Clément Cormi, Arnaud Allemang--Trivalle, Olivia Chevalier, Matthieu Robert, Eric Vibert (2025): From Enhancing Individual to Team Capacities: Stakeholders' Insights on Generative-AI and Digital-Twin Adoption in Surgery. In: Proceedings of the 10th International Conference on Infrastructures in Healthcare - Paper, (ISSN 2510-2591), DOI: 10.48340/ihc2025_p005

From Enhancing Individual to Team Capacities: Stakeholders' Insights on Generative-AI and Digital-Twin Adoption in Surgery

Clément Cormi [1,*†], Arnaud Allemang--Trivalle [1, 2, 3,†],
Olivia Chevalier[1, 4, 5], Matthieu Robert[6], Eric Vibert[1, 7, 8]
[1] Chaire Innovation Bloc Opératoire Augmenté (BOPA), AP-HP, Institut Mines Telecom, Université Paris Saclay, Villejuif, France
[2] Haptic Intelligence Department, Max Planck Institute for Intelligent Systems, Stuttgart, Germany
[3] IMT Atlantique, Lab-STICC, UMR CNRS 6285, Brest, France
[4] Institut Mines Télécom Business School (IMT-BS), Évry-Courcouronnes, France
[5] PEPR O2R
[6] GHU AP-HP.Sorbonne Université, Mission expérience patient, Direction de la Qualité, Paris, France
[7] Paul Brousse Hospital – AP-HP, Villejuif, France
[8] Université Paris-Saclay, Inserm, UMR-S 1193, Paris, France
[*]*Corresponding author: Clément Cormi, clement.cormi@inserm.fr*
[†]*Contributed equally to this work and joint first authors.*

Abstract. This paper distills insights from two three-hour workshops involving 34 multidisciplinary participants that explored the individual and collective applications of generative AI (GenAI) and digital twins (DTs) in surgery. Participants began by clarifying the meaning of GenAI and DTs in surgical contexts. GenAI was perceived as an individual aid to generate clinical notes, referral letters, and explanations to patients, while DTs were praised to render patient-specific anatomy, forecast trocar placement, and provide teams a live common operating-room view. Together the technologies create a layered data infrastructure that toggles between solo efficiency and collective sense-making. Participants highlighted adoption challenges such as ethical considerations, data accuracy, and infrastructural readiness. Design imperatives emerge for a nuanced integration of these technologies into surgical practices, such as modeling uncertainty, preserving clinical reasoning, and introducing tools through existing hospital systems, highlighting the essential role of human-computer interaction experts.

1 Introduction and Background

Hospitals are under increasing pressure to deliver complex surgical care efficiently despite staffing shortages and rising patient expectations. Against this backdrop, generative artificial intelligence (GenAI) and digital twins (DTs) are being explored as potential tools not only for automating clinical tasks but also as socio-technical resources that could redistribute work within surgical teams. GenAI, especially through large language models (LLMs), has shown promise in reducing clinicians' administrative burden by assisting in the generation of medical reports, patient letters, and informed consent forms, as well as supporting communication during consultations (Hassan et al., 2023; Ayers et al., 2023; You and Gui, 2021). In parallel, DTs have been leveraged to simulate patient-specific anatomy and physiology, offering predictive insights for surgical planning and intraoperative guidance (Croatti et al., 2020; Sun et al., 2023; Mendizabal et al., 2023).

While these technologies are advancing, their adoption in surgical practice remains limited, less by technical immaturity than by misalignment with organizational routines, practical workflow constraints, trust norms, and team dynamics (Riedel et al., 2022; Gilbert et al., 2023). DTs, for instance, are often categorized into physiological, anatomical, and organizational types, each promising precision and personalization but raising questions about data fidelity, visualization complexity, and patient comprehension (Braun, 2021; Pascual et al., 2023). Similarly, although GenAI can speed up text production, its uneven performance across clinical content types and lack of explainability pose barriers to safe integration in complex care settings (Balel, 2023; Bumgardner et al., 2023).

Beyond technical concerns, surgical teamwork adds a layer of socio-technical complexity. Evidence from trauma-resuscitation studies shows that surgical outcomes hinge on real-time coordination, error recovery, and information hand-off among team members (Sarcevic et al., 2012; Webman et al., 2016). Successful operations therefore depend on synchronized mental models, shared spatial reasoning, and fluid turn-taking among surgeons, anesthesiologists, and nurses (Suliburk et al., 2019). Emerging CSCW research warns that even well-designed AI tools often create "data work" rather than relief when deployed without regard for coordination dynamics or accountability structures (Bjørnstad and Ellingsen, 2019). This makes surgery a critical site for studying how GenAI and DTs might redistribute cognitive as well as coordinative labor.

This paper examines how surgical stakeholders represent and perceive the potential roles of GenAI and DTs into their everyday work practices, with a particular focus on how these technologies may shift responsibilities between individual clinicians and the surgical team as a whole. While technical capabilities are advancing rapidly, there is limited understanding of how these systems are interpreted, trusted, and expected to function within complex surgical environments. This gap is especially relevant in contexts where high-stakes decisions, division of labor, and collaboration unfold in real time.

To address this, the main research question is: *How do surgical team members represent the potential roles of generative AI and digital twins in enhancing individual and team activities, and what integration models do they believe could make these technologies practically useful the surgical pathway?*

2 Methods

At the 2023 *Week-end de l'innovation chirurgicale*, a two-day research–practice forum on the future of surgery, we ran two three-hour workshops in parallel, one on generative AI and the other on Digital Twins, to explore anticipated uses, risks and organizational requirements for the two technologies. Participants were free to choose their session, resulting in 19 stakeholders attending the GenAI discussion and 15 the DT discussion, balancing clinical, technical, and managerial viewpoints while still allowing for deep, interactive debate. All participants gave their verbal consent to take part in the workshops and agreed to the recording of discussions for research purposes.

2.1 Participants

Thirteen aliases are used for verbatim quoting.

GenAI_P1 , ***GenAI_P2*** , ***GenAI_P3*** are male senior surgeons (40–55 y, public hospitals); ***GenAI_P4*** and ***GenAI_P5*** are male junior surgeons (25–35 y, public hospitals).

DT_P1 , ***DT_P2*** , ***DT_P3*** are male senior surgeons (45–55 y, private or public hospitals); ***DT_P4*** is a male full-professor sociologist (60–65 y); ***DT_P5*** a male sociology PhD candidate (25–30 y); ***DT_P6*** a female lecturer (45–50 yrs, public university).

The remaining participants comprised 14 additional GenAI attendees (11 M / 3 F: resident surgeons, anesthesiologists, engineers, student designers, tech CEOs) and 9 additional DT attendees (5 M / 4 F: resident surgeon, OR staff, designer, engineer, researcher, head of division).

2.2 Workshop Structure

Each workshop, co-facilitated by two domain researchers, followed a three-step script: (i) analysis of a brief GenAI clinical case or joint definition of DT concepts, (ii) a moderated brainstorm on use cases and concerns, and (iii) mapping of prerequisites for safe adoption. Sessions were held in separate rooms to prevent cross-pollination of ideas. Each room ended with a short wrap-up, followed by a joint debrief the next morning. All participants joined a shared WhatsApp group to capture additional reflections that were included in the data corpus.

2.3 Analysis

All sessions were video- and audio-recorded with two non-participating researchers in each room taking field notes. Recordings were transcribed and combined with the notes and chat logs to create the analysis corpus. Using thematic analysis (Braun and Clarke, 2006), three authors independently coded the data, clustered the codes into themes, and labeled them *major* when raised by at least one-third of the speakers in either workshop and *minor* otherwise. Coding differences were resolved in a consensus meeting, and the final thematic map was reviewed and approved by all co-authors.

3 Findings

The four most practice-relevant themes that emerged from our workshop are presented in the following sections, namely (1) conceptual ambiguity, (2) patient information and autonomy, (3) administrative automation and OR logistics, and (4) surgical planning and training.

3.1 Conceptual ambiguity: "What is GenAI / DT?"

Difficulties in defining the concepts of GenAI and DTs emerged in both workshop groups, sometimes to the point of confusing them. The following conversation between two surgeons from the GenAI group illustrates it well:

> -"*ChatGPT could suggest the best surgical strategy based on a patient's information. When autonomous vehicles were trained to take a turn, all the curves were tested, and the one where the car didn't end up in the wall was selected.*" ***GenAI_P1***.
>
> -"*I am not sure ChatGPT works like that. I think it can provide all the possible surgical strategies, but it is not specific enough to select the best one. That responsibility falls to the surgeon.*" ***GenAI_P4***

In the second group, definition challenges mainly emerged from the real-time nature inherent in DTs, and the delimitation between DT and simulation raised a palpable skepticism concerning the starkness of these differences:

> *"While all DTs inherently involve simulation, not all simulations are DTs."* ***DT_P6***

Participants emphasized that without shared, operational definitions of GenAI and DT, it is difficult to make informed decisions about procurement, regulatory oversight, or professional training.

3.2 Patient information and autonomy

LLM chatbots were seen as ideal for routine questions between visits, allowing patients to have their medical conditions, necessary treatments, and post-operative aftermath comprehensively explained to them. "*After a transplant having a tool* *that can answer patients' questions like 'Can I eat this?' or 'Can I ride my bike yet?' and so on, that would be fabulous!*"***GenAI_P1***. However, all questions must be logged so that clinicians can review any misconceptions during follow-up visits or use the inquiries submitted to the chatbot to guide future discussions.

In the clinic, DTs were portrayed as visual bridges for consent helping to build a mutual understanding and support collaborative therapeutic decisions. Fidelity, however, must vary with patient preference. "*Not all patients want to know the details of their illness, nor know the details of their care pathway and intervention.*"***DT_P1***. However, concerns were raised about whether excessive realism could have legal repercussions for healthcare professionals. "*If the surgeon deviates from what was planned and something goes wrong... would the patient sue the surgeon?*"***DT_P4***.

Effective empowerment, therefore, hinges on adjustable detail, auditable logs, and clear liability.

3.3 Administrative automation & OR logistics

A live demonstration of an AI-powered scribe generating medical consultation reports—initiated by a participant—sparked divided opinions. Senior clinicians cautioned, " *We may be missing the point by seeking to automate it.*" ***GenAI_P2***, while junior staff argued that most clinical notes are repetitive. A consensus emerged: GenAI should be limited to low-risk documentation and always require human validation.

For preparing instrument trays, participants preferred simple usage analytics over a full organizational twin: "*We could generate a list of necessary instruments based on the operating program, and we would stop opening boxes that end up in the trash*" ***GenAI_P1***. DTs were judged over-engineered for this task. In sum, GenAI can cut paperwork and basic software can cut waste, freeing time for clinical work.

The overarching sentiment was that LLMs have the potential to significantly reduce paperwork time, allowing healthcare providers more time for clinical duties. Yet, the specific administrative tasks these changes affect require further research and should be connected to physicians' roles, such as clinical work, training, etc.

3.4 Surgical planning & training

Spatial tasks pulled DTs to the fore. Surgeons want a patient-specific twin that merges imaging with deformation models: "*The idea behind the digital twin is to recreate the three-dimensional anatomy, the physiological behavior of the organs, any deformations, and possibly the adhesions that might be encountered during surgery.*"***DT_P1***.

Some participants are involved in multiple digital twin projects aimed at improving liver surgery planning, supporting both pre-operative and intraoperative discussions to determine optimal resection strategies or grafting approaches. Their goals were brought to the fore through the development of two key projects. The first is an anatomical digital twin of the liver, designed to assist in identifying the regions to be resected based on the localization of tumors, vasculature, and other anatomical structures. The liver is an opaque and deformable organ, and intraoperative navigation typically relies on ultrasound, which remains highly operator-dependent. The anatomical digital twin allows for the superimposition of pre-operative CT data onto the physical liver during surgery, helping to localize target areas while accounting for deformations caused by surgical manipulation or artificial ventilation. This tool supports collaborative decision-making among surgeons to determine the most appropriate surgical approach. A second line of research focuses on hemodynamic digital twins. Derived from one-dimensional to three-dimensional models, these simulations are intended to guide surgical planning for liver grafts or hepatectomies. Both procedures require careful assessment of the future liver remnant to avoid postoperative liver failure. Hemodynamic digital twins simulate the functional remnant liver volume, thereby supporting collaborative and informed surgical decision-making.

Despite these advancements, participants acknowledged that achieving a perfect model remains an aspirational goal: "*To create an almost perfect digital twin, you would ideally have to put sensors everywhere, which seems impossible.*"***DT_P4***.

For education, GenAI was praised for generating rich case variations, while VLM images were deemed "*At least it's pretty, even if it's far from a hepatectomy.*" ***GenAI_P3***. Participants expect DT-based rehearsal once model creation is faster, offering immersive practice beyond today's static phantoms.

4 Discussion

Our workshops revealed persistent conceptual ambiguity. Clinicians, engineers and researchers do not share stable definitions of either GenAI or DTs. This gap suggests a self-limiting approach to exploring and utilizing DTs and GenAI. Participatory-design methods can narrow the gap by turning users into co-designers and demystifying AI "black boxes" (Wagner, 2018). This diversity in perceptions highlights the need for a standardized definition and categorization of DTs and GenAI to bridge the comprehension gap and leverage their full potential (Hassani et al., 2022).

As summarized in Table I, participants considered GenAI as tools that boost individual efficiency, whereas DTs support team coordination and shared understanding.

Table I: Contrasting stakeholder visions of the applications of GenAI and DT.

GenAI → Enhances individual work	**DT → Enhances team work**
AI scribe for notes	Shared 3-D anatomy for orientation
Patient Q&A chatbot	Joint selection of graft strategy
Automatic instrument tray list	Live common OR view

GenAI could alleviate the burden of low-risk auxiliary tasks for healthcare professionals by simplifying report writing and streamlining administrative processes. In contrast, DTs inherently promote a team-oriented approach in healthcare, serving as shared representations that enhance collaborative discussions among practitioners. They are increasingly recognized for their potential to act as virtual meeting grounds, allowing professionals to deliberate on patient cases and focusing on optimizing patient selection.

However, our findings also indicate a palpable skepticism from healthcare professionals, particularly regarding the current reliability of GenAI and DTs and practical integration into ORs. GenAI, for instance, has shown contrasted performance: it appears to be adequate for responding to patient inquiries (Samaan et al., 2023) but less reliable for supporting professional medical decision-making (Ali, 2023; Xie et al., 2023). This contrast highlights the critical need to train LLMs on domain-specific data sources, such as medical publications from PubMed and structured clinical records, to improve their reliability and applicability in medical contexts. Similarly, intraoperative visualizations of digital twins lack sufficient precision, especially for deformable organs (Acidi et al., 2023; Gavriilidis et al., 2022). Movements and deformations caused by mechanical ventilation or manual manipulation introduce variability that existing systems struggle to account for. Although several recalibration techniques have been proposed, few are fully automated, and most require additional equipment (e.g.,RGB-D cameras), making widespread adoption technically challenging.

Beyond reliability concerns, participants highlighted deeper structural challenges to integrating GenAI and DTs into hospital settings. While these technologies promise workflow efficiencies, they also introduce new burdens, especially around data generation and management. Clinical data must be continuously produced, cleaned, annotated, and maintained to support AI systems. This hidden "data work" often becomes a cumbersome and undervalued task for medical practitioners affecting adoption and sustainability (Bonde et al., 2019; Arndt et al., 2017; Sinsky et al., 2016; Bjørnstad and Ellingsen, 2019; Cabitza et al., 2019; Vallo Hult et al., 2019). One promising approach is the implementation of binary classifiers, which could streamline the process and reduce the burden on healthcare providers (Oukelmoun et al., 2023). But even with better tooling, running GenAI and DT applications at scale requires high-performance computing, often provisioned through commercial American cloud services, raising concerns over data sovereignty and regulatory compliance in European contexts. Early integration of policy considerations is crucial to address the risks of failure (Centivany, 2016).

The use-cases envisioned in our workshops inevitably mirror today's OR infrastructure. One of the most promising enablers for realizing the potential of both GenAI and DTs in surgery is the integration of OR "black boxes': comprehensive monitoring systems inspired by aeronautics that capture synchronized data streams such as audio, video, physiological data, and motion throughout procedures (Allemang--Trivalle, 2023). While GenAI use cases were oriented towards augmenting individuals, coupling it with black box data opens new avenues for team-oriented support. Access to continuous, multimodal OR streams could enable GenAI to automate tasks such as generating material lists or summarizing key intraoperative events, while also helping reduce waste from unused surgical tools (Guetter et al., 2018; Chasseigne et al., 2018). For DTs, these same data streams would provide real-time inputs for updating anatomical or physiological models, supporting dynamic planning, retrospective analysis of surgical errors (Suliburk et al., 2019; Jung et al., 2020), and the generation of objective operative reports (Wauben et al., 2011; Eryigit et al., 2019). "Black boxes" could also help streamline OR logistics (Uhl et al., 2022; Augusto et al., 2018, 2010), transforming DTs into a bridge between surgical simulation and real-time coordination.

5 Conclusion

Based on two workshops with 34 participants from various backgrounds (surgery, sociology, engineering, health management, and tech entrepreneurship), this study investigates stakeholders' representations of generative AI (GenAI) and digital twins (DTs) in surgery. Four identified major themes structure these representations: conceptual ambiguity, (2) patient information and autonomy, (3) administrative automation and OR logistics, and (4) surgical planning and training.

Stakeholders framed GenAI as a valuable tool for relieving healthcare professionals of individual low-risk auxiliary tasks like writing reports and managing administrative processes. Conversely, DTs were highly regarded for their ability to support team-oriented approaches, acting as catalysts for collaborative discussions and optimizing patient care strategies.

However, GenAI's potential in augmenting team dynamics has been poorly explored, although innovative applications like surgical "black boxes" could create new possibilities for improving teamwork and operational efficiency in operating rooms. Hence, further work should explore a less dichotomized vision than the *individual-GenAI vs. collaborative-DTs* we highlight here.

Persistent confusion about what constitutes an LLM, a DT, or even "real-time" emerged as a barrier. A concise, task-oriented glossary is a necessary starting point for procurement, regulation, and training.

Finally, doubts emerged about the reliability of such tools, particularly regarding the generated information's accuracy and *integration-in-practices*, and our work highlighted the need for interdisciplinary solid collaborations between surgical, AI, and CSCW/HCI experts. Future studies should observe OR teams piloting early GenAI scribes and anatomical twins to test whether the perceived split between individual and team support materializes in practice.

Acknowledgments

We thank the Fondation AP-HP, the Fondation Mines Telecom and BOPA's sponsors for funding the BOPA Chair and the WIC; the International Max Planck Research School for Intelligent Systems (IMPRS-IS) and the AI@IMT program of the French National Research Agency (ANR) for supporting Arnaud Allemang--Trivalle.

References

Acidi, B., M. Ghallab, S. Cotin, E. Vibert, and N. Golse (2023): 'Augmented reality in liver surgery'. *Journal of Visceral Surgery*, vol. 160, no. 2, pp. 118–126.

Ali, M. J. (2023): 'ChatGPT and lacrimal drainage disorders: Performance and scope of improvement'. *Ophthalmic Plastic & Reconstructive Surgery*, vol. 39, no. 3, pp. 221–225.

Allemang--Trivalle, A. (2023): 'Enhancing surgical team collaboration and situation awareness through multimodal sensing'. In: *Proceedings of the 25th International Conference on Multimodal Interaction*. New York, NY, USA, pp. 716–720, Association for Computing Machinery.

Arndt, B. G., J. W. Beasley, M. D. Watkinson, J. L. Temte, W.-J. Tuan, C. A. Sinsky, and V. J. Gilchrist (2017): 'Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations'. *Annals of Family Medicine*, vol. 15, no. 5, pp. 419–426.

Augusto, V., M. Murgier, and A. Viallon (2018): 'A modelling and simulation framework for intelligent control of emergency units in the case of major crisis'. In: *2018 Winter Simulation Conference (WSC)*. pp. 2495–2506.

Augusto, V., X. Xie, and V. Perdomo (2010): 'Operating theatre scheduling with patient recovery in both operating rooms and recovery beds'. *Computers & Industrial Engineering*, vol. 58, no. 2, pp. 231–238.

Ayers, J. W., A. Poliak, M. Dredze, E. C. Leas, Z. Zhu, J. B. Kelley, D. J. Faix, A. M. Goodman, C. A. Longhurst, M. Hogarth, and D. M. Smith (2023): 'Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum'. *JAMA Internal Medicine*, vol. 183, no. 6, pp. 589–596.

Balel, Y. (2023): 'Can ChatGPT be used in oral and maxillofacial surgery?'. *Journal of Stomatology, Oral and Maxillofacial Surgery*, pp. 101471.

Bjørnstad, C. and G. Ellingsen (2019): 'Data work: A condition for integrations in health care'. *Health Informatics Journal*, vol. 25, no. 3, pp. 526–535.

Bonde, M., C. Bossen, and P. Danholt (2019): 'Data-work and friction: Investigating the practices of repurposing healthcare data'. *Health Informatics Journal*, vol. 25, no. 3, pp. 558–566.

Braun, M. (2021): 'Represent me: please! Towards an ethics of digital twins in medicine'. *Journal of Medical Ethics*, vol. 47, no. 6, pp. 394–400. Publisher: Institute of Medical Ethics Section: Feature article.

Braun, V. and V. Clarke (2006): 'Using thematic analysis in psychology'. *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101.

Bumgardner, V. K. C., A. Mullen, S. Armstrong, C. Hickey, and J. Talbert (2023): 'Local large language models for complex structured medical tasks'. arXiv:2308.01727 [cs].

Cabitza, F., G. Ellingsen, A. Locoro, and C. Simone (2019): 'Repetita iuvant: Discovering and supporting redundancy in health information technology'. *Computer Supported Cooperative Work (CSCW)*, vol. 28, pp. 61–94.

Centivany, A. (2016): 'Policy as embedded generativity: A case study of the emergence and evolution of HathiTrust'. In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. San Francisco, California, USA, pp. 926–940, Association for Computing Machinery.

Chasseigne, V., G. Leguelinel-Blache, T. L. Nguyen, R. de Tayrac, M. Prudhomme, J. M. Kinowski, and P. Costa (2018): 'Assessing the costs of disposable and reusable supplies wasted during surgeries'. *International Journal of Surgery*, vol. 53, pp. 18–23.

Croatti, A., M. Gabellini, S. Montagna, and A. Ricci (2020): 'On the integration of agents and digital twins in healthcare'. *Journal of Medical Systems*, vol. 44, no. 9, pp. 161.

Eryigit, , F. W. van de Graaf, and J. F. Lange (2019): 'A systematic review on the synoptic operative report versus the narrative operative report in surgery'. *World Journal of Surgery*, vol. 43, no. 9, pp. 2175–2185.

Gavriilidis, P., B. Edwin, E. Pelanis, E. Hidalgo, N. de'Angelis, R. Memeo, L. Aldrighetti, and R. P. Sutcliffe (2022): 'Navigated liver surgery: State of the art and future perspectives'. *Hepatobiliary Pancreatic Diseases International*, vol. 21, no. 3, pp. 226–233.

Gilbert, S., H. Harvey, T. Melvin, E. Vollebregt, and P. Wicks (2023): 'Large language model AI chatbots require approval as medical devices'. *Nature Medicine*, pp. 1–3. Publisher: Nature Publishing Group.

Guetter, C. R., B. J. Williams, E. Slama, A. Arrington, M. C. Henry, M. G. Möller, J. E. Tuttle-Newhall, S. Stein, and M. Crandall (2018): 'Greening the operating room'. *The American Journal of Surgery*, vol. 216, no. 4, pp. 683–688.

Hassan, A. M., J. A. Nelson, J. H. Coert, B. J. Mehrara, and J. C. Selber (2023): 'Exploring the potential of artificial intelligence in surgery: Insights from a conversation with ChatGPT'. *Annals of Surgical Oncology*, vol. 30, no. 7, pp. 3875–3878.

Hassani, H., X. Huang, and S. MacFeely (2022): 'Impactful digital twin in the healthcare revolution'. *Big Data and Cognitive Computing*, vol. 6, no. 3, pp. 83.

Jung, J. J., P. Jüni, D. W. Gee, Y. Zak, J. Cheverie, J. S. Yoo, J. M. Morton, and T. Grantcharov (2020): 'Development and evaluation of a novel instrument to measure severity of intraoperative events using video data'. *Annals of Surgery*, vol. 272, no. 2, pp. 220–226.

Mendizabal, A., A. Odot, and S. Cotin (2023): 'Chapter 5 - Deep learning for real-time computational biomechanics'. In: F. Chinesta, E. Cueto, Y. Payan, and J. Ohayon (eds.): *Reduced Order Models for the Biomechanics of Living Organs*, Biomechanics of Living Organs. Academic Press, pp. 95–126.

Oukelmoun, A., N. Semmar, G. De Chalendar, E. Habran, E. Vibert, E. Goblet, M. Oukelmoun, and M.-A. Allard (2023): 'A study on the relevance of generic word embeddings for sentence classification in hepatic surgery'. In: *2023 20th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA)*. pp. 1–8.

Pascual, H., X. M. Bruin, A. Alonso, and J. Cerdà (2023): 'A Systematic Review on Human Modeling: Digging into Human Digital Twin Implementations'. arXiv:2302.03593 [cs].

Riedel, P., M. Riesner, K. Wendt, and U. Aßmann (2022): 'Data-driven digital twins in surgery utilizing augmented reality and machine learning'. In: *2022 IEEE International Conference on Communications Workshops (ICC Workshops)*. pp. 580–585. ISSN: 2694-2941.

Samaan, J. S., Y. H. Yeo, N. Rajeev, L. Hawley, S. Abel, W. H. Ng, N. Srinivasan, J. Park, M. Burch, R. Watson, O. Liran, and K. Samakar (2023): 'Assessing the accuracy of responses by the language model ChatGPT to questions regarding bariatric surgery'. *Obesity Surgery*, vol. 33, no. 6, pp. 1790–1796.

Sarcevic, A., I. Marsic, and R. S. Burd (2012): 'Teamwork Errors in Trauma Resuscitation'. *ACM Trans. Comput.-Hum. Interact.*, vol. 19, no. 2.

Sinsky, C., L. Colligan, L. Li, M. Prgomet, S. Reynolds, L. Goeders, J. Westbrook, M. Tutty, and G. Blike (2016): 'Allocation of physician time in ambulatory practice: A time and motion study in 4 specialties'. *Annals of Internal Medicine*, vol. 165, no. 11, pp. 753.

Suliburk, J. W., Q. M. Buck, C. J. Pirko, N. N. Massarweh, N. R. Barshes, H. Singh, and T. K. Rosengart (2019): 'Analysis of Human Performance Deficiencies Associated With Surgical Adverse Events'. *JAMA Network Open*, vol. 2, no. 7, pp. e198067.

Sun, T., X. He, and Z. Li (2023): 'Digital twin in healthcare: Recent updates and challenges'. *DIGITAL HEALTH*, vol. 9, pp. 20552076221149651. Publisher: SAGE Publications Ltd.

Uhl, L., V. Augusto, V. Lemaire, Y. Alexandre, F. Jardinaud, P. Bercelli, and S. Aloui (2022): 'Progressive prediction of hospitalisation and patient disposition in the emergency department'. In: *2022 IEEE International Conference on Big Data (Big Data)*. pp. 1719–1728.

Vallo Hult, H., A. Hansson, L. Svensson, and M. Gellerstedt (2019): 'Flipped healthcare for better or worse'. *Health Informatics Journal*, vol. 25, no. 3, pp. 587–597.

Wagner, I. (2018): 'Critical reflections on participation in design'. In: V. Wulf, V. Pipek, D. Randall, M. Rohde, K. Schmidt, and G. Stevens (eds.): *Socio-Informatics*. Oxford University Press, p. 0.

Wauben, L. S. G. L., W. M. U. van Grevenstein, R. H. M. Goossens, F. H. van der Meulen, and J. F. Lange (2011): 'Operative notes do not reflect reality in laparoscopic cholecystectomy'. *The British Journal of Surgery*, vol. 98, no. 10, pp. 1431–1436.

Webman, R. B., J. L. Fritzeen, J. Yang, G. F. Ye, P. C. Mullan, F. G. Qureshi, S. H. Parker, A. Sarcevic, I. Marsic, and R. S. Burd (2016): 'Classification and Team Response to Nonroutine Events Occurring during Pediatric Trauma Resuscitation'. *Journal of Trauma and Acute Care Surgery*, vol. 81, no. 4, pp. 666–673.

Xie, Y., I. Seth, D. J. Hunter-Smith, W. M. Rozen, R. Ross, and M. Lee (2023): 'Aesthetic surgery advice and counseling from artificial intelligence: A rhinoplasty consultation with ChatGPT'. *Aesthetic Plastic Surgery*.

You, Y. and X. Gui (2021): 'Self-diagnosis through AI-enabled chatbot-based symptom checkers: User experiences and design considerations'. *AMIA Annual Symposium Proceedings*, vol. 2020, pp. 1354–1363.