Systems & Control Transactions

Original Research Article
Peer Reviewed Conference Proceeding
ESCAPE 36 - European Symposium on Computer Aided Process Engineering
Sheffield, UK. 21-24 June 2026

PSE PRESS

Assessing Workflow Automation Platforms in Engineering Education: Towards an Ethical, Technical, and Pedagogical Framework

Daniela Galatro[a*], Stuart Grey[b], and Sourojeet Chakraborty[c]

[a] University of Toronto, Department of Chemical Engineering & Applied Chemistry, Toronto, Ontario, Canada
[b] University of Glasgow, James Watt School of Engineering, Glasgow, United Kingdom
[c] Johns Hopkins University, Whiting School of Engineering, Department of Chemical and Biomolecular Engineering, Baltimore, Maryland, United States
* Corresponding Author: daniela.galatro@utoronto.ca.

ABSTRACT

Workflow automation platforms connect applications and services to automate data transfer and multistep processes. Although widely used in engineering research and institutional administration, including engineering institutions, they are rarely integrated into undergraduate engineering curricula, and their educational adoption introduces ethical, technical, and pedagogical risks. This paper proposes a practical framework for developing, deploying, and assessing workflow-automation-enabled learning tools coupled with generative AI, with explicit attention to institutional constraints and learning outcomes. As a conceptual case study, we present it in a second-year chemical engineering course (Heat and Mass Transfer) to support learning of heat conduction. The platform includes instructor-approved assets such as content slides, solved problems, pre-prompts, and a validated question database, through an automation pipeline that issues structured API calls to a generative AI system and returns explanations and worked examples. A subject matter expert (SME) authors and tests prompt modules aligned with learning outcomes mapped to Bloom's taxonomy levels (understand, apply, analyze, evaluate) and aligned with accreditation expectations. Effectiveness metrics are proposed to estimate learning gains, and a tool quality audit and an SME review are used to verify conceptual correctness and computational validity. To assess risks, we apply a likelihood–impact risk matrix and define five categories: data privacy and protection, transparency, pedagogical impact, institutional compliance, and ethical considerations. Preliminary analysis suggests potential concept gains from structured prompt/content integration but indicates that numerical problem reliability depends strongly on question-database coverage. The highest perceived risks relate to privacy and transparency, followed by pedagogical impact, motivating mitigations such as anonymization protocols, restricted data pathways, and human verification throughout the lifecycle.

Keywords: Workflow automation platforms, Generative AI, Bloom taxonomy, risk assessment, effectiveness assessment

INTRODUCTION

Workflow automation platforms and tools can enhance, and even reduce, administrative tasks and provide structured learning as well as instantaneous follow-up support across many settings.

This paper proposes a framework to guide the design, deployment, and assessment of workflow automation platforms used for engineering education in the teaching and learning domains. As a conceptual case study, we present an integrated platform for a second-year chemical engineering course (Heat and Mass Transfer) focused on steady-state and transient heat conduction. The platform includes instructor-approved content and a validated question database powered by an automation pipeline that delivers structured API calls to a

generative AI tool, with learning outcomes mapped to Bloom's taxonomy and aligned with accreditation expectations.

The inherent risks of using this platform were assessed using a likelihood–impact risk matrix across five categories, including data privacy and protection, transparency, pedagogical impact, institutional compliance, and ethical considerations. We later propose mitigation measures for the identified and quantified high-risk areas.

BACKGROUND

Workflow Automation Platforms in Engineering

Workflow automation platforms (WAPs) are complex integration tools designed to connect multiple applications and execute multistep processes automatically via a flexible, centralized layer. These platforms allow "unskilled people to produce applications with little or no coding" through low-code approaches [1]. They have transformative potential, as they can drive operational efficiency, process consistency, and scalability across diverse industries [2]. The main capabilities of these platforms include: (i) defining workflows via triggers, actions, and conditional logic; (ii) supporting API integration; (iii) routing/connecting data across heterogeneous tools; (iv) maintaining execution histories, and (v) standardization and scaling of routine tasks. These platforms are efficient for streamlining complex operational processes across engineering, academic, and business environments.

Workflow automation platforms have been increasingly adopted in engineering and education to streamline routine/repetitive processes and enhance data consistency and reproducibility. An example of a recent use and implementation of WAPs includes the use of Robotic Process Automation to transform pedagogical practices by automating repetitive tasks, shifting the focus of education toward high-value instructional strategies [3]. Thus, automated design systems reshape learning outcomes by enhancing conceptual modeling, accelerating prototyping, and improving interdisciplinary collaboration [4]. Though this automation has proven beneficial in different settings, caution is advised when implemented, with warnings of an "automation abyss" predicting the erosion of core learning processes when overreliance on this technology [5]. Design automation should be used as a supporting tool rather than replacing human agency and critical thinking.

Risks for Education Automation

When workflow automation platforms integrate with generative AI, risks escalate, creating systemic issues that can pose challenges in educational settings. Thus, a single learning interaction can now involve multiple services, data stores, and decision points, increasing potential vulnerability points [6], including logs and cached prompts that can retain information for longer.

Another inherent risk of WAP adoption and implementation is institutional sensitivity, as many tasks, such as submissions, performance signals, and identifiers, may cross boundaries into external APIs outside the institution's "official" and intrinsically protected environment. A governance-risk approach is then required to set compliance constraints, such as protections for learning management systems (LMSs) and data residency requirements. Risk management, therefore, requires data transfer permissions, clear policies, consent mechanisms, and retention practices.

Generative AI (GenAI) in educational settings introduces transparency and accountability risks that can distort learning practices [6] and even erode teacher-student trust. AI's non-deterministic outputs create substantial challenges in tracing decision-making processes, potentially leading to unfair perceptions [7]. Addressing these issues and risks requires audits/accountability mechanisms, versioned prompts, clear pedagogical transparency, and institutional oversight [8].

In general, AI in education poses critical pedagogical risks by promoting shortcut behaviors (bypassing) that might undermine the learning process and diminish problem-solving and critical thinking skills [9] .Thus, students might prioritize quick answers over deep conceptual and/or fundamental understanding, again, undermining analytical reasoning and consequently, critical thinking [10]. Pedagogical safeguards are then designed and implemented to mitigate these risks, requiring students to reason and cross-validate AI responses and "reasoning".

Bloom's taxonomy can be used to explain why WAPs introduce pedagogical risks. Thus, when coupled with AI, they provide students with immediate outputs, potentially bypassing the lower-to-mid levels of Bloom (remembering, understanding, applying, and analyzing), which, when scaffolded in our learning process, lead to higher-order performance [11]. The "shortcut" behaviour then appears as students achieve a solution that has not been reasoned, so learning is not durable due to overreliance and an "automation abyss" [5]. The central risk is not simply incorrect AI output, but misaligned cognition, as metacognitive monitoring is paused and learning transfer is broken, as students do not build the mental models required to analyze variants [12]. To minimize these risks, a Bloom-aligned mitigation strategy is required to design "gated" assistance strategies that lead students to demonstrate understanding, reasoning, and justify choices while automation speeds up the solution-generation process, in a way that automation is seen and used as a supportive tool, rather than replacing the cognitive work that the curriculum is meant to develop [13].

FRAMEWORK AND CASE STUDY

In this work, we propose a framework to guide the development, implementation, and assessment of AP-based learning tools coupled with Gen AI. The framework includes four phases: (i) design and policies, (ii) build and integration, (iii) implementation and monitoring, and (iv) assessment and iteration. This lifecycle allows decisions to be fully auditable.

In the design and policies phase, the teaching team defines the learning scope and outcomes, mapped to Bloom's cognitive levels, and estimates the required student data (if any). Constraints/rules are specified, including approved data and consent requirements, which limit the pipeline's inputs. In the build and integration phase, the WAP is used to connect learning assets and establish tool interactions. The platform executes tasks using approved content and structured prompts to a generative AI API, returning outputs to an interface. A fixed set of question databases and prompt templates is provided. Some basic technical safeguards are also implemented at the workflow level, such as error handling, to keep the system "error-free". In the implementation and monitoring phase, the system is constrained within the provided context and guidance to students on appropriate use. Monitoring is performed on both technical performance (for instance, drift in responses) and off-track patterns (such as overreliance, bypass, or scaffolding use). A subject-matter expert (SME) oversees the review process, especially when solving numerical problems. In the assessment and iteration phase, metrics and risk assessments are used to determine whether the platform enhances learning while aligned with ethical/institutional standards. Feedback from this phase is returned to the design phase, including refined prompts, expanded question databases, and enhanced risk mitigation strategies. Figure 1 shows the conceptual architecture of the WAP for our case study.

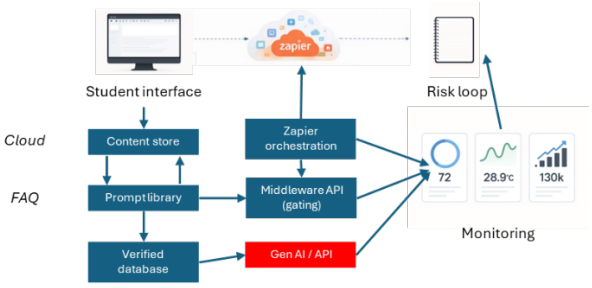
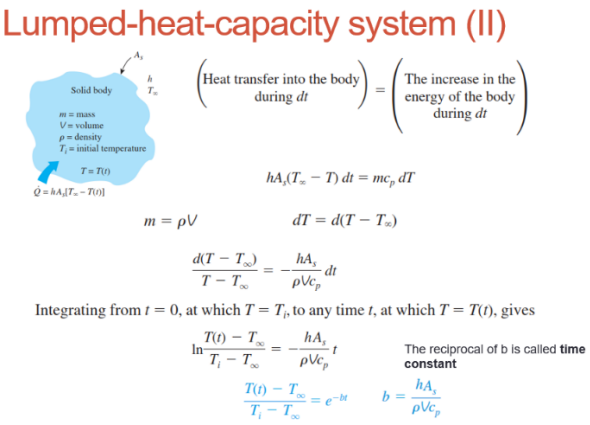


Figure 1. Conceptual architecture of the WAP

In the architecture, students are "directed" (via Zapier or similar) into a request package built from the teaching team assets, prompt templates, and a question database. A middleware API layer applies governance and pedagogical controls (such as gating, validation, and logging) before calling the GenAI API. Outputs are then returned as modular learning components and are monitored through SME verification and a risk-assessment feedback loop to support transparency, safety, and iterative refinement.

Our conceptual case study is drawn from a core second-year undergraduate course in chemical engineering, Heat and Mass Transfer, focusing on heat conduction. Heat conduction is well-suited for implementing this framework because it combines conceptual understanding (including physical meaning of temperature gradients and boundary conditions) with numerical problem-solving (steady versus transient conduction, 1D conduction with various geometries, and thermal resistance analogies). The learning design process, therefore, requires both accurate understanding and reliable computations.

Figure 2 shows an example of content, template prompts, and the question database input in our platform.



(a)

```
You are a Heat Transfer TA.
Problem: {{problem_text}}
Student work: {{student_work}}
Course snippets: {{retrieved_snippets}}
Return exactly:
1) Assumptions (max 4)
2) Model choice + why (mention Bi/Fo if relevant)
3) Next 3 steps (no full solution)
If the student says "full solution", then provide it with units.
```

(b)

A solid sphere (radius $r_0 = 2.0$cm) is initially at a uniform temperature and is suddenly exposed to convection in air with h = 60 W/m²K. The sphere has k = 0.20 W//mK. Should the lumped-capacitance method be used?

A) Yes, because $Bi \approx 0.02$ (lumped is valid)
B) Yes, because Fo is always the only criterion
C) No, because $Bi \approx 6$ (internal gradients are significant)
D) No, because lumped is only valid for plane walls

Correct answer: C

(c)

Figure 2. Example of (a) content, (b) template prompts, and (c) question database input in out platform.

The platform uses instructor-controlled learning assets such as lecture slides and notes, solved problems, and a verified question database. These assets are paired with a library of "pre-prompts" that define the "instructor" role and required output workflow. The SME is responsible for maintaining prompt modules to ensure alignment with course/subject learning outcomes.

Learning outcomes are organized in the cognitive domain and mapped to Bloom's taxonomy at the levels of understand, apply, analyze, and evaluate for the studied topic. This mapping is crucial because it ensures the platform supports students in developing understanding and structuring solution scaffolds, as supporting tasks for their learning. Table 1 presents the learning outcomes mapped to Bloom levels and platform links.

Table 1. Learning outcomes mapped to Bloom levels and platform links.

Bloom level	Outcome	Platform support
Understand	Explain Fourier's law physically; interpret temperature gradients; state key assumptions	Brief concept explanation
Understand	Identify boundary conditions; state implications	Method outline through boundary conditions checklist
Apply	Compute heat flow for steady 1D conduction and use thermal resistance analogies	Method outline, steps, and fully worked solution
Apply	Apply lumped-capacitance and oner-term approximation	Method outline, steps, and fully worked solution
Analyze	Select the correct model (transient versus steady state)	Method outline through decision tree
Evaluate	Compare solution approaches and support assumptions	Conceptual explanation

The workflow automation pipeline is designed to treat generative AI as one component within a controlled instructional system. First, the teaching team's assets are "uploaded" to a structured content store, organized by topic and subtopic. Second, the question database is included, with relevant examples and constraints (e.g., geometry, boundary conditions, steady/transient, required units). Third, the automation layer assembles a request package that includes the appropriate prompt template and the selected question or student query. This package is then sent to the GenAI API to improve repeatability.

This workflow is also designed to support graduate attributes, aiming to produce evidence through auditable outputs aligned with outcomes, while maintaining a risk layer.

Students' interaction is modular, with a conceptual module that emphasizes assumptions and physical interpretation, a practical module that generates problems from the database, and an assessment module that focuses on quality checks (e.g., boundary conditions, alternative solution pathways). The pipeline includes a verification layer, in which the SME reviews a random/representative sample of outputs across modules to verify correctness and validity.

Despite our case study being conceptual, we have preliminary proposed effectiveness metrics to assess the future implementation of this tool. Thus, conceptual learning gains can be quantified using, for instance, pre/post multiple choice concept checks (aligned with the Bloom outcomes). Skill gains from numerical problems can be assessed using specific rubrics that focus on reasoning, model/calculation description, etc. The platform can also quantify the correct solutions and explain the sources of errors. Indicators of overreliance are also proposed, including the number of repeated requests for solutions, the omission of intermediate reasoning, etc. These flags will eventually allow adjusting the scaffolding and determining whether the platform supports learning.

Finally, the proposed framework and case study pipeline motivate a risk assessment, discussed in the next section, as the platform navigates across different systems (content stores, automation tooling, and external APIs) and modifies students' learning experience.

RISK ASSESSMENT

We use a risk matrix to score risks based on likelihood and impact. This assessment supports course planning decisions while aligning with all stakeholders. The scope of this exercise included the platform lifecycle described in the previous section (content, automation workflow, Gen AI/API calls, output, monitoring, and validation practices). Risk scenarios are defined by the teaching team and mapped to the following categories: (I) data privacy and protection, (II) transparency, (III) pedagogical impact, (IV) institutional compliance, and (V) ethical considerations.

In the data privacy and protection category, we consider risks/scenarios related to the disclosure, leakage, retention, or misuse of student data during transfers between systems.

In the transparency category, we consider risks/scenarios incorrect/unclear/difficult to audit GenAI outputs, unusual workflow logic, etc.

In the pedagogical impact category, we consider risks and scenarios in which the platform might undermine learning outcomes or promote overreliance.

In institutional compliance, we consider risks/scenarios in which institutional policies, as well as ethics/consent processes, are or might be violated.

In ethical considerations, we evaluate risks and scenarios in which automation (e.g., tool availability) might affect bias, fairness, and/or inequitable access.

Each risk scenario is scored along two dimensions

on a 1–5 scale: likelihood (L) or probability that the risk scenario occurs during the tool deployment, and impact (I) or severity of harm if the scenario occurs (to students, teaching team, and the institution).

The overall risk score (R) is estimated as R = L x I, producing values from 1 to 25. Scores are binned as follows: low (1-5), moderate (6-10), high (11-15), and critical (16-25). Scoring is conducted by the teaching team and the SME.

Mitigation actions are then defined, a responsible owner is assigned, and progress of implementation is tracked. Mitigations are prioritized by risk score and feasibility, and the assessment records the residual risk after mitigation (e.g., changes from high risk to moderate). Mitigation actions include design, technical, and human controls.

A summary of the risk assessment is shown in Table 2.

Table 2. Summarized risk assessment

Scenario	Category	L	I	R	Mitigation
Data transmitted to external services	I	4	5	20	Anonymization, protect LMS, retention policy
Incorrect numerical solutions	III	4	4	16	Human verification, include limitations and disclosure
Overreliance	III	4	4	16	Define scaffolding; provide pre-assumptions
Protected data rerouted	IV	3	5	15	Increase restrictions
Inequitable outcomes	V	3	3	9	Standardization of templates; provide other references beyond Gen AI
Drift after prompt	II	3	3	9	Maintain prompt templates and question database

RESULTS AND DISCUSSION

While the course is still not deployed, a preliminary analysis by the teaching team and the SME of the heat conduction module suggests that structuring GenAI interactions through instructor-suggested prompts and verified content can enhance conceptual clarity. For instance, slide-aligned definitions and worked examples support coherent narratives around core concepts such as Fourier's law, boundary conditions, and thermal resistance analogies. This framework is therefore expected to yield measurable benefits in concept checks targeting "understand" outcomes (e.g., interpreting the meaning of a temperature gradient) and "apply" outcomes (e.g., selecting the correct governing equation and solving a standard 1D conduction problem).

However, early testing highlights an important limitation for numerical problem solving, since solution reliability depends strongly on the diversity/rep–resentativeness of the question database and, hence, the benchmark set. When the database is small or limited (e.g., includes little or no information about mixed boundary conditions, changing geometry assumptions, or unit conventions), the system is more likely to produce apparently "reasonable" but incorrect steps and to omit assumptions. Numerical outputs should be treated as scaffolded rather than "linear", as flagged by the SME review.

When applying the risk matrix, risks are mostly clustered in three categories: data privacy and protection, transparency, and pedagogical impact. Privacy risk is quite elevated because automation platforms can unintentionally "leak" sensitive information through connectors, logs, intermediate files, or external APIs. Transparency risk is also elevated because it might be difficult for students to interpret limitations and for instructors to audit pipeline behavior. Finally, the pedagogical risk is elevated because the platform can indeed encourage shortcut-like behaviors unless scaffolding is provided and students are guided toward reasoning and assumptions.

Other categories, such as institutional compliance risk and ethical considerations, are moderate and depend on a specific scenario. Thus, with minimal data, institutional compliance declines, while LMS data routed through unapproved connectors may pose elevated risks. Ethical considerations can also be relative, as prompting skills might differ (e.g., expertise), leading to perceived inequities. Some recommended mitigation actions include: (i) minimal data pathways (e.g., student data anonymization, protected LMS routing, clear document retention and logging policies and practices), (ii) SME constant verification for numerical problems and periodic review of questions database, (iii) maintain/update prompt templates and content, and provide students with clear limitations on expected results, (iv) implement pedagogical scaffolding (provide several examples), and (v) standardize templates to ensure equity in terms of prompting expertise.

Although the heat conduction module presents specific risks, these can be easily extrapolated to other subjects; hence, the risk categories and scoring approach presented in this work are useful for transfer to other engineering education contexts where automation and coupled WAP/GenAI are introduced. The specific likelihood and impact values will, inevitably, vary by deployment details and risk team. Nonetheless, the evaluation's structure provides a reliable mechanism for adoption decisions across courses and departments.

The results are preliminary and primarily conceptual, relying on SME-designed testing rather than large-scale classroom trials. Therefore, learning gains are proposed rather than empirically established, and risk scores reflect expert judgment under stated assumptions only.

Future work should validate effectiveness metrics in a controlled pilot, test the platform with representative student use patterns, and refine mitigation strategies based on observed behaviors.

CONCLUSIONS

This paper presents a GenAI-aligned workflow automation platform designed to support modular, instructor-aligned learning outputs for undergraduate engineering courses. By combining validated content, assistance, and verification/monitoring, our approach aims to improve consistency and transparency when delivering outputs while reducing risks related to correctness, over-reliance, and data privacy. The framework also provides a practical pathway to document evidence aligned with CEAB Graduate Attributes.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the use of ChatGPT 5.2 Thinking for supporting the structuring of this paper's content. The conceptualization, analysis, validation, and writing have been performed by the authors.

AUTHOR IDENTIFIERS

Author ORCIDs:
Daniela Galatro: 0000-0003-0934-7492
Stuart Grey: 0000-0001-9274-0601
Sourojeet Chakraborty: 0000-0002-9461-880X

REFERENCES

1. Tuyishime A, Basciani F, Iovino L, Izquierdo JLC, Cabot J, Pierantonio A. Bridging workflow automation tools and EMF modeling ecosystems. 2023 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C) :893-897 (2023). https://doi.org/10.1109/models-c59198.2023.00140
2. Siva Prakash Bikka . Workflow automation engines: driving innovation in cloud-native and ai-enhanced business processes. Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol 11:282-289 (2025). https://doi.org/10.32628/cseit25111227
3. Marchena Sekli G, Godo A. Transforming teaching and learning with robotic process automation: a systematic review of pedagogical applications. JITE:Research 24:039 (2025). https://doi.org/10.28945/5655
4. Abduvohid X. The role and importance of using automated design systems in the educational process. eijp 5:75-79 (2025). https://doi.org/10.55640/eijp-05-09-16
5. Hawkinson E. Is education falling into the automation abyss? why struggle and human connection is essential for learning in an ai-driven world. together 2: (2025). https://doi.org/10.62883/ijpj5817
6. Ma Y, Li X, Hu S, Liu S, Cheong KH. Trustworthy AI in education: framework, cases, and governance strategies. Innov. Emerg. Technol. 12: (2025). https://doi.org/10.1142/s2737599425500264
7. Dabis A, Csáki C. AI and ethics: investigating the first policy responses of higher education institutions to the challenge of generative AI. Humanit Soc Sci Commun 11: (2024). https://doi.org/10.1057/s41599-024-03526-z
8. Geethanjali KS, Umashankar N. Enhancing educational outcomes with explainable AI: bridging transparency and trust in learning systems. 2025 International Conference on Emerging Systems and Intelligent Computing (ESIC) :325-328 (2025). https://doi.org/10.1109/esic64052.2025.10962568
9. Vivian R. From artificial intelligence to real dummies. Digital Technol. Res. Appl. 4:1-13 (2025). https://doi.org/10.54963/dtra.v4i2.1205
10. Westover JH. AI in education: building learning systems that elevate rather than erode human capability. HCLReview 27: (2025). https://doi.org/10.70175/hclreview.2020.27.4.4
11. Gonsalves C. Generative ai's impact on critical thinking: revisiting bloom's taxonomy. Journal of Marketing Education 48:4-19 (2024). https://doi.org/10.1177/02734753241305980
12. Tian J, Zhang R. Learners' AI dependence and critical thinking: the psychological mechanism of fatigue and the social buffering role of AI literacy. Acta Psychologica 260:105725 (2025). https://doi.org/10.1016/j.actpsy.2025.105725
13. Qian K, Liu S, Li T, Rakovi? M, Li X, Guan R, Molenaar I, Nawaz S, Swiecki Z, Yan L, Gaševi? D. Towards reliable generative ai-driven scaffolding: reducing hallucinations and enhancing quality in self-regulated learning support. Computers & Education 240:105448 (2026). https://doi.org/10.1016/j.compedu.2025.105448

© 2026 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See https://creativecommons.org/licenses/by-sa/4.0/