Federal Register Notice: 89 FR 51554, Federal Register :: Networking and Information Technology Research and Development Request for Information on Digital Twins Research and Development, June 18, 2024.

Request for Information on the National Digital Twins R&D Strategic Plan

Digital Twins for Health (DT4H) Consortium

A Strategic Plan for Research and Development of Human Digital Twins

Digital Twins for Health (DT4H) Consortium

I. Introduction

Although the concept of a Digital Twin and virtual simulation dates back to early work in aerospace vehicles by NASA in the 1960's, the DT concept has been gaining increasing popularity in a variety of industries and areas of science as a promising platform for integrating data-driven and mechanistic modeling with artificial intelligence (AI) [1-3]. In early 2023, the National Academies of Sciences, Engineering, and Medicine appointed an ad-hoc committee to identify the needs and opportunities to advance the mathematical, statistical, and computational foundations of digital twins in applications across science, medicine, engineering, and society [4].

Among a wide range of applications, the concept of digital twin can play a particularly important role in modeling and simulation of the human body and its intricate subsystems. Known as a Human Digital Twin (HDT), this branch of digital twin research holds the potential to profoundly impact human society and enhance lives by enabling personalized health monitoring, improving the detection, screening, and prevention of adverse health conditions, and facilitating virtual testing and clinical trials. It is acknowledged, however, that significant research and development efforts are still needed to fully realize the potential of digital twins in healthcare. In this response, we present our vision for advancing the emerging field of HDT and propose strategies, discuss challenges, and make recommendations for its future applications in the scientific community.

II. Definition of a Human Digital Twin (HDT)

In its most general form, a human digital twin (HDT) is a virtual representation of an individual human, that can encompass real-time simulations of multiple sub-systems operating at multiple length scales ranging from microscopic cells and molecules to tissues, organs, organ systems, and ultimately the entire human body. In general, the biological class of digital twins share the following basic properties [5]:

- **Individualized** - An HDT is highly personalized, at the level of an individual human, or a specific genotype or phenotype.
- **Interconnected** - An HDT is informed by a real-time connection to a living biological system.
- **Interactive** - An HDT enables a real-time closed loop feedback between the physical and virtual systems.
- **Informative** - In addition to the physical and virtual components, an HDT platform must also provide a means for third-party observers to view, control, test, and interpret the behavior and response of the virtual system.
- **Impactful** - An HDT can contribute significantly to better health and well-being through continuous monitoring and analysis, early detection of potential health issues, and improved treatment strategy.

III. Strategies for Advancing Human Digital Twins for Better Health

In March 2022, the Digital Twins for Health (DT4H) Consortium was established, comprising a global network of professionals and researchers with diverse domain expertise who share a common goal of advancing HDTs [6]. Since its inception, the DT4H Consortium has conceptualized a research cyberinfrastructure, i.e., a DT4H Gateway, to advance R&D for HDTs by addressing the challenges faced by researchers, developers, and users and facilitating their navigation of the HDT landscape across multiple disciplines for the first time. As shown in **Figure 1**, the DT4H Gateway includes five infrastructure modules: (1) computing; (2) standardizing; (3) learning; (4) modeling; and (5) training, operating under four guiding principles: (i) collaborative scientific teamwork; (ii) ethical and trustworthy digital twins; (iii) commitment to diversity, equity, and inclusion; and (iv) active community involvement and partnership.

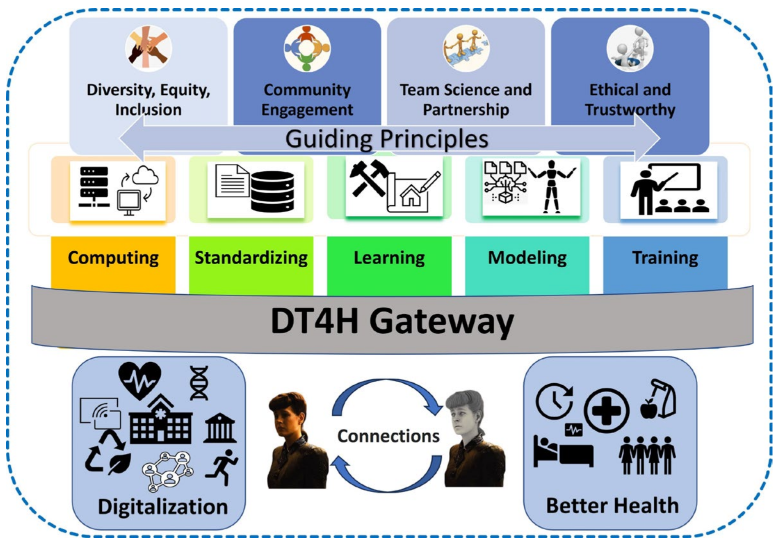

Figure 1. Our overarching vision of advancing HDTs for better health via a DT4H Gateway.

As a consortium, we have identified several fundamental strategies for advancing the field of human Digital Twin research. A basic summary of the strategic areas is listed below.

III.1 Concerted Research Experience Integrated through Computing

The vision of the computing module is to provide a concerted research experience by orchestrating a seamless integration of diverse datasets, tools, models, and computing resources to provide researchers with an efficient and cohesive workflow management environment [7-9]. Providing interfaces between data and AI algorithms allowing continuous model enhancements and interfaces for comparing user data to existing models is a significant technical undertaking and a necessity for data and model standardization. Consistent well-documented and well-preforming application programming interfaces (APIs) will allow reproducible HDT workflows which include a mixture of data, algorithm, model, sharing, and visualization components.

Specifically, the environment needs to manage user accounts with precision and foster a collaborative spirit by enabling users to work effectively in groups. Tools and datasets need to be managed efficiently to facilitate team science. Users should be able to write and execute code remotely, and user interfaces should be intuitive for users to design, track, and manage virtual experiments. Additionally, the computing environment must support collaborative scientific inquiries through metadata-rich virtual experiments. Adherence to metadata standards can ensure those experiments are FAIR (findable, accessible, interoperable, and reusable), enhancing collaborative scientific inquiries. Adopting industry standards on task and workflow execution, data management, privacy and security management is necessary to foster translation of user research environments with diverse backend resources.

III.2 Enhancing HDT Data Standardization and Integration under an Ethical Framework

The standardizing module focuses on providing resources, best practices, and tooling to facilitate data standardization and integration for HDTs, as summarized in **Table 1** above. This will ensure the HDT data are accurate, reliable, and interoperable across different systems and platforms.

Table 1. The multimodal data types used in HDTs and the potential standardization strategies.

Multimodal data types	Standardization strategies
Electronic Health Record (EHR) data	Observational Medical Outcomes Partnership (OMOP) Common Data Model [10] and HL7 FHIR (Fast Healthcare Interoperability Resources) [11]
Mobile health data	HL7 consumer mobile health application function framework guideline [12]; Open mHealth tools [13]
Physiological and biomedical imaging data	Digital Imaging and Communications in Medicine (DICOM) [14]; Hierarchical Data Format version 5 (HDF5) [15]
Genetic testing data (-omics)	Variation Representation Specification (VRS) [16]; MIGS-MIMS for genomics [17]; MIAPE for proteomics [18]; CIMR for metabolomics [19]; MIAME for transcriptomics [20]
Social Determinants of Health data	Social Determinants of Health Ontology (SDoHO) [21]

III.3 Establishing Advanced Machine Learning Framework for Building HDTs

The vision of the learning module is to provide a comprehensive and advanced framework for building HDTs, encompassing a range of tools and methodologies to extract, augment, integrate, and interpret data effectively, as summarized in **Table 2** below.

Table 2. Various approaches to extract, augment, integrate, and interpret multimodal datasets.

Multimodal data types	Learning strategies
Genetic/genomic data	Pararead [2], COCOA [23], AIList [24], GATK's FilterIntervals [25], LOLAWeb [26], DeepChrome [27]
Imaging data and textual feature	AALIM [28], large language models (LLM) [29]
Behavior and health data	Machine learning tools [30]
Social media data	Natural language processing (NLP) tools [31]

Moreover, tensor fusion networks [32] and multiplexed graph neural networks [33] can be used for modality integration. For data visualization and interpretability, a variety of methods can be applied, such as AI Explainability 360 [34], Boolean Decision Rules [35], Generalized Linear Rule Models [36], LIME [37], SHAP [38], TED [39], t-SNE [40], PCA [41], and UMAP [42].

III.4 Developing HDT Capabilities through Modeling and Simulation

The modeling and simulation module, integral to building specific HDTs, is responsible for developing retrospective restructuring, monitoring, analytical, and predictive capabilities of HDTs by integrating data-driven and mechanistic approaches with advanced AI technologies and data science methods. The type of modeling employed can depend on the specific type of data being used. **Table 3** lists various data types and the corresponding modeling and simulation strategies.

Table 3. Modeling and simulation strategies for various data types and scenarios.

Multimodal data types and scenarios	Modeling and simulation strategies
Longitudinal Data modeling and forecasting	Various dynamical system models [43], including neural dynamical systems, neural integral-differential equation models, RNN models, transformers, GNNs, diffusion models, and latent variable approaches
Omics data	Gene regulatory networks [44] and functional protein network models
Single-cell, cell aggregates	Multiphase materials models, level-set models, phase field models, and Cellular Potts Model (CPM) for modeling single-cell and cell aggregates [45]
Organs and tissues	Spatial-temporal dynamical models based on non-equilibrium thermodynamics and network models, including level-set models, phase-field models, etc. [46]
Multimodality data fusion	Graph-based toolkits and graph neural networks [47]
Environmental particulates and drug interactions	Pharmacokinetics/Pharmacodynamics modeling tools [48]
Human behavior, mental health	Various temporal models for simulating norms in online social networks and cross-platform prediction and simulation [49]
Human-environment interactions	Agent-Based Models [50]
Integration of physics-based, agent-based, and statistical models with generative AI such as ChatGPT, GPT-4, Claude 3, and DALL-E	Large language models (GPT-4, Claude 3, etc.), multimodal generative AI model based on stable diffusion models [51-52]

III.5 Building a Sustainable Community Through Training and Workforce Development

Our vision for the training module is to facilitate the training of the next generation of leaders in engineering, science, and technology to become HDT creators, builders, and users. As listed in **Table 4** above, it is imperative to build a sustainable program focusing on career development support and mentoring across all career stages, emphasizing underrepresented and minority groups.

Table 4. Various programs to enhance training and workforce development for HDTs.

Programs	Training and workforce development strategies
Interdisciplinary Mentoring	Dual mentorship in biomedical science and AI/ML, leveraging online forms, PubMed knowledge graph [53], and mentorship databases [54] for diverse pairing and flipped mentorship

Recruitment and Diversity	Utilizing various channels for recruitment and focusing on increasing diversity through lectures and mentorships by experts from academic partners and government organizations [55]
Outreach and Partnership	Forming partnerships with communities, industries, and advocate groups, offering internships, practicum opportunities, and HDT data and tool challenges
Ethical and Trustworthy DT	Emphasizing ethical AI algorithm and model development, data privacy, and security technologies

IV. Specific Challenges for Human Digital Twins

While HDTs represent a groundbreaking advancement in personalized healthcare and offer immense potential for improving health outcomes, realizing the full potential of HDTs in healthcare requires addressing several formidable challenges that span across various domains.

Challenges of inherent complexities and uncertainties in human bodies: Human bodies exhibit significant heterogeneities at the molecular, cellular, and organ levels, resulting in substantial differences in biological responses among individuals. Biological processes can change over time. The dynamic nature of biological systems often exhibits non-linear behaviors and emergent properties, making it difficult to predict responses accurately. Incorporating the temporal aspect into an HDT is complicated, especially considering the various factors that influence human health over a person's lifetime, the incomplete understanding of underlying mechanisms, and often incomplete or limited available biological data from various sources.

Challenges in advanced algorithms, computational resources, and validation: The current state of knowledge regarding human physiology, disease mechanisms, and treatment responses remains incomplete and continuously evolving. Integrating this fragmented understanding and theoretical frameworks into HDT models, while simultaneously accounting for inherent uncertainties and facilitating seamless model updates as novel insights emerge, poses a critical challenge that demands innovative solutions. Furthermore, HDT simulations and analyses often involve processing massive volumes of data and performing computationally intensive operations, necessitating the development of highly efficient computational algorithms and the strategic harnessing of high-performance computing resources to enable real-time simulations and analyses. Moreover, ensuring the accuracy, reliability, and robustness of HDT models and simulations is of paramount importance for their practical applications in healthcare settings. Developing rigorous validation and verification frameworks, encompassing virtual clinical trials and comparative studies with real-world data represents a significant challenge that requires focused efforts and interdisciplinary collaborations.

Privacy and regulatory challenges: The wealth of personal and sensitive healthcare data embedded in HDTs raises significant privacy concerns, necessitating stringent measures to safeguard individuals' information. HDT data collection and sharing must also adhere to existing regulations and standards related to health data protection. This includes compliance with frameworks like the Health Insurance Portability and Accountability Act (HIPAA) in the United States or the General Data Protection Regulation (GDPR) in Europe to ensure the lawful and ethical use of digital twin technology in healthcare. Informed consent in the HDT context, particularly with continuous data collection and updates, necessitates a nuanced approach that prioritizes ongoing communication, transparency, and respect for individuals' autonomy over their health data. Informed consent in HDT involves providing individuals with comprehensive information regarding the purpose, risks, benefits, and potential uses of creating a dynamic digital representation of their health based on personal data. Challenges arise due to the complexity of HDT technology, requiring efforts to ensure individuals fully grasp the implications of this technology.

Security and safety challenges: The reliability of an HDT hinges on the accuracy and fidelity of its digital representation, which, if compromised, can lead to erroneous conclusions and potentially harmful consequences in fields such as healthcare and engineering. Ensuring the reliability of data inputs, calibration processes, and the overall model becomes crucial to maintaining the integrity of the HDT's predictions. Simultaneously, security-related issues pose significant ethical challenges, as the vast amount of personal and sensitive data incorporated into HDTs requires robust protection mechanisms. Unauthorized access or malicious tampering with HDT data could not only jeopardize individual privacy but also result in inaccurate representations and recommendations, potentially impacting real-world entities connected to the HDT.

Data heterogeneity and quality challenges: One of the primary challenges of healthcare data is its inherent heterogeneity and varied quality. Health-related information frequently exists in disparate systems and diverse formats, making integration a complex task. This data, sourced from EHRs, wearable devices, and other digital health tools, often varies significantly in accuracy, completeness, and reliability. The technical intricacies involved in harmonizing this data are considerable, as it necessitates sophisticated methods to ensure consistent and accurate interpretation across different platforms and data types. Moreover, the reliability of HDT models in healthcare heavily relies on the quality of the underlying data. Therefore, establishing robust protocols for data quality assurance is crucial. These protocols must address the nuances of health data, ensuring that the integrated data is not only interoperable but also maintains a high standard of precision and validity. Such meticulous attention to data quality is essential for the successful implementation and effectiveness of HDT models in healthcare.

Data representation and bias challenges: The potential for bias in healthcare data, often reflective of historical disparities and systemic inequities, can be perpetuated in HDTs, leading to unequal outcomes. If not meticulously addressed, this bias may result in disparities in healthcare recommendations and interventions. Furthermore, the development and utilization of AI predictive models within the HDT context carries the risk of encoding and amplifying existing biases present in the training data. Equitable access to HDT technology is another critical concern, as disparities in access may exacerbate existing healthcare inequalities. Ensuring that the benefits of HDTs are accessible across diverse populations becomes imperative to prevent the technology from inadvertently reinforcing societal disparities.

V. Recommendations for Human Digital Twin Research

Addressing these challenges requires a combination of advanced computational techniques, interdisciplinary collaboration, improved data standards, and ongoing refinement of biological models as our understanding of human biology advances. Collaboration with relevant stakeholders, adherence to ethical guidelines, establishment of standards for data integration, model development, and system benchmarking, and a commitment to data privacy are crucial.

Regulatory framework for data security and privacy: The regulatory framework for data security and privacy must evolve to address the unique considerations of HDT in healthcare. It involves collecting diverse and relevant information about individuals to create accurate and representative models. Potential sources include EHRs, wearable devices, genetic data, surveys, mobile apps, and more. It is critical to implement robust data privacy and security measures to protect sensitive health information and adhere to relevant regulations, such as HIPAA or GDPR, and obtain informed consent from individuals. Meanwhile, longitudinal data collection is necessary to capture changes over time. This is particularly important for understanding health trends, monitoring interventions, and predicting future health states.

Standardization: By offering personalized insights, early disease detection, and treatment optimization, HDTs are capable of revolutionizing healthcare. However, there is a lack of standards and guidelines for modeling humans as part of the system, and data standardization is set to become an issue. Thus, it is important to architect data standards that can provide robust information to support human modeling. Additionally, a global perspective should be considered to allow advances in HDTs to have a worldwide impact.

Ethical implications of human digital twins: The HDT technology for personalized medicine may not be accessible to each individual or community, highlighting the unequal distribution of technology. This can cause an additional form of 'digital divide' among persons and populations. It is therefore important to ensure digital equality to advance HDTs for better health. Moreover, unacceptable segmentation and discrimination/injustice may be triggered by patterns identified across a population of HDTs. Thus, there is a need for governance mechanisms to safeguard the rights of individuals who own HDTs, ensure data security and privacy, and foster transparency and fairness of data usage, health equity, and all derived benefits at both individual and wider societal levels.

Increase community engagement: Efforts should ensure community input in the development and implementation of HDTs. HDT research should promote ongoing bi-directional engagement, ensuring the inclusive involvement of diverse community perspectives. Some of these strategies include focus groups, town hall meetings, dissemination forums, and interactive digital platforms tailored to the different community stakeholders. Further, these forms of engagement should endure cultural sensitivity and inclusivity to avoid biases and ensure relevance across different communities and stakeholders.

Funding implications of human digital twins: Funding agencies would have a significant amount of impact on advancing HDTs. Specifically, they can offer financial support, promote research prioritization, interdisciplinary collaboration, data sharing, and standards, ethical and regulatory frameworks, invest in educational programs and public outreach efforts, invest in the development of enabling technologies such as HPC and ancillary resources, help ensure that there is sustainable support for maintaining and updating HDTs, continuously monitor the progress and impact of HDT, and encourage collaboration and information sharing on a global scale.

Specifically, a substantial initial investment is required to develop the necessary cyberinfrastructure. This includes hardware, software, and network capabilities to support complex data processing and simulation tasks. Leveraging collaborative funding from various sources, such as government grants, private sector investments, the healthcare industry and academic institutions, can provide a more robust financial foundation. Ongoing operational costs, including maintenance of technology infrastructure, data storage, and security, require sustainable funding sources. This might involve subscription models, partnerships with healthcare providers, or government support. Investing in training programs to develop a skilled workforce capable of building, maintaining, and utilizing DT technology is essential. This includes funding for educational programs, workshops, and certification courses. Such goals can be embedded in the request for proposals (RFP) and broad agency announcements (BAA) as the NSF and NIH routinely do in their portfolio.

VI. Conclusions

The emerging HDT technology offers tremendous opportunities for personalized healthcare, predictive interventions, remote monitoring, and medical research advancements. It has the potential to revolutionize healthcare by integrating with the healthcare sector, information technology, AI industries, the government and private sector stakeholders. While there are still many obstacles and challenges in implementing human digital twins for healthcare, we envision a bright future for HDT with cross-disciplinary collaborations and efforts by all the stakeholders including the government, industry, academia, and private sectors.

References

1. Tao F, Xiao B, Qi Q, Cheng J, Ji P. Digital twin modeling. Journal of Manufacturing Systems. 2022;64:372-389.
2. Pylianidis C, Osinga S, Athanasiadis IN. Introducing digital twins to agriculture. Computers and Electronics in Agriculture. 2021;184:105942.
3. Fuller A, Fan Z, Day C, Barlow C. Digital twin: Enabling technologies, challenges and open research. IEEE access. 2020;8:108952-108971.
4. National Academies of Science E, and Medicine. Foundational Research Gaps and Future Directions for Digital Twins. https://www.nationalacademies.org/our-work/foundational-research-gaps-and-future-directions-for-digital-twins
5. Katsoulakis E, Wang Q, Wu H, Shahriyari L, Fletcher R, Liu J, Achenie L, Liu H, Jackson P, Xiao Y, Syeda-Mahmood T, Tuli R, and Deng J. Digital Twins for Health: A Scoping Review. npj Digital Medicine, 7, 77 (2024). https://doi.org/10.1038/s41746-024-01073-0
6. Digital Twins for Health Consortium. https://dt4h.org
7. Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support. https://access-ci.org
8. Partnership to Advance Throughput Computing. https://path-cc.io/
9. Hancock DY, Fischer J, Lowe JM, et al. Jetstream2: Accelerating cloud computing via Jetstream. Practice and Experience in Advanced Research Computing. 2021:1-8.
10. OMOP Common Data Model. https://ohdsi.github.io/CommonDataModel/
11. Peterson KJ, Jiang G, Liu H. A corpus-driven standardization framework for encoding clinical problems with HL7 FHIR. J Biomed Inform. Oct 2020;110:103541. doi:10.1016/j.jbi.2020.103541
12. The HL7 consumer mobile health application function framework (cMHAFF) guideline https://www.hl7.org/implement/standards/product_brief.cfm?product_id=476
13. Open mHealth Tools. https://www.openmhealth.org/

14. Mustra M, Delac K, Grgic M. Overview of the DICOM standard. IEEE; 2008:39-44.
15. Folk M, Heber G, Koziol Q, Pourmal E, Robinson D. An overview of the HDF5 technology suite and its applications. 2011:36-47.
16. Freimuth RR, Milius RP, Murugan M, Terry M. Clinical genomic data on FHIR®: Case studies in the development and adoption of the Genomics Reporting Implementation Guide. Genomic Data Sharing. Elsevier; 2023:91-110.
17. Kottmann R, Gray T, Murphy S, et al. A standard MIGS/MIMS compliant XML Schema: toward the development of the Genomic Contextual Data Markup Language (GCDML). Omics a journal of integrative biology. 2008;12(2):115-121.
18. Taylor CF, Paton NW, Lilley KS, et al. The minimum information about a proteomics experiment (MIAPE). Nature biotechnology. 2007;25(8):887-893.
19. Haug K, Salek RM, Steinbeck C. Global open data management in metabolomics. Current opinion in chemical biology. 2017;36:58-63.
20. Deutsch EW, Ball CA, Berman JJ, et al. Minimum information specification for in situ hybridization and immunohistochemistry experiments (MISFISHIE). Nature biotechnology. 2008;26(3):305-312.
21. Dang Y, Li F, Hu X, et al. Systematic Design and Evaluation of Social Determinants of Health Ontology (SDoHO). arXiv preprint arXiv:221201941. 2022;
22. Introduction to Pararead. http://code.databio.org/pararead/
23. Lawson JT, Smith JP, Bekiranov S, Garrett-Bakelman FE, Sheffield NC. COCOA: coordinate covariation analysis of epigenetic heterogeneity. Genome Bio. 2020;21(1):1-23.
24. Feng J, Ratan A, Sheffield NC. Augmented Interval List: a novel data structure for efficient genomic interval search. Bioinformatics. 2019;35(23):4907-4911.
25. Van der Auwera GA, O'Connor BD. Genomics in the cloud: using Docker, GATK, and WDL in Terra. O'Reilly Media; 2020.
26. Nagraj V, Magee NE, Sheffield NC. LOLAweb: a containerized web server for interactive genomic locus overlap enrichment analysis. Nucleic acids research. 2018;46(W1):W194-199.
27. Singh R, Lanchantin J, Robins G, Qi Y. DeepChrome: deep-learning for predicting gene expression from histone modifications. Bioinformatics. 2016;32(17):i639-i648.
28. Syeda-Mahmood T, Wang F, Beymer D, Amir A, Richmond M, Hashmi S. AALIM: Multimodal mining for cardiac decision support. IEEE; 2007:209-212.
29. Shi L, Syeda-mahmood T, Baldwin T. Improving neural models for radiology report retrieval with lexicon-based automated annotation. 2022:3457-3463.
30. Bompelli A, Wang Y, Wan R, et al. Social and behavioral determinants of health in the era of artificial intelligence with electronic health records: A scoping review. Health Data Science. 2021;2021
31. Eichstaedt JC, Kern ML, Yaden DB, et al. Closed-and open-vocabulary approaches to text analysis: A review, quantitative comparison, and recommendations. Psychological Methods. 2021;26(4):398.
32. Zadeh A, Chen M, Poria S, Cambria E, Morency L-P. Tensor fusion network for multimodal sentiment analysis. arXiv preprint arXiv:170707250. 2017;
33. Fu S, Chen D, He H, et al. Clinical concept extraction: A methodology review. J Biomed Inform. Sep 2020;109:103526. doi:10.1016/j.jbi.2020.103526
34. Arya V, Bellamy RK, Chen P-Y, et al. AI explainability 360: hands-on tutorial. 2020:696-696.

35. Dash S, Gunluk O, Wei D. Boolean decision rules via column generation. Advances in neural information processing systems. 2018;31
36. Wei D, Dash S, Gao T, Gunluk O. Generalized linear rule models. PMLR; 2019:6687-6696.
37. Ribeiro MT, Singh S, Guestrin C. " Why should i trust you?" Explaining the predictions of any classifier. 2016:1135-1144.
38. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Advances in neural information processing systems. 2017;30
39. Hind M, Wei D, Campbell M, et al. TED: Teaching AI to explain its decisions. 2019:123-129.
40. Van der Maaten L, Hinton G. Visualizing data using t-SNE. Journal of machine learning research. 2008;9(11)
41. Abdi H, Williams LJ. Principal component analysis. Wiley interdisciplinary reviews: computational statistics. 2010;2(4):433-459.
42. McInnes L, Healy J, Melville J. Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:180203426. 2018;
43. Hou J, Deng J, Li C, Wang Q. Tracing and forecasting metabolic indices of cancer patients using patient-specific deep learning models. Journal of personalized medicine. 2022;12(5):742.
44. Lopes-Ramos CM, Kuijjer ML, Ogino S, et al. Gene regulatory network analysis identifies sex-linked differences in colon cancer drug metabolism. Cancer research. 2018;78(19):5538-5547.
45. Szabó A, Merks RM. Cellular potts modeling of tumor growth, tumor invasion, and tumor evolution. Frontiers in oncology. 2013;3:87.
46. Hormuth DA, Jarrett AM, Feng X, Yankeelov TE. Calibrating a predictive model of tumor growth and angiogenesis with quantitative MRI. Annals of biomedical engineering. 2019;47:1539-1551.
47. Zhou J, Cui G, Hu S, et al. Graph neural networks: A review of methods and applications. AI open. 2020;1:57-81.
48. He H, Yuan D, Wu Y, Cao Y. Pharmacokinetics and pharmacodynamics modeling and simulation systems to support the development and regulation of liposomal drugs. Pharmaceutics. 2019;11(3):110.
49. Hunter DR, Goodreau SM, Handcock MS. Goodness of fit of social network models. Journal of the american statistical association. 2008;103(481):248-258.
50. Smith ER, Conrey FR. Agent-based modeling: A new approach for theory building in social psychology. Personality and social psychology review. 2007;11(1):87-104.
51. Salakhutdinov R. Learning deep generative models. Annual Review of Statistics and Its Application. 2015;2:361-385.
52. Xu J, Kim S, Song M, et al. Building a PubMed knowledge graph. Scientific data. 2020;7(1):205.
53. Frei E, Stamm M, Buddeberg-Fischer B. Mentoring programs for medical students-a review of the PubMed literature 2000-2008. BMC medical education. 2010;10:1-14.
54. Hardy TM, Hansen MJ, Bahamonde RE, Kimble-Hill AC. Insights Gained into the Use of Individual Development Plans as a Framework for Mentoring NIH Postbaccalaureate Research Education Program (PREP) Trainees. Journal of chemical education. 2021;99(1):417-427.