Bridging the Gap in Competency Training for Students in AI and GenAI Technologies in the Biotech Industry: Exploring Biodigital Twins

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. Alyssa Catherine Taylor, University of California San Diego

Alyssa C. Taylor is a Teaching Professor in bioengineering with thirteen years of teaching experience across introductory, laboratory, and capstone design courses. Her teaching career began in 2010 when she joined the University of Washington as an assistant teaching professor. She was a faculty member at the University of Washington from 2010 - 2022, after which she joined the University of California San Diego. Areas of interest include industry engagement, communication, reflection, universal design, and fostering a sense of community, connection, and inclusion in engineering education. Dr. Taylor aims to foster the development of inclusive, thoughtful engineering graduates who will integrate their technical and professional skills to positively impact society and she is excited to contribute to the educational journey of engineering students.

Work in Progress:

Bridging the Gap in Competency Training for Students in AI and GenAI Technologies in the Biotech Industry: Exploring Biodigital Twins

Reem Khojah, Alyssa Catherine Taylor, Isgard S. Hueck
Bioengineering, University of California San Diego

Abstract
Bioengineers have long served as a vital bridge between healthcare and technology, adapting innovations from interdisciplinary fields to advance patient care and medical solutions. One such emerging innovation is generative AI (GenAI) (e.g., large language models (LLMs)) [1] , a technology with significant potential to impact the biotechnology industry [2]. As artificial intelligence (AI) continues to reshape research and development in this field, it is imperative for biomedical engineering programs to equip students with AI competencies that align with evolving industry demands [3], [4].

To address this need, this work aims to survey leading biotechnology companies to identify the most essential AI tools, techniques, and skills currently employed or actively being taught to their employees, to gain insights on whether they can be integrated into the biomedical engineering curriculum. By gathering insights directly from industry professionals, the study seeks to bridge the gap between academic instruction and practical applications in key areas such as machine learning, data analysis, and AI-driven automation. The survey is designed to focus on AI tools currently used in biotechnology research while exploring the specific skills valued in future biomedical engineers.

Additionally, this work explores the potential for leveraging biodigital twin systems, developed and provided by industry, as a means to close the gap between academic training and industry practices. Biodigital twins, which provide virtual representations of biological entities, can serve as a valuable resource for students [5]. By making the data from these systems accessible, students can train and apply AI or GenAI applications to analyze, simulate, and interpret biological processes, offering hands-on, simulation-based learning experiences aligned with real-world workflows [6].

The outcomes of this study will inform curriculum design, ensuring that graduates are well-prepared to navigate the rapidly evolving biotech landscape. By aligning educational programs with industry needs and exploring innovative tools like biodigital twins, this research contributes to shaping a forward-thinking approach to biomedical engineering education, preparing students for the next generation of AI-driven healthcare solutions and fostering innovation at the intersection of technology and medicine

Motivation

The integration of AI into biotechnology is driving significant changes in research and development. However, academic programs often lag in equipping students with practical AI skills, creating a mismatch between graduate competencies and employer expectations. As the demand for AI expertise continues to grow, it is imperative to revise curricula to prepare biomedical engineering students for AI-driven industry roles. This work seeks to address the following challenges:

1- Industry-Academia Gap: A disconnect exists between theoretical instruction and practical applications of AI and biodigital twin systems in the biotechnology field, limiting students' exposure to cutting-edge tools used in industry workflows.

2- Competency Deficits: Graduates often lack experience with advanced AI tools, techniques, and biodigital twin systems, which are becoming essential for real-world problem-solving in areas such as simulation, data analysis, and predictive modeling.

3- Emerging Trends: The rise of generative AI tools and the increasing exploration of biodigital twins introduce new opportunities for innovation, as well as ethical considerations, emphasizing the need for targeted education that incorporates these advancements into the curriculum.

Methodology

A survey was developed to gather insights from biotechnology professionals regarding the AI and generative AI (GenAI) tools and techniques they currently use or train their employees to utilize. The objective is to identify the most critical AI skills and tools required in the industry and assess whether there are gaps in the current biomedical engineering curriculum that need to be addressed. The survey design avoids directly soliciting advice from participants; instead, it focuses on understanding industry practices to infer actionable insights for curriculum improvement.

The survey consists of the following key components:

Demographic Information:

- Participants are asked about their roles (e.g., research scientist, data scientist, product manager) and the size and focus areas of their organizations (e.g., drug development, diagnostics, medical devices). This information helps contextualize responses and ensures a diverse representation of perspectives.

Current Use of AI and GenAI Tools:

- Questions are designed to identify the AI and GenAI tools being actively used in biotechnology workflows, such as machine learning frameworks, generative design platforms, and data analysis tools.
- Participants are also asked about the types of tasks or problems these tools address, including automation, synthetic data generation, biomedical imaging, and drug discovery.

Employee Training and Skill Development:

- The survey explores whether companies provide training for their employees on AI tools and, if so, the specific skills or techniques emphasized during this training.
- Questions probe for insights into the level of proficiency expected from new hires versus what is developed through internal training in the workplace. Skills expected from recent graduates can be prioritized in curriculum design.

Importance of AI Skills:

- Participants rate the relevance of various foundational AI skills (e.g., programming, data analysis) and advanced GenAI skills (e.g., prompt engineering, fine-tuning models) to their organization's operations.
- This helps identify which skills are considered critical, important, or supplementary.

Barriers and Challenges:

- Open-ended questions allow respondents to highlight any challenges they face in adopting AI tools or training employees. This includes identifying bottlenecks in implementation, skill gaps, or the need for ethical considerations and regulatory compliance.

Biodigital Twin Adoption and Potential:
Participants are asked about their familiarity with biodigital twins and whether their organization is actively exploring or implementing such models. Questions are designed to understand:

- **Current or Planned Applications of Biodigital Twins**:
  How biodigital twins, enhanced by AI and GenAI technologies, are used or planned to be used in workflows such as virtual organ simulation, drug-device interaction testing, and personalized medicine. For example, the use of machine learning models for real-time physiological simulation or GenAI tools for generating synthetic biological datasets to improve twin accuracy.
- **Feasibility as a Training Tool**:
  The potential of biodigital twins, integrated with AI and GenAI capabilities, to serve as a platform for training bioengineering students. This includes exploring how tools like AI-driven simulation frameworks and generative design algorithms can enhance the educational value of these models.
- **Benefits and Challenges for Education**:
  The perceived advantages of developing AI-enhanced biodigital twin models for education, such as providing students with hands-on experience in simulation-based learning. Additionally, identifying challenges such as the cost of implementing AI and GenAI in twin models, intellectual property concerns, and the technical complexity of integrating these technologies into educational programs.
- **Data Sharing for Training**:
  Participants are asked about their willingness to share data to support the creation of biodigital twin models for educational purposes. Questions focus on:
    - **Types of Data**:
      Which types of data they would feel comfortable sharing, such as anonymized patient data, preclinical testing data, synthetic datasets, or AI-generated model parameters. More specifically, several biomedical engineering companies have developed biodigital twin modules designed for training and education purposes [7] [8] [9].
    - **Data Protection Measures**:
      The kinds of safeguards or agreements (e.g., data anonymization, non-disclosure

agreements) that would make them more inclined to share data for educational applications.

- **Collaborative Potential**:
  Whether they would consider sharing AI or GenAI-enhanced data (e.g., simulated physiological datasets or generative models) with academic institutions and under what conditions this collaboration might occur.

Data Collection and Analysis

The survey is being distributed to professionals across a range of biotechnology companies, including startups, mid-sized organizations, and large corporations, to capture diverse industry perspectives. Specifically, the data is being collected from ~15 international and national companies with local branches with whom we have previous interactions and partnerships or active collaborations in student education and networking events. These companies are also part of our Industrial Advisory Board (IAB) and serve as advisors in our academic programs. Responses are collected through an online confidential survey platform to facilitate broad participation and standardize data collection.

Quantitative responses will be analyzed using statistical methods to identify trends in tool usage, skill priorities, and training practices. Qualitative responses from open-ended questions will be thematically coded to uncover patterns and insights regarding industry expectations and challenges.

The findings will be used to map existing educational offerings against the identified industry requirements. This comparison may highlight gaps in the current biomedical engineering curriculum, providing a basis for recommending specific updates to better align academic training with industry needs.

By focusing on real-world applications and industry practices, this methodology ensures that the study captures actionable insights to inform the development of a forward-thinking, AI-integrated biomedical engineering curriculum.

Conclusion

This study aims to inform the gap between academic training and industry requirements in biomedical engineering by identifying critical AI and GenAI tools and skills used in the biotechnology sector. By surveying industry professionals about their current use of AI technologies and employee training practices, we seek to uncover actionable insights that can guide curriculum enhancements.

The anticipated findings will provide a clearer understanding of the specific AI competencies valued by employers, from foundational programming and data analysis to advanced GenAI applications such as synthetic data generation and prompt engineering. Additionally, the study

will highlight areas where new graduates may lack sufficient preparation, enabling educators to address these gaps through targeted curriculum updates.

Aligning biomedical engineering education with industry advancements is essential to preparing students for the rapidly evolving biotech landscape. This research will contribute to the development of a workforce equipped with the AI skills necessary to excel in AI-driven healthcare and biotechnology. By fostering stronger connections between academia and industry, this work represents a critical step toward creating a more relevant and impactful biomedical engineering education.

References:

[1] J. P. Guilford, “The nature of human intelligence,” *N. Y. Macgraw Hill*, 1967.

[2] “Pal, Soumen, et al. ‘A domain-specific next-generation large language model (LLM) or ChatGPT is required for biomedical engineering and research.’ Annals of biomedical engineering 52.3 (2024): 451-454.”.

[3] “ASEE PEER - Impact of AI Tools on Engineering Education.” Accessed: Jan. 13, 2025. [Online]. Available: https://peer.asee.org/impact-of-ai-tools-on-engineering-education?utm_source=chatgpt.com

[4] “ASEE PEER - Revolutionizing Engineering Education: The Impact of AI Tools on Student Learning.” Accessed: Jan. 13, 2025. [Online]. Available: https://peer.asee.org/revolutionizing-engineering-education-the-impact-of-ai-tools-on-student-learning?utm_source=chatgpt.com

[5] B. R. Barricelli, E. Casiraghi, and D. Fogli, “A survey on digital twin: Definitions, characteristics, applications, and design implications,” *IEEE Access*, vol. 7, pp. 167653–167671, 2019.

[6] L. Kartashova, A. Gurzhii, V. Zaichuk, T. Sorochan, and F. Zhuravlev, “Digital Twin of an Educational Institution: An Innovative Concept of Blended Learning:,” in *Proceedings of the 1st Symposium on Advances in Educational Technology*, Kyiv, Ukraine: SCITEPRESS - Science and Technology Publications, 2020, pp. 300–310. doi: 10.5220/0010931100003364.

[7] 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, doi: 10.1038/s41551-018-0282-2.

[8] P. Hunter *et al.*, “A vision and strategy for the virtual physiological human: 2012 update,” *Interface Focus*, vol. 3, no. 2, p. 20130004, Apr. 2013, doi: 10.1098/rsfs.2013.0004.

[9] “Living Heart Project | SIMULIA™ - Dassault Systèmes®.” Accessed: Jan. 11, 2025. [Online]. Available: https://www.3ds.com/products-services/simulia/solutions/life-sciences-healthcare/the-living-heart-project/