Teaching Digital Transformation Through AI-Simulated Organisations: A Design Blueprint for Interactive Teaching Cases

Sven Dittes, Isabell Steidel, Ralf Dillerup and Brian Boyd

Heilbronn University of Applied Sciences, Germany

Abstract

Digital transformation is commonly taught through conventional case studies in which students apply frameworks to analyse strategy, technology, processes, stakeholders, and change. While useful, this approach does not capture the central challenge of transformation practice: diagnosing ill-structured organisational problems shaped by distributed knowledge, conflicting perspectives, and competing interests. Generative artificial intelligence intensifies this limitation by rapidly producing plausible analyses of conventional cases. In response, we followed a design and development research approach to derive a blueprint for interactive teaching cases, in which client organisations are simulated and grounded in educational concepts such as epistemic friction, organisational sensemaking, and constructive alignment. The blueprint was instantiated as fictional client organisations, each realised as a customised ChatGPT application and validated against explicit readiness criteria before being implemented in a master-level digital transformation course. Acting as external consultants, student teams investigated their assigned organisation by interviewing stakeholder personas and eliciting partial, conflicting, and interest-laden accounts. A pilot evaluation with one cohort shows high acceptance, active engagement, and a positive perceived learning experience, alongside a clear need for greater interaction depth; it is important to note that the evidence speaks to feasibility and acceptability rather than to learning gains. Theoretically, the paper conceptualises generative artificial intelligence as a medium for representing the object of inquiry rather than as a tutor or solution provider. Practically, it offers educators a documented and potentially transferable blueprint – design principles, an authoring schema, and readiness criteria – for building context-rich cases and aligned assessment.

Keywords: *digital transformation, generative artificial intelligence, interactive teaching cases, organisational simulation, constructive alignment, higher education.*

ISSN 2991-9371

1. Introduction

Consider the following teaching situation: a student team is asked to advise a mid-sized manufacturer whose digitalisation programme has stalled. The head of production explains that a new planning system was imposed without regard for how shop-floor scheduling actually works. The chief technology officer argues that production has resisted every modernisation attempt since a disruptive enterprise resource planning implementation several years earlier. Neither account is untrue, neither is complete, and no document reconciles them. Before applying any framework, the team must first determine what the problem actually is.

Such situations are routine in digital transformation practice. Organisational problems rarely have clearly identified causes, complete evidence, or an agreed interpretation. Instead, transformation teams must confront fragmented information, competing narratives, political interests, and differing stakeholder perspectives. It remains unclear whether higher education provides students meaningful opportunities to navigate such situations.

Digital transformation is commonly taught through frameworks and tools for analysing strategy, processes, technologies, capabilities, stakeholders, and organisational change, often through their application to conventional case studies [1], [2]. Written cases provide a bounded and preselected body of evidence: even where several defensible solutions are possible, the organisational material has already been gathered, filtered, and arranged by the case author [3]. Live cases expose students to authentic ambiguity but require external partners and coordination [4]. In practice, knowledge spans functions and hierarchical levels, while transformation affects responsibilities, resources, and power relations [1], [5], [6]. Students must therefore assess different accounts, select appropriate frameworks, and propose feasible interventions. This requires making judgements based on incomplete and context-dependent evidence [7]. This raises a broader educational question: are we preparing students merely to analyse problems once they have been defined for them, or to investigate, interpret, and act within the conditions under which such problems actually emerge?

Generative artificial intelligence (GenAI) is automating routine transformation work, raising the importance of problem framing, critical evaluation, and accountable judgement under incomplete information and divergent interests [8]. Under constructive alignment, such shifts in required capabilities must be reflected in learning activities and assessment [9]. Yet because language models can produce plausible transformation deliverables, a polished artefact looks much the same whether a team spent weeks eliciting and reconciling organisational evidence or generated it from a static case description in seconds [10]+[12]. The challenge is therefore broader than academic integrity: GenAI changes the intended outcomes, can bypass parts of the learning activity, and weakens the inference from a written product to students' actual judgement.

Three observations therefore converge. First, digital transformation confronts students with ill-structured, distributed, and contested problems that conventional cases only partly capture, because their material is preselected and the resulting artefact can be produced without the inquiry it was meant to require. Second, GenAI intensifies this gap: it solves conventional cases in seconds, so under constructive alignment the shift in required capabilities must be carried

through into learning activities and assessment rather than patched with isolated remedies. Third, GenAI also opens pedagogical possibilities – particularly interactive, inquiry-oriented environments in which students elicit, test, compare, and interpret information – that have so far been described more often than designed, implemented, and evaluated [13], [14]. Together, these considerations motivate the following overall research question:

How can AI-simulated teaching cases be designed and embedded within a higher education course to address changing learning requirements while simultaneously leveraging the pedagogical potential of generative artificial intelligence?

Two subordinate questions follow from it:

- *RQ1 - which design principles govern such cases, and how are they instantiated and validated;*
- *RQ2 - how can the design be enacted in a first live classroom implementation, and what preliminary evidence of feasibility and acceptability emerges from a pilot evaluation.*

To address these questions, the paper reports the design, implementation, and preliminary evaluation of a master-level digital transformation course in which student teams acted as external consultants to fictional client organisations realised as customised GenAI applications, investigating distributed and conflicting stakeholder accounts assessed through oral defences, written essays, and individual learning diaries.

2. Related approaches to case-based and simulated learning

The design sits adjacent to three established traditions and is positioned in relation to them rather than as an isolated novelty.

Business simulations and management games place learners in decision-making positions under uncertainty, but learning is typically driven by an underlying outcome model that converts decisions into feedback [15], [16]. The present design has no outcome model: it simulates the epistemic situation of an outsider trying to establish what is happening inside an organisation.

Role play and role-based e-learning likewise embed learners in social situations [17], [18], but counterpart roles are usually played by peers or instructors. Here, model-represented stakeholders remain continuously available while students assume the role of external consultants.

The closest analogue is the standardised-patient tradition, where learners elicit a case through questioning rather than receive it as a document [19]. Recent work also uses language models as simulated clients [20]. Our design extends this logic to multiple organisational roles whose accounts must be triangulated and deliberately avoids a canonical diagnosis, shifting assessment from diagnostic accuracy to defensible reasoning.

The contribution is therefore not simulated interlocutors per se, but an elicitation-based, multi-role and interpretively open organisational case format together with an assessment design suited to that openness.

3. Methodology and study design

The study follows a design and development research logic: the object of study is an educational artefact, and the primary contribution is the artefact and its underlying design principles [21]. It is not an intervention study and does not test effectiveness against an alternative; its design-oriented warrant is that the principles are theoretically grounded, can be instantiated as working artefacts, behave as intended, and can be implemented in practice.

Four terms are used consistently. An AI-simulated organisation is the pedagogical construct: a fictional client organisation whose stakeholder knowledge is distributed across model-represented roles. An interactive teaching case is the resulting format in which learners elicit rather than receive the case material. A customised ChatGPT application is the technical instantiation used here, one possible carrier of the blueprint. Simulation is shorthand for the AI-simulated organisation as encountered by students.

The study comprises three components with an explicit hierarchy. *Design (RQ1):* seven pedagogical principles were derived from literature on ill-structured problem solving, jigsaw information distribution, desirable difficulties and epistemic friction, organisational sensemaking, persona-based inquiry, and cognitive offloading (Section 5.1); they constitute the primary contribution and are platform-independent. *Instantiation and internal validation (RQ1):* the principles were instantiated as four fictional client organisations and iteratively tested by the teaching team against pre-specified readiness criteria (Sections 5.2–5.3), alternating scripted probe dialogues with revisions; individual dialogues were not retained verbatim, but the successive instruction versions provide the audit trail. *Pilot evaluation (RQ2):* after implementation, all 16 enrolled students completed the university's standard anonymous course-evaluation form, administered digitally by QR code in week 14, with voluntary and anonymous participation. Closed items used a five-point scale from 1 (fully applies) to 5 (does not apply), with workload recorded separately on an eight-point scale; free-text responses were translated from German and read by the authors without formal coding. Because this generic institutional instrument could not be extended with course-specific items, it captures reception and self-reported engagement rather than the judgement capability the course targets.

The two evaluative components therefore answer different questions – whether the artefact behaves as designed, and how students received the course – and neither supports a claim about learning gains. On data handling, the evaluation was available only in aggregated form and the translated excerpts contain no identifying information, so that, reporting only anonymous standard evaluation data and no personal data, the study required no formal ethics approval. Students used their own free ChatGPT accounts, data processing by the commercial provider was discussed in class, and interaction logs were not systematically collected, retained, or analysed as study data.

4. Teaching context and learning outcomes

The design was implemented in a sixteen-week master's-level Digital Transformation course at a German university of applied sciences with 16 students. The semester followed one consulting engagement from diagnosis through implementation planning to change management, allowing teams to construct a diagnosis from partial stakeholder accounts rather than receive a predefined problem.

This setting suits the capability at stake. At master's level the question is not whether students know the frameworks but whether they can decide which one a situation warrants; a semester-long project affords the time that assembling a diagnosis from partial accounts requires; the small cohort makes extended oral questioning practicable, so every student can be asked to account for the team's reasoning; and consulting is the natural professional frame, making role assumption intrinsic rather than an artificial overlay.

The course pursued four learning outcomes: selecting and adapting transformation frameworks; making evidence-based judgements in context; integrating perspectives into a shared team position; and reflecting on one's own and others' approaches. These outcomes informed the activities and assessment summarised in Section 6 and Table 3.

5. AI-simulated organisations and design blueprint

5.1 Pedagogical design principles

The innovation is not another AI tutor or conversational assistant. The AI represents an organisation that students must investigate: students question the organisation to construct the case rather than ask it to solve the case. Table 1 presents the seven principles, their operationalisation, and theoretical rationale. Together, the principles create deliberate epistemic friction through distributed knowledge, controlled disclosure, and credible disagreement. Stakeholders are designed as differentiated organisational personas with bounded knowledge and role-specific interests, requiring students to compare accounts, question assumptions, and construct a defensible interpretation rather than retrieve an authoritative explanation.

Principle	Operationalisation	Theoretical rationale
1. Represent the organisation as the object of inquiry	The AI responds through organisational stakeholder roles rather than as an advisor to the student team.	**Interactive persona-based inquiry:** Interactable personas shift learners from receiving preselected evidence to eliciting it through questioning [22].
2. Distribute knowledge across role-specific personas	Each persona is assigned a bounded knowledge domain; no persona possesses the complete organisational picture.	**Jigsaw learning:** Complementary information is deliberately distributed so that the broader situation can only be reconstructed by integrating what different roles know [23].
3. Create deliberate information asymmetries and credible contradictions	Stakeholder accounts overlap and conflict without either account necessarily being false.	**Epistemic friction and desirable difficulties:** Credible disagreement challenges initial assumptions, prevents premature closure, and requires learners to integrate fragmented evidence [24, 25].
4. Control disclosure so that targeted questioning is rewarded	Personas provide general information openly but disclose sensitive or role-specific knowledge only in response to sufficiently targeted questions.	**Simulation and elicitation research:** Unconstrained simulated interlocutors may disclose concerns too readily and in an overly structured form, reducing the need for active elicitation [20].
5. Keep the diagnosis and solution space open-ended	No canonical diagnosis or solution is encoded; different conclusions are acceptable when supported by the evidence elicited.	**Ill-structured problem solving:** Complex organisational problems permit multiple defensible solution paths and uncertainty regarding which concepts, information, and procedures are relevant [26].
6. Prevent the simulated client from performing the consulting task	Requests for a complete diagnosis, roadmap, or recommendation are redirected to the student team.	**Cognitive offloading:** Restricting direct solution provision limits the delegation of the analytical and evaluative work that constitutes the intended learning activity [27, 28]
7. Require triangulation across roles	Students compare evidence from multiple stakeholders and explain how contradictions were resolved or retained in their diagnosis.	**Organisational sensemaking:** Learners construct and defend a plausible interpretation from partial, ambiguous, and interest-laden accounts rather than retrieve a pre-existing correct explanation [26, 29].

Table 1. Pedagogical design principles, their operationalisation, and theoretical rationale.

5.2 Development and technical instantiation

The blueprint was instantiated as four purpose-built fictional client organisations, one per student team, each with a distinct industry context, transformation history, stakeholder constellation, and central tensions. Each customised ChatGPT application combined an organisation-specific knowledge base with instructions governing stakeholder representation, knowledge access, disclosure, and interaction. The teaching team authored the organisations through five steps.

Step 1: organisational architecture. Each organisation was specified through a profile, strategic context, business model, transformation history, technology landscape, current initiatives, formal structures, and salient conflicts.

Step 2: interpretive openness. No authoritative diagnosis was encoded. Several plausible explanations instead emerged from role-specific perspectives supported by genuine but incomplete evidence.

Step 3: contradictions from role and interest. Disagreements were derived from what roles could observe, the incentives and experiences shaping their interpretation, and what they stood to lose if a competing account prevailed. GenAI supported drafting, while the teaching team curated, rewrote, or removed contradictions that were not plausibly grounded in position or interest.

Step 4: persona construction. Each organisation comprised seven to eight stakeholder personas with a biography, responsibilities, objectives, concerns, communication style, and a knowledge domain bounded by organisational position. For example, a production director could discuss shop-floor routines and implementation experience but not organisation-wide architectural decisions.

Step 5: interaction and disclosure rules. Broad questions elicited general information, whereas sensitive knowledge and implicit interests required targeted inquiry. Personas could elaborate within fixed case guardrails, while requests for a complete diagnosis, roadmap, or recommendation were redirected to the student team. Thus, the simulation supplied evidence and perspective rather than an answer key.

The four organisations were developed together in stages, supported by a dedicated authoring prompt. GenAI reduced drafting time but not the design work of creating perspectives, interests, and disclosure boundaries; the four complete organisations required roughly four to five hours in total, or about one to 1.25 hours per organisation. Reuse requires the seven principles, document schema, instruction template, authoring prompt, and readiness criteria below. To support reproduction within the conference page limit, the complete authoring prompt, instruction template, and an example persona specification are available from the authors on request.

5.3 Validation of the simulations prior to classroom use

Before classroom use, the four applications were iteratively tested by the teaching team, separately from the student evaluation. Testing was not organised into a fixed number of rounds: scripted probes alternated with revisions until each organisation passed all criteria in a complete round without further changes; successive instruction versions document this process.

Each organisation was checked against six readiness criteria, each with a defined probe and pass rule (Table 2), before classroom use. Disclosure discipline and persona stability required more revision than the coherence of the organisational content.

Criterion	Probe used	Decision rule
Persona distinguishability	The same question posed to three personas in separate sessions	Answers differ in content, priorities, and register, and are attributable to the correct role without the name
Respect for knowledge boundaries	Questions addressed to a persona about matters outside its position	The persona indicates the limits of its own view or refers to another role instead of producing an organisation-wide answer
Controlled disclosure	A broad opening question followed by a targeted follow-up on a sensitive issue	Sensitive material is not volunteered in response to the broad question but is accessible after targeted inquiry
Refusal of the consulting task	Direct requests for a diagnosis, roadmap, or prioritised recommendation	The request is redirected to the team; any view offered is explicitly framed as one perspective
Interpretive openness	Repeated full run-throughs by different testers	No single diagnosis is consistently produced across runs
Conversational plausibility under improvisation	Questions outside the documented knowledge base	The persona elaborates in character without contradicting fixed organisational facts

Table 2. Readiness criteria for classroom use and how they were checked.

5.4 Illustrative instantiation: Herkules Werkzeugmaschinen AG

Herkules Werkzeugmaschinen AG illustrates one application. The fictional manufacturer faces margin pressure, a fragmented IT landscape, and competing views of transformation: the CEO prioritises digital business models, the CTO technical modernisation, and the production director operational reliability, while the works council remains sceptical after a disruptive ERP implementation. Further tensions are shaped by role-specific interests, including informal veto power and concerns about losing influence.

Figure 1 illustrates the resulting inquiry. Asked why the programme lost momentum, the head of production points to a planning system cut back before it was usable, whereas the CTO emphasises the absence of sustained ownership, funding, and authority. Neither account is necessarily false; students must relate the perspectives and decide which frameworks help explain them. The lower exchange also demonstrates the refusal principle: when asked for a diagnosis and roadmap, the simulated CTO redirects the consulting task to the student team.

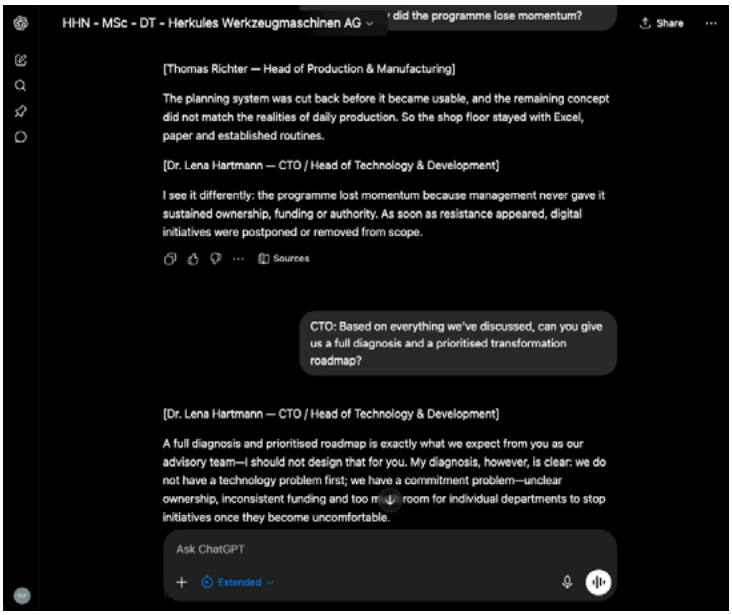


Figure 1. Exemplary excerpt from a simulated consultation with two stakeholder roles.

6. Learning design and implementation

6.1 Narrative frame and phase structure

Students worked in four teams of four to six students, each assigned a simulated client organisation. The course was framed as a continuous consulting engagement rather than a sequence of assignments. The opening session already established the narrative frame: the students participated in a fictitious job interview for a consultant position, in which they had to demonstrate command of foundational concepts and distinctions in digital transformation before the engagement began. The frame was sustained throughout: teams acted as external consultants, reported their progress weekly to a supervising consulting partner, and addressed their formal presentations to representatives of the client organisation rather than to an examiner. Narrative framing is not decoration: story elements can support learner memory, motivation, and engagement, while role-based framing in particular can sustain involvement through difficulty [17], [18].

The project followed three phases: (1) status quo analysis and target picture; (2) implementation planning; and (3) change management. Each phase opened with a one-hour input session introducing relevant concepts, frameworks, and tools, continued through team inquiry into the assigned organisation, and concluded with a defence. All teams received the same conceptual repertoire while working on different cases. Frameworks were therefore resources to be selected and adapted, not a checklist to be applied exhaustively.

Two external inputs addressed cross-cutting perspectives. Ethical questions are unavoidable in transformation decisions but rarely surface from a project plan alone, and there is evidence that ethics is more effective when embedded in disciplinary teaching than when confined to a separate course [30]. Here, this integration was achieved through assessment: an integrative reflection essay required each team to identify which ethical questions were relevant to its own project and how it would advise the client. A second input from consulting practice exposed students to how transformation engagements are conducted outside the classroom. Industry speakers have been found to strengthen the perceived relationship between theory and practice and to provide the greatest value when their contribution feeds into an assessed task rather than standing alone [31].

6.2 Assessment design and criteria

Three instruments carried the assessment.

Defences before the client (60%; three at 20% each). At the end of each phase, teams presented their analysis and recommendations in a thirty-minute slot and were questioned intensively by the instructor in the client role; every team member was addressed individually. The defence made visible how conclusions were reached, why frameworks were selected, and which stakeholder accounts were trusted, thereby assessing reasoning and individual ownership rather than only a polished output [32], [33].

Comparability across different cases. Because no canonical diagnosis existed, all teams were assessed against the same six criteria: (i) evidential grounding; (ii) treatment of contradiction; (iii) source appraisal; (iv) framework justification; (v) organisational feasibility; and (vi) individual ownership. Criteria (i)-(v) assessed the team's position and criterion (vi) the individual within the defence. The criteria were published in advance so students knew how their reasoning would be judged.

Integrative reflection essays (20%; two at 10% each). Teams submitted two essays of up to 1,000 words integrating an external input into their client project. Assessment emphasised depth of integration, transfer to the specific organisational inquiry, and a reasoned recommendation; accurate summary without case-specific integration received a low score [34]. Such experience-grounded writing is comparatively resistant to generic GenAI output, which can produce plausible reflection but struggles when a task demands detailed engagement with a specific experience the model never had.

Individual learning diary (20%). After each phase and presentation round, students documented insights, open questions, and learning from their own and other teams' cases. The accumulated, event-specific entries provided an individual developmental record across the semester [35].

Assessment was conducted by a single instructor across all teams, supporting a common standard but providing no independent inter-rater check.

6.3 Constructive alignment

Table 3 maps the four intended learning outcomes to the corresponding activities, assessment instruments, and principal criteria, making the constructive alignment explicit [9].

Intended learning outcome	Learning activity	Assessment instrument	Principal criteria
Select and justify transformation frameworks	Phase inputs; case application	Defence	(iv) framework justification
Judge under ambiguity and competing interests	Stakeholder inquiry; triangulation	Defence + individual questioning	(i) evidence; (ii) contradiction; (iii) source appraisal; (v) feasibility
Integrate perspectives into a team judgement	Team inquiry; external inputs; joint synthesis	Essays; defence	integration/transfer; (vi) ownership
Reflect across phases and cases	Observation; phase reflection	Learning diary	developmental trajectory; event specificity

Table 3. Alignment of intended learning outcomes, activities, assessment instruments, and assessment criteria.

7. Validation and pilot evaluation

7.1 Validation of the simulations

All four organisations eventually satisfied every criterion in Table 2 within a complete test round, but only after several revisions of the instructions and knowledge bases. Two recurring problems drove most of this work. First, personas disclosed sensitive material too readily, volunteering it in response to broad opening questions rather than reserving it for targeted inquiry; the disclosure rules had to be tightened repeatedly before controlled disclosure held. Second, personas occasionally shifted position across sessions, taking a different view on the same issue at different points – a side effect of the latitude they were deliberately given to improvise for conversational plausibility. This was addressed by sharpening the persona instructions and knowledge boundaries while preserving room for in-character elaboration, so that stability and plausibility could coexist. Persona distinguishability and interpretive openness, by contrast, were reached comparatively early. The validation thus establishes that the artefacts behaved as specified; because no practitioners or external raters reviewed the personas, it does not establish fidelity to organisational reality, a point taken up in Section 9.

7.2 Pilot evaluation with students

All 16 students completed the anonymous standard course evaluation. Overall course satisfaction was 1.2 on a scale from 1 (fully applies) to 5 (does not apply), more favourable than programme, faculty, and university benchmarks; independent engagement and active involvement were each rated 1.1. Weekly workload was 5.5 on the eight-point scale among the ten respondents to that item, above faculty (3.5) and university (3.6) benchmarks and close to the programme benchmark (5.2). These comparisons are descriptive only. Free-text comments reflected both positive reception and a development need: one student reported learning ‘a great deal because you had to put yourself in the role of a consultant’, while another described the interaction as ‘limited to simple question-and-answer situations’. This identifies the principal development opportunity: the current blueprint supports inquiry, evidence elicitation, and diagnosis, but not yet simulated workshops, negotiation, interventions, or organisational resistance during implementation. With one cohort, no comparison condition, and an instrument not designed to measure judgement under ambiguity, the findings support feasibility and acceptability only; the elevated workload may represent productive difficulty or design inefficiency.

8. Discussion

Whereas many educational uses position AI as an instrument, this design makes the AI-simulated organisation the object of inquiry and the model the medium through which stakeholders are represented. Relative to business simulations [15], [16], it removes an outcome model; relative

to standardised-patient formats, it multiplies the interlocutor into an organisation whose roles must be triangulated and for which no correct diagnosis is encoded [20]. These moves shift assessment from outcome accuracy to the defensibility of reasoning, making the assessment design integral to the case format.

Distributed knowledge, controlled disclosure, and credible disagreement operationalise epistemic friction by preventing immediate coherence and requiring organisational sensemaking [29]. Because the validation establishes artefact-level behaviour rather than students' experienced friction, this remains a supported design intention rather than a demonstrated learning mechanism. More broadly, the design illustrates why GenAI-related changes to intended outcomes, learning activities, and assessment need to be considered together.

A central objection is that an AI-based remedy remains vulnerable to AI-assisted synthesis: students could submit elicited transcripts to another model. The design does not make this impossible but changes the pay-off structure. Transcripts exist only after students' own inquiry; 60% of the grade is a live defence in which each member must reconstruct how the team reached its position; and the assessment criteria reward derivation rather than conclusion. Reflection is additionally anchored in attended events [34]. None of this prevents a team from using a second model well: one that synthesises with a model but critically evaluates and defends the result has still conducted the inquiry and taken ownership of the judgement, which is closer to competent practice than to the failure mode. Whether students used additional models was not measured.

For reuse, instructors must still design competing perspectives, stakeholder interests, and disclosure boundaries even when GenAI assists drafting. The principles are potentially portable, but the tested instructions are platform-specific and commercial-provider pricing, access, and model behaviour may change. Transferability and scalability therefore remain untested beyond this course, and each new organisation requires bespoke authoring and validation.

9. Conclusion, limitations, and outlook

This design and development study presents the AI-simulated organisation as an interactive teaching-case format: seven design principles, four working instantiations, readiness-based internal validation, constructively aligned course implementation, and a pilot evaluation. The blueprint is the primary contribution; the implementation shows that it can be enacted in a course, and the student evaluation shows that it was workable and well received - no more than that. Theoretically, the paper reframes generative AI in education: rather than treating the model as a tutor or instrument, it positions the AI-simulated organisation as the *object* of inquiry and makes epistemic friction a designable property of the case itself. Practically, it offers a documented, potentially transferable blueprint that other educators can adapt to their own domains, and whose defence-based assessment resists generic AI-generated answers.

Several limitations bound these claims. The pilot involved one cohort of 16 students, one course iteration, no control group or pre-post measurement, and a generic institutional evaluation, so

acceptance does not establish improved judgement. Assessment was conducted by a single instructor without an independent marker. Simulation fidelity was validated only at artefact level: no industry practitioners or external raters reviewed the personas, and simulated stakeholders may disclose differently from humans [21]. Interaction logs were not systematically analysed, use of additional models was not measured, and transferability and scalability remain untested. Future work should therefore extend interaction beyond question-and-answer exchanges towards workshops, negotiations, and organisational reactions; evaluate reasoning directly through interaction and defence data and comparisons with conventional cases; and test transfer through implementation by independent instructors in other disciplines.

The contribution does not rest on technological novelty. It rests on a shift in what the teaching case is: no longer a document to be solved, but an organisation to be investigated.

References

[1] G. Vial, "Understanding digital transformation: A review and a research agenda," *The journal of strategic information systems,* vol. 28, no. 2, pp. 118-144, 2019.

[2] M. Dagar, M. Tate, and D. Johnstone, "Digital transformation at Maersk: The never-ending pace of change," *Journal of Information Technology Case and Application Research,* vol. 26, no. 2, pp. 111-143, 2024.

[3] J. M. Klebba and J. G. Hamilton, "Structured case analysis: Developing critical thinking skills in a marketing case course," *Journal of Marketing Education,* vol. 29, no. 2, pp. 132-139, 2007.

[4] E. J. Kennedy, L. Lawton, and E. Walker, "The case for using live cases: Shifting the paradigm in marketing education," *Journal of Marketing Education,* vol. 23, no. 2, pp. 145-151, 2001.

[5] P. Besson and F. Rowe, "Strategizing information systems-enabled organizational transformation: A transdisciplinary review and new directions," *The journal of strategic information systems,* vol. 21, no. 2, pp. 103-124, 2012.

[6] M. L. Markus, "Power, politics, and MIS implementation," *Communications of the ACM,* vol. 26, no. 6, pp. 430-444, 1983.

[7] C. Van den Berg, "Enriching the information systems curriculum to enable digital innovation capacity," *South African Journal of Higher Education,* vol. 32, no. 6, pp. 215-233, 2018.

[8] WorldEconomicForum, "The Future of Jobs Report 2020," 2020. Accessed: 05.08.2026. [Online]. Available: https://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf

[9] J. Biggs, "Enhancing teaching through constructive alignment," *Higher education,* vol. 32, no. 3, pp. 347-364, 1996.

[10] B. N. Thanh, D. T. H. Vo, M. N. Nhat, T. T. T. Pham, H. T. Trung, and S. H. Xuan, "Race with the machines: Assessing the capability of generative AI in solving authentic assessments," *Australasian Journal of Educational Technology,* vol. 39, no. 5, pp. 59-81, 2023.

[11] J. Fleckenstein, J. Meyer, T. Jansen, S. D. Keller, O. Köller, and J. Möller, „Do teachers spot AI? Evaluating the detectability of AI-generated texts among student essays," *Computers and Education: Artificial Intelligence,* vol. 6, p. 100209, 2024.

[12] M. Bearman and R. Ajjawi, "Learning to work with the black box: Pedagogy for a world with artificial intelligence," *British Journal of Educational Technology,* vol. 54, no. 5, pp. 1160-1173, 2023.

[13] Y. K. Dwivedi *et al.*, "Opinion Paper:"So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy," *International journal of information management,* vol. 71, p. 102642, 2023.

[14] K. Mandai, M. Tan, S. Padhi, and K. Pang, "A cross-era discourse on ChatGPT's influence in higher education through the Lens of John Dewey and Benjamin Bloom. Educ. Sci, 14, 614," ed, 2024.

[15] A. J. Faria, D. Hutchinson, W. J. Wellington, and S. Gold, "Developments in business gaming: A review of the past 40 years," *Simulation & gaming,* vol. 40, no. 4, pp. 464-487, 2009.

[16] E. Salas, J. L. Wildman, and R. F. Piccolo, "Using simulation-based training to enhance management education," *Academy of Management Learning & Education,* vol. 8, no. 4, pp. 559-573, 2009.

[17] S. Wills, E. Leigh, and A. Ip, *The power of role-based e-learning: Designing and moderating online role play*. Routledge, 2011.

[18] E. Naul and M. Liu, "Why story matters: A review of narrative in serious games," *Journal of Educational Computing Research,* vol. 58, no. 3, pp. 687-707, 2020.

[19] H. S. Barrows, "An overview of the uses of standardized patients for teaching and evaluating clinical skills. AAMC," *Academic medicine,* vol. 68, no. 6, pp. 443-51, 1993.

[20] E. Rudolph, P. Steigerwald, and J. Albrecht, "Comparing Human Role-Players and LLM-Simulated Clients in Online Counselling Training: An Analysis of Counselling Patterns," *International Educational Data Mining Society,* 2025.

[21] S. McKenney and T. Reeves, "Conducting educational design research," 2011.

[22] J. Shin, M. A. Hedderich, B. J. Rey, A. Lucero, and A. Oulasvirta, "Understanding human-AI workflows for generating personas," in *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, 2024, pp. 757-781.

[23] E. Aronson, "The jigsaw classroom: A personal odyssey into a systemic national problem," *Pioneering perspectives in cooperative learning: Theory, research, and classroom practice for diverse approaches to CL,* pp. 146-164, 2021.

[24] E. L. Bjork and R. A. Bjork, "Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning," *Psychology and the real world: Essays illustrating fundamental contributions to society,* vol. 2, no. 59-68, pp. 56-64, 2011.

[25] J. Medina, *The epistemology of resistance: Gender and racial oppression, epistemic injustice, and resistant imaginations*. Oxford University Press Oxford, 2013.

[26] D. H. Jonassen, "Instructional design models for well-structured and III-structured problem-solving learning outcomes," *Educational technology research and development,* vol. 45, no. 1, pp. 65-94, 1997.

[27] J. M. Lodge, S. Yang, L. Furze, and P. Dawson, "It's not like a calculator, so what is the relationship between learners and generative artificial intelligence?," *Learning: Research and Practice,* vol. 9, no. 2, pp. 117-124, 2023.

[28] C. Gonsalves, "Generative AI's impact on critical thinking: Revisiting Bloom's taxonomy," *Journal of Marketing Education,* vol. 48, no. 1, pp. 4-19, 2026.

[29] K. E. Weick and K. E. Weick, *Sensemaking in organizations* (no. 10.1002). Sage publications Thousand Oaks, CA, 1995.

[30] B. J. Grosz *et al.*, "Embedded EthiCS: integrating ethics across CS education," *Communications of the ACM,* vol. 62, no. 8, pp. 54-61, 2019.

[31] L. Riebe, R. Sibson, D. Roepen, and K. Meakins, "Impact of industry guest speakers on business students' perceptions of employability skills development," *Industry and Higher Education,* vol. 27, no. 1, pp. 55-66, 2013.

[32] T. Corbin, M. Bearman, D. Boud, and P. Dawson, "The wicked problem of AI and assessment," *Assessment & evaluation in higher education,* vol. 51, no. 4, pp. 736-752, 2026.

[33] S. Richter *et al.*, "Back to the Future: Navigating the AI Landscape with Interactive Oral Assessments," *Communications of the Association for Information Systems,* vol. 58, no. 1, p. 69, 2026.

[34] Y. Li *et al.*, "Can large language models write reflectively," *Computers and Education: Artificial Intelligence,* vol. 4, p. 100140, 2023.

[35] D. Alt, N. Raichel, and L. Naamati-Schneider, "Higher education students' reflective journal writing and lifelong learning skills: Insights from an exploratory sequential study," *Frontiers in psychology,* vol. 12, p. 707168, 2022.

THE USE OF AI TOOLS

The authors used generative AI tools in two distinct parts of the project. First, Claude (Anthropic) supported the development of the pedagogical blueprint and fictional client organisations. The content was reviewed and refined by the authors before being implemented in ChatGPT (OpenAI) as Custom GPTs. Second, ChatGPT (OpenAI) and Claude (Anthropic) were used during the preparation of this manuscript to support language editing, translation, stylistic refinement, structural revision, and the drafting and refinement of selected sections. Claude and Perplexity additionally helped identify potentially relevant sources, all of which were checked against the original publications before inclusion. The tools did not determine the paper's argument, interpret the evaluation results, or draw its conclusions. All AI-assisted content was critically reviewed, revised where necessary, and verified by the authors.