [35], which in turn shape how biodigital twins can be adapted for educational use when the goal is to reflect operational practice rather than simplified instructional abstractions.

Programmable Cloud Laboratories

Educational digital and biodigital twins require grounding in physical experimentation to support meaningful learning, particularly in biomedical and bioengineering education where experimental validation is central to disciplinary practice. Cloud-based biology experimentation platforms designed for education demonstrate how students can engage with real biological systems through remote interaction, enabling inquiry-driven learning that connects computational models to physical outcomes [36], [37]. Educational laboratory digital twins that integrate real experimental workflows with computational representations further illustrate how students can experience authentic system behavior rather than purely simulated environments [38].

Programmable Cloud Laboratories support engineering education by enabling students to design, execute, and observe real experiments through software-mediated laboratory systems, removing constraints related to physical lab capacity, safety, and location [39] [40]. Robotic cloud laboratories used in instructional contexts demonstrate how students can engage with state-of-art experimental workflows remotely while retaining hands-on learning characteristics [41], [42]. Open-source programmable laboratory platforms further expand educational access to automated experimentation, allowing students to interact with contemporary bioengineering tools and practices early in their training [43].

In biomedical engineering education, remote and cloud-enabled laboratories have been shown to support cognitive, behavioral, and affective learning outcomes by enabling guided interaction with real experimental systems [44], [45]. Cloud-enabled live-cell biotechnology platforms further demonstrate how remote experimentation can reduce educational inequities by expanding access to authentic laboratory experiences for diverse student populations [46]. Importantly, programmable cloud laboratories complement educational biodigital twins by closing the loop between in silico models and physical experimentation, enabling student-generated experimental data to inform model validation and iterative refinement in ways that mirror modern bioengineering practice [47], [48].

Scaling Situated Learning via Educational Biodigital Twins and Programmable Cloud Laboratories

Educational biodigital twins and programmable cloud laboratories have increasingly been framed as scalable situated learning environments that embed learners within authentic representations of physical–digital systems while preserving system behavior, data relationships, and operational constraints characteristic of professional engineering practice [49], [50]. Foundational work on situated learning emphasizes that learning is most effective when knowledge is acquired through participation in meaningful, context-rich activities rather than through abstract or decontextualized instruction [51]. From this perspective, educational digital twins can be interpreted not as visual simulations, but as environments that situate learners within realistic systems, workflows, and problem spaces that reflect professional practice. When coupled with biodigital twins and programmable cloud laboratories, these environments have the potential to support industry-aligned learning experiences that contribute to workforce development by enabling scalable exposure to authentic data, tools, and constraints [26].

Recent scholarship further connects digital twin–based instructional approaches to situated and practice-based learning in engineering education. Large-scale reviews of digital twin–enabled education emphasize that instructional value arises when task structure, system interdependencies, and operational constraints are preserved, rather than simplified for instructional convenience [52]. Related work in open and distance learning [53], industrial robotics training [54], Internet of Things (IoT) education using digital twins [55], and educational robotic agents [56] similarly points to learning environments consistent with situated learning principles, in which knowledge acquisition is strengthened through engagement with realistic system contexts that preserve workflow structure and data relationships. Accordingly, AI and generative AI training, particularly in biomedical engineering, must be grounded in context-aware data that reflect the provenance, regulatory constraints, and decision structures of real biomedical systems, since model generalization, safety, and professional competence depend critically on how domain-specific context is encoded rather than on algorithmic sophistication alone [7], [57].

Use Case

In a bioengineering programming course [blinded], we implemented an educational biodigital twin use case that mirrors a real organoid culture imaging pipeline, emphasizing workflow context over isolated coding skills. Students were first given access to a shared organoid image set captured under intentionally varied illumination and exposure settings, then asked to select the images they planned to process. This learner-driven selection step surfaced an industry-relevant reality: some organoids are inherently easier to segment because acquisition conditions create cleaner foreground–background separation, while others become ambiguous when lighting introduces shadows, saturation, or low signal. Students then executed a structured, end-to-end analysis workflow that linked acquisition quality to downstream outcomes, moving from inspection and preprocessing decisions to AI-assisted segmentation and quantitative feature extraction to interpret growth and morphology patterns. By comparing processing steps and outputs across their chosen images, students discovered that segmentation performance and biological interpretability were constrained less by the sophistication of the code and more by upstream data context, specifically contrast and acquisition consistency. This experience positioned AI as a context-dependent tool within an operational pipeline, helping students internalize when AI-enabled analysis is reliable, when it fails, and why workflow-aware decisions (such as choosing higher-contrast images or standardizing lighting) are prerequisite to meaningful organoid phenotyping.

Motivation

Although many educational digital twin initiatives have been developed, a substantial proportion remain disconnected from industry data, workflows, and decision-making contexts [3], [58]. As AI adoption in biomedical and bioengineering continues to accelerate, data context matters. Without industry partnership, educational implementations risk misalignment with workforce realities, particularly in domains governed by regulation, data governance, and ethical constraints. This study is motivated by the need to understand from industry how to collaborate with academic institutions and inform the design and use of biodigital twins and programmable cloud laboratories for workforce-oriented biomedical engineering education.

Methods

Data Collection Instrument and Analysis

Data were collected using a structured online survey administered via Google Forms to capture industry practices and perspectives related to artificial intelligence (AI) and generative AI (genAI) training and collaboration opportunities, and views on educational mechanisms such as biodigital twins and programmable cloud laboratories. The survey included multiple-choice, Likert-scale, and open-ended questions and was distributed to professionals working in bioengineering or closely related sectors that employ bioengineering graduates. Quantitative responses were analyzed to identify dominant trends, while qualitative responses were listed in **table 1**. Findings were interpreted through a continuous improvement lens to inform curriculum alignment and workforce-relevant training interventions. The analysis is industry-informed, designed to characterize trends and relationships in industry practices that are relevant for evidence-informed curriculum development.

Survey Design

The survey was designed in two stages to capture both current industry practices and potential pathways for addressing identified gaps. The first stage focused best industry collaboration practices and partnerships for AI and generative AI workforce training. The second stage focused on understanding how academic–industry collaboration could help bridge observed gaps, with questions probing collaboration mechanisms, workforce training approaches, and perceptions of educational tools. To support consistent interpretation of emerging concepts, the survey introduced biodigital twins and programmable cloud laboratories using a brief definition, a summary description, and an explanatory figure illustrating bioengineering examples. This approach ensured that responses reflected informed perspectives rather than assumed familiarity. Overall, the survey avoided prescriptive or hypothetical prompts and instead emphasized capturing existing industry behaviors, practices, and constraints related to AI, generative AI, biodigital twins, and programmable cloud laboratories.

Ethical Approval

This study reports findings from educational assessment and stakeholder feedback activities designed to guide instructional improvement. Participants, including students, faculty, and industry partners, provided informed consent under a protocol approved by the [institution name] Institutional Review Board (IRB 813454). The study was deemed exempt under 45 CFR 46.104(d)(2)(iii) following review by the [institution name] IRB, with a limited review conducted under 46.111(a)(7).

Participants

The dataset comprises responses from 25 organizations operating in biotechnology or closely adjacent sectors, with all participants confirming active employment of bioengineering graduates (**Figure 2**). Organizations ranged from startups (1–50 employees) to medium and large enterprises (50+ employees), including five leading international companies, four of which report annual revenues exceeding USD 1 billion. Respondents represented diverse professional roles spanning executive leadership, research and engineering, product development and management, software and systems engineering, and data science or AI specialization, enabling triangulation of workforce perspectives across organizational and functional levels. The organizations operate across multiple bioengineering domains, including medical devices, diagnostics, pharmaceutical and drug development, biomanufacturing, and digital health. Participants were recruited through institutional collaborators, professional networking outreach, industry-facing venues, including the Biomedical Engineering Society (BMES) annual meeting industry exhibition and outreach to companies participating in bioengineering career fairs. Collectively, this participant pool provides industry-informed input suitable for continuous improvement processes, supplying external stakeholder perspectives that inform curriculum alignment, workforce relevance, and iterative refinement of educational interventions.

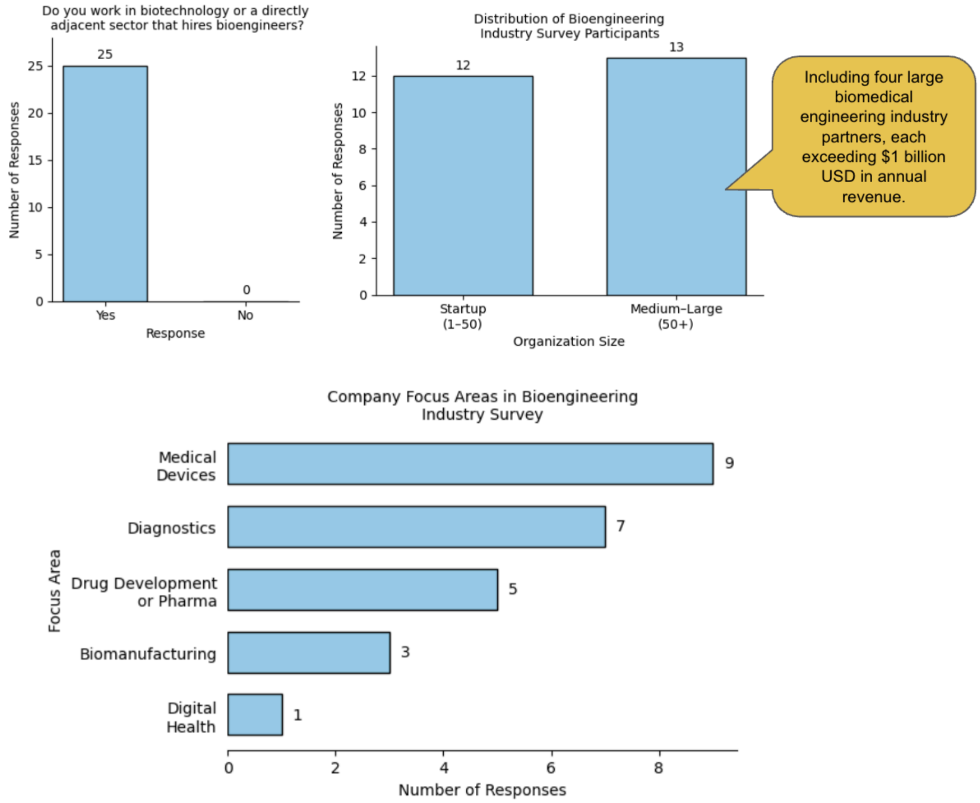


Figure 2. Biomedical and bioengineering industry survey participants. The survey includes respondent confirmation of employment in biotechnology or closely adjacent sectors that hire bioengineers (top left), organization size categorized as startup (1–50 employees) and medium to large organizations (50+ employees) (top right), and primary company focus areas within bioengineering (bottom).

Results and discussion

How can bioengineering departments collaborate with industry to help close the AI and genAI gap in biomedical engineering education?

To examine how bioengineering departments can collaborate with industry to help close gaps in AI and genAI preparation, Figure 3 summarizes industry perspectives on academic–industry collaboration mechanisms. The results indicate that most organizations are open to collaboration, although interest is commonly conditional on available resources and organizational capacity. Among the collaboration models considered, student internships and co-op programs were most frequently identified as beneficial, followed by joint AI research projects, while guest lectures, workshops, and access to university computing or simulation resources were viewed as less impactful. Overall, these findings suggest that industry places greater value on sustained, applied forms of engagement that integrate students into authentic workflows, rather than short-term or infrastructure-focused collaborations, when supporting AI and genAI workforce preparation.

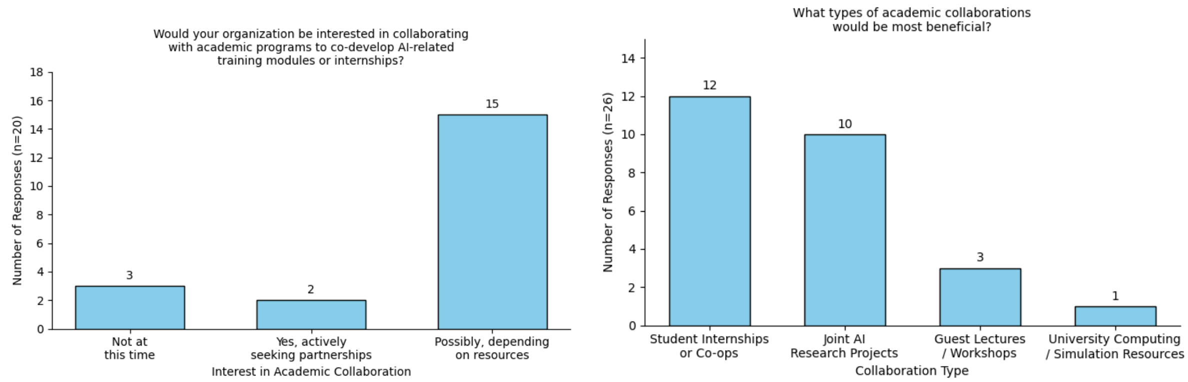
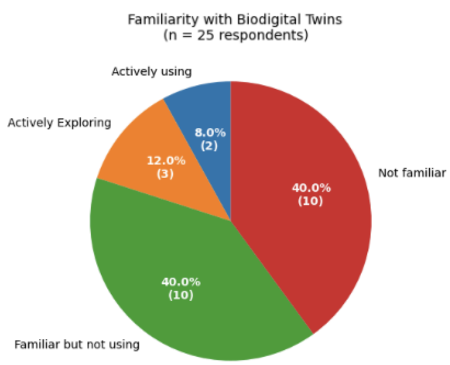


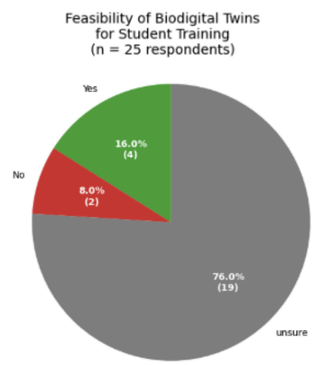
Figure 3. Academic–industry collaboration for AI and genAI training with industry data from Industry perspectives. The top-left panel shows organizational interest in collaborating with academic programs to co-develop AI-related training modules or internships. The top-right panel reports which forms of academic collaboration are perceived as most beneficial.

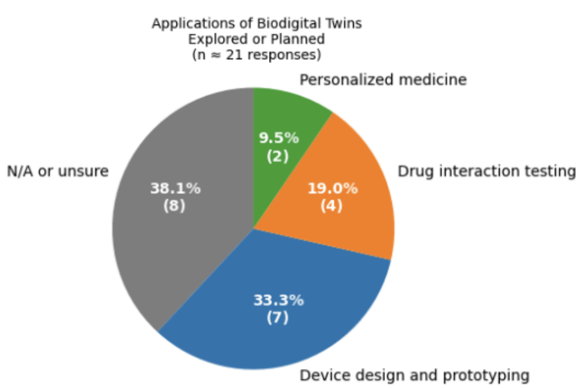
How does industry perceive the use of biodigital twins (BDT) for workforce preparation?

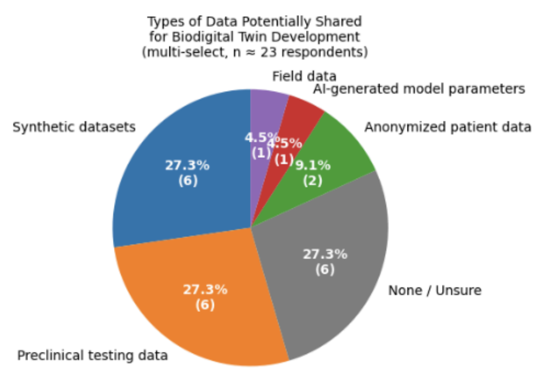
To understand industry perceptions of biodigital twins (BDT) as an emerging educational and workforce development tool, we examined respondent familiarity, perceived feasibility for student training, anticipated application areas, data-sharing considerations, and perceived impact on workforce readiness (**Figure 4**). Results indicate that familiarity with biodigital twins is limited, with a substantial proportion of respondents reporting no familiarity or only conceptual awareness without active use. Most expressed uncertainty rather than opposition when asked about feasibility for student training. This pattern suggests that limited familiarity, rather than resistance, underlies hesitation toward adoption. Among respondents who had explored or considered biodigital twins, reported applications clustered around high-impact biomedical engineering workflows, including device design and prototyping, drug interaction testing, and personalized medicine. However, a large fraction remained unsure about applications, reinforcing the interpretation that the technology is still emerging in practice. Data-sharing responses further contextualize these findings: respondents were most willing to consider synthetic and preclinical datasets, while sharing of patient-derived or field data was limited or uncertain. Interest in sharing anonymized data for educational biodigital twin development was similarly cautious, with most respondents indicating uncertainty rather than clear willingness or refusal, reflecting regulatory, ethical, and infrastructural constraints rather than lack of perceived value. Importantly, despite uncertainty around familiarity, feasibility, and data sharing, respondents expressed generally positive views regarding the potential of educational biodigital twins to improve workforce readiness. A majority agreed or strongly agreed that such tools could enhance student preparation, indicating recognition of their pedagogical value even in the absence of widespread industrial deployment.

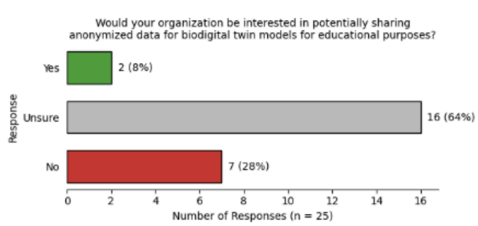
These findings suggest that biodigital twins occupy a transitional space between emerging technology and educational opportunity. An important practical implication of these findings is that data sharing remains a major barrier for AI-driven organizations, even when nondisclosure agreements are in place, due to concerns about competitive risk and proprietary advantage. Consequently, broad access to industrial datasets is unlikely. More viable academic–industry collaborations may instead center on educational biodigital twins built using synthetic or non-sensitive data, or on targeted partnerships in which academic teams derive insights from narrowly scoped dataset subsets, enabling meaningful workforce preparation while minimizing data exposure.











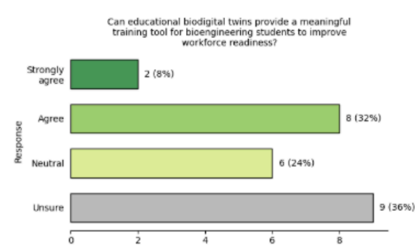


Figure 4. Industry perspectives on awareness, feasibility, applications, data considerations, and educational value of biodigital twins (BDT). Panels report respondent familiarity with biodigital twins, perceived feasibility for student training, applications explored or planned, types of data potentially shared for biodigital twin development, interest in sharing anonymized data for educational purposes, and perceived value of educational biodigital twins for improving workforce readiness. Bars and pie segments indicate the number and proportion of respondents selecting each option.

How does industry perceive the value of programmable cloud laboratory (PCL) training for improving workforce readiness?

To examine industry perceptions of programmable cloud laboratories as a workforce development tool, we analyzed responses related to educational value, expected workforce impact, preferred training approaches, and conditions for early adoption (**Figure 5**). Results indicate generally positive perceptions of PCL-based training, with a majority of respondents agreeing or strongly agreeing that such training would improve workforce readiness by providing students with realistic expectations around experimental constraints, documentation practices, and project timelines. Consistent with this view, most respondents agreed that graduates trained in a PCL environment would strengthen workforce readiness faster than typical new hires. This finding aligns with earlier results emphasizing productivity, automation, and workflow familiarity as central industry priorities, and suggests that exposure to realistic, end-to-end laboratory workflows may reduce onboarding time.

When asked about effective approaches for training and onboarding teams to use a Programmable Cloud Laboratory, respondents most frequently selected one day onsite workshops with hands-on labs and self-paced modules with sandbox accounts, indicating a preference for practical, low-barrier training formats. A smaller number of respondents selected live virtual bootcamps consisting of four 2-hour sessions, suggesting interest in synchronous training when time commitments are clearly bounded. Less frequently selected approaches included role-based microlearning plus weekly office hours, shadowed pilot projects using an organization's own data, and train-the-trainer packages for internal leads, which require greater coordination, longer time horizons, or deeper organizational commitment. Collectively, these responses indicate that industry favors onboarding models that are modular, time-limited, and easy to integrate into existing workflows, while more intensive or customized training structures are viewed as secondary or situational rather than default pathways.

To identify conditions that would encourage early adoption of a Programmable Cloud Laboratory, respondents reported which factors would most increase their likelihood of adoption. Respondents most frequently selected no-cost sandbox access for 60 days and published case studies that show return on investment, indicating that low-risk evaluation opportunities and evidence of tangible value are primary drivers of interest. A smaller number of respondents selected dedicated support engineers during onboarding and tailored pilot deployments on their own workflows with a defined service level agreement (SLA), suggesting that while hands-on support is valued, it is not the default expectation. Less frequently selected factors included ready security and compliance review packages, native integration with existing laboratory information management systems (LIMS), and native integration with electronic laboratory notebooks (ELN), reflecting the higher organizational and regulatory effort required for such commitments. Collectively, these responses indicate that industry interest in Programmable Cloud Laboratory adoption is driven less by novelty and more by opportunities for low-risk exploration, demonstrated return on investment, and minimal upfront integration burden.

Overall, the results suggest that programmable cloud laboratories are viewed as a promising educational mechanism for improving workforce readiness, particularly by exposing students to realistic operational constraints and documentation practices that align with industry expectations. At the same time, adoption is contingent on practical considerations, including ease of access, clear evidence of benefit, and flexible training models. These findings support the use of PCL environments in academic settings as a scalable, low-risk approach to preparing students for industry workflows, while highlighting the importance of modular design and evidence-driven deployment to encourage broader industry engagement.

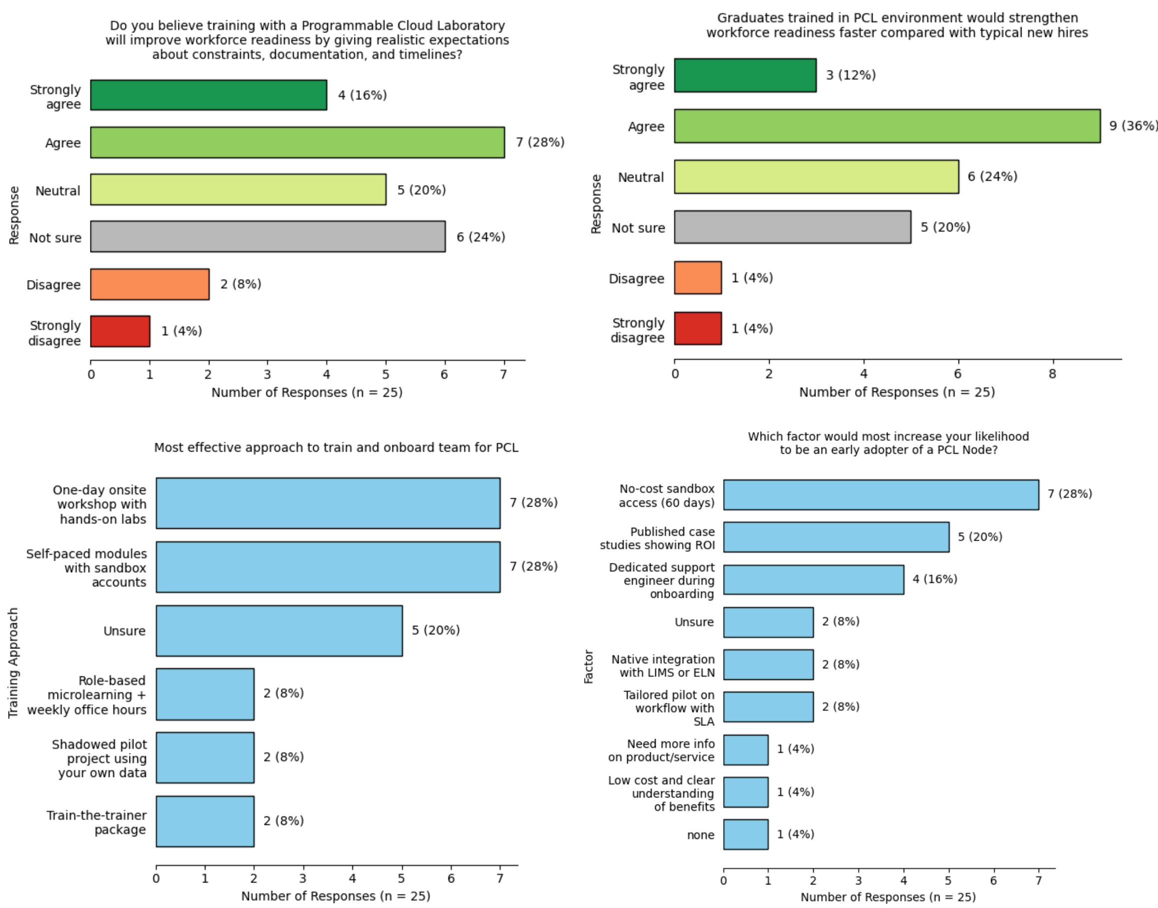


Figure 5. Industry perspectives on the educational value, workforce impact, and adoption conditions of a Programmable Cloud Laboratory (PCL). The top-left panel reports whether respondents believe PCL-based training would improve workforce readiness by providing realistic exposure to constraints, documentation, and timelines. The top-right panel reports perceptions of whether graduates trained in a PCL environment would strengthen workforce readiness faster than typical new hires. The bottom-left panel summarizes preferred approaches for training and onboarding teams to use PCL. The bottom-right panel reports factors most likely to increase industry willingness to become early adopters of a PCL node. Bars indicate the number of respondents selecting each option (n = 25).

Industry recommendations on bridging academia with industry using BDT and PCL

To better understand the conditions under which industry would engage in academic collaboration, respondents were asked the following open-ended question: "Under what conditions would you consider collaborating with academic institutions to develop biodigital twins and programmable cloud laboratories? (**Table. 1**). The open-ended responses reveal that industry views collaboration on biodigital twins and programmable cloud laboratories as a selective product and workflow partnership evaluated through the same criteria used for internal R&D investments. Interest is highest when collaboration directly advances automation, infrastructure integration, or a product roadmap, and drops sharply when data exposure, staffing burden, or unclear ownership is introduced. Data security concerns persist even under nondisclosure agreements, indicating that trust is not contractual but structural. As a result, the most realistic collaboration pathway is not broad data sharing, but narrowly scoped engagements in which academic teams derive insights from constrained or synthetic datasets, often supported by external funding to offset internal costs. This reframes the role of academia: universities are not expected to replicate industrial systems, but to serve as low-risk testbeds where workforce skills, automation workflows, and biodigital twin concepts can be developed and validated before full industrial adoption.

Limitations

This work is intended to explore industry expectations and preferences to determine a future role of universities to collaborate with industry in preparing bioengineering graduates for the changing workforce. However, the limited number of surveyed companies is not intended to represent a universal industry opinion or demand, but rather can be interpreted as an initial approach to gather industry perspectives on the use of new AI-driven BDT and PCL technologies as educational tools. Furthermore, the findings reflect current industry trends, which may change rapidly as BDT and PCL technologies continue to emerge and mature. As observed in the results, many established biotechnology companies reported limited familiarity with these technologies at the time of the survey. Finally, this paper does not yet report student learning outcomes or curricular intervention effects; its contribution is to characterize industry perspectives that can guide the design of subsequent educational implementations, which are outlined in the Future Work section.

Theme	Under what conditions would you consider collaborating with academic institutions to develop BDT and PCL?
Mutual Benefit / Alignment with Company Goals / Partnership Quality / Security	PCL would be most interesting. We've built infrastructure and tools for this, so some sort of development partnership would be interesting. Mutually beneficial advancement of our product roadmap As long as the data are handled securely and there is mutual benefit, we would be open to collaborating. I think there's value when initiatives can align with company values and interests. Good project. Good partnership. Collaboration would be considered when: There is clear alignment in scientific goals around automation, data standards, and predictive biology. The partner institution has infrastructure for robotic or high-throughput experimentation that integrates with programmable lab environments. Data sharing follows FAIR principles and supports interoperable, real-time digital twin updates. IP, data ownership, and publication policies are transparent and mutually beneficial. There is a joint funding or sustainability plan with measurable translational outcomes beyond academic publication. Clear indication of how data will be secured for biodigital twins. Similar concerns for PCL- industry standards for data security are high, clear indications of how PCL will manage the data generated in labs will need to be determined. Data anonymization and access to university facilities and instruments
Funding / Resources / ROI	Funding for staff One where there would be cost sharing potentially through an external Grant

	High ROI
Internal Process / Leadership / Coordination / Integration	I need buy-in from the leadership team. I have ideas that I would need to pitch to stakeholders. Need to have a cadence of meetings with lab team, IT and external partner PCL would be of interest if they can integrate our hardware Perhaps looking at contrived samples i.e. cell lines might be good for initial proof of concept

Table 1. Open-ended industry responses describing conditions under which they would consider collaborating via biodigital twins (BDT) and programmable cloud laboratories (PCL) with academic institutions. Responses were listed under themes, including mutual benefit and strategic alignment, funding and resource considerations, infrastructure and technical integration, data security and intellectual property governance, internal leadership and coordination requirements, and the use of low-risk proof-of-concept approaches.

Future Work

To strengthen the industry-grounded evidence base and capture the rapidly evolving nature of AI and generative AI adoption in biomedical engineering, additional biomedical and bioengineering companies across the United States are invited to complete the survey on an ongoing basis. Expanding participation will allow for tracking shifts in workforce expectations, governance practices, and adoption barriers as AI-driven tools mature and regulatory landscapes evolve. In addition, follow-up conversations with selected industry organizations are planned to collect deeper qualitative insights into collaboration conditions, partnership models, and data governance strategies relevant to biodigital twins and programmable cloud laboratories. These discussions will focus on identifying practical mechanisms through which academic–industry collaboration can support industrial advancement while informing educational decision-making, particularly under constraints related to data sensitivity, privacy, and scalability. Collectively, these future efforts aim to refine the synthesis of industry requirements and constraints presented in this study and to inform evidence-based curriculum planning for preparing future generations of work-ready bioengineering graduates.

Conclusion

This study examined industry perspectives on preparing an AI- and genAI-ready biomedical engineering workforce and identified key barriers to scaling workforce-aligned training through academia–industry collaboration. Guided by a continuous feedback framework, the findings indicate that while industry broadly recognizes the potential educational value of biodigital twins and programmable cloud laboratories, their scalability is constrained less by technical feasibility and more by structural factors, including data sensitivity, regulatory compliance, intellectual property protection, and internal resource limitations. Industry interest in collaboration was consistently conditional, favoring low-risk engagement models such as internships, narrowly scoped applied projects, synthetic or non-sensitive datasets, and externally funded partnerships that minimize staffing and governance burden. These constraints limit the viability of large-scale data sharing or replication of industrial systems within academic settings, particularly in regulated biomedical contexts. Programmable cloud laboratories were viewed as a promising complementary mechanism for scaling exposure to realistic experimental workflows, documentation practices, and timelines, but adoption similarly depends on modular design, ease of access, and clear evidence of return on investment. Collectively, the results suggest that scalable AI and genAI training in biomedical engineering will require educational infrastructures that function as boundary mechanisms, mediating between academic constraints and industry realities, rather than attempting to replicate proprietary industrial environments, with scalability shaped primarily by governance, data, and resource considerations rather than instructional ambition alone. Once these constraints are addressed in future work, we envision the possibility of an accessible, scalable, workforce-aligned training opportunity via academia-industry collaborations, which could be available across institutions and incorporated into the existing curricula.

References

[1] G. C. Fleming, M. Klopfer, A. Katz, and D. Knight, “What engineering employers want: An analysis of technical and professional skills in engineering job advertisements,” *J. Eng. Educ.*, Feb. 2024, doi: 10.1002/jee.20581.

[2] N. Jones, “AI race in 2025 is tighter than ever before,” *Nature*, Apr. 2025, doi: 10.1038/d41586-025-01033-y.

[3] E. J. Topol, “High-performance medicine: the convergence of human and artificial intelligence,” *Nat. Med.*, vol. 25, no. 1, pp. 44–56, Jan. 2019.

[4] J. L. Mason, E. Johnston, S. Berndt, K. Segal, M. Lei, and J. S. Wiest, “Labor and skills gap analysis of the biomedical research workforce,” *FASEB J.*, vol. 30, no. 8, pp. 2673–2683, Aug. 2016.

[5] G. Lam, I. Hueck, C. Rivera, and P. Widder, “A future-focused lens to equipping biomedical engineering graduates for an evolving field,” *Biomed. Eng. Educ.*, vol. 5, no. 2, pp. 177–188, Jul. 2025.

[6] R. Khojah, A. Werth, K. W. Broadhead, L. W. Dobrucki, C. Geiger, and D. A. Rubenstein, “Correction: Integrating generative artificial intelligence tools and competencies in biomedical engineering education,” *Biomed. Eng. Educ.*, vol. 5, no. 2, pp. 153–153, Jul. 2025.

[7] A. Arora *et al.*, “The value of standards for health datasets in artificial intelligence-based applications,” *Nat. Med.*, vol. 29, no. 11, pp. 2929–2938, Nov. 2023.
[8] M. Grieves and J. Vickers, “Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems,” in *Transdisciplinary Perspectives on Complex Systems*, Cham: Springer International Publishing, 2017, pp. 85–113.
[9] E. Katsoulakis *et al.*, “Digital twins for health: a scoping review,” *NPJ Digit. Med.*, vol. 7, no. 1, p. 77, Mar. 2024.
[10] National Science Foundation, “NSF to invest in new national network of AI-programmable cloud laboratories,” National Science Foundation. Accessed: Jan. 18, 2026. [Online]. Available: https://www.nsf.gov/news/nsf-invest-new-national-network-ai-programmable-cloud
[11] I. S. Hueck, A. Guével, R. S. MacLeod, and K. Billiar, “Integration of co-curricular experiential learning in BME programs to increase student success,” *Biomed. Eng. Educ.*, vol. 5, no. 2, pp. 189–202, Jul. 2025.
[12] M. Chinyeaka, K. Agwu, M. Adelodun, G. Igwama, and E. Anyanwu, “Enhancing Biomedical Engineering Education: Incorporating Practical Training in Equipment Installation and Maintenance,” *International Journal of Engineering Research and Development (IJERD)*, vol. 20, no. 9, pp. 178–184, 2024.
[13] D. Jackson and R. Bridgstock, “What actually works to enhance graduate employability? The relative value of curricular, co-curricular, and extra-curricular learning and paid work,” *High. Educ. (Dordr.)*, vol. 81, no. 4, pp. 723–739, Apr. 2021.
[14] M. Hirudayaraj, R. Baker, F. Baker, and M. Eastman, “Soft skills for entry-level engineers: What employers want,” *Educ. Sci. (Basel)*, vol. 11, no. 10, p. 641, Oct. 2021.
[15] A. J. DiMeo *et al.*, “Biomedical engineering professional skills development: The RADxSM Tech impact on graduates and faculty,” *IEEE Open J. Eng. Med. Biol.*, vol. 2, pp. 163–169, Apr. 2021.
[16] J. Lave and E. Wenger, “Legitimate Peripheral Participation,” in *Situated Learning*, Cambridge: Cambridge University Press, 1991, pp. 27–44.
[17] A. Johri and B. M. Olds, “Situated engineering learning: Bridging engineering education research and the learning sciences,” *J. Eng. Educ.*, vol. 100, no. 1, pp. 151–185, Jan. 2011.
[18] B. Gray, G. Fragoso-Diaz, O. Ogunrinde, and T. Rider, “Implementation of a project-based learning approach in an upper level course in engineering technology,” in *2024 ASEE Annual Conference & Exposition Proceedings*, ASEE Conferences, 2024. doi: 10.18260/1-2--47575.
[19] A. Fuller, Z. Fan, C. Day, and C. Barlow, “Digital twin: Enabling technologies, challenges and open research,” *IEEE Access*, vol. 8, pp. 108952–108971, 2020.
[20] M. Viceconti, M. De Vos, S. Mellone, and L. Geris, “From the digital twins in healthcare to the Virtual Human Twin: a moon-shot project for digital health research,” *arXiv [physics.med-ph]*, Mar. 27, 2023. [Online]. Available: http://arxiv.org/abs/2304.06678
[21] S. A. Niederer, M. S. Sacks, M. Girolami, and K. Willcox, “Scaling digital twins from the artisanal to the industrial,” *Nat. Comput. Sci.*, vol. 1, no. 5, pp. 313–320, May 2021.
[22] J. Wu and V. H. Koelzer, “Towards generative digital twins in biomedical research,” *Comput. Struct. Biotechnol. J.*, vol. 23, pp. 3481–3488, Dec. 2024.
[23] A. Prakosa *et al.*, “Personalized virtual-heart technology for guiding the ablation of infarct-related ventricular tachycardia,” *Nat. Biomed. Eng.*, vol. 2, no. 10, pp. 732–740, Oct. 2018.

[24] K. P. Venkatesh, G. Brito, and M. N. Kamel Boulos, "Health digital twins in life science and health care innovation," *Annu. Rev. Pharmacol. Toxicol.*, vol. 64, pp. 159–170, Jan. 2024.
[25] P.-H. Huang, K.-H. Kim, and M. Schermer, "Ethical issues of digital twins for personalized health care service: Preliminary mapping study," *J. Med. Internet Res.*, vol. 24, no. 1, p. e33081, Jan. 2022.
[26] M. A. Hazrat, N. M. S. Hassan, A. A. Chowdhury, M. G. Rasul, and B. A. Taylor, "Developing a skilled workforce for future industry demand: The potential of Digital Twin-based teaching and learning practices in engineering education," *Sustainability*, vol. 15, no. 23, p. 16433, Nov. 2023.
[27] L. Bondaryk, A. Kyle, I. Davidesco, C. Dorsey, and B. Montrosse-Moorhead, "Embodied sensors and digital twins as an introduction to microprocessor programming for middle and high school non-CS majors," in *2025 ASEE Annual Conference & Exposition Proceedings*, ASEE Conferences, 2025. doi: 10.18260/1-2--56330.
[28] K. Sakkas *et al.*, "Multiplayer virtual labs for electronic circuit design: A digital twin-based learning approach," *Electronics (Basel)*, vol. 14, no. 16, p. 3163, Aug. 2025.
[29] H. Xu, K. Chen, H. Xu, and X. Wang, "Application of digital twin technology in nuclear engineering teaching," *Advances in Education, Humanities and Social Science Research*, vol. 12, no. 1, p. 356, Dec. 2024.
[30] C. Appl, A. Moser, F. Baganz, and V. C. Hass, "Digital Twins for bioprocess control strategy development and realisation," *Adv. Biochem. Eng. Biotechnol.*, vol. 177, pp. 63–94, 2021.
[31] B. Zhao *et al.*, "BioDT: An integrated digital-twin-based framework for intelligent biomanufacturing," *Processes (Basel)*, vol. 11, no. 4, p. 1213, Apr. 2023.
[32] T.-C. Lin, C.-N. Chiu, P.-T. Wang, and L.-D. Fang, "VisFactory: Adaptive multimodal digital twin with integrated visual-haptic-auditory analytics for industry 4.0 engineering education," *Multimedia*, vol. 1, no. 1, p. 3, Aug. 2025.
[33] X. Lv, Y. Tao, Y. Zhang, and Y. Xue, "Design of an immersive basketball tactical training system based on digital twins and federated learning," *Appl. Sci. (Basel)*, vol. 15, no. 7, p. 3831, Mar. 2025.
[34] "Artificial Intelligence in Science," OECD. Accessed: Jan. 19, 2026. [Online]. Available: https://www.oecd.org/en/publications/artificial-intelligence-in-science_a8d820bd-en/full-report/ai-and-scientific-productivity-considering-policy-and-governance-challenges_da5bf9a6.html
[35] N. Soveizi, F. Turkmen, and D. Karastoyanova, "Security and privacy concerns in cloud-based scientific and business workflows: A systematic review," *Future Gener. Comput. Syst.*, vol. 148, pp. 184–200, Nov. 2023.
[36] Z. Hossain *et al.*, "Interactive and scalable biology cloud experimentation for scientific inquiry and education," *Nat. Biotechnol.*, vol. 34, no. 12, pp. 1293–1298, Dec. 2016.
[37] H. G. Martin *et al.*, "Perspectives for self-driving labs in synthetic biology," *Curr. Opin. Biotechnol.*, vol. 79, no. 102881, p. 102881, Feb. 2023.
[38] S. D. Rihm, J. Bai, A. Kondinski, S. Mosbach, J. Akroyd, and M. Kraft, "Transforming research laboratories with connected digital twins," *Nexus*, vol. 1, no. 1, p. 100004, Mar. 2024.
[39] R. B. Canty *et al.*, "Science acceleration and accessibility with self-driving labs," *Nat. Commun.*, vol. 16, no. 1, p. 3856, Apr. 2025.
[40] M. A. Torres-Acosta, G. J. Lye, and D. Dikicioglu, "Automated liquid-handling operations

for robust, resilient, and efficient bio-based laboratory practices," *Biochem. Eng. J.*, vol. 188, no. 108713, p. 108713, Dec. 2022.

[41] C. Arnold, "Cloud labs: where robots do the research," *Nature*, vol. 606, no. 7914, pp. 612–613, Jun. 2022.

[42] A. I. Cooper *et al.*, "Accelerating discovery in natural science laboratories with AI and robotics: Perspectives and challenges," *Sci. Robot.*, vol. 10, no. 106, p. eadv7932, Sep. 2025.

[43] P. Dettinger *et al.*, "Open-source personal pipetting robots with live-cell incubation and microscopy compatibility," *Nat. Commun.*, vol. 13, no. 1, p. 2999, May 2022.

[44] L. S. Post, P. Guo, N. Saab, and W. Admiraal, "Effects of remote labs on cognitive, behavioral, and affective learning outcomes in higher education," *Comput. Educ.*, vol. 140, no. 103596, p. 103596, Oct. 2019.

[45] I. Dubovi and V. R. Lee, "Instructional support for learning with agent-based simulations: A tale of vicarious and guided exploration learning approaches," *Comput. Educ.*, vol. 142, no. 103644, p. 103644, Dec. 2019.

[46] S. Vera-Choqqueccota, B. E. Y. Belmekki, M.-S. Alouini, M. Teodorescu, D. Haussler, and M. A. Mostajo-Radji, "Reducing education inequalities through cloud-enabled live-cell biotechnology," *Trends Biotechnol.*, vol. 43, no. 1, pp. 43–60, Jan. 2025.

[47] N. Hillson *et al.*, "Building a global alliance of biofoundries," *Nat. Commun.*, vol. 10, no. 1, p. 2040, May 2019.

[48] M. Abolhasani and E. Kumacheva, "The rise of self-driving labs in chemical and materials sciences," *Nat. Synth.*, Jan. 2023, doi: 10.1038/s44160-022-00231-0.

[49] A. I. Multi-Scale, "A Multi-Scale AI Framework for InformalSTEM Learning: Paramorphic DigitalTwins for Underserved Communities".

[50] *Development of an Immersive, Digital Twin-Supported Smart Reconfigurable Educational Platform for Manufacturing Training: A Proof of Concept*.

[51] L. Dawley and C. Dede, "Situated learning in virtual worlds and immersive simulations," in *Handbook of Research on Educational Communications and Technology*, New York, NY: Springer New York, 2014, pp. 723–734.

[52] V. Chamola *et al.*, "Metaverse for education: Developments, challenges, and future direction," *Comput. Appl. Eng. Educ.*, vol. 33, no. 3, May 2025, doi: 10.1002/cae.70018.

[53] N. R. W. Mohamed, N. Mahmud, and A. F. Arham, "Exploring digital curation's role in enhancing student engagement in Open Distance Learning among Gen Z," *International Journal of Research and Innovation in Social Science*, vol. VIII, no. IIIS, pp. 4842–4852, 2024.

[54] J. Osborne and J. Williams, "WIP: Virtual reality as a tool for reinforcing real-world robot programming skills," in *2025 ASEE Annual Conference & Exposition Proceedings*, ASEE Conferences, 2025. doi: 10.18260/1-2--57465.

[55] Z. Zhu *et al.*, "LearnIoTVR: An end-to-end virtual reality environment providing authentic learning experiences for internet of things," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: ACM, Apr. 2023, pp. 1–17.

[56] V. G. A. Hakim, S.-H. Yang, J.-H. Wang, H.-H. Lin, and G.-D. Chen, "Digital twins of pet robots to prolong interdependent relationships and effects on student learning performance," *IEEE Trans. Learn. Technol.*, vol. 17, pp. 1883–1897, 2024.

[57] L. Goetz, N. Seedat, R. Vandersluis, and M. van der Schaar, "Generalization-a key challenge for responsible AI in patient-facing clinical applications," *NPJ Digit. Med.*, vol. 7,

no. 1, p. 126, May 2024.
[58] B. van Breugel, T. Liu, D. Oglic, and M. van der Schaar, “Synthetic data in biomedicine via generative artificial intelligence,” *Nat. Rev. Bioeng.*, vol. 2, no. 12, pp. 991–1004, Oct. 2024.