PROCEEDINGS

CITRENZ 2023 Conference
Auckland, 27–29 September

CITRENZ
COMPUTING AND INFORMATION TECHNOLOGY RESEARCH AND EDUCATION NEW ZEALAND

ePress
Research with impact
Unitec

Unitec is a business division of Te Pūkenga –
New Zealand Institute of Skills and Technology

sas

IT Professionals
NEW ZEALAND
Te Pou Hangarau Ngaio

TOI MAI
Workforce
Development
Council

ISBN: 978-1-99-118344-6

Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September

The CITRENZ 2023 Conference was held at Unitec's Te Whare Wānanga o Wairaka Campus from 27 to 29 September. All papers published in these proceedings have been double-blind peer-reviewed by two referees. The papers in this publication comprise the proceedings of the CITRENZ 2023 Conference. They reflect the authors' opinions and their inclusion in this publication does not necessarily constitute endorsement by the editor, ePress, Unitec or CITRENZ.

Editor: Hamid Sharifzadeh
Cover design: Penny Thomson

This publication may be cited as:

Sharifzadeh, H. (Ed.). (2024). *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September*. ePress, Unitec. https://doi.org/10.34074/proc.2401

An ePress publication

epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025
Victoria Street West, Auckland 1010
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

ePress
Research with impact
Unitec

Unitec is a business division of Te Pūkenga –
New Zealand Institute of Skills and Technology

Contents

Foreword

Professor Hamid Sharifzadeh
CITRENZ 2023 Technical Programme Chair
School of Computing, Unitec

https://doi.org/10.34074/proc.240101
Correspondence: hsharifzadeh@unitec.ac.nz

This publication may be cited as:

Sharifzadeh, H. (2024). Foreword. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ Conference 2023, Auckland, 27–29 September* (pp. 5–6). ePress, Unitec. https://doi.org/10.34074/proc.240101

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga – New Zealand Institute of Skills and Technology

The School of Computing at Unitec was proud to host the CITRENZ 2023 conference in Auckland from 27 to 29 September 2023.

> Computing and Information Technology Research and Education New Zealand (CITRENZ) was formed as an organisation in 2010 as a result of an expansion of the National Advisory Committee on Computing Qualifications (NACCQ), which dates back to 1988. CITRENZ has supported academic staff across Aotearoa New Zealand in various ways, including workshops, panels, research seminars, assessment of prior learning, moderation, and by organising an annual conference. The philosophy of CITRENZ is the encouragement and support of new, emerging and established researchers while encouraging excellence and academic discourse. (CITRENZ, n.d., https://citrenz.itp.nz/about-citrenz)

CITRENZ 2023 was the 14th annual CITRENZ conference, and the theme was *Collective Intelligence: Exploring Mahi Kotahitanga (Collaboration) in Computing*. The conference featured three keynote speakers, three panel discussions, two special sessions, five paper presentation sessions and two parallel poster sessions. The 2023 conference was sponsored by SAS analytics Australia, Toi Mai Workforce Development Council, and IT Professionals NZ (ITPNZ).

We received 71 submissions for the conference, with 45 submissions accepted after peer review. This included 21 papers and 24 posters. The papers covered a wide range of applied computing topics, such as teaching and learning in the era of generative AI, institutional capacity-building and student success, research reports and more. These proceedings present the full text of 20 accepted papers that were presented at the conference. The technical committee also selected the following two papers for the best paper awards:

> *Embracing the Use of Generative AI in a First-Year Information Systems Course*
> by Angela Martin and Trevor Nesbit

Feedback from the Technical Committee: The concept of engaging students not only in using ChatGPT but also in encouraging peer discussions, critical analysis of its responses, and relating them to individual cultural contexts is highly commendable. Your initiative has the potential to serve as a valuable model for diverse courses and assignments across different educational levels.

> *Assessment Validity in the Era of AI Generative Tools*
> by Eltahir Kabbar and Bashar Barmada

Feedback from the Technical Committee: This paper systematically explores the evolving education landscape in the era of generative AI. Although in its early stages, the presented risk-identification framework offers a direction for educators navigating the integration of AI tools.

On behalf of the conference organising committee, I would like to thank the team of academics who contributed their time to peer review the papers and posters. We are also grateful for the outstanding work of the Unitec ePress staff, without which we could not have offered this publishing opportunity, and for all the conference participants who seized the opportunity and submitted their contributions.

Ngā mihi nui kia koutou katoa.

Embracing the Use of Generative AI in a First-Year Information Systems Course

Angela Martin and Trevor Nesbit

https://doi.org/10.34074/proc.240102
Correspondence: trevor.nesbit@canterbury.ac.nz

This publication may be cited as:

Martin, A., & Nesbit, T. (2024). Embracing the Use of Generative AI in a First-Year Information System Course. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ Conference 2023, Auckland, 27–29 September* (pp. 7–14). ePress, Unitec. https://doi.org/10.34074/proc.240102

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga – New Zealand Institute of Skills and Technology

ABSTRACT

The use of generative artificial intelligence (AI) tools in higher education has gained a lot of attention in the media. Much of this narrative is about the use of tools such as ChatGPT as a way for students to cheat on the assessments they submit. This paper presents an analysis of how ChatGPT has been used in the design and delivery of a first-year information systems course. The background of the course and the institution that it is delivered in is presented, along with the design of the research, which is consistent with the early phases of an action research project.

A literature review is presented that highlights the use of tools such as ChatGPT for developing hypothetical scenarios; the importance of students developing AI literacy skills; preparing the students for workplaces where these tools are being used; and the importance of considering the context of students when seeking to enhance student engagement.

The design and delivery of the course is described, analysed and compared with the outcomes of the literature review and a revised version of Bloom's Taxonomy of Learning. It is concluded that tools such as ChatGPT can be used for the effective design and delivery of courses in ways that enhance student engagement, increase student AI literacy, enhance the employability of graduates, and address all levels of a revised version of Bloom's Taxonomy of Learning.

Areas in which to extend the research include conducting focus groups with students from the first and subsequent iterations of the course examined in this study.

KEYWORDS

Generative AI, ChatGPT, student engagement, socio-cultural learning, scaffolding

INTRODUCTION

The year 2023 has seen a significant amount of interest in the media relating to the emergence of generative artificial intelligence (AI) tools such as ChatGPT. Much of the media narrative has been around how students can use tools such as ChatGPT to, in essence, cheat in their assignments, essays and other submitted work. Meanwhile, some higher-educational institutions have set about embracing generative AI tools as a method for enhancing teaching and learning, with some of this including the use of generative AI tools by staff in preparing resources, students evaluating the use of generative AI tools, and preparing students for employment in roles where generative AI is being used.

Much research is emerging that highlights this, including the work of Bowles and Kruger (2023), who comment that "with work, educators can minimize its genuine disadvantages for assessment while leveraging potential advantages to continue to deliver enhanced critical thinking and employment outcomes for students" (Bowles & Kruger, 2023, p. 76).

The purpose of this paper is to describe and evaluate how an information systems course at the University of Canterbury International College (UCIC) has been redesigned to incorporate the use of ChatGPT in the design and delivery of the course.

The structure of the paper includes background information about the course and UCIC, a short literature review, description of the redesigned course, analysis and discussion, and conclusions and next steps in the research. The findings and outcomes will be useful to educators considering the adoption of generative AI tools as an integral part of their courses.

BACKGROUND OF UCIC AND THE COURSE

The University of Canterbury International College (UCIC) runs as a pathway college into the University of Canterbury (UC). All students are from an international background and are part of UCIC's Foundation programme or the University Transfer Programme. During the University Transfer Programme, students complete the first year of their degree choice with UCIC then move on to UC to follow through with further years of study in the degree areas and faculties they have chosen to study in. The course that is the focus of this paper is Information Systems (INF001). The course gives students credit for a first-year information systems course that is a required course for the Bachelor of Commerce at UC.

AIMS AND RESEARCH DESIGN

The aim of this paper is to describe how ChatGPT was used in the design and delivery of INF001 in its first iteration of 2023 and how this relates to the literature of student engagement, the use of generative AI tools in higher education, and other literature relating to teaching and learning.

This paper has its basis partly in aspects of action research (Stringer, 2013), with this first iteration of the course forming the observation and reflection phases, and future iterations relating to revising and implementing the course coming later as consideration is given to changes that could be made in future iterations. The importance of reflective processes proposed in the PARA model (Ganly, 2018) could be used to connect reflection with practice. In the PARA model there are four phases making up reflective practice: Pause, Attend, Revise, Adopt/Adapt (Ganly, 2018). In this paper, the authors are pausing to attend to what has been experienced so that the practice could potentially be revised for future iterations.

LITERATURE REVIEW

The topics covered in this literature review include the importance of student AI literacy when it comes to tools such as ChatGPT, the use of generative AI tools for generating hypothetical scenarios to be used in courses, the importance of preparing students to work in roles where generative AI tools are being used, the importance of context when considering student engagement in higher education, and a revised version of Bloom's Taxonomy of Learning.

The idea of using generative AI tools for developing hypothetical scenarios has been identified in a growing number of studies, with the ability of tools such as ChatGPT to generate text being a major focus in the study conducted by Ivanov and Soliman (2023). This includes the ability to generate short stories based on relatively simple instructions from users. These short stories could take the form of short case-studies or scenarios that are provided to students. The concept of using generative AI tools to generate a range of scenarios that could be utilised in a learning context was identified by Hayes (2023), with this concept of generating hypothetical scenarios also being identified by Nguyen et al. (2023).

The importance of students developing AI literacy, particularly as it relates to the use of generative AI, has also been receiving growing attention in the literature, with Chan and Hu (2023) commenting on the importance of student AI literacy and how this is essential for the responsible use of generative AI tools. The need to develop AI literacy amongst students (and staff) was part of the findings of the study conducted by Ivanov and Soliman (2023), with this including the importance of students and staff not putting too much trust in the output of generative AI tools but needing to rely on their expertise, other (academic) sources and common sense (Ivanov & Soliman, 2023).

The need to prepare graduates for roles in industry where generative AI tools are being used is attracting growing attention in the literature. In addition to the comments about AI literacy in Chan and Hu (2023,) they go on to comment that employers may raise their recruitment requirements, and those that do not develop the right skills may fall behind. Ivanov and Soliman (2023) also comment that tools such as ChatGPT will enable students to gain digital skills that are sought after by employers. The study conducted by Bowles and Kruger (2023) includes the comment that the ability to use generative AI tools is an employable skill, and they go on to comment that "employers can work to inform education providers with a clearer sense of the skills that will be needed in the workplace of the future" (Bowles & Kruger, 2023, p. 75).

The importance of considering the context of students when designing a course with the aim of increasing their level of engagement has been commented on in the literature. In a study conducted by Sun and Holt (2022) it is noted that those disseminating information should take student context into consideration when making decisions about how a course should be designed. Connolly et al. (2022), when reporting on a study into increasing student engagement in an introductory information systems course, comment that the key is to follow the philosophies of socio-cultural learning. The concept of socio-cultural learning is related to how individuals learn and construct meaning with a social context, in which the instructor provides a starting point for learning through social interactions (Connolly et al., 2022).

Bloom's Taxonomy of Learning was revised by Anderson and Krathwohl (2001) and is summarised in Table 1.

Table 1. Revised Bloom's Taxonomy of Learning (Anderson & Krathwohl, 2001).

Level	Description
Create	Producing new or original work
Evaluate	Justifying a stand or decision
Analyse	Drawing connections among ideas
Apply	Using information in a new situation
Understand	Explaining ideas or concepts
Remember	Recalling facts and basic concepts

In summary, key factors to emerge from the literature review are:

- The concept of using generative AI tools for developing hypothetical scenarios for learning materials (Ivanov & Soliman, 2023; Hayes, 2023; Nguyen et al., 2023).
- The importance of students developing AI literacy skills (Chan & Hu, 2023; Ivanov & Soliman, 2023).
- The importance of preparing graduates for the workplace, where generative AI tools will grow in use (Chan & Hu, 2023; Bowles & Kruger, 2023).
- Making sure that the contexts of students are considered when developing strategies for increasing student engagement (Sun & Holt, 2022; Connolly et al., 2022).

THE REDESIGNED COURSE

Given that the need to prepare students for using generative AI in industry has emerged (Bowles & Kruger, 2023), it was decided to incorporate the use of ChatGPT into the course in a structured and purposeful way to give students some experience in its use; as part of this, aspects of the course (specifically the tutorials) were reorganised.

In the past, these tutorials were delivered in such a way that every student had the same set of discussion questions they were required to answer. Attempts had been made to incorporate group work into the tutorials, with this resulting in little engagement. Each tutorial was designed to use group work to aid student understanding and to have individual students submit work by the end of the tutorial.

Reorganisation of Course

For this iteration of the course, changes were made to incorporate generative AI content. However, this is not as simple as generating content and letting the students read it. It was decided to approach this both horizontally (across the length of the course), and vertically (delving deeper into understanding in each week). The tutorials were reorganised to incorporate the use of questions, scenarios and group work.

Departmental Scenarios

It was decided to base the tutorials around a number of departments in a hypothetical organisation, which for this iteration of the course were accounting, sales, marketing, social club, human-resource management, logistics and inventory.

Each department needed a scenario that would include aspects that related to each topic in the course. These departmental scenarios were generated for the hypothetical organisation using ChatGPT, with the scenarios being shared with the students in the second week of the course. The students were asked to select their preferred departmental scenario, and were allocated to groups based on their preferences for the duration of the course.

Many of the students selected preferred departmental scenarios based on what they planned to major in for their undergraduate degree at UC. Three more scenarios were created than were needed for the numbers of students in the course, with this being done to optimise the allocation of students to scenarios that they were interested in.

Internationalisation

In the first week of the course, an audience response system (Socrative) was used to gain an understanding of what other courses the students were enrolled in and what their home country was, with this information being able to be revisited in the three topics of the course – competitive advantage, information systems development, and e-commerce.

The students worked in the groups based on their departmental scenarios, with this work including asking the students questions about how aspects of the topic differed in their home country and how this might be applied to the departmental scenario that their group had been allocated.

Use of Generative AI

During the tutorials across the course, the students worked in the groups based on their departmental scenarios and were asked to complete five tasks or activities that were associated with ChatGPT:

1. Discuss their departmental scenario in their groups to ensure that they could remember and understand the scenario (Week 1).
2. Work in their groups to use ChatGPT to generate a description of their departmental scenario of approximately 100 words (Week 3).
3. Use ChatGPT to create a multiple-choice question about big data and business intelligence (the topic for the week) and provide four possible answers to the question. Each student created their own question, and these were then discussed in the group and refined so that everyone in the group was happy with the result (Week 7).
4. Make a group presentation that included, among other things, what they had learned about the use of ChatGPT during the course (Week 10).

5. Take the original departmental scenario that they were provided with at the start of the course and evaluate its content, after having been told that it had been generated by ChatGPT. This included discussing whether they had the knowledge at the start of the course to evaluate the scenario compared with the knowledge they had at the end of the course, and whether this concept could be applied to other courses that the students had studied or would study in the future. As part of this, the students were asked to re-create the departmental scenario (Week 11).

Students Generating Questions

The students were asked to work in their groups to use PeerWise to generate a range of questions from listing to describing, then comparing, followed by comparing and contrasting. This was set up in the tutorials in the first week, with the questions being about how things could be different in the students' home countries. By the end of the course the students were generating substantial questions relating to their departmental scenarios.

Summary

In summary, the course is designed to get students to analyse and use their critical-thinking skills in a context that is known to them (their home country) and applied to the departmental scenarios. This is fleshed out within the groups the students are in. The students are given permission to use generative AI (ChatGPT), but within the bounds of good understanding.

ANALYSIS AND DISCUSSION

The use of the departmental scenarios for the student groups was consistent with a number of aspects in the literature, including considering the context of students (Sun & Holt, 2022), with this connecting with the idea of socio-cultural learning (Connolly et al., 2022) when it came to students being able select their preferred departmental scenarios based on what they had the most interest in. This also connects with the internationalisation aspects, where students were able to reflect on how things differed in their home countries.

The use of ChatGPT to generate the initial departmental scenarios that the students were provided with, along with the students using ChatGPT to generate questions, is consistent with the idea of using generative AI tools to generate new or hypothetical content that has been referred to by a number of authors (Ivanov & Soliman, 2023; Hayes, 2023; Nguyen et al., 2023).

That the students were given exposure to the use of several tasks in the tutorials across the course was done with the intention of increasing the AI literacy of students, which has been identified in the literature as being important (Chan & Hu, 2023; Ivanov & Soliman, 2023). Those tasks included the evaluation of content generated by ChatGPT and the use of ChatGPT to generate descriptions and questions.

The tasks and activities associated with the use of ChatGPT can be mapped onto the revised Bloom's Taxonomy of Learning (Anderson & Krathwohl, 2001):

- The first task/activity, in Week 1, of students discussing their departmental scenarios in groups so that they could remember and understand the scenarios, maps directly onto the levels of 'Remember' and 'Understand' in the revised taxonomy as shown in Table 1.
- The second task/activity, in Week 3, of generating a 100-word description of their departmental scenario, maps onto the level of 'Apply' in the revised taxonomy as shown in Table 1. This is because the students are using ChatGPT to generate the description.
- The third task/activity, in Week 7, of each student using ChatGPT to create a multiple-choice question and having the group analyse all the questions and revise them, maps onto the level of 'Analyse' in the revised taxonomy as shown in Table 1.

- The fourth task/activity, in Week 10, of making a group presentation that included some aspects about the use of ChatGPT in the course, maps onto the level of 'Evaluate' in the revised taxonomy as shown in Table 1. Some of the requirements for the presentation required the students to evaluate the use of ChatGPT.
- The fifth task/activity, in Week 11, of recreating the departmental scenario, maps onto the level of 'Create' in the revised taxonomy as shown in Table 1.

At the time of submitting this paper, six of the 12 weeks of the course had been completed. At this stage the only likely change to the course that had been identified for its next iteration was a possible stronger connection being made to the internationalisation aspect. Once the first iteration of the course has been completed there may be other changes that are contemplated, with this process being consistent with the action research model (Stringer, 2012) and the PARA model (Ganly, 2018) as outlined earlier in the paper.

CONCLUSION AND NEXT STEPS IN THE RESEARCH

The design of this iteration of the course addressed some of the key aspects surrounding the use of generative AI tools (such as ChatGPT) in higher education, approaches to increasing student engagement in higher education, and creating more AI literacy amongst students, with all of this happening in the context of international students.

- The importance of considering the context of students, which can include the home countries of the students, and the concept of socio-cultural learning.
- How tools such as ChatGPT can be used to generate hypothetical scenarios that can be used as the basis for tasks and activities in teaching and learning, and can serve to enhance the engagement of students by creating scenarios that students have a personal level of interest in.
- That incorporating the use of tools such as ChatGPT into the teaching and learning process can serve to increase the AI literacy of students, and potentially enhance the employability of the students.
- That tools such as ChatGPT can be incorporated into a teaching and learning process that addresses all levels of the revised version of Bloom's Taxonomy of Learning (Anderson & Krathwohl, 2001).

The next step that is planned in this research is to conduct focus groups with the students at the end of this and future iterations of the course. The aim of the focus groups is to gain an in-depth understanding of the student perceptions of the processes that were used in the course, and the impact of any changes that are made in future iterations of the course. Another step that is planned is to interview lecturers about their experiences with generative AI tools such as ChatGPT in their teaching. At the time of writing, an ethics application has been made to the UC Human Ethics Committee for approval to conduct these steps.

REFERENCES

Anderson, L., & Krathwohl, D. (2001). *A taxonomy for learning, teaching, and assessing*: *A revision of Bloom's Taxonomy of Educational Objectives*. Longman.

Bowles, D. C., & Kruger, J. S. (2023). Generating employable, intelligent graduates in a world with generative AI: Thoughts for educators. *Pedagogy in Health Promotion*, *9*(2), 75–77. https://doi.org/10.1177/23733799231175171

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *arXiv*. https://doi.org/10.48550/arXiv.2305.00290

Connolly, A. J., Mutchler, L. A., & Rush, D. E. (2022). Teaching tip: Socio-cultural learning to increase student engagement in Introduction to MIS. *Journal of Information Systems Education*, *33*(2), 113–126. https://jise.org/Volume33/n2/JISE2022v33n2pp113–126.html

Ganly, T. (2018). Taking time to pause: Engaging with a gift of reflective practice. *Innovations in Education and Teaching International*, *55*(6), 713–723. https://doi.org/10.1080/14703297.2017.1294492

Hayes, A. (2023). “Conversing” with qualitative data: Enhancing qualitative sociological research through large language models (LLMs). *SocArXiv Papers*. https://osf.io/preprints/socarxiv/yms8p

Ivanov, S., & Soliman, M. (2023). Game of algorithms: ChatGPT implications for the future of tourism education and research. *Journal of Tourism Futures*, *9*(2), 214–221. https://doi.org/10.1108/jtf-02-2023-0038

Nguyen, T., Cao, L., Nguyen, P., & Nguyen, P. (2023). Capabilities, benefits, and role of ChatGPT in chemistry teaching and learning in Vietnamese high schools. *EdArXiv Preprints*. https://osf.io/preprints/edarxiv/4wt6q

Stringer, E. T. (2013). *Action research*. Sage Publications.

Sun, X., & Holt, D. (2022). Student engagement and voice in higher education: Students’ perceptions. *Journal of Learning Development in Higher Education*, *23*, 1–21. https://doi.org/10.47408/jldhe.vi23.838

AUTHORS

Angela Martin is a Programme Convenor and Lecturer at the University of Canterbury Te Whare Wānanga o Waitaha International College. Her research interests include enhancing student experience and engagement. angela.martin@ucic.ac.nz

Dr Trevor Nesbit is a Senior Lecturer in Computer Science and Software Engineering at the University of Canterbury Te Whare Wānanga o Waitaha. His research interests are in educational technologies, including the use of audience-response systems to enhance student engagement and in the uses of generative AI in higher education. trevor.nesbit@canterbury.ac.nz

ChatGPT 3.5: Another Tool or a New Collaborative Project Team Member?

Robert Nelson, Samuel Mann and Ruth Myers

https://doi.org/10.34074/proc.240103
Correspondence: Robert Nelson, rob.nelson001@gmail.com

This publication may be cited as:

Nelson, R., Mann, S., & Myers, R. (2024). ChatGPT 3.5: Another Tool or a New Collaborative Project Team Member? In H. Sharifzadeh (Ed.), Proceedings: *CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 15–19). ePress, Unitec. https://doi.org/10.34074/proc.240103

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga –
New Zealand Institute of Skills and Technology

ABSTRACT

The increasing use and capabilities of the large language model ChatGPT 3.5 are indicative of its potential for knowledge-work tasks such as web searching and generating various types of creative literature. This makes it a useful tool for project teams working on collaborative tasks. It also raises the question of whether ChatGPT 3.5 can be more than a tool – could human team members collaborate with it in a similar way to collaborating with each other, and what might this collaboration look like? This paper discusses a team-based project simulation in which team members interacted with ChatGPT 3.5 in a team-member role to produce a series of text-based deliverables. Members' reflections and researcher observations showed that the working relationship was a positive one, and that training in effective use of ChatGPT 3.5, together with more opportunities to use it in a collaborative role, could lead to production of more fit-for-purpose deliverables in shorter timeframes.

KEYWORDS

ChatGPT, collaborative projects, teamwork

INTRODUCTION

One of our most important current and future challenges, whether as industry practitioners or learning facilitators, is the rapid development and deployment of AI-powered language models; specifically, in the case of this paper, ChatGPT 3.5. When deployed as part of a collaborative team project, ChatGPT 3.5 may be more than just another technical tool. Despite its ability to generate new content based on content it has been trained with (Gill & Kaur, 2023), the limitations of its training content may restrict the relevance and accuracy of its output (Marr, 2023). There is still uncertainty around its optimal use (Hosseini et al., 2023).

ChatGPT 3.5 is, therefore, currently positioned between the innovation and early adoption stages of Rogers' (1962) Diffusion of Innovation model. In this paper we describe an experiment in which a project team carried out a simulated collaborative project, using ChatGPT 3.5 in the role of a team member to provide creative input to address transdisciplinary project deliverables. The experiment was carried out as a part of the lead researcher's work towards a Doctor of Professional Practice (DPP) qualification at Otago Polytechnic, investigating holistically integrated collaborative project design and practice.

AIMS AND RESEARCH DESIGN

The aim of the research was to explore how ChatGPT 3.5 could be used as a project team member, rather than simply as a research tool, and how this influenced team members' interactions with ChatGPT and with each other. No limits were placed on how the team could interact with ChatGPT 3.5, allowing for naturalistic behaviour to occur.

A team of three members was tasked to collaborate with ChatGPT 3.5 to solve a wicked, transdisciplinary management problem. The team comprised staff working within the College of Work-Based Learning, Otago Polytechnic; two in academic roles and one in an administrative role, and who had differing experience levels using ChatGPT 3.5 prior to the experiment.

Team members were provided with scenario details five days before the simulation took place, to familiarise themselves with broad task requirements. Because time for the experiment was operationally limited to two hours, this allowed possible queries over process to be identified and addressed before the substantive project work commenced.

The project team members were working in the same meeting room within their normal workplace, while the simulation was observed remotely by the lead researcher, using Microsoft Teams.

The experiment was divided into three task groups. Firstly, the team had to create an identity and set of standard operating procedures for the team. Secondly, the main objectives of the task were to create a project proposal to diversify the product/service offerings of a small suburban service station at risk of losing its oil company supply contract and, therefore, the majority of its livelihood. Primary and secondary research would lead to a formal business plan, including a marketing plan, together with appropriate appendices/artefacts, and an oral presentation. Thirdly, an oral debrief was led by the lead researcher to capture first impressions on the experience of working with ChatGPT as a team member while the experience was fresh in participants' minds. These impressions were supplemented by more detailed written reflections completed by team members within the following 72 hours.

Thematic analysis (Clarke & Braun, 2013) allowed for different theoretical perspectives to be considered. Surface themes were determined at a semantic level (there is nothing beyond what participants said), and at a latent (deeper) level, where ideas and assumptions underlying what participants said were identified (Braun & Clarke, 2006).

KEY FINDINGS, CONCLUSIONS AND RECOMMENDATIONS

Based on lead researcher observation and project participants' reflections, the following issues and themes were noted:

- Differing prior levels of experience with ChatGPT 3.5 meant an extended learning curve for some team members, impacting their ability to contribute as fully as others (Wrike, n.d.).
- Participants seemed to realise early on in the process that what they got back from ChatGPT 3.5 was highly dependent on the prompts they gave it – the more specific these were, the more specific its responses were (Gerwitz, 2023; Cook, 2023).
- Conflict observed between team members during the exercise was positive, aimed at generating options in response to the brief. The member with the least amount of prior ChatGPT 3.5 experience seemed to take longer analysing ideas before further developing and contributing them than the others, but this may also have been a manifestation of their natural working style, independent of this particular learning situation. There was no obvious conflict noted or reported between team members and ChatGPT 3.5; nor were there any responses from ChatGPT 3.5 to changing instructions that might have been interpreted as emotional/subjective rather than objective and compliant.

A particular challenge noted was learning to "think at the speed of ChatGPT", which may have produced content more quickly, but there was a feeling this didn't allow sufficient time for human reflection and debate.

Having learned in the specific context of this project, participants were unanimous in the belief that working on future projects would be easier as a result of this experience, and they would likely produce more detailed deliverables (in text form) in a shorter time and with greater confidence.

Limitations of the research and recommendations for future research:

- There are three significant limitations to the research undertaken to date. The first is that the simulation used ChatGPT 3.5 rather than ChatGPT 4. ChatGPT 3.5 produces output more quickly than ChatGPT 4, although with a writing style that may not be as polished (Mast, 2023). In comparison, ChatGPT 4 offers more concise and relevant responses, with improved comprehension and ability to understand more nuanced instructions (*The Times of India*, 2023; Umair, 2023). **It is recommended that** further simulations be carried out using both ChatGPT 3.5 and ChatGPT 4 to better understand the advantages and disadvantages of each as a collaborative-project team member.

- Secondly, the findings are based on a single simulation exercise. To draw more robust conclusions about their representativeness, and about the possibilities of ChatGPT 3.5 as a collaborative team member, **it is recommended that** further simulations need to be carried out to better understand the challenges, limitations, and opportunities for integrating ChatGPT 3.5 as a team member rather than as a tool.
- Thirdly, while the project scenario was made available to participants several days ahead of the simulation being carried out, no pre-training was carried out in how to use ChatGPT 3.5. This may have disadvantaged participants who were either new to ChatGPT 3.5, or who had limited experience in its use, by limiting the extent to which they could play a full part within the team. This disadvantage may have been exacerbated by the relatively short timeframe for completing the project. **It is recommended that** training be provided in the use of ChatGPT 3.5 and ChatGPT 4 as a precursor to future experiments of this type.

REFERENCES

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*, 77–101. https://doi.org/10.1191/1478088706qp063oa

Clarke, V., & Braun, V. (2013). Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *The Psychologist*, *26*(2), 120–123. https://uwe-repository.worktribe.com/output/937596

Cook, J. (2023, July 19). 7 ChatGPT prompts to write better content (And not sound like a robot). *Forbes*. https://www.forbes.com/sites/jodiecook/2023/07/19/7-chatgpt-prompts-for-writing-better-content/?sh=194c6b802952

Gerwitz, D. (2023, October 5). *7 advanced ChatGPT prompt-writing tips you need to know*. ZD Net. https://www.zdnet.com/article/7-advanced-chatgpt-prompt-writing-tips-you-need-to-know/

Gill, S. S., & Kaur, R. (2023). ChatGPT: Vision and challenges. *Internet of Things and Cyber-Physical Systems*, *3*, 262–271. https://doi.org/10.48550/arXiv.2305.15323

Hosseini, M., Gao, C. A., Liebovitz, D., Carvalho, A., Ahmad, F. S., Luo, Y., MacDonald, N., Holmes, K., & Kho, A. (2023). An exploratory survey about using ChatGPT in education, healthcare, and research. *medRxiv*. https://doi.org/10.1101/2023.03.31.23287979

Marr, B. (2023, March 3). The top 10 limitations of ChatGPT. *Forbes*. https://www.forbes.com/sites/bernardmarr/2023/03/03/the-top-10-limitations-of-chatgpt/?sh=761780b38f35

Mast, J. (2023, May 7). *"Outwitting the devil" – a comparison between ChatGPT3.5 and ChatGPT 4.0*. Jonathan Mast. https://jonathanmast.com/outwitting-the-devil-a-comparison-between-chatgpt-3-5-and-chatgpt-4-0/

Rogers, E. M. (1962). *Diffusion of innovations*. Free Press of Glencoe.

The Times of India. (2023, March 15). ChatGPT-4 vs ChatGPT-3.5: Five things that are new and different. https://timesofindia.indiatimes.com/gadgets-news/chatgpt-4-vs-chatgpt-3-5-5-things-that-are-new-and-different/articleshow/98661079.cms

Umair, S. A. (2023). *ChatGPT4 vs ChatGPT 3.5: A concise comparison*. LinkedIn. https://www.linkedin.com/pulse/chatgpt-4-vs-35-concise-comparison-syed-ali-umair/

Wrike. (n.d.). *What are the disadvantages of collaborating?* https://www.wrike.com/collaborative-work-guide/faq/what-are-the-disadvantages-of-collaboration/

AUTHORS

Robert Nelson is a Doctor of Professional Practice ākonga in the College of Work-Based Learning, Otago Polytechnic Te Kura Matatini ki Otago. His research interests are in collaborative project work and collaborative assessment practice. rob.nelson001@gmail.com

Dr Samuel Mann is a Professor in the College of Work-Based Learning, Otago Polytechnic Te Kura Matatini ki Otago. His research interests are in computer science education and sustainability. Samuel.mann@op.ac.nz

Dr Ruth Myers is a Facilitator in the College of Work-Based Learning, Otago Polytechnic Te Kura Matatini ki Otago. Her research interests are in performative body practice and early film histories. Ruth.myers@op.ac.nz

Pilot Study of Using ChatGPT in a First-Year Programming Assignment

Minjie Hu, Tony Assadi and Hamid Mahroeian

https://doi.org/10.34074/proc.240104
Correspondence: Minjie.Hu@weltec.ac.nz

This publication may be cited as:

Hu, M., Assadi, T., & Mahroeian, H. (2024). Pilot Study of Using ChatGPT in a First-year Programming Assignment. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 20–25). ePress, Unitec. https://doi.org/10.34074/proc.240104

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga –
New Zealand Institute of Skills and Technology

ABSTRACT

ChatGPT can be a useful aid in computer programming education, due to its cutting-edge functionality of generating program code, debugging, etc. This research focuses on both ethical considerations and the impact on student performance of using ChatGPT. To ensure students used ChatGPT ethically, guidance was provided together with a declaration form that stated that ChatGPT was used ethically by students in each phase of the assignment. Next, we collected and analysed a survey and the declarations from students, and compared student effort, time spent and performance outcomes from those who were and were not using ChatGPT. Based on the findings from the survey and the comparison of data about student efforts and learning outcomes in both 2022 and 2023, we concluded that ChatGPT provides an opportunity for first-year students to learn programming through analysis, synthesis, and evaluation. Many students still preferred the conventional way of learning programming in terms of comprehension and application. We argue that because our students in the programming course are from different academic background levels, we should continue to use both ChatGPT and conventional e-learning resources to meet different learning requirements.

KEYWORDS

ChatGPT, first-year programming, divide-and-conquer strategy, Bloom's Taxonomy

INTRODUCTION

The emergence of the ChatGPT (Chat Generative Pre-Trained Transformer) chatbot has impacted how first-year programming education is taught and assessed. This is mostly due to ChatGPT's ability to carry out various complex tasks, such as generating and improving computer program code, debugging code, and providing comments and explanations (Sadik et al., 2023). As a result, it can enhance students' problem solving, critical thinking, code performance and reliability (OpenAI, 2023). However, it raises some challenges and potential risks of academic dishonesty because students might use ChatGPT to write code without adequately acknowledging the source (Villasenor, 2023).

Furthermore, ChatGPT has difficulty dealing with complex logic when there is insufficient context and background knowledge (Sadik et al., 2023), because the quality of its outputs entirely depends on the AI training data (Surameery & Shakor, 2023). Subsequently, ChatGPT may generate inaccurate and irrelevant code, which could lead to poor learning outcomes (Diaz, 2023). Hence, there is a risk to students who may rely too heavily on ChatGPT and fail to develop their own coding skills or understanding (Roos, 2023).

However, there is a lack of published research on the topic of using ChatGPT in programming education, especially for first-year students at tertiary level such as Bachelor of Information Technology (BIT) and the Level 5 Diploma in Information Systems (DipIS L5). Therefore, this research aims to explore the impact of ChatGPT on students' learning in computer programming and to develop teaching strategies that leverage ChatGPT's capabilities to improve the effectiveness and efficiency of computer programming education.

AIMS AND RESEARCH DESIGN

Research Questions:

- RQ1: What are the ethical considerations and solutions for first-year BIT and DipIS L5 students who use ChatGPT to write computer programs in their assignments?
- RQ2: What is the impact of ChatGPT on the programming competencies and learning outcomes of first-year BIT and DipIS L5 students compared to those who do not use ChatGPT?

This study examined how students used ChatGPT to help them complete an assignment in the first-year programming course in the BIT and DipIS L5 at Whitireia/WelTec | Te Pūkenga. The assignment required students to develop a C# GUI project using object-oriented programming (OOP) techniques, including arrays, lists and files, for a car-dealership business application. This application contained four phases: (1) validated input and output; (2) processing complex data with arrays and methods; (3) basic file management; and (4) searching data under complex conditions.

To support students on this integrated application, a prototype of a wage process application system was introduced in two workshops. It included a simple version of documents and code, such as simple functions without validation control for inputting, saving and searching values, without using array and static method, simple condition for searching, etc.

The two primary components for evaluating this study were: (1) ensuring students used ChatGPT ethically; and (2) collecting relevant data impacted by using ChatGPT. First and foremost, guidance was provided to guarantee that students were utilising ChatGPT in an ethical manner, specifically including the divide-and-conquer strategy (Wegner, 1990) to divide the assignment project into many small sub-programs for using ChatGPT.

Consequently, the evaluating, analysing and synthesising skills were attempted to integrate these generated sub-programs into a fully functional program. Therefore, our proposed method provided students a chance to learn programming through the methods of analysis, synthesis, and evaluation (Thompson et al., 2008).

Moreover, a declaration form regarding the ethical use of ChatGPT in each phase of the assignment was mandatory for all students. Students were required to declare whether they did use ChatGPT ethically in four phrases of this project, including: (1) their comprehension of the generated code; (2) the usage of generated details; (3) the changes and the rationale of changes to the generated code; and (4) the gain being established from per phase.

To analyse the data impacted by using ChatGPT, several methods were employed, including evaluation of data relevant to the assignment, such as time spent, course materials usage and assignment results; comparison of this data from both 2022 and 2023, when ChatGPT was not available and was available, respectively, to use in the same assignment briefs; and a survey of 2023 students to evaluate the use of ChatGPT from their perspectives. The survey contained four sections: (1) Demographic and General Information; (2) ChatGPT Usage; (3) Programming Competencies measured by Likert scale (Tan, 2009); and (4) Open Questions.

ANALYSIS AND FINDINGS

Survey Results

A survey invitation was sent via email to all 26 students enrolled in this course before the final exam, and 14 of them responded. In the survey, only three respondents claimed that they had used ChatGPT for their assignment. All these three respondents agreed that ChatGPT helped them learn new skills and improved their ability to solve problems and believed that ChatGPT would help them better understand programming code. However, one of them commented that it was not at all beneficial, while the other two believed that the generated coding and comments by ChatGPT were beneficial for their assignment.

Another thought that ChatGPT fulfilled a great deal of the requirements of his learning new skills, while others felt it did so to a moderate extent. Finally, they all had different opinions about the value of adopting ChatGPT to advance their professional competence, in terms of "Agree", "Somewhat", and "Neutral", respectively.

For the students who did not use ChatGPT in this assignment, some believed that ChatGPT could aid in their learning of programming by experimenting with new technology and receiving feedback. They also believed that it would inspire them to learn program coding. Others were not sure whether it could be used in learning programming or in completing this assignment. Instead, they preferred to rely on the course materials on Moodle.

Declaration Form

The Declaration Form for Ethically Using ChatGPT was collected from 21 students when they submitted their assignments. There were four students who admitted to using ChatGPT to assist them in this assignment (one of them missed out the survey). During the lab practice, the attendees were asked three times during the last five weeks whether anyone had attempted to use or had used ChatGPT in this assignment, but none of them indicated that they would be eager to. They justified their claim by saying that the course materials were sufficient for them to finish this project. They did not want to bother with ChatGPT because they had to spend time learning the newly generated coding.

Comparison of Students' Efforts and Outcomes

We collected data about students' efforts and outcomes in 2022 and 2023, including information about course attendance rate (%), Moodle site access (times), extension rate (%), and assignment results (marks %) (see Figure 1). The analysis results indicate no statistically significant difference in all the collected data between these two years. However, the mean values from all students in Figure 1 suggest that the lower average attendance rate in 2023 led to more online access times and greater extension rate than in 2022. There was a lower average performance rate in 2023 than in 2022.

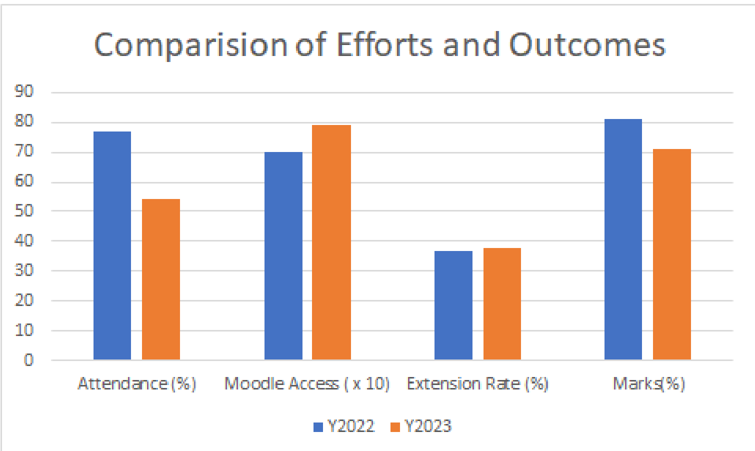


Figure 1. Comparison of students' efforts and learning outcomes in 2022 and 2023.

CONCLUSION AND RECOMMENDATIONS

We conclude that ChatGPT could positively support students in learning new features in programming and assist them in catching up on some missed course content. The divide-and-conquer strategy of using ChatGPT leads students to attempt an experience at a higher-order thinking-skill level in terms of analysis, synthesis and evaluation. However, the low attendance rate – which was due to various circumstances such as health issues, personal issues, family commitments, etc. – could be the main reason that some students used ChatGPT in the assignment. Additionally, due to ChatGPT's limitations (using the free GPT 3.5 rather than GPT 4) and the poor quality of input prompts provided by students, ChatGPT may produce unoptimised code or irrelevant solutions. Consequently, this might lead students to quit using ChatGPT and return to the conventional method while working at the lower-order thinking-skill level, such as comprehension and application. Therefore, ChatGPT is more appropriate for experienced students than for beginners.

Finally, because our students are always from varied academic backgrounds, we would continue to use both ChatGPT and standard e-learning resources to meet the requirements at all the thinking-skill levels. We would also encourage educators to employ various innovative methods for utilising ChatGPT as an engaging and effective educational tool.

REFERENCES

Diaz, M. (2023). How to use ChatGPT: What you need to know now. *ZDNET Tech Today*. https://www.zdnet.com/article/how-to-use-chatgpt

OpenAI. (2023). *Introducing ChatGPT*. https://openai.com/blog/chatgpt

Roos, J. (2023, February 1). ChatGPT: The next firestorm in education. *AACSB*. https://www.aacsb.edu/insights/articles/2023/02/chatgpt-the-next-firestorm-in-education

Sadik, A., Ceravola, A., Joublin, F., & Patra, J. (2023). Analysis of ChatGPT on source code. *arXiv*. https://arxiv.org/abs/2306.00597

Surameery, N. M. S., & Shakor, M. Y. (2023). Use ChatGPT to solve programming bugs. *International Journal of Information Technology and Computer Engineering*, *3*(1). https://doi.org/10.55529/ijitc.31.17.22

Tan, P-H., Ting, C-Y., & Ling, S-W. (2009). Learning difficulties in programming courses: Undergraduates' perspective and perception. *Proceedings of International Conference on Computer Technologies and Development*, *1*, 42–46. https://ieeexplore.ieee.org/document/5359977

Thompson, E., Luxton-Reilly, A., Whalley, J. L., Hu, M., & Robbins, P. (2008). Bloom's Taxonomy for CS assessment. In S. Hamilton & M. Hamilton (Eds.), *Proceedings of The Tenth Australasian Computing Education Conference* (pp.155–162). ACE. https://www.researchgate.net/publication/303809631_Bloom's_taxonomy_for_CS_assessment

Villasenor, J. (2023, February). How ChatGPT can improve education, not threaten it. *Scientific American*. https://www.scientificamerican.com/article/how-chatgpt-can-improve-education-not-threaten-it/

Wegner, P. (1990). Concepts and paradigms of object-oriented programming. *ACM SIGPLAN OOPS Messenger*, *1*(1), 7–87. https://doi.org/10.1145/382192.383004

AUTHORS

Dr Minjie (Min) Hu is a Senior Lecturer in IT at the School of Innovation, Design and Technology, Whitireia-WelTec. Having published extensively at domestic and international conferences in the last two decades, Min's research interests are teaching programming to novices, e-learning, data analytics in business and financial information, software testing and postgraduate education. Minjie.Hu@weltec.ac.nz

Tony Assadi is a Programme Manager at the School of Innovation, Design and Technology, Whitireia-WelTec. He is responsible for the undergraduate programmes at Whitireia and the postgraduate IT programmes at the Wellington campus. Tony's research interests are in work-integrated learning, computational thinking in tertiary education, and digital divide research. Tony.Assadi@wandw.ac.nz

Dr Hamid Mahroeian is a Senior Lecturer in IT at the School of Innovation, Design and Technology, Whitireia-WelTec. His research interests cover areas such as learning analytics, big data, and educational technology. Hamid.Mahroeian@whitireia.ac.nz

Assessment Validity in the Era of Generative AI Tools

Eltahir Kabbar and Bashar Barmada

https://doi.org/10.34074/proc.240105
Correspondence: Eltahir Kabbar, ekabbar@unitec.ac.nz

This publication may be cited as:

Kabbar, E., & Barmada, B. (2024). Assessment Validity in the Era of Generative AI Tools. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 26–33). ePress, Unitec. https://doi.org/10.34074/proc.240105

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga – New Zealand Institute of Skills and Technology

ABSTRACT

Generative AI tools, a recent disruptive educational technology, are expected to change how education is delivered and administered. This study proposes a risk identification framework to support educators in identifying assessment integrity risks caused by generative AI tools. The framework also suggests possible actions to mitigate these risks. The proposed framework uses four factors (Assessment Type, AI Knowledge, Course Level, and Bloom's Taxonomy Cognitive Domain Level) to identify the risks associated with an assessment resulting from the usage of generative AI tools. It is critical to have such a framework to ensure the integrity of assessments while the education industry adapts to the generative AI tools era.

KEYWORDS

Artificial intelligence, generative AI tools, assessment validity, assessment evaluation framework

INTRODUCTION

Over the past few decades, various information communication technologies (ICTs) have significantly impacted teaching and learning in education. In recent years, many educators and educational institutes have been interested in applying educational technologies as they improve the learner experience (Flavin, 2012). Our working definition of educational technologies is informed by Bajpai and Leedham (1970), who take a holistic view of applying different systems to support teaching and learning.

The recent adoption of several educational technologies by educational institutes has disrupted the conventional way of learning and forced changes to their practices (Cassidy, 2023). These technologies have profoundly impacted education, transforming how students learn and teachers teach. As technology evolves, education will likely continue to be disrupted and changed in new and exciting ways.

This paper proposes a framework to support educators in designing and developing valid assessments that meet the regulatory authorities' requirements while addressing the challenges posed by the most recent disruptive technology (generative AI tools) and its risk to the integrity of assessments. The proposed framework is discussed in line with the assessment types approved by the New Zealand Qualifications Authority (NZQA) while being generic enough to be used as a blueprint for other education systems. This framework aims to enable educators to produce valid assessments that meet the learning-outcome requirements while being resilient to the threat posed by misusing generative AI tools.

DISRUPTIVE TECHNOLOGIES IN EDUCATION

In the 1980s, personal computers (PCs) became increasingly affordable and accessible, leading to a widespread adoption of PCs by educational institutes. Learners had to acquire computer skills to access digital educational resources quickly. Simultaneously, educators had to change how they delivered course material and assessment practices to take advantage of PC technology. The emergence of the internet in the 1990s transformed education by providing access to vast amounts of information, enabling online communication and collaboration. The widespread use of the internet led to the development of online courses and distance learning programmes (Flavin, 2012; Gejendhiran et al., 2020).

In the early 2000s, learning management systems (LMSs) were introduced, providing a centralised platform for managing and delivering educational content and allowing for more personalised and interactive learning experiences (Buckley & Doyle, 2014; Oliveira et al., 2016). In addition, the use of LMSs enhanced course resource accessibility by students (Hew, 2015), enabling course instructors to customise students' learning experiences, moving away from the one-size-fits-all model (Berking & Gallagher, 2016). The introduction of LMSs also enabled educators to create streamlined assessment and feedback processes (Dahlstrom et al., 2014), and allowed students better visibility and control of their progress (Zabolotniaia et al., 2020), among other benefits.

Furthermore, the evolution of internet-enabled handheld devices, such as smartphones and tablets in the 2010s, enhanced learners' access to learning resources (Ait-Hroch & Ibrahimi, 2024). This disruptive technology introduced a mobile-learning (m-learning) aspect to digital learning (Sharples et al., 2010). In addition, it enabled learners to collaborate and communicate better with peers and educators (Hrastinski, 2008), and to have instant access to educational resources (Sheng et al., 2010), and better engagement with course content (Hamari et al., 2014; Salhab & Daher, 2023). An exciting pedagogical impact of this technology is the gamification of learning and the anticipated impact this may have on education (Flavin, 2012; Hamari et al., 2014).

In recent years, artificial intelligence and machine learning have developed adaptive learning platforms to personalise each student's learning experience (Chen et al., 2019), based on individual learning patterns and preferences (Kuppusamy, 2019). With the surge of devices enabled by the internet of things (IoT), more data can be collected and analysed to boost the learning capabilities of learners, based on a better understanding of their behaviours.

GENERATIVE AI TOOLS

Generative AI tools, also called AI chatbots, are the most recent types of disruptive technology in education. These tools are not standard search engines; they take the user prompt and use an intelligent algorithm based on natural language modelling to analyse the request, then they search through an extensive database of information, such as articles, books and websites, to formulate a response (Agar, 2023). Because they rely on AI, generative AI tools can learn over time and continuously improve their responses. Further, Abd-Elaal et al. (2019) assert that AI chatbots, such as grammar tools, use other assistive tools to enhance their decisions.

How Organisations React to Disruptive Technologies

Some academic institutes consider the use of AI without explicit permission to be a breach of assessment integrity. The response from such institutes is to ignore or fight AI usage. For example, an educational institute may develop tools that can detect AI-generated content, (Abd-Elaal et al., 2019); the University of Melbourne, the University of New South Wales and Waikato University have already adopted tools to detect the use of AI in student work (University of Waikato, n.d.).

However, others, such as the University of Sydney, who acknowledge generative AI as a tool to improve the student learning experience but are reluctant to adopt such tools before thorough testing and verification, have taken a different approach and adopted a wait-and-see strategy before committing either way.

A third group, including such institutes as the University of Wollongong and Massey University (Massey University, 2023), seem to embrace AI and consider its use in education to be beneficial to the learning process, emphasising that students need to learn how to use it in the right way (Saunders, 2023).

NZQA Assessment Types

The New Zealand Qualifications Authority (NZQA) is the official New Zealand body that awards secondary- and tertiary-level qualifications. They ensure that the education providers' programmes match corresponding qualification requirements and are recognised nationally and internationally. NZQA lists various assessment tools and approaches, classified into five main categories: Oral, Written, Practical, Verification and Other. Table 1 shows the NZQA categories and some of their assessments. For the complete list of NZQA assessments, the reader is advised to check (NZQA, n.d.). Depending on the qualification level, assessments are assigned to learners, and with the help of educational frameworks, such as Bloom's Taxonomy (Forehand, 2005), the tasks of assessments are defined.

Table 1. NZQA assessment categories and some of their assessment methods.

Written evidence	Practical evidence	Oral evidence	Verification evidence	Other evidence
- Reports - Portfolio - Tests - Assignments - Tables/Charts/Forms - Booklets - Worksheet	- Project - Simulation - Demonstration - Model - Posters	- Presentations - Questions/Answers - Interviews - Speech	- Feedback	- Prior Knowledge

THE PROPOSED APPROACH: A GENERIC FRAMEWORK

Generative AI tools are here to stay, disrupting the education process and current assessment methods. However, there is no unified approach to managing the impact of generative AI tools on the education process, particularly on assessment integrity. This paper proposes a framework to support educators in deciding the risk level AI poses for existing assessments. Furthermore, this paper will identify changes that can be used to improve assessment methods and practices, and increase resilience against the recent development of AI. The framework shown in Figure 1 encourages educators to consider four factors when assessing the level of risk posed by generative AI tools to the current assessment practices and take appropriate action to address the identified risk level. The four factors to consider are Assessment Type, AI Knowledge, Course Level and Bloom's Taxonomy Cognitive Domain Level.

Assessment Type

NZQA assessment types fall within five categories, as shown in Table 1. The risks associated with assessments vary depending on the assessment type. For example, the risks associated with the integrity of written reports as evidence of learning due to the illegitimate use of AI tools are potentially higher than in-class presentations or tests carried out in a controlled environment. Therefore, educators must consider the assessment type in their risk-assessment evaluation.

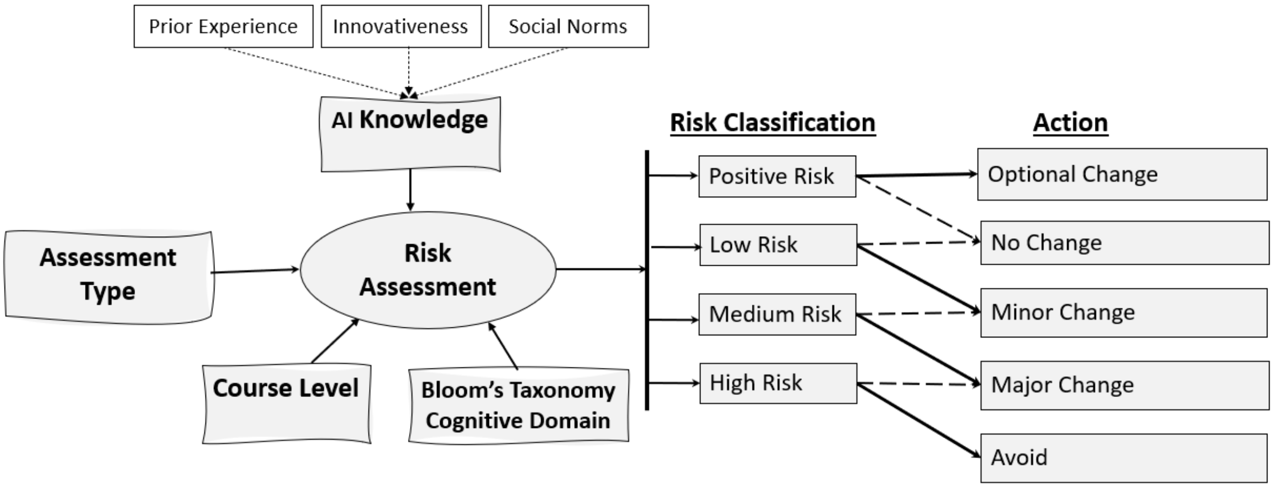


Figure 1. Proposed risk-evaluation framework to mitigate the impact of generative AI tools on assessment validity.

AI Knowledge

The second factor in the framework is AI Knowledge, which focuses on the educator's awareness of the different generative AI tools, their usage in an educational context, and their functionality. The educators' AI knowledge is also shaped by their prior experience, innovativeness, and the social norms and practices of their organisation. An adequate AI knowledge level is critical for educators in reaching accurate risk-assessment outcomes.

Course Level

The third factor in the framework is Course Level. Academic institutes offer different qualifications at different levels. The assessments associated with each course are usually designed to assess learning at the designated level using the appropriate assessment type. While some assessment types pose a significant risk at certain levels, the same assessment type could be less risky at a different level; therefore, the course level must be considered during the risk assessment. For example, a written report at Level 5, where students are expected to recall facts and basic concepts, poses a relatively lower risk compared to a written report at Level 7, where students are expected to draw connections among ideas in a literature review domain.

Bloom's Taxonomy Cognitive Domain Level

The fourth factor to be considered is the Bloom's Taxonomy Cognitive Domain Level embedded in the course learning outcome(s). Often, a course has different learning outcomes, and various assessment types are used to collect evidence of learning. Therefore, the course assessment's integrity risks can vary depending on the cognitive domain level associated with each assessed task. Generally, the lower the Bloom's Taxonomy domain level, the higher the risk, as various generative AI tools can be illegitimately used to produce work that would require students to work at the Remember/Understand levels compared to the Evaluate/Create levels. Therefore, educators must consider the Bloom's Taxonomy level during the risk assessment.

While considering that each of the four factors discussed provides a reasonable risk-level assessment, the accuracy of the risk-level assessment can only be achieved by considering all elements.

The outcome of the risk-assessment exercise leads to one of the four risk-classification statuses described in Figure 1. Assessments with a 'Positive Risk' status are valid; however, educators may opt to change the requirements to take advantage of the opportunities presented by generative AI tools. We recommend educators make changes to their assessments to improve the overall student learning experience and quality of the assessment with the existence of generative AI tools while maintaining the validity of the assessment.

In addition, assessments with a 'Low Risk' status are also valid; therefore, no changes are required because generative AI tools have little to no impact on the assessment validity. We recommend educators make minor changes to their assessments, leading to more resilient assessments to mitigate the risk of generative AI tools.

Assessments with a 'Medium Risk' status require greater changes, such as modifying/updating the assessment instrument from written reports to written assessments to be completed under a controlled environment, to ensure the assessment is resilient against generative AI tools (as shown in Figure 2). While the required changes may not affect the core of the assessment requirements, we believe these changes will significantly improve the assessment's resilience against generative AI tools, moving it to the Green Zone.

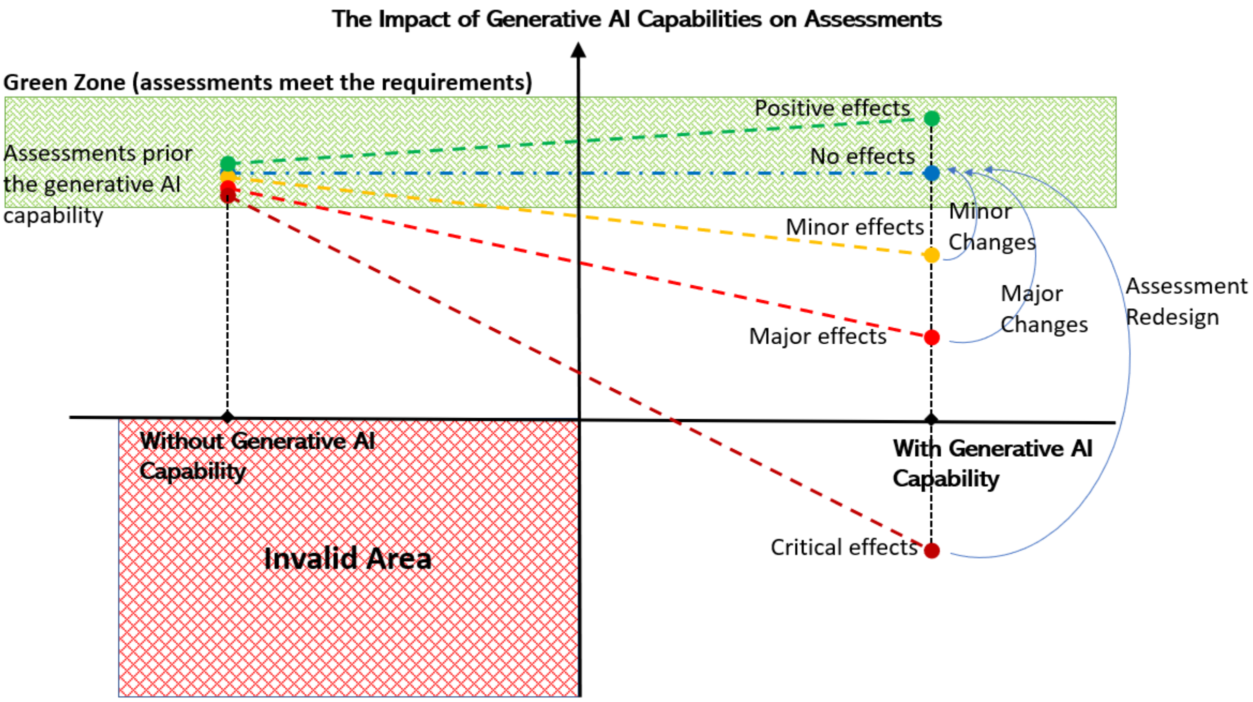


Figure 2. The impact of generative AI capabilities on assessments.

Finally, we recommend avoiding the assessments with 'High Risk' status, or at least undertaking significant changes to the assessment requirements and type, leading to a complete redesign. Educators are encouraged to change assessments that can be easily solved by generative AI tools, such as written research reports or even developing software, and include presentation/viva components in these assessments to allow assessors to judge better the authenticity of the work submitted. We acknowledge that the approval of changes that impact a programme can be complex, and the redesign takes time; hence, the option exists to make significant changes.

CONCLUSION

Disruptive technologies have constantly challenged the norms and brought exciting opportunities to many industries, including education. Generative AI tools are one of the most recent disruptive technologies in education. This technology is here to stay, and students around the globe are taking advantage of it to learn and complete the required assessments. Academic institutes react differently to the challenges posed by generative AI tools, particularly to assessment integrity. Some educational institutes have banned the use of these technologies in assessments, while others consider them to be valuable tools for learning and actively set policies to promote them.

Some academic institutes have yet to decide on a clear approach to the use of generative AI tools, leaving it to the academic staff to decide. Nevertheless, the impact of generative AI tools on assessments makes educators nervous, because of the potential implications on assessment integrity and ensuring that an assessment's learning outcomes are met without compromising its integrity.

The paper has introduced a framework to assist educators in navigating and managing the risks posed by generative AI tools to ensure assessment validity during our transition into the new education norm. We believe that generative AI tools are a game changer; the proposed model will help educators maintain assessment validity and integrity during the transition, but ultimately, a rethink of how we assess course learning outcomes is required.

REFERENCES

Abd-Elaal, E-S., Gamage, S. H. P. W., & Mills, J. E. (2019). Artificial intelligence is a tool for cheating academic integrity. *AAEE Conference Proceedings*. https://aaee.net.au/wp-content/uploads/2020/07/AAEE2019_Annual_Conference_paper_180.pdf

Agar, N. (2023). *The AIs are coming: Will ChatGPT create a future of bullsh*t (jobs)?* The University of Waikato. https://www.waikato.ac.nz/news-opinion/media/2023/the-ais-are-coming-will-chatgpt-create-a-future-of-bullsht-jobs

Ait-Hroch, A., & Ibrahimi, A. (2024). The place of e-learning in language learning: Its contribution to the enhancement of language proficiency. In M. Khaldi (Ed.), *Innovative instructional design methods and tools for improved teaching* (pp. 341–362). IGI Global. https://doi.org/10.4018/979-8-3693-3128-6.ch017

Bajpai, A. C., & Leedham, J. (1970). *Aspects of educational technology* (vol. 4). Association for Programmed Learning and Educational Technology.

Berking, P., & Gallagher, S. (2016). *Choosing a learning management system*. Advanced Distributed Learning. https://adlnet.gov/assets/uploads/ChoosingAnLMS.pdf

Buckley, P., & Doyle, E. (2014). Gamification and student motivation. *Interactive Learning Environments*, *24*(6), 1162–1175. https://doi.org/10.1080/10494820.2014.964263

Cassidy, C. (2023). Australian universities split on using new tool to detect AI plagiarism. *The Guardian*. https://www.theguardian.com/australia-news/2023/apr/16/australian-universities-split-on-using-new-tool-to-detect-ai-plagiarism

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, *8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Dahlstrom, E., Brooks, D. C., & Bichsel, J. (2014). The current ecosystem of learning management systems in higher education: Student, faculty, and IT perspectives. *Educause*. https://doi.org/10.13140/RG.2.1.3751.6005

Flavin, M. (2012). Disruptive technologies in higher education. *Research in Learning Technology, ALT-C 2012 Conference Proceedings*. https://doi.org/10.3402/rlt.v20i0.19184

Forehand, M. (2005). Bloom's Taxonomy. In M. Orey (Ed.), *Emerging perspectives on learning, teaching and technology* (pp. 41–47). Global Text. https://textbookequity.org/Textbooks/Orey_Emergin_Perspectives_Learning.pdf

Gejendhiran, S., Arokia Anicia, S., Selvam, V., & Kalaimani, M. (2020). Disruptive technologies – A promising key for sustainable future education. *9th World Engineering Education Forum, Elsevier, Procedia Computer Science*, *172*, 843–847. https://doi.org/10.1016/j.procs.2020.05.121

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. *IEEE*, 3025–3034. https://ieeexplore.ieee.org/document/6758978

Hew, K. F. (2015). Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCS. *British Journal of Educational Technology*, *46*(2), 288–301. https://doi.org/10.1111/bjet.12235

Hrastinski, S. (2008). What is online learner participation? A literature review. *Computers & Education*, *51*(4), 1755–1765. https://doi.org/10.1016/j.compedu.2008.05.005

Kuppusamy, P. (2019). Smart education using internet of things technology. In L. Y. Zongmin Ma (Ed.), *Emerging technologies and applications in data processing and management* (pp. 385–412). IGI Global. https://doi.org/10.4018/978-1-5225-8446-9.ch017

Massey University. (2023). *Use of artificial intelligence in assessment policy*. Retrieved April 3, 2024, from https://www.massey.ac.nz/documents/1445/Use_of_Artificial_Intelligence_in_Assessment_Policy.pdf

NZQA. (n.d.). *Assessment tools and approaches*. Retrieved July 24, 2023, from https://www.nzqa.govt.nz/providers-partners/assessment-and-moderation-of-standards/assessment-of-standards/generic-resources/assessment-tools-and-approaches/#:~:text=There%20are%20two%20different%20types,task%20assessment%20and%20evidence%20assessment

Oliveira, P. C. D., Cunha, C. J. C. D. A., & Nakayama, M. K. (2016). Learning management systems (LMS) and e-learning management: An integrative review and research agenda. *Journal of Information Systems and Technology Management*, *13*, 157–180. https://www.scielo.br/j/jistm/a/TT3Pk4mwkp5Cmmbf4NbfPvw/

Salhab, R., & Daher, W. (2023). The impact of mobile learning on students' attitudes towards learning in an educational technology course. *Multimodal Technologies and Interaction*, *7*(7), 74. https://doi.org/10.3390/mti7070074

Saunders, S. (2023, February). Rather than ban generative AI, universities must learn from the past. *University World News*. https://www.universityworldnews.com/post.php?story=20230221075136410

Sharples, M., Taylor, J., & Vavoula, G. (2010). A theory of learning for the mobile age. In B. Bachmair (Ed.), *Medienbildung in neuen Kulturräumen* (pp. 87–99). VS Verlag für Sozialwissenschaften. https://www.researchgate.net/publication/226417865_A_Theory_of_Learning_for_the_Mobile_Age

Sheng, H., Siau, K., & Nah, F. F. H. (2010). Understanding the values of mobile technology in education: A value-focused thinking approach. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, *41*(2), 25–44. https://doi.org/10.1145/1795377.1795380

University of Waikato. (n.d.). *Academic integrity – Student assessment guide*. Retrieved April 3, 2024, from https://www.waikato.ac.nz/students/student-assessment/assessment-at-university/academic-integrity

Zabolotniaia, M., Cheng, Z., Dorozhkin, E., & Lyzhin, A. (2020). Use of the LMS Moodle for an effective implementation of an innovative policy in higher educational institutions. *International Journal of Emerging Technologies in Learning (iJET)*, *15*(13), 172–189. https://www.learntechlib.org/p/217602/

AUTHORS

Dr Eltahir Kabbar is a Senior Lecturer at the School of Computing Electrical and Applied Technology at Unitec. His research interests include diffusion of innovation (DOI), technology adoption and e-government systems. He has several publications, including a book chapter, journal articles and international conferences. ekabbar@unitec.ac.nz

Dr Bashar Barmada is a Senior Lecturer and Academic Programme Manager in the School of Computing Electrical and Applied Technology at Unitec. His research areas are mainly in IoT applications and network security. bbarmada@unitec.ac.nz

A Chitchat on Using ChatGPT for Cheating

Sreenivas Sremath Tirumala and Chris Mayhew

https://doi.org/10.34074/proc.240106
Correspondence: Sreenivas Sremath Tirumala, Sreenivas.Tirumala@manukau.ac.nz

This publication may be cited as:

Tirumala, S. S., & Mayhew, C. (2024). A Chitchat on Using ChatGPT for Cheating. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023, Auckland, 27–29 September* (pp. 34–39). ePress, Unitec. https://doi.org/10.34074/proc.240106

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga –
New Zealand Institute of Skills and Technology

ABSTRACT

With the rapid increase in the usage of artificial intelligence, there has been considerable impact on how AI tools are utilised for academic misconduct. Catching this behaviour has been a big challenge, as learners have developed approaches to remove traces and footprints that indicate possible usage of AI tools. Tools such as Turnitin can check for originality or plagiarism, but provide very little assistance in differentiating AI- and non-AI-generated works. Generative tools such as ChatGPT, Bard and Copilot are posing challenges to academic staff in identifying misconduct. This research aims to emphasise the importance of investigating various approaches to tackle academic misconduct, particularly in computer science courses that involve a considerable amount of coding. This research looks at various resources that learners may use and go undetected, particularly in coding and presentations.

KEYWORDS

ChatGPT, academic misconduct, AI-based cheating

INTRODUCTION

Academic misconduct by students has been an ongoing issue for decades, if not centuries. Catching academic misconduct has been challenging, as learners have developed several ways to remove any footprints or indications in their work that might be suspicious. Tools such as Turnitin provide an evidence-based report to address this issue. However, Turnitin is limited to a set of rules that can be manipulated.

Recent advances in artificial intelligence (AI), particularly with generative AI tools such as ChatGPT, Bard, etc., mean staff are faced with the challenge of identifying AI-generated content within a learner's submission. The key issue is identifying criteria to differentiate AI-generated content from human-written content. Turnitin provides a new feature for AI detection. However, a disclaimer by Turnitin states "Caution: Percentage may not indicate academic misconduct. Review required."

Turnitin acknowledges the possibility of false positives and provides the following three-step approach to tackle that. All three tips involve human intervention and judgement, which might lead to a confrontational interaction that could damage the relationship with learners (point three) (Chechitelli, 2023).

Top three tips for addressing false positives:

- Know before you go – make sure you consider the possibility of a false positive upfront and have a plan for what your process and approach will be for determining the outcome. Even better, communicate that to students so that you have a shared set of expectations.
- Assume positive intent – in this space of so much that is new and unknown, give students the strong benefit of the doubt. If the evidence is unclear, assume students will act with integrity.
- Be open and honest – it is important to acknowledge that there may be false positives upfront, so both the instructor and the student should be prepared to have an open and honest dialogue. If you don't acknowledge that a false positive may occur, it will lead to a far more defensive and confrontational interaction that could ultimately damage relationships with students. (Chetichelli, 2023, para. 9)

Academic misconduct in content is just the tip of the iceberg, as there are several areas of computer science, particularly programming, where plagiarism detection is very difficult. Considering the number of online resources available (such as Stack Overflow), detecting a copied code from an online resource has been a challenge for programming lecturers.

AIMS AND RECENT RESEARCH

This research aims to emphasise the importance of investigating various approaches to tackle academic misconduct, particularly in computer science courses that involve a considerable amount of coding. This research looks at various resources that learners may use and go undetected, particularly in coding and presentations.

In this process, various resources including ChatGPT are used to seek possible ways to handle misconduct. There is no known technical solution at this stage. However, the research tries to look into available tools or the implementation of new regulations. The importance of handling learners with compassion when lecturers are unable to provide affirmative evidence of misconduct is also considered. A review of existing works provides an overview of how other researchers are looking into misconduct and will provide readers with the latest developments.

Generative AI applications such as ChatGPT have several advantages over search engines such as Google. They are useful as assistive AI that aids in learning easily and sometimes effectively. For instance, ChatGPT may be able to generate several answers or examples to solve a particular problem. ChatGPT can also be used to generate challenging questions or problems. This is one side of the coin; on the other side, researchers such as Cotton et al. (2023) emphasise its misuse and reiterate that there can be no automated process to detect cheating.

The outcomes of a comprehensive review presented by Lo (2023) provide insights into the good and the bad about generative AI tools: ChatGPT-like tools can be very useful, with a rating of 'outstanding' in the field of economics and 'unsatisfactory' in mathematics (Lo, 2023). For the programming domain, Lo's (2023) outcomes state 'satisfactory', considering the ability to generate multiple answers for the same question. From Lo's (2023) research, it is also noted that academic faculty are using ChatGPT extensively to generate course material and assessments.

A recent survey conducted by Cecilia Ka Yuk Chan (2023) of the University of Hong Kong reiterates the importance of comprehensive AI practices that need to be incorporated into university education policy. The recommendations provided by Chan are similar to others; i.e., providing necessary processes for human intervention and ethical practices.

OBSERVATIONS

The recent research discussed in the previous section converges on the following three domains:

1. **Lecturers/teachers:** Rewriting assessments based on assessable components that could determine the ability of the learners. This includes framing clear rules and regulations, including what is allowed and what is not.
2. **Ethical practices:** Educating learners about misconduct and its implications. Preparing a learning environment that discourages misconduct practices.
3. **Conversations with learners:** In case of possible AI detection, creating a healthy environment for discussion that will enable learners to present their views.

The practices mentioned above have important implications for both lecturers and learners.

Implications for Lecturers

1. Academic practices such as course delivery and assessment are governed by sets of rules in line with institutional policies. Therefore, any modifications require high-level approvals, moderation and reconciliation.
2. Lecturing by itself involves several responsibilities. Changes to assessments require a considerable amount of time, particularly when dealing with computer science programming assessments that involve coding.
3. As stated earlier, tools such as Turnitin may not be reliable, considering the difficulty in differentiating AI content from human-written content. When challenged by learners, lecturers might end up in a situation wherein they may not be able to provide sufficient evidence with confidence. This will end up creating friction between the learner and the lecturer.

Implications for Learners

1. The learner's confidence will be impacted by the implementing of more rules and regulations, which will create an environment of stress during tests and assessments.
2. Learners will be reluctant to seek help from good resources that, if chosen, could improve their learning ability.
3. Not every learner is confident in explaining their work, particularly when confronted on the basis of AI detection tools that cannot guarantee the results to be accurate. This will impact the learner psychologically, particularly learners from some cultural backgrounds.

DISCUSSION

The practices and observations presented in the previous section reiterate the fact that human intervention is required at each and every stage of identifying AI-based misconduct. However, there are several untouched elements that require attention before proposing any stringent policies or procedures to overcome this misconduct.

Considering the statements and disclaimers presented by reputed tools such as Turnitin, it is the right time to question the trustworthiness of these AI-based detection tools. Turnitin continues to emphasise the issue of false positives, which create more pressure on lecturers and learners, in articles concerning their tools. The following statements are quoted from the Turnitin blog post *How to prepare for and discuss the possibility of false positives* (West-Smith, 2023).

- "Remember that the indicator is not a replacement for educator judgment"
- "Make sure students know to cite any use of AI writing tools"
- "Suggest that students be prepared to share evidence"
- "Avoiding assumptions, being explicit, and creating trusting relationships are the only ways to get to a place of shared understanding"

All the articles presented in the literature, including documentation from Turnitin, talk about the use of generative AI tools for writing reports and content, ignoring the fact that these tools can be used to generate code that would enable learners to create software projects, web applications, slides and content for their presentations, as well as code correction and optimisation, among other tasks.

ChatGPT can be used to generate and debug programming code, implementing best practices and producing line-wise comments. For an assessment that involves coding, learners tend to use websites such as Stack Overflow. A common practice in this area is to copy the code and rename variables. Can this be considered as plagiarism? Technically, yes, but what are the criteria to prove it? Is it feasible to invite each and every learner to a face-to-face meeting to discuss their assignment?

Assessing the ability of a learner could be possible when a lecturer encounters a complex code that was never taught in class. However, there are some situations in which a learner who is already working and is looking for a qualification could come up with advanced coding practices. Can this be ignored?

It is interesting to see how generative AI tools such as ChatGPT and Google Bard react to the questions about preventing the misuse of generative AI tools. Some tips generated using ChatGPT and Google Bard are presented in Figure 1 and Figure 2 respectively.

Establish clear guidelines: Educators should establish clear guidelines for the use of ChatGPT in the classroom. This can include rules for appropriate language, topics of conversation, and expectations for behavior. By setting clear expectations and boundaries, educators can help prevent misuse of ChatGPT.

Monitor ChatGPT conversations: Educators should actively monitor ChatGPT conversations to ensure that students are using the platform appropriately. This can involve reviewing chat logs, participating in conversations, and intervening if necessary

Figure 1. Tips generated by ChatGPT.

Educate students about the risks of using ChatGPT for academic purposes: Students should be aware that using ChatGPT to generate essays, reports, or other assignments can lead to academic dishonesty charges. They should also be aware of the potential for ChatGPT to be used to spread rumors, make threats, or harass other students.

Monitor student use of ChatGPT: Educators can monitor student use of ChatGPT by using software that tracks student activity online. This software can be used to identify students who are using ChatGPT for academic purposes or who are using it to engage in inappropriate behavior.

Develop clear policies about the use of ChatGPT: Educators should develop clear policies about the use of ChatGPT in the classroom. These policies should state that ChatGPT cannot be used for academic purposes and that it cannot be used to engage in inappropriate behavior.

Figure 2. Tips generated using Google Bard.

CONCLUSION AND RECOMMENDATIONS

This research investigates the misuse of generative AI-based tools such as ChatGPT leading to misconduct. The practices that are stated above could create an environment of doubt, in which trust will diminish, particularly when the lecturer relies on tools for detecting AI content – especially Turnitin, which clearly mentions human intervention at each and every stage.

Creating awareness of the negative impacts of generative AI tools needs to be considered and an education framework is required to be built that would provide a clear direction for lecturers and learners. From here, the investigation will be moving towards a comprehensive survey in various institutes, to understand the practical implications and resilience in implementing a new framework around AI-based misconduct.

This is an ongoing research project and requires the immediate attention of education providers and government agencies.

ACKNOWLEDGEMENT

The authors would like to thank Geri Harris, Samah Ridha, Zarrar Javaid and Michael Andrews of Manukau Institute of Technology for the fruitful discussions on this topic.

REFERENCES

Chan, C. K. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 1–25. https://doi.org/10.1186/s41239-023-00408-3

Chechitelli, A. (2023, March 16). *Understanding false positives within our AI writing detection capabilities* [Blog post]. Turnitin. https://www.turnitin.com/blog/understanding-false-positives-within-our-ai-writing-detection-capabilities

Cotton, D., Cotton, P., & Shipway, J. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1–12. https://doi.org/10.1080/14703297.2023.2190148

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, *13*(4). https://doi.org/10.3390/educsci13040410

West-Smith, P. (2023, March 24). *How to prepare for and discuss the possibility of false positives* [Blog post]. Turnitin. https://www.turnitin.com/blog/how-to-prepare-for-and-discuss-the-possibility-of-false-positives

AUTHORS

Dr Sreenivas Tirumala is a Senior Lecturer in the School of Digital Technologies at Manukau Institute of Technology Te Whare Takiura o Manukau. His research areas centre on cybersecurity – privacy, security practices and implementing low-cost AI solutions for cybersecurity. Sreenivas.Tirumala@manukau.ac.nz

Dr Chris Mayhew MITP is a Senior Lecturer in the School of Digital Technologies at Manukau Institute of Technology Te Whare Takiura o Manukau. His interest areas are system implementation, software testing and, in his role of Academic Lead – Students, academic misconduct and its effect on ākonga. Chris.Mayhew@manukau.ac.nz

Charting the Course: A Framework for Integrating Industry 5.0 Technologies into Higher Education. A Case Study in a Cybersecurity Class

Amr Adel

https://doi.org/10.34074/proc.240107
Correspondence: amrvandenadel@gmail.com

This publication may be cited as:

Adel, A. (2024). Charting the Course: A Framework for Integrating Industry 5.0 Technologies into Higher Education. A Case Study in a Cybersecurity Class. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 40–51). ePress, Unitec. https://doi.org/10.34074/proc.240107

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga – New Zealand Institute of Skills and Technology

ABSTRACT

This paper explores the transformative potential of Industry 5.0 in higher education, presenting a conceptual framework for its effective integration. Emphasising immersive learning experiences (ILX), hyper automation and cybersecurity, the framework provides a roadmap for leveraging advanced technologies to enhance learning outcomes and student engagement. The framework was validated via an innovative, game-based learning experiment for cybersecurity students at Torrens University Australia, called Safe Passage. The game created an immersive, interactive environment in which students grappled with real-world cybersecurity challenges, thereby augmenting their theoretical understanding with practical application.

Results from the implementation demonstrated the framework's potential in fostering deeper comprehension, enhancing student engagement, and facilitating practical-skill acquisition. The success of this approach within cybersecurity education suggests its potential applicability across various disciplines, indicating a promising future for Industry 5.0 in higher education. This paper contributes to the growing discourse on technology's role in education, underlining the transformative potential of Industry 5.0 in creating more dynamic, engaging and relevant learning environments.

KEYWORDS

Industry 5.0, higher education, immersive learning experience, automation, cybersecurity, game-based learning

INTRODUCTION

In the contemporary world, the symbiotic relationship between technology and education continues to evolve, fuelling a paradigm shift in the way knowledge is imparted and received. Over the last two decades, the educational landscape has witnessed significant transformations with each industrial revolution. Industry 4.0, characterised by automation and data exchange in manufacturing technologies, led to the emergence of innovative digital learning platforms, collaborative online communities, and the global democratisation of education (Xu et al., 2021).

Entering the era of Industry 5.0, also known as the Fifth Industrial Revolution, a wave of unprecedented change stands on the horizon. Industry 5.0 transcends the previous phase's boundaries, intertwining human intelligence with machine learning more intricately to unlock unique, personalised and efficient ways to facilitate learning (Al-Emran & Al-Sharafi, 2022). The increased collaboration between humans and smart systems in Industry 5.0 is set to revolutionise the education sector, particularly higher education.

Higher education forms the crucible in which young adults are equipped with knowledge and skills to participate in and contribute to society. It serves as a bridge to the future, preparing students for careers that are still emerging. Therefore, its evolution is integral to societal advancement. As the integration of Industry 5.0 into higher education is explored, a perspective is gained on how this fusion will enhance the effectiveness and accessibility of educational systems and shape the higher-learning landscape of the future (Adel, 2022).

The integration of Industry 5.0 in higher education offers a broad spectrum of potential advancements. From adaptive and personalised learning experiences to a more in-depth, data-driven understanding of individual student behaviour and needs, the possibilities for innovation are endless (Leng et al., 2022). Furthermore, it provides a platform for greater access to higher education across geographical and socioeconomic boundaries, thereby promoting the democratising effect of education.

However, the integration of Industry 5.0 in higher education also presents a significant set of challenges. Institutions must adapt their curricula and infrastructure to effectively incorporate the new technologies (Huang et al., 2022). Simultaneously, educators must be trained to utilise these tools while preserving the critical human element in teaching. Moreover, ethical and privacy considerations need to be addressed as more data is collected and machine learning algorithms become increasingly integral to the educational process (Özdemir & Hekim, 2018).

Industry 5.0 signifies the next phase of the industrial revolution, where the human touch is reintroduced and integrated with technological advancements. The personalised and collaborative nature of Industry 5.0 makes it especially applicable in the context of higher education. Here are seven typical applications:

1. **Personalised learning:** Industry 5.0 technologies allow for the creation of adaptive learning environments that tailor educational content and delivery based on individual student needs and capabilities, leading to a more effective and personalised learning experience (Rachmawati et al., 2021).
2. **Enhanced collaboration:** With Industry 5.0, higher education can leverage technology to promote international and cross-disciplinary collaborations. For instance, using virtual reality and augmented reality, students from different geographical locations can participate in shared learning experiences and projects (Rachmawati et al., 2021).
3. **Advanced research capabilities:** Industry 5.0 tools can help automate and expedite research processes, giving higher-education institutions the means to drive cutting-edge research and innovation (Akundi et al., 2022). Machine learning algorithms, for instance, can process vast amounts of data more efficiently, helping researchers identify patterns and trends more quickly.
4. **Immersive learning experiences:** Technologies such as virtual reality (VR) and augmented reality (AR), central to Industry 5.0, can provide immersive learning experiences, helping students gain a deeper understanding of complex concepts and procedures, especially in fields such as medicine, engineering and environmental sciences (Alexa et al., 2022).
5. **Real-time assessment and feedback:** Advanced analytics tools can provide real-time feedback on students' progress and understanding, allowing for immediate adjustment and enhancement of teaching methods and content delivery (Bakir & Dahlan, 2022).
6. **Enhanced accessibility:** Industry 5.0 has the potential to make higher education more accessible, regardless of physical location or personal circumstances. Technologies such as AI-powered chatbots can provide immediate responses to student queries, while digital learning platforms can bring the classroom to the student (Gürdür Broo et al., 2022).
7. **Career preparation:** Industry 5.0 in higher education can help students better prepare for their future careers. By integrating Industry 5.0 technologies into the curriculum, students can develop hands-on experience and skills that are crucial in the future job market (Bouezzeddine, 2022).

The primary aim of this paper is to offer a roadmap for the integration of Industry 5.0 in higher education. By delving into the intersection of Industry 5.0 and higher education, the opportunities and challenges it presents are explored. This roadmap outlines the steps and procedures for institutions to adopt Industry 5.0 technologies, helping strategically navigate this technological shift.

THE IMPACT OF INDUSTRY 5.0 TECHNOLOGIES IN HIGHER EDUCATION

The implementation of Industry 5.0 in higher education stands to fundamentally revolutionise the traditional educational model, creating a more personalised, dynamic and accessible learning environment (Bouezzeddine, 2022). Through personalisation and human–machine collaboration, educational experiences can be tailored to suit individual learning styles, speeds and preferences. Such customisation is not only anticipated to increase student engagement and satisfaction, but also to improve overall academic outcomes.

Moreover, with technologies such as artificial intelligence, machine learning, augmented reality and virtual reality becoming integral to the teaching and learning process, complex concepts can be made more tangible and engaging (Zahabi & Abdul Razak, 2020). These tools offer immersive, interactive experiences that significantly enhance students' understanding, leading to deeper knowledge retention.

By enabling real-time feedback and assessment, these technologies can also help educators swiftly identify students' strengths and weaknesses, allowing for timely interventions and support. This fosters an environment conducive to continuous improvement and growth (Farrokhnia et al., 2023).

Industry 5.0 can also democratise higher education by removing geographical and socioeconomic barriers. Digital learning platforms and AI-powered tools can deliver high-quality education resources to students irrespective of their location or economic background, thus making education more equitable and inclusive (Xing, 2023).

Furthermore, the integration of Industry 5.0 into higher education equips students with essential skills for the future job market (Sikora et al., 2023). As Industry 5.0 technologies permeate various sectors, hands-on experience with these tools during education will prepare students for the workforce of tomorrow. Hence, the impact of Industry 5.0 on higher education extends beyond immediate educational outcomes, paving the way for lifelong learning and career readiness.

In Figure 1, the proposed conceptual framework for the integration of Industry 5.0 in higher education is displayed. This framework is divided into four key stages: Input, Process, Output and Outcome.

Input: The first stage identifies the necessary components required for implementing Industry 5.0 technologies in higher education. These include the actual Industry 5.0 technologies, such as artificial intelligence (AI), machine learning, augmented reality (AR) and virtual reality (VR). Additionally, human capital is critical, which includes educators who are trained in using these technologies, IT support staff to maintain the technical infrastructure, etc. Infrastructural requirements encompass all the physical and digital necessities, including hardware, software and high-speed internet connectivity.

Process: This stage reflects how the inputs are transformed within the educational setting. It includes development and implementation of AI-based personalised curricula, establishment of AR/VR-enabled virtual classrooms, and the application of machine learning algorithms for assessing students' performance and progress.

Output: This is the immediate result of the implementation processes. The outputs include adaptive learning experiences customised to individual students' needs, enhanced levels of student engagement and interaction, improved and timely assessment and feedback mechanisms, and creation of a more collaborative and inclusive learning environment.

Outcome: This final stage represents the long-term impacts and benefits resulting from the integration of Industry 5.0 technologies. These include improved academic performance due to personalised learning, increased accessibility of higher-education resources for students regardless of their geographical location or socioeconomic status, and preparing students for the future job market by equipping them with the skills and experience to work with advanced technologies.

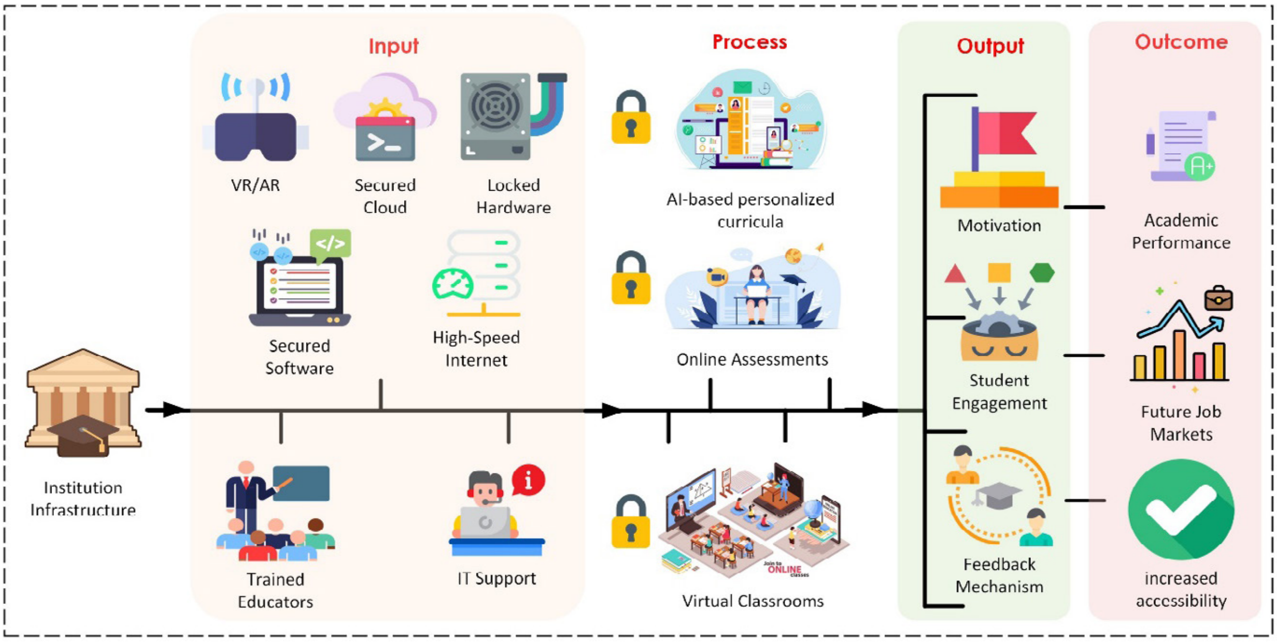


Figure 1. Conceptual framework of Industry 5.0 integration in education.

Arrows are used to demonstrate the flow from one stage to the next, indicating that each stage builds upon the previous. This progression illustrates how the strategic integration of resources (Input), coupled with effective execution (Process), leads to immediate benefits (Output) and, over the long term, results in significant positive impacts (Outcome) on students and the overall higher education landscape.

Table 1 shows the impact of Industry 5.0 on education, and how it can contribute efficiently to making creative environments that encourage both students and teachers to create knowledge.

Table 1. Impact of Industry 5.0 on education.

Stakeholder	Impact of Industry 5.0
Researchers	Enhanced data-collection and analysis capabilities due to AI and machine learning, leading to more accurate and faster research findings.
	Automation of time-consuming tasks allows more focus on high-value aspects of research.
	Advanced collaboration tools foster global, multidisciplinary research partnerships.
	Improved access to diverse data-sets, enabling more comprehensive and nuanced research.
Teachers	Personalised learning tools help in customising and adapting curricula to individual student needs.
	Real-time assessment tools provide immediate feedback on student understanding, aiding in timely interventions.
	AR/VR technologies offer immersive teaching tools, enhancing student engagement.
	Digital platforms facilitate the sharing of resources, ideas and best practices among educators.
Students	Personalised learning experiences cater to individual learning styles and speeds, enhancing understanding and retention.
	AR/VR and other immersive technologies offer hands-on, interactive learning experiences.
	Real-time feedback allows for immediate identification of learning gaps and areas for improvement.
	Increased access to higher-education resources, regardless of geographical or socioeconomic constraints.

ADOPTION STRATEGY OF INDUSTRY 5.0 IN HIGHER EDUCATION: A ROADMAP

The successful integration of Industry 5.0 in higher education requires a well-defined strategy and roadmap. Here is a step-by-step guideline for institutions looking to embark on this journey:

Step 1: Strategic Planning. Begin with the end in mind. Define clear objectives for integrating Industry 5.0 technologies. What outcomes are expected? This could range from improved student-learning outcomes, increased accessibility, better preparation for the job market, or enhanced research capabilities. This step should also involve a thorough analysis of the current state of technological integration in the institution to identify areas of improvement.

Step 2: Resource Allocation. Once clear objectives are set, allocate the necessary resources for the integration. This includes financial investments for procuring technology and updating infrastructure, as well as time and personnel for training and implementation.

Step 3: Staff Training. Human capital is key in the successful implementation of Industry 5.0. Invest in comprehensive training programmes for educators and IT support staff. This would equip them with the necessary skills to effectively utilise and manage Industry 5.0 technologies.

Step 4: Technology Implementation. Start implementing the chosen Industry 5.0 technologies in a phased manner. Beginning with pilot programmes or specific courses can help in understanding potential challenges and refining the approach before a full-scale rollout.

Step 5: Continuous Assessment and Feedback. Once implemented, continuously assess the effectiveness of the integration. Use data-driven insights to understand how well the technologies are meeting the defined objectives. Gather feedback from students and staff to identify areas of improvement.

Step 6: Iteration and Improvement. Based on the assessment and feedback, make necessary adjustments and improvements. The implementation of Industry 5.0 should be seen as an iterative process that evolves with time, technological advancements and changing needs.

Step 7: Scaling and Evolution. Once the implementation has proven successful, consider scaling it to other areas of the institution. Also, stay abreast of emerging technologies and trends in Industry 5.0 to continually evolve and enhance the integration.

ESSENTIAL CRITERIA OF INDUSTRY 5.0 SOLUTIONS

Industry 5.0 solutions are multifaceted and complex, requiring thoughtful consideration across various factors. Figure 2 shows the essential criteria that should be taken into account while selecting and implementing Industry 5.0 solutions:

Figure 2. Framework of Essential Criteria for Industry 5.0 Solutions.

1. Adaptability: The solution should be adaptable to the specific needs and context of the institution. It should provide flexibility to tailor applications based on individual course requirements, teaching methodologies and student needs.

2. Scalability: The chosen solutions should be scalable, i.e., they should be capable of expanding in accordance with the growth of the institution or the student body. This ensures that the technology can support increasing user loads and expanded functionalities over time.

3. Integration: The solution should be able to integrate seamlessly with existing systems and processes. This minimises disruptions, reduces complexities in implementation, and enhances overall user experience.

4. Usability: The user interface and experience should be intuitive and easy to navigate for all users, including students, faculty and administrators. This promotes higher adoption rates and ensures that the technology serves as an enabler rather than a hindrance.

5. Security: Given the digital nature of Industry 5.0 solutions, they should have robust cybersecurity measures in place to protect sensitive data. Additionally, they should be compliant with all relevant data-protection and privacy regulations.

6. Cost-effectiveness: While quality and functionality are paramount, the cost-effectiveness of the solution should also be considered. This involves looking at both upfront and ongoing costs, and weighing them against the value and benefits the solution provides.

7. Support and training: The solution provider should offer comprehensive support and training services. This ensures smooth implementation and assists users in maximising the value of the solution.

8. Future-proof: Given the rapid pace of technological advancements, the chosen solution should be future-proof, i.e., it should be able to accommodate or adapt to technological advancements or trends.

IMPLEMENTATION AND VALIDATION: IMMERSIVE LEARNING EXPERIENCE (ILX)

Immersive learning experience (ILX) represents a groundbreaking shift in education, utilising advanced technologies such as virtual reality (VR), augmented reality (AR) and mixed reality (MR) to create highly interactive and engaging learning environments. ILX goes beyond traditional learning methods by enabling students to explore, interact with, and become part of the subject matter. This innovative approach supports experiential learning, where students gain knowledge and skills through firsthand experiences within an artificial, yet highly realistic, environment. Whether it's a student virtually exploring ancient ruins for a history class, or a medical student performing a complex surgical procedure in a safe, simulated setting, ILX enhances understanding, retention and application of knowledge. By creating a sense of presence and involvement, ILX not only enriches the learning process but also makes it more enjoyable and memorable for students.

The proposed conceptual framework, designed to integrate Industry 5.0 into higher education, has undergone a practical test of its efficiency and effectiveness. As part of this validation process, a specialised game was developed and launched for undergraduate students in the cybersecurity class. The game, informed by the principles of immersive learning experience (ILX), was strategically designed to support and enhance the students' learning journey. It leverages advanced technologies inherent in Industry 5.0, creating an interactive, engaging and realistic learning environment where theoretical concepts are transformed into practical, hands-on experiences. The following section, accompanied by Figure 3, presents an overview of this novel experiment and its role in validating the proposed framework's applicability and benefits.

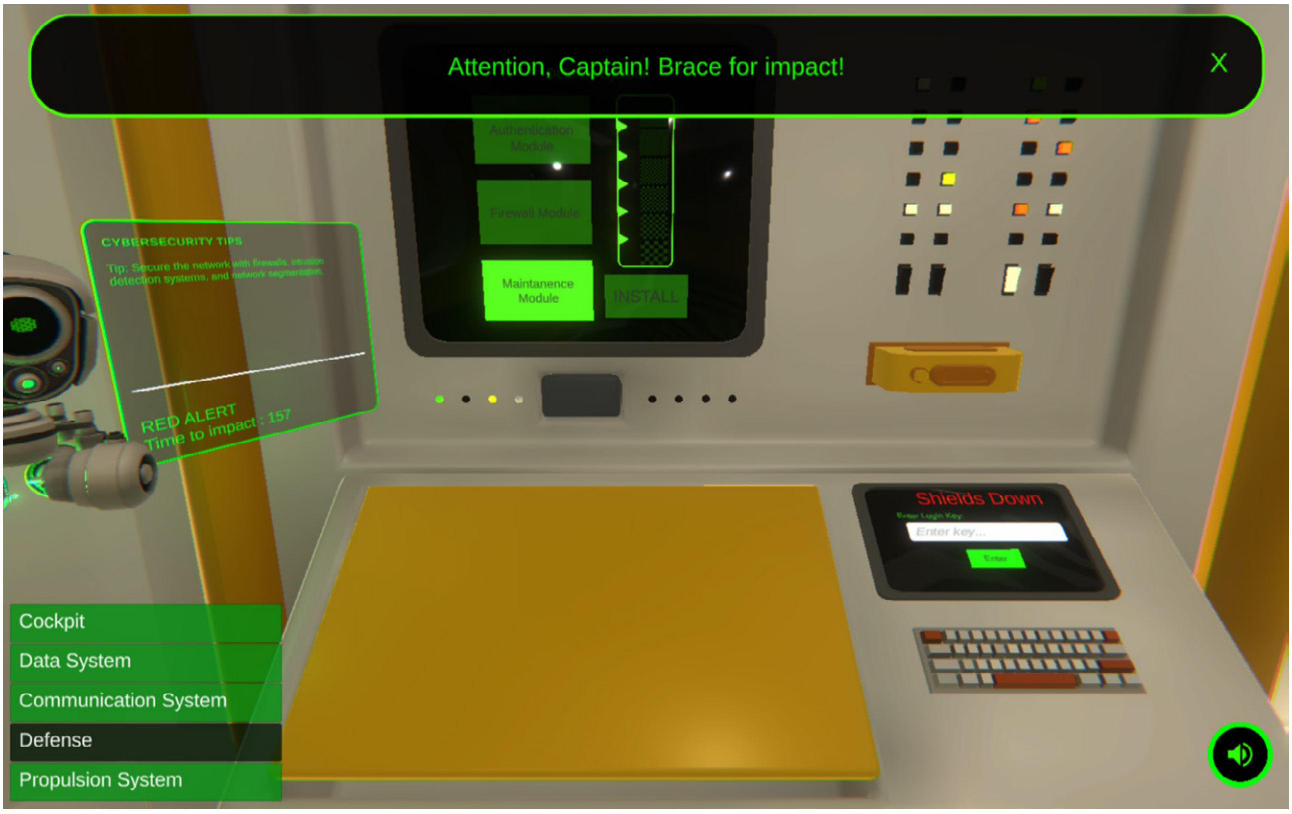


Figure 3. Safe Passage game. Image: Torrens University, reproduced with permission.

The novel game for undergraduate students in cybersecurity, Safe Passage, marks a unique departure from conventional teaching methods, offering a fully immersive and interactive learning environment. This cosmic journey propels students from the role of passive learners into the active position of a Chief Security Officer. Their mission? Safeguarding digital data from a diverse array of cyber threats, primarily the menacing Mal.

Safe Passage aligns seamlessly with the course curriculum, weaving thrilling scenarios into the fabric of each module. As students venture into this digital universe, they encounter real-world cybersecurity challenges that range from investigating system vulnerabilities and implementing sophisticated user-access controls to combating malware infections and decrypting complex codes. Moreover, students ensure the secure exchange of data between third parties, fostering a deep understanding of communication protocols in a digital landscape.

Operated through a web browser, Safe Passage provides learners with an adventure that extends far beyond the physical boundaries of a classroom. With each new mission, students delve deeper into the pulsating heart of cybersecurity. Instead of merely reading about it, they live it, breathe it and master it. This innovative learning approach transforms cybersecurity education into an exciting journey of challenges, victories and valuable insights. With Safe Passage, students not only navigate through their coursework but also skilfully steer their way through the vast cosmos of cybersecurity, emerging as proficient, battle-ready digital defenders.

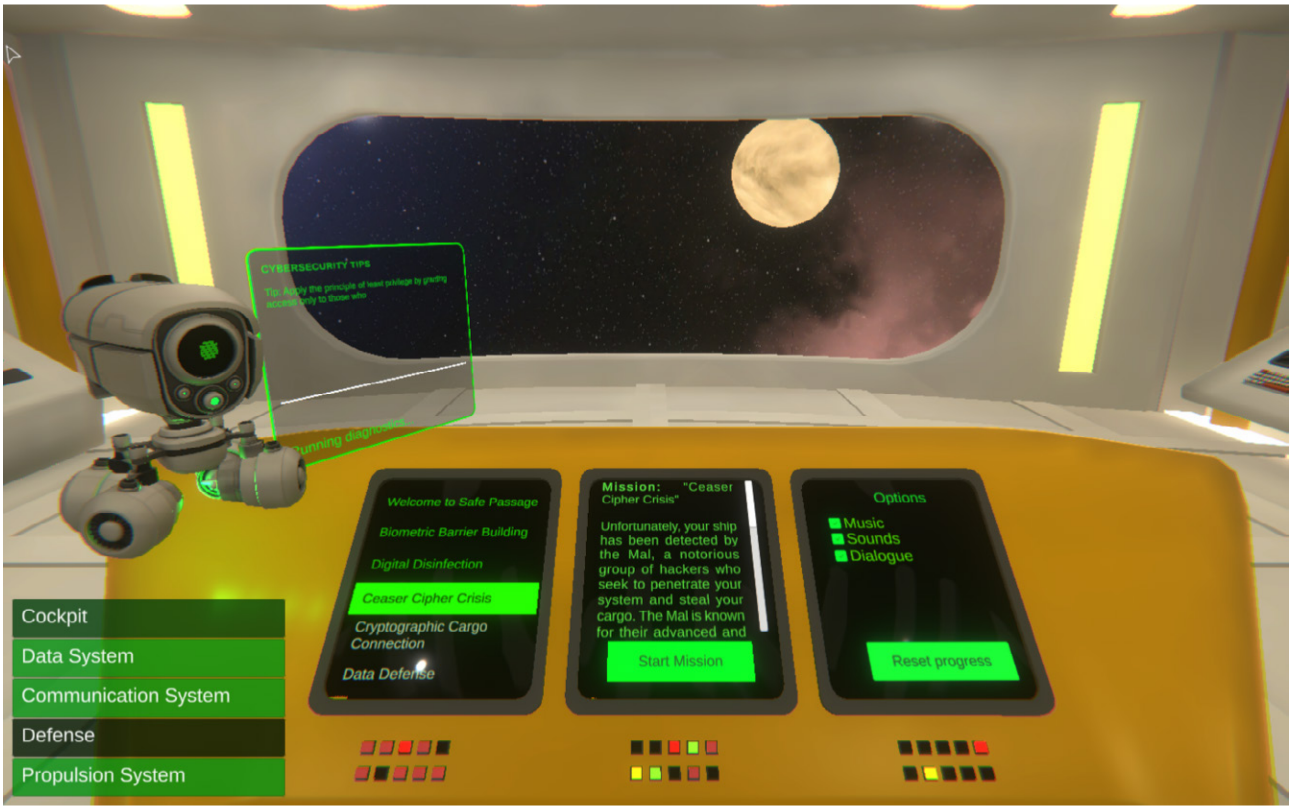


Figure 4. Safe Passage missions. Image: Torrens University, reproduced with permission.

In Figure 4, Safe Passage casts the players in the pivotal role of Chief Security Officer on a spaceship, entrusting them with the responsibility of delivering the ship's cargo securely despite the looming threats of system vulnerabilities and space pirates, notably the malicious Mal. The game comprises six sequential missions, each designed to challenge and educate the players in various aspects of cybersecurity.

Mission 1: Welcome to Safe Passage. In the inaugural mission, players are acquainted with the ship's primary systems by their android assistant, Bingo. This provides a foundation for understanding the environment they need to protect.

Mission 2: Biometric Barrier Building. This mission focuses on access control and authentication, challenging players to secure their system and data against potential breaches, setting the initial line of defence against Mal.

Mission 3: Digital Disinfection. Players face a robust cybersecurity challenge involving malware threats. They must respond swiftly and comprehensively, demonstrating their ability to maintain the security integrity of their ship.

Mission 4: Caesar Cipher Crisis. The fourth mission plunges players into the world of cybersecurity and cryptography. They must decrypt a crucial code to defend their ship, fostering an understanding of the interplay between cryptography and cybersecurity.

Mission 5: Cryptographic Cargo Connection. This mission revolves around the concept of secure data transfer. Players are required to utilise symmetric encryption to protect the ship's vital data, emphasising the importance of secure communication in the digital realm.

Mission 6: Data Defence. The final mission is all about data backup and threat anticipation. Players are tasked with mastering the art of secure data backups, thereby ensuring the safety of their digital assets from future threats.

Each mission offers a practical and immersive learning experience, helping players to develop a well-rounded understanding of the key aspects of cybersecurity.

DISCUSSION

The conceptual framework's efficacy, designed for the integration of Industry 5.0 into higher education, has been put to the test through the development and implementation of the Safe Passage game. This pedagogical experiment has demonstrated the successful amalgamation of advanced technology with immersive learning experiences, resulting in a dynamic and engaging education platform for cybersecurity students.

Notably, the game embodies key elements of the conceptual framework, such as the inclusion of immersive learning experience (ILX) and hyper automation, and their effective utilisation in a real-world setting. For instance, the game creates an immersive, interactive environment where learners encounter realistic cybersecurity scenarios, emphasising the practical application of theoretical knowledge. This aligns with the ILX component of the framework, highlighting its effectiveness in transforming the traditional learning experience.

In terms of hyper automation, the game introduces learners to the importance of automating various cybersecurity processes, including threat detection and response, data backup and secure data transfer. This real-world application of hyper automation within an educational setting further validates its essential role as outlined in the conceptual framework.

Additionally, the iterative nature of the missions in the game reinforces the concept of continuous learning and skill development, a core principle of Industry 5.0. Each mission builds on the preceding one, mirroring the rapid pace of technological advancements and the corresponding need for ongoing learning and adaptation.

The overwhelmingly positive response from the students further attests to the framework's effectiveness. The students reported not only an enhanced understanding of complex cybersecurity concepts but also a significant increase in engagement and interest in the subject matter. This positive feedback underscores the potential of the proposed framework in creating a more stimulating, effective and relevant learning environment in the era of Industry 5.0.

In conclusion, the Safe Passage game experiment has served as a successful practical validation of the proposed conceptual framework, demonstrating the immense potential of Industry 5.0 technologies in enhancing higher education. While this framework has proven effective in the cybersecurity course, it holds the promise to be scaled and adapted across various disciplines, further revolutionising the learning experience in higher education.

CONCLUSION

The advent of Industry 5.0 presents a transformative opportunity for higher education, offering new avenues for interactive, immersive and technologically advanced learning experiences. The proposed conceptual framework in this paper provides a strategic blueprint for integrating Industry 5.0 technologies into educational settings, with an emphasis on immersive learning experiences, hyper automation and cybersecurity.

The implementation of this framework via the Safe Passage game for cybersecurity students showcased the immense potential of this approach. The game's success highlights the powerful role of Industry 5.0 technologies in enhancing student engagement, fostering deeper understanding, and facilitating the practical application of theoretical concepts.

While the game served as an effective proof-of-concept within the realm of cybersecurity education, the principles and strategies outlined in the framework can be applied across a range of disciplines. This versatility underscores the universality of Industry 5.0's potential impact on higher education.

In conclusion, Industry 5.0 represents a major advancement in higher education's ongoing journey of evolution and innovation. By embracing this shift, higher-education institutions can better prepare their students for a future characterised by digital transformation and technological proficiency. As educators and administrators, the mission is clear: harness the power of Industry 5.0 to create a vibrant, engaging, and effective learning environment for the digital age.

ACKNOWLEDGEMENT

Torrens University Australia has granted approval for my request to publish and disseminate the findings of this experiment.

REFERENCES

Adel, A. (2022). Future of Industry 5.0 in society: Human-centric solutions, challenges and prospective research areas. *Journal of Cloud Computing*, *11*(1), 40. https://doi.org/10.1186/s13677-022-00314-5

Akundi, A., Euresti, D., Luna, S., Ankobiah, W., Lopes, A., & Edinbarough, I. (2022). State of Industry 5.0: Analysis and identification of current research trends. *Applied System Innovation*, *5*(1), Article 1. https://doi.org/10.3390/asi5010027

Al-Emran, M., & Al-Sharafi, M. A. (2022). Revolutionizing education with Industry 5.0: Challenges and future research agendas. *International Journal of Information Technology and Language Studies*, *6*(3), 1–5. https://www.researchgate.net/publication/367077900_Revolutionizing_Education_with_Industry_50_Challenges_and_Future_Research_Agendas

Alexa, L., Pîslaru, M., & Avasilcăi, S. (2022). From Industry 4.0 to Industry 5.0: An overview of European Union enterprises. In A. Draghici & L. Ivascu (Eds.), *Sustainability and innovation in manufacturing enterprises: Indicators, models and assessment for Industry 5.0* (pp. 221–231). Springer. https://doi.org/10.1007/978-981-16-7365-8_8

Bakir, A., & Dahlan, M. (2022). Higher education leadership and curricular design in Industry 5.0 environment: A cursory glance. *Development and Learning in Organizations: An International Journal*, *37*(3), 15–17. https://doi.org/10.1108/DLO-08-2022-0166

Bouezzeddine, M. (2022). *Insights into global engineering education after the birth of Industry 5.0*. Books on Demand.

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, *0*(0), 1–15. https://doi.org/10.1080/14703297.2023.2195846

Gürdür Broo, D., Kaynak, O., & Sait, S. M. (2022). Rethinking engineering education at the age of Industry 5.0. *Journal of Industrial Information Integration*, *25*, 100311. https://doi.org/10.1016/j.jii.2021.100311

Huang, S., Wang, B., Li, X., Zheng, P., Mourtzis, D., & Wang, L. (2022). Industry 5.0 and Society 5.0: Comparison, complementation and co-evolution. *Journal of Manufacturing Systems*, *64*, 424–428. https://doi.org/10.1016/j.jmsy.2022.07.010

Leng, J., Sha, W., Wang, B., Zheng, P., Zhuang, C., Liu, Q., Wuest, T., Mourtzis, D., & Wang, L. (2022). Industry 5.0: Prospect and retrospect. *Journal of Manufacturing Systems*, *65*, 279–295. https://doi.org/10.1016/j.jmsy.2022.09.017

Özdemir, V., & Hekim, N. (2018). Birth of Industry 5.0: Making sense of big data with artificial intelligence, "The Internet of Things" and next-generation technology policy. *OMICS: A Journal of Integrative Biology*, *22*(1), 65–76. https://doi.org/10.1089/omi.2017.0194

Rachmawati, I., Multisari, W., Triyono, T., Simon, I. M., & da Costa, A. (2021). Prevalence of academic resilience of social science students in facing the Industry 5.0 era. *International Journal of Evaluation and Research in Education*, *10*(2), 676–683. https://doi.org/10.11591/ijere.v10i2.21175

Sikora, A., Titova-Kosturkova, T., Janevska, G., Kostov, M., Halunga, S., Suciu, G., & Georgiev, G. (2023). EREMI: An innovative interdisciplinary approach for higher education in resource efficient manufacturing environments. *Sustainability, 15*(13), Article 13. https://doi.org/10.3390/su151310244

Xing, H. (2023). Application study of AI-powered non-translatables function in computer-aided translation. In *Proceedings of the 2022 5th International Conference on Education Technology Management* (pp. 104–109). Association for Computing Machinery. https://doi.org/10.1145/3582580.3582597

Xu, X., Lu, Y., Vogel-Heuser, B., & Wang, L. (2021). Industry 4.0 and Industry 5.0: Inception, conception and perception. *Journal of Manufacturing Systems*, *61*, 530–535. https://doi.org/10.1016/j.jmsy.2021.10.006

Zahabi, M., & Abdul Razak, A. M. (2020). Adaptive virtual reality-based training: A systematic literature review and framework. *Virtual Reality*, *24*(4), 725–752. https://doi.org/10.1007/s10055-020-00434-w

AUTHOR

Dr Amr van den Adel is a Senior Lecturer in Computing at Eastern Institute of Technology Te Aho a Māui. His research interests are smart education, cybersecurity and Industry 5.0.
amrvandenadel@gmail.com

Recommendation of Topics and Practical Labs to Teach the Semantic Web in Current Bachelor of Computing Systems

Daniel Dang

https://doi.org/10.34074/proc.240108
Correspondence: ddang@eit.ac.nz

This publication may be cited as:

Dang, D. (2024). Recommendation of Topics and Practical Labs to Teach the Semantic Web in Current Bachelor of Computing Systems. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 52–61). ePress, Unitec. https://doi.org/10.34074/proc.240108

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga – New Zealand Institute of Skills and Technology

ABSTRACT

The role of the Semantic Web in Web 3.0, the next evolution of the internet, is significant. Web 3.0 is often referred to as the 'intelligent web' or the 'Semantic Web'. The Semantic Web is an essential topic in Web 3.0 because it enables data integration, enhances data search on the World Wide Web, and promotes the development of knowledge graphs that allow computers to derive new insights and generate knowledge. Teaching the Semantic Web to undergraduate students offers numerous benefits. Students gain a deep understanding of cutting-edge web technologies, technical skills in data integration (the ability to develop systems with diverse datasets and linked data), and capabilities in information retrieval and knowledge management (data modelling and linked data querying). Moreover, they are able to design and develop different types of web-based applications such as intelligent data-driven decision-making apps, AI apps, or personalised-recommendation apps. In this article, we present ways to effectively integrate the Semantic Web as a subject in the current Computing Systems curriculum offering at Eastern Institute of Technology (EIT). We also propose a recommended list of theory topics, practical labs and practical projects that should be designed and developed to teach this paper effectively for undergraduate students.

KEYWORDS

Semantic Web, Web 3.0, course design, theory topics, practical labs

INTRODUCTION: THE SEMANTIC WEB IN THE WEB 3.0 ERA

The Semantic Web refers to a vision of the World Wide Web (WWW) in which information is not only human-readable but also machine-readable. It is an extension of the existing web that adds a layer of data and metadata, enabling computers to understand and process content on the internet. The concept of the Semantic Web was first introduced by Sir Tim Berners-Lee, the inventor of the WWW. He envisioned a web where data is not simply presented in a human-readable format, but also structured and organised to facilitate automated processing and knowledge discovery. The Semantic Web aims to bridge the gap between human understanding and machine interpretation of information, enabling more efficient data integration, search and automation. The Semantic Web technologies are based on the use of standardised formats and technologies such as RDF (resource description framework), OWL (web ontology language) and SPARQL (a query language for RDF data), as shown in Figure 1, the complete stack of Semantic Web technologies proposed by Tim Berners-Lee (Heldler et al., 2002).

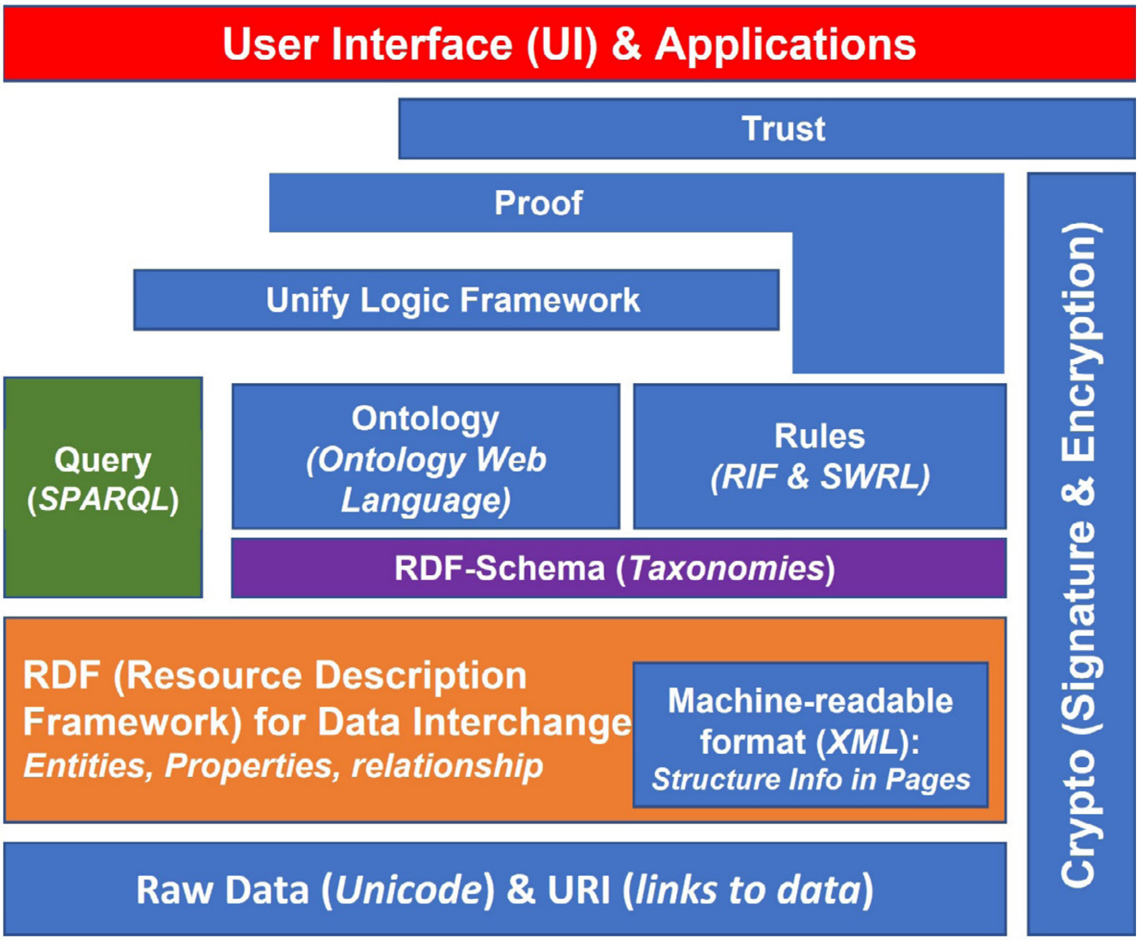


Figure 1. The complete stack of Semantic Web technologies as proposed by Tim Berners-Lee and W3C, according to Heldler et al. (2002).

Figure 2 shows a timeline of web technologies evolving from version 1.0 to the current version 3.0. While Web 1.0 was mainly webpage linked and Web 2.0 was application linked, Web 3.0 is data linked. Web 3.0 is often referred to as the 'intelligent web' or the 'Semantic Web'. The idea of the Semantic Web is to create a more interconnected and intelligent web, where data can be linked, shared and interpreted in meaningful ways (Cano-Benito et al., 2019). The role of the Semantic Web in Web 3.0 is significant because:

- It enables data integration: different sources of data will be connected and integrated seamlessly. By providing a common framework for representing data, it allows information from various websites and databases to be easily linked and combined, leading to a more comprehensive and holistic understanding of the data. As the internet of things continues to grow, the Semantic Web becomes essential in making sense of the vast amounts of data generated by IoT devices. It allows machines to understand the data and take appropriate actions based on the context and relationships between different data points. The Semantic Web's use of standardised data formats ensures that data can be easily exchanged and used across different applications and platforms, promoting data interoperability and reducing data silos.
- It enhances search and discovery on the WWW: search engines can go beyond keyword matching and understand the actual meaning of the content. This leads to more accurate search results and better discovery of related information, improving the overall user experience.
- It promotes the development of knowledge graphs, which are a common way to organise and connect data in a meaningful way, allowing computers to derive new insights and generate knowledge.

It is important to note that the field of Semantic Web technology continues to evolve rapidly, and new advancements are expected in coming years.

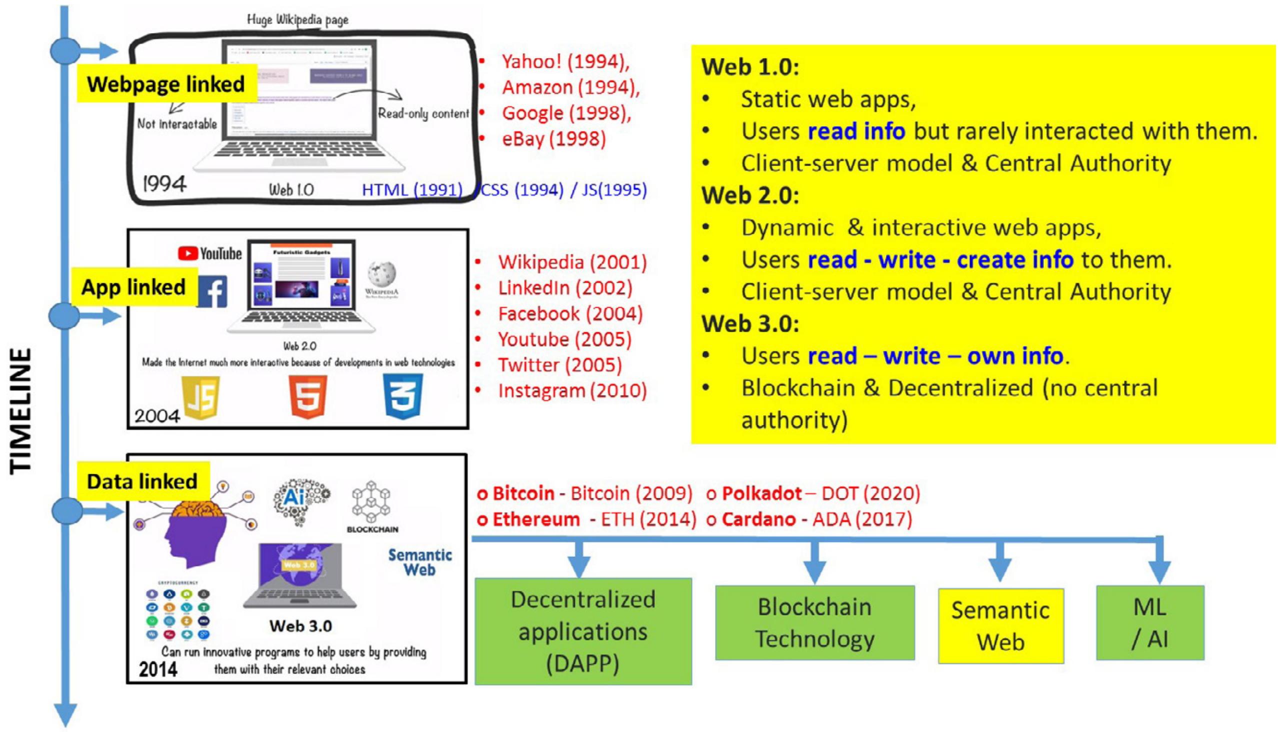


Figure 2. The evolution of web technologies over last 30 years. The Semantic Web is an important component in the Web 3.0 era.

The Semantic Web has found numerous applications in the Web 3.0 era. In the domain of information retrieval and search engines, it has enabled more accurate and context-aware search results, facilitating efficient information discovery (Hitzler, 2021). In e-commerce applications, the Semantic Web has facilitated improved product recommendations and personalised shopping experiences by leveraging user preferences and semantic descriptions of products. It has also been employed in the healthcare sector for better integration of medical data, enabling the development of intelligent decision support-systems and personalised medicine.

Despite its potential, the Semantic Web still faces challenges in its widespread adoption. One of the primary obstacles is the scalability and performance of semantic technologies. Processing large volumes of data, especially in real time, can be computationally expensive. Another challenge lies in the creation and maintenance of ontologies, which require expertise and continuous updates to remain accurate and relevant (Hitzler, 2021). Additionally, privacy and security concerns associated with sharing and integrating diverse data sources need to be addressed to foster trust and collaboration.

AIMS AND RESEARCH DESIGN: HOW TO DESIGN AND TEACH THE SEMANTIC WEB AS A TOPIC IN UNDERGRADUATE PROGRAMMES?

Teaching web technologies and programming is crucial in today's digital world, because web technologies have become an integral part of various industries. In the Bachelor of Computing Systems (BCS) programme currently offered at EIT, we focus more on Web 2.0 technologies and programming, as shown in Figure 3.

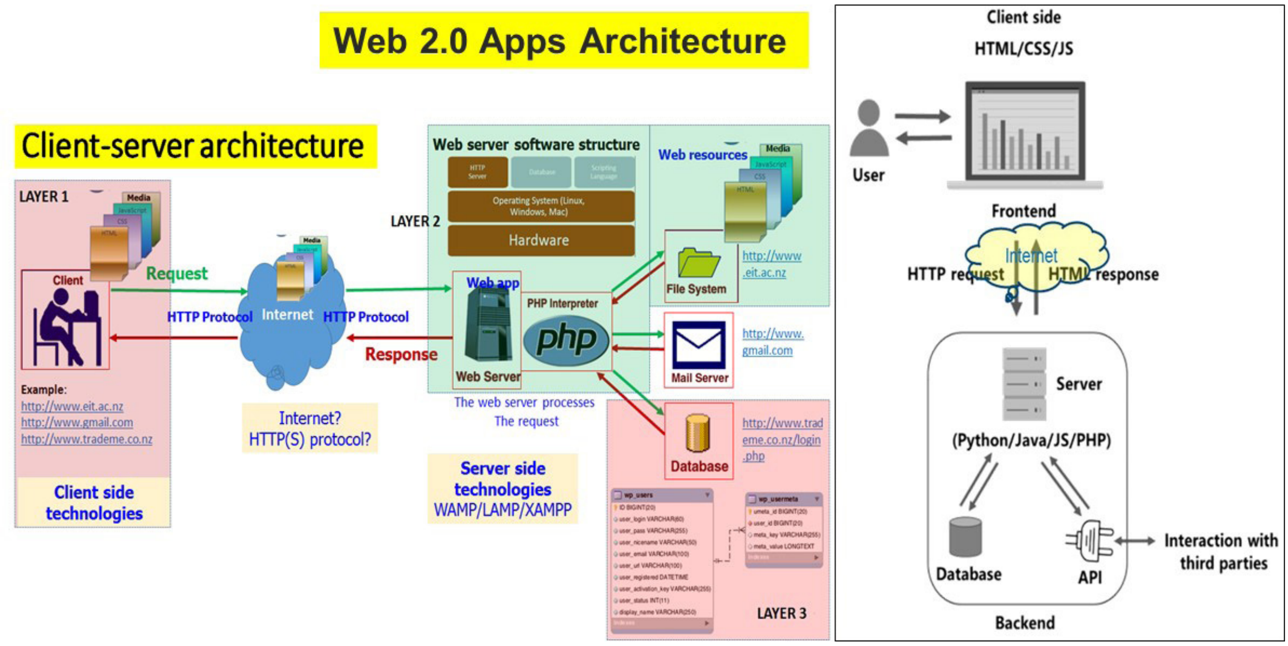


Figure 3. Web 2.0 technologies.

The main topics covered in this paper are front end and back end, as shown in Table 1.

Table 1. Summary of topics and web technologies taught in the at Bachelor of Computing Systems (BCS) programme, EIT.

	Topics	**Web technologies**
Front-end web development	+ Static & Dynamic Webpage Programming + Web Design (CSS3) + Web Content (HTML5) + Web Responsiveness (HTML5+CSS3) + Web Interactivity (JS) + Web Accessibility (HTML/CSS/JS) + Web UX: Usability & Compatibility + Client-Side Technologies + Web Publishing, SEO, Maintenance, Legal (Copyright, Creative Common Licenses, Privacy)	Languages: HTML5, CSS3, JS (ES6) Data Transportation: JSON, XML, AJAX Frameworks: Bootstrap, JQuery, ReactJS
Back-end web development	+ Data-Driven Web App Programming + Database Design & Implementation + Server–Client Centralised Model + Internet and HTTP Protocol + MVC Design Pattern in Web App + Single-Page vs Multiple-Page vs Progressive Web App + Bac-End Frameworks + Web/Cyber Security	Server Languages: PHP, SQL Frameworks: WordPress, Laravel

Teaching the Semantic Web as a subject to undergraduate students offers numerous benefits. Students gain a deep understanding of cutting-edge web technologies, technical skills in data integration (the ability to develop systems with diverse dataset and linked data), and capabilities in information retrieval and knowledge management (data modelling and linked data querying). Moreover, they are able to design and develop different types of web-based applications, such as intelligent data-driven decision-making apps, AI apps, or personalised recommendation apps. Students also gain insights into the ethical considerations and best practices for ensuring data privacy and security.

As a result, students with expertise in the Semantic Web have various job opportunities in e-commerce, healthcare, finance and government, sectors which are increasingly relying on semantic technologies for data integration, knowledge management and intelligent systems.

Three research questions need to be investigated:

1. What are the challenges to teaching the Semantic Web as a subject?
2. What is the best strategy to design and integrate the Semantic Web as a subject into the current curriculum at EIT?
3. What are the theory concepts, practical labs and practical projects recommended to teach the Semantic Web effectively?

ANALYSIS AND FINDINGS

Teaching the Semantic Web as a subject to undergraduate students can present several challenges. By recognising these challenges and proactively addressing them, we can create an engaging and effective learning environment for teaching the Semantic Web. It is also important to be aware of these challenges and address them effectively to ensure successful learning outcomes. Below are some common challenges in teaching the Semantic Web as a subject (Jensen, 2019; Narayanasamy et al., 2022):

1. **Complexity and abstractness:** The Semantic Web involves complex concepts, technologies, and standards that can be initially challenging for students to grasp. The abstract nature of the ontologies, RDF, OWL and SPARQL, may require additional effort to explain and illustrate effectively. It is crucial to break down complex ideas into simpler concepts and provide concrete examples to enhance understanding.
2. **Technical prerequisites:** Teaching the Semantic Web assumes a certain level of technical prerequisites, including knowledge of programming, databases and web technologies. Students may struggle if they lack a solid foundation in these areas. It is essential to assess students' prior knowledge and provide necessary background materials or supplementary resources to bridge any gaps. Creating and managing ontologies requires expertise in domain knowledge and modelling. Students may struggle with ontology engineering, including defining classes, properties and relationships, and applying reasoning.
3. **Lack of familiarity:** The Semantic Web is still a relatively new and evolving field, and students may have limited exposure to it. They may not be familiar with the relevant technologies, tools and applications. This challenge is addressed by providing comprehensive and up-to-date learning materials, demonstrations and hands-on activities to familiarise students with the Semantic Web ecosystem.
4. **Keeping pace with advancements:** The Semantic Web field is constantly evolving, with new technologies, frameworks and standards being introduced. Staying current with the latest advancements can be challenging for educators, and it is crucial to ensure the curriculum reflects current trends and practices.
5. **Limited resources and tools:** Availability of resources and tools for teaching the Semantic Web may be limited, especially in educational settings. Finding appropriate tools, software and datasets can pose a challenge. Processing and managing large-scale semantic data can be computationally expensive and pose challenges in terms of scalability and performance. Students may encounter difficulties when dealing with real-world datasets or when applying semantic technologies to handle substantial amounts of data. Consider providing simplified datasets or exercises that focus on core concepts before gradually introducing more complex scenarios.
6. **Ethical and social implications:** The Semantic Web raises ethical and social considerations related to data privacy, security and fairness. Addressing these implications can be challenging in the classroom setting. Discussions on ethical dilemmas and case studies can help students develop a critical awareness of the societal impact of the Semantic Web and foster responsible and ethical use of semantic technologies.

Integrating the topic of the Semantic Web into the current curriculum of the BCS offered at EIT requires a structured and engaging approach. Establishing clear learning objectives, designing effective learning materials and incorporating hands-on projects can help students to develop a solid understanding of the Semantic Web's concepts, technologies and applications. Encouraging collaboration, critical thinking and topic exploration will empower students to stay up-to-date with the latest advancements. Below is a guide for educators to design and then effectively teach the Semantic Web subject (Sigalov et al., 2023).

- **Establish learning objectives:** Begin by defining clear and measurable learning objectives for the Semantic Web subject. These objectives should align with the broader goals of the course and specify the knowledge and skills students should acquire. In our case, learning objectives are: understanding the fundamental concepts of the Semantic Web; becoming proficient in RDF and OWL technologies; and being able to apply Semantic Web principles to solve real-world problems.
- **Design engaging materials:** Develop a variety of learning materials to engage students and cater to different learning styles. These materials can include lecture slides, online resources, practical labs, interactive tutorials and practical projects. Incorporate examples and real-world applications to demonstrate the practical relevance of the Semantic Web. Encourage students to explore additional resources independently to deepen their understanding.
- **Design assessments (hands-on projects):** The assessments require students to apply their learned knowledge and technologies effectively into projects, such as developing semantic applications, designing and querying ontologies, and analysing and improving existing Semantic Web systems.
- **Stay updated and continuously improve:** Given the evolving nature of the Semantic Web, it is crucial to stay updated on the latest advancements, research and industry trends. The course content should be continuously enhanced by incorporating new case studies, examples and technologies.

When teaching the Semantic Web as a subject, it is important to cover a range of topics and concepts that provide students with a comprehensive understanding of this field. The Semantic Web encompasses various foundational concepts, technologies and applications that enable intelligent data-processing and data-integration. Practical labs are invaluable for learning the Semantic Web, as they provide students with hands-on experience and a deeper understanding of the concepts and technologies involved. These labs allow students to apply their knowledge in real-world scenarios, reinforcing their understanding and fostering critical-thinking skills. We recommend a range of practical lab ideas that can be implemented to teach the Semantic Web effectively, described in Table 2 (Sigalov et al., 2023; Zlatareva, 2021).

Table 2. A recommended list of theory topics, practical labs and practical projects to teach the Semantic Web as a subject.

Topics	Detail content and practical labs
Introduction to the Semantic Web	+ Intro of the basic concept of the Semantic Web and its goals + Need for adding semantics to web resources and the benefits + Comparison: Semantic Web vs traditional web technologies
Resource Description Framework (RDF)	+ RDF – a flexible data model for representing knowledge + Triples model: SPO = Subjects – Predicates – Objects + Create semantic representations of resources using RDF
Ontologies and Web Ontology Language (OWL)	+ Intro of ontologies: knowledge nodes and relationships + Create knowledge graphs + Web ontology language (OWL): different versions and syntax **Practical Lab:** Developing Ontologies with Protégé – a free open-source web-based tool: https://protege.stanford.edu/ In this lab, students can learn to use the Protégé tool to develop their own ontologies: Create classes, properties and individuals. Define relationships. Set constraints.

Linked Data Principles and RDF Schema	+ Intro to linked data principles: interlink and integrate datasets from various sources of datasets across the web + Principles of referenceable URIs, RDF linking and using HTTP as the protocol for data access + RDF Schema (RDFS): a basic vocabulary for representing schemas and defining simple relationships between classes and properties **Practical Lab:** Semantic Web Integration with Linked Data. In this lab, students can work with linked data sources to understand how different datasets can be combined. Provide multiple datasets. Integrate datasets. Two main sources of datasets: + Wikidata: https://www.wikidata.org/ + Google Knowledge Graph Search: https://developers.google.com/knowledge-graph Use RDF links to develop SPARQL queries to retrieve integrated information. + SPARQL: https://www.w3.org/TR/rdf-sparql-query/ + RDF: https://www.w3.org/RDF/
SPARQL Query Language	+ SPARQL: Query language for RDF data and extracting information from datasets + Demonstrate SPARQL queries to retrieve specific data, perform filtering and traverse relationships within the graph + Demonstrate SPARQL for data exploration, integration and analysis **Practical Lab:** Creating and Querying RDF Data. In this lab, students can learn the basics of creating RDF data and querying it using SPARQL. Provide an available dataset or ask students to create their own small dataset using RDF triples. Write SPARQL queries to retrieve specific information from the dataset, perform filtering and explore relationships within the data.
Ontology Development and OWL Reasoning	+ The process of ontology development using OWL + Key concepts: classes, individuals, properties and restrictions + Define class hierarchies, property domains and ranges + Intro of reasoning in OWL: the role of inference engines in inferring implicit knowledge and verifying ontology consistency **Practical Lab:** Ontology Reasoning and Inference. This lab focuses on reasoning capabilities in ontologies using OWL reasoning engines. Create an inference scenario, set up an OWL reasoner and apply it to infer implicit knowledge from the ontology. + OWL: https://www.w3.org/2001/sw/wiki/OWL/Implementations
Semantic Web Applications	+ Explore a range of Semantic Web applications + Applications in information retrieval, e-commerce, healthcare, social networks and government services + Applications in intelligent search, recommendation systems, data integration and decision support using the Semantic Web **Practical Lab:** Semantic Web Applications in Information Retrieval: In this lab, students can explore how the Semantic Web enhances information retrieval. • Provide a dataset of web resources with semantic annotations and metadata. • Develop a simple search engine that leverages these annotations to provide more accurate and context-aware search results. • Use SPARQL queries to retrieve relevant resources based on user queries and present the results. Semantic Web and Internet of Things (IoT): In this lab, students can explore how semantic technologies can be applied in IoT scenarios. • Provide a set of IoT sensor data. • Develop an application that leverages semantic annotations to enhance data integration and reasoning. • Use RDF, SPARQL and ontologies to process and analyse IoT data effectively. Semantic Web and Natural Language Processing (NLP): In this lab, students can work with NLP tools and Semantic Web technologies to extract semantic information from textual data. • Provide a corpus of text document. • Use NLP techniques to identify entities, relationships and sentiments. • Create RDF representations and query the data using SPARQL. Semantic Web and Recommendation Systems: In this lab, students can explore how the Semantic Web enhances recommendation systems. • Provide a dataset of user preferences and product information with semantic annotations. • Develop a recommendation system that utilises semantic knowledge to generate personalised recommendations.
Challenges and Future Directions	+ Address the challenges faced by the Semantic Web: scalability, ontology engineering, data privacy and security concerns + Emerging technologies: decentralised approaches, blockchain integration, and advanced machine learning techniques + Future direction and potential impact of the Semantic Web

CONCLUSION AND RECOMMENDATIONS

In conclusion, the Semantic Web is a fundamental component of Web 3.0, providing a standardised framework for machine-readable data and enabling the development of an intelligent and interconnected web. It empowers machines to understand, process, and reason about information, leading to more efficient data integration, personalised experiences and improved knowledge discovery. As the internet continues to evolve, the Semantic Web will play a vital role in shaping the future of the digital landscape.

Teaching the Semantic Web as a subject to undergraduate students can present several challenges. By recognising these challenges and proactively addressing them, we can create an engaging and effective learning environment for teaching the Semantic Web. It is also important to be aware of these challenges and address them effectively to ensure successful learning outcomes. Encouraging active learning, providing hands-on experiences, and promoting collaborative discussions will help students overcome these challenges and develop a solid understanding of the subject matter.

Teaching the Semantic Web as a subject involves covering a wide range of topics, hands-on practical labs and projects to ensure students gain a comprehensive understanding of this field. We recommend a list of topics and practical labs that can be implemented to teach the Semantic Web effectively, described in Table 2. By engaging students in creating and querying RDF data, developing ontologies, working with linked data and exploring various applications, they gain practical skills and insights into real-world scenarios. These labs foster critical thinking, problem-solving abilities, and an appreciation for the practical applications of the Semantic Web.

REFERENCES

Cano-Benito, J., Cimmino, A., & García-Castro, R. (2019). Towards blockchain and Semantic Web. *International Conference on Business Information Systems*. https://doi.org/10.1007/978-3-030-36691-9_19

Heldler, J., Berners-Lee, T., & Miller, E. (2002). Integrating applications on the Semantic Web. *Journal of the Institute of Electrical Engineers of Japan*, *122*(10), 676–680. https://doi.org/10.1541/ieejjournal.122.676

Hitzler, P. (2021). A review of the Semantic Web field. *Communications of the ACM (2021)*, *64*(2), 76–83. https://cacm.acm.org/research/a-review-of-the-semantic-web-field/

Jensen, J. (2019). A systematic literature review of the use of Semantic Web technologies in formal education. *British Journal of Educational Technology*, *50*(2), 505–517. https://doi.org/10.1111/bjet.12570

Narayanasamy, S. K., Srinivasan, K., Hu, Y. C., Masilamani, S. K., & Huang, K. Y. (2022). A contemporary review on utilizing Semantic Web technologies in healthcare, virtual communities, and ontology-based information processing systems. *Electronics*, *11*(3), 453. https://doi.org/10.3390/electronics11030453

Sigalov, S. E., & Nachmias, R. (2023). Investigating the potential of the semantic web for education: Exploring Wikidata as a learning platform. *Education and Information Technologies*, *28*, 12565–12614. https://doi.org/10.1007/s10639-023-11664-1

Zlatareva, N. P. (2021). A hands-on project for teaching semantic web technologies in an undergraduate AI course. *Journal of Machine Intelligence and Data Science*, *2*, 1–11. https://doi.org/10.11159/jmids.2021.001

AUTHOR

Dr Daniel Dang is an ICT Lecturer in the School of Computing at Eastern Institute of Technology Te Aho a Māui. His research interests are sharing best practices in teaching ICT subjects, and using digital technologies to tackle social and environmental issues. ddang@eit.ac.nz

Assessing Web Application Security Through Vulnerabilities in Programming Languages and Environments

Thomas Desmond Kerr-Smith, Sreenivas Sremath Tirumala and Michael Andrews

https://doi.org/10.34074/proc.240109
Correspondence: Sreenivas Sremath Tirumala, Sreenivas.Tirumala@manukau.ac.nz

This publication may be cited as:

Kerr-Smith, T. D., Tirumala, S. S., & Andrews, M. (2024). Assessing Web Application Security Through Vulnerabilities in Programming Languages and Environments. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 62–69). ePress, Unitec. https://doi.org/10.34074/proc.240109

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

ePress Research with impact Unitec

Unitec is a business division of Te Pūkenga – New Zealand Institute of Skills and Technology

ABSTRACT

Cybersecurity has become increasingly important for organisations. At present, there are several software development companies that employ strategies to eliminate code-based security risks. It is estimated that 95% of businesses use web applications that are the most vulnerable among all other types of applications. Application security company Veracode checked out 759,445 applications over a year using different security scans and found that 74% of these apps had at least one security flaw. It is important to note that Veracode also pointed out that 69% of the applications had at least one vulnerability from the Open Worldwide Application Security Project's Top Ten list.

This research is divided into two parts. First, two web applications were developed using ASP.NET and Node.js respectively, using security best practices. This was followed by analysing the security of these applications for some of the OWASP's Top 10 vulnerabilities using cybersecurity software and tests. The outcomes are then reported, which provides an understanding of the vulnerabilities of the applications developed using ASP.NET and Node.js, in spite of using security best practices. The conclusions of the report emphasise the importance of incorporating cybersecurity practices at the earlier stages of the software development lifecycle (SDLC) and providing sufficient awareness about cybersecurity vulnerabilities to the developers.

KEYWORDS

Web application vulnerabilities, programming best practices, security in programming languages

INTRODUCTION

Cybersecurity has become a key requirement for any organisation. However, due to existing vulnerabilities in applications, particularly web applications, it is always a challenge to implement a zero-tolerance approach. At present, over 95% of businesses have a presence on the internet through web applications. In 2023, Veracode, an application security company, performed an independent analysis on 759,445 applications over a period of 12 months using static and dynamic vulnerability scans. According to the analysis, over 74% of the applications had at least one security flaw within their application due to the code or programming language they used for the development. Also, Veracode highlighted the fact that about 69% of the applications had at least one vulnerability from the Open Worldwide Application Security Project's (OWASP) Top 10 of 2021. Given how prevalent security risks were in the Veracode study, it raises the question of why there are so many risks found within these applications (Pattison-Gordon, 2023).

The Veracode study also shows the rate and age at which applications generate security flaws. Out of the 759,445 applications that were studied, 20% of them presented new flaws at the deployment stage. These risks can occur for many reasons, such as code quality, server configuration, insecure design, or code/SQL injection. This shows how important the overall design and quality of code is for handling growth of an application over time (Veracode, 2023). Given how prevalent security risks are in modern-day applications, it is important to ask why these risks develop and what is the best strategy to mitigate these risks for both short- and long-term security of an application.

Shift-left strategies have been successful in reducing the number of bugs in software programs. However, this process may not be sufficient to provide a higher level of cybersecurity. Firstly, shifting the responsibility of securing an application onto developers with little or no knowledge of cybersecurity may produce over-caution and will slow down the software development lifecycle (SDLC), as stated by Laura Bell (2022). Bell also asserts that shifting the responsibility to developers individually may not have the same impact as a collaborative high-level strategy or framework.

This research is divided into two parts. Firstly, two web applications were developed using ASP.NET and Node.js respectively, using security best practices. This was followed by analysing the security of these applications for some of the OWASP's Top 10 vulnerabilities using cybersecurity software and tests. The outcomes are then reported, which provides an understanding of the vulnerabilities of the applications developed using ASP.NET and Node.js, in spite of using security best practices. The conclusions of the report emphasise the importance of incorporating cyber-security practices at the earlier stages of the SDLC and providing sufficient awareness about cybersecurity vulnerabilities to the developers.

AIMS AND RESEARCH DESIGN

The research project aims to answer the following two questions:

- How big an impact does the choice of a programming language or programming environment have on the overall mitigation of security risks?
- Will a testing shift in the software development lifecycle improve overall application security?

This research uses a deductive approach to conduct research and work towards answering these research questions. It is also grounded in the practice of software security and its increasing importance in modern-day web applications. This affords insights into how security is handled and what the comparative cybersecurity risks are when two similar web applications are developed using Node.js v18.16.0 and ASP.NET version 7.0 respectively.

The research tries to capture information on how different web applications introduce standard security features, such as hashed passwords for password security or validation of user input, to avoid injection and cross-site scripting.

Security Features

The following security features were considered for the testing the web applications:

- **HTML encoding:** This is used to prevent malicious scripts or content being pushed through web applications in a process called cross-site scripting (XSS).
- **Password management:** Password hashing was tested for both Node.js and ASP.NET. Node.js uses a library called bcrypt, which contains methods and functions that allow the user to hash passwords using the bcrypt algorithm.
- **Database connectivity:** When connecting to a database it is best practice to hide database credentials from users to prevent unauthorised access and to protect user data.
- **Token validation** is vital in web applications to protect a user's sessions and reduce the risk of cross-site request forgery (CSRF). CSRF is the process of an attacker tricking a user's browser into making malicious requests that the user does not intend.
- **User roles** are an important feature for any web application for managing access to certain areas or features of a web application. Restricting users from certain actions or from accessing sensitive data is vital to the integrity of a web application.

- **Session management:** Sessions are used to retain information about a user over multiple requests within a web application. Sessions are always an important aspect of applying security features to cookies and session IDs.

Testing Vulnerabilities and Outcomes

The applications were developed and various tests were conducted on the web applications.

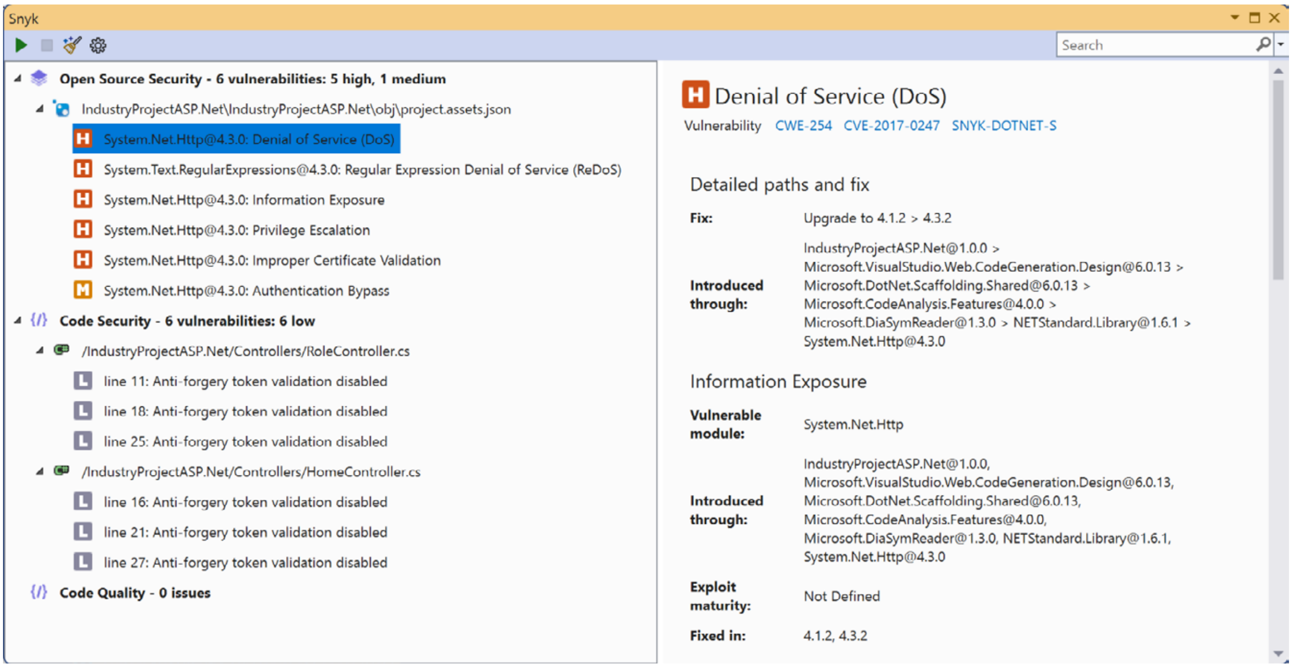
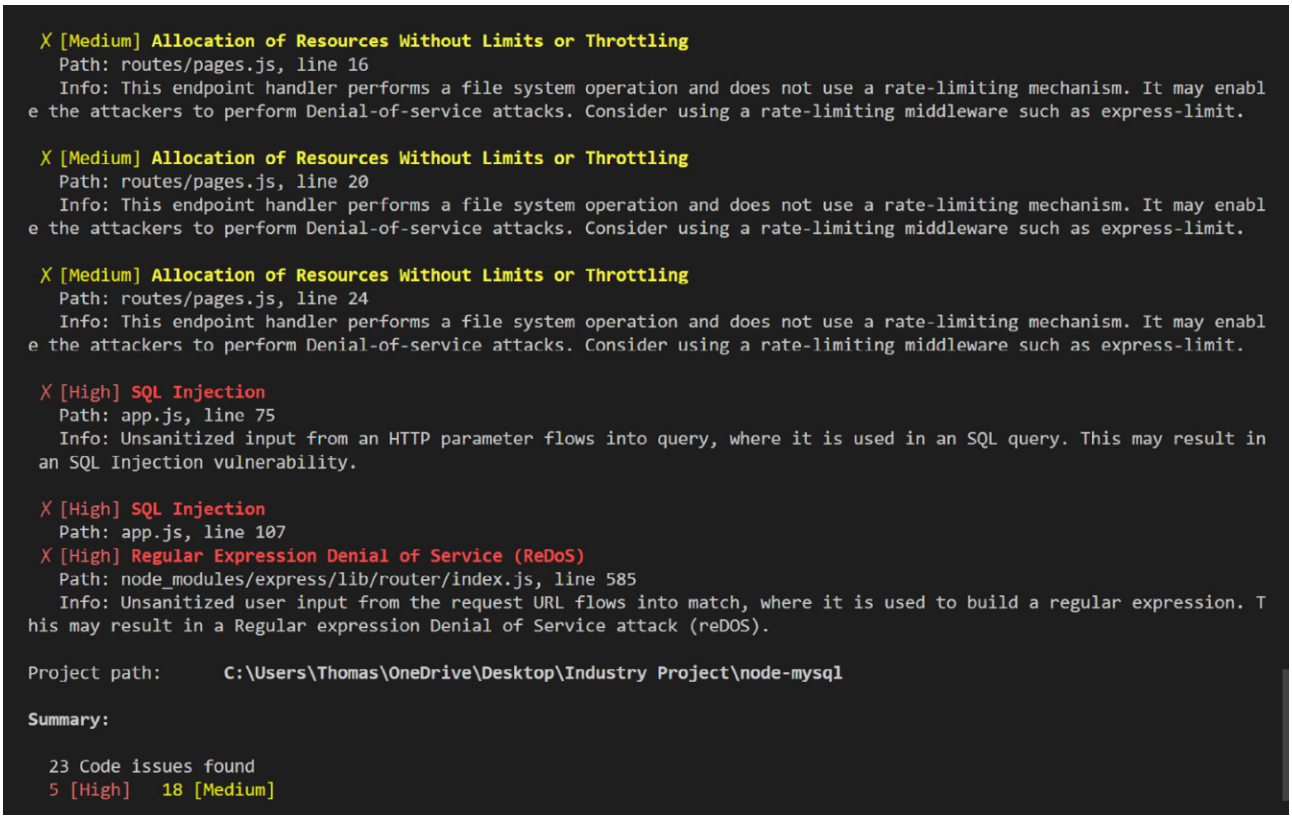


Figure 1. ASP.NET first Snyk security scan.

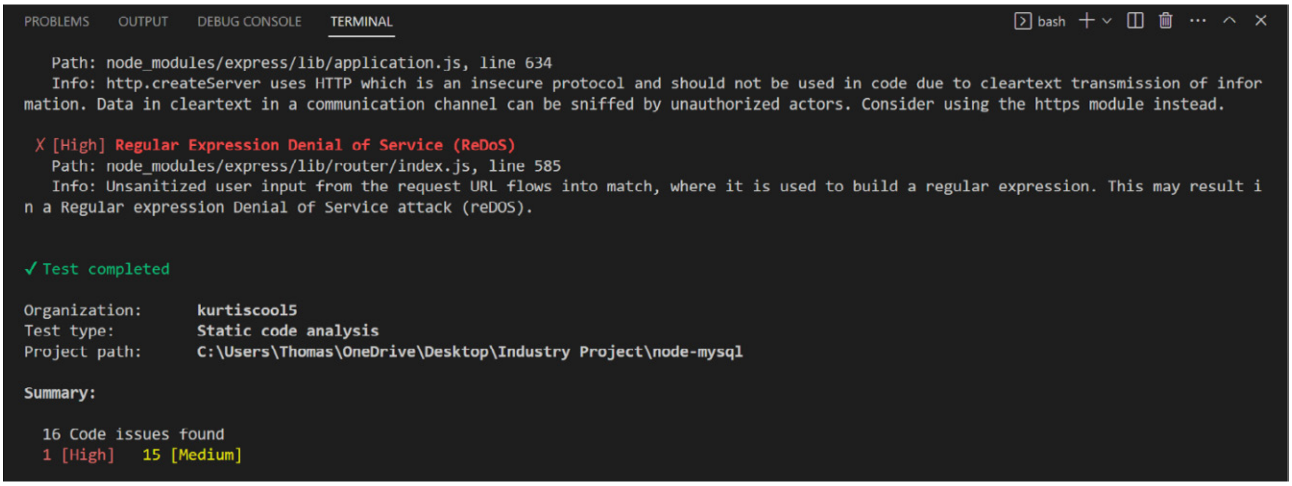
The first scan of the ASP.NET web application returned five high, one medium and six low vulnerabilities, as shown in Figure 1. The first scan of the Node.js application returned five high-risk vulnerabilities and 18 medium-risk vulnerabilities, as shown in Figure 2.



```
 X [Medium] Allocation of Resources Without Limits or Throttling
   Path: routes/pages.js, line 16
   Info: This endpoint handler performs a file system operation and does not use a rate-limiting mechanism. It may enabl
e the attackers to perform Denial-of-service attacks. Consider using a rate-limiting middleware such as express-limit.

 X [Medium] Allocation of Resources Without Limits or Throttling
   Path: routes/pages.js, line 20
   Info: This endpoint handler performs a file system operation and does not use a rate-limiting mechanism. It may enabl
e the attackers to perform Denial-of-service attacks. Consider using a rate-limiting middleware such as express-limit.

 X [Medium] Allocation of Resources Without Limits or Throttling
   Path: routes/pages.js, line 24
   Info: This endpoint handler performs a file system operation and does not use a rate-limiting mechanism. It may enabl
e the attackers to perform Denial-of-service attacks. Consider using a rate-limiting middleware such as express-limit.

 X [High] SQL Injection
   Path: app.js, line 75
   Info: Unsanitized input from an HTTP parameter flows into query, where it is used in an SQL query. This may result in
 an SQL Injection vulnerability.

 X [High] SQL Injection
   Path: app.js, line 107
 X [High] Regular Expression Denial of Service (ReDoS)
   Path: node_modules/express/lib/router/index.js, line 585
   Info: Unsanitized user input from the request URL flows into match, where it is used to build a regular expression. T
his may result in a Regular expression Denial of Service attack (reDOS).

Project path:      C:\Users\Thomas\OneDrive\Desktop\Industry Project\node-mysql

Summary:

  23 Code issues found
  5 [High]   18 [Medium]
```

Figure 2. Vulnerabilities in the application developed using Node.js.

The first high-risk vulnerability of the web application developed using Node.js occurred in app.js and it was a function in risk of SQL injection. The reason for this vulnerability to SQL injection is due to non-sanitisation of the parameters, which is a default configuration. The second end-to-end scan on the Node.js app provided a high-risk vulnerability that is prone to denial-of-service attack, as shown in Figure 3.



```
PROBLEMS   OUTPUT   DEBUG CONSOLE   TERMINAL                                              bash

   Path: node_modules/express/lib/application.js, line 634
   Info: http.createServer uses HTTP which is an insecure protocol and should not be used in code due to cleartext transmission of infor
mation. Data in cleartext in a communication channel can be sniffed by unauthorized actors. Consider using the https module instead.

 X [High] Regular Expression Denial of Service (ReDoS)
   Path: node_modules/express/lib/router/index.js, line 585
   Info: Unsanitized user input from the request URL flows into match, where it is used to build a regular expression. This may result i
n a Regular expression Denial of Service attack (reDOS).

✓ Test completed

Organization:      kurtiscool5
Test type:         Static code analysis
Project path:      C:\Users\Thomas\OneDrive\Desktop\Industry Project\node-mysql

Summary:

  16 Code issues found
  1 [High]   15 [Medium]
```

Figure 3. End-to-end scan of the application developed using Node.js.

Brute force was performed on both applications to exploit passwords, which was successful, with a matching length of 1550 and 1183. The XSS attack was unsuccessful, due to HTML encoding practices used in the development of the web applications.

SQL injection attack and 'comment SQL injection attack' were performed on both the applications, as shown in Figure 4 and Figure 5. Also, a wordlist-based SQL injection was also performed using a penetrating testing software. All three were unsuccessful due to coding practices. It is surprising to see that the vulnerability test shown in Figure 3 indicates that the Node.js application is prone to SQL injection.

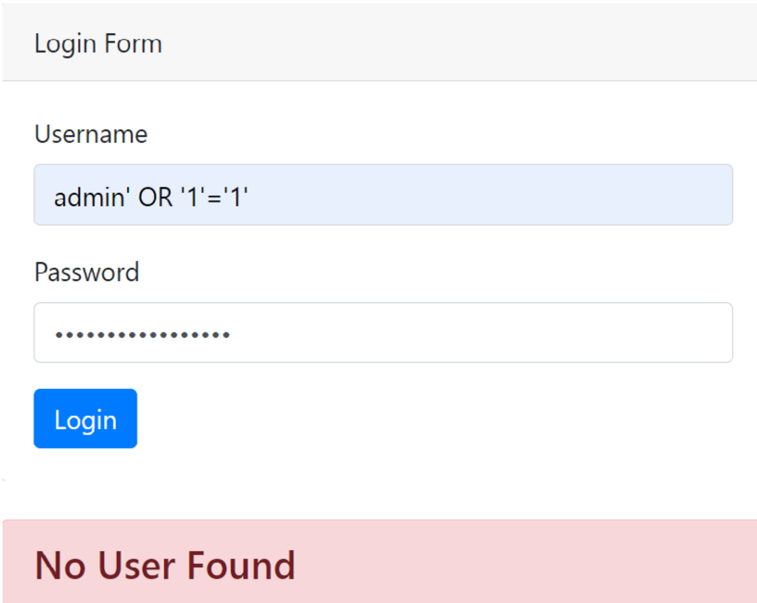


Figure 4. SQL injection test on the application developed using ASP.NET.

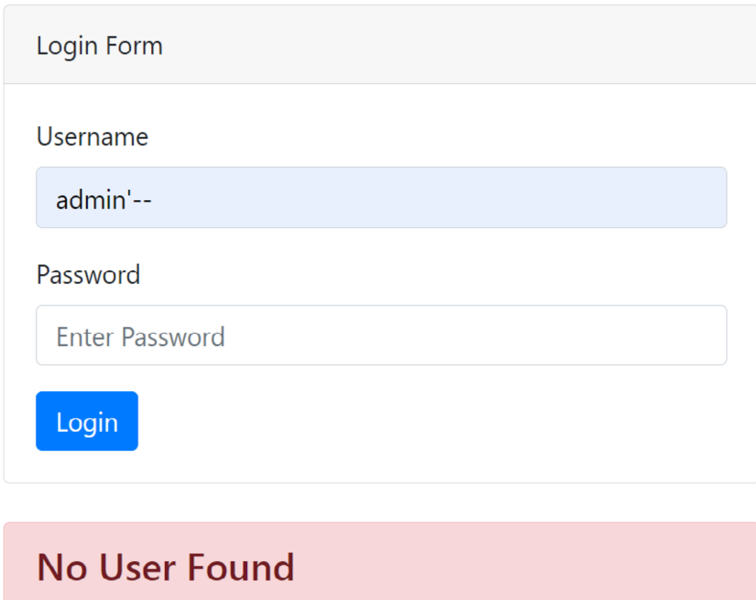


Figure 5. SQL injection test on the application developed using Node.js.

There was a vulnerability discovered in the code that was not written by the developer. During the end-to-end testing it was noticed that a high-risk vulnerability called "Regular expression denial of service (ReDoS)" was occurring in code that was generated in the backend automatically when the applications were executed. This vulnerability was defined as "Unsanitised user input from the request URL flows into match, where it is used to build a regular expression."

CONCLUSION AND RECOMMENDATIONS

Identifying the importance of detecting vulnerabilities of web applications before deployment is the key aspect of this research. Two web applications were developed using ASP.NET and Node.js respectively, also incorporating security best practices. When the applications were tested for vulnerabilities, the outcomes provided a clear indication of high-risk and medium-risk vulnerabilities in those applications. Some vulnerabilities, but not all, could be fixed by updating the framework or using similar techniques.

Both applications were good enough to tackle SQL injection and code injection through HTML, due to mandatory coding practices that are inbuilt into the framework itself. Regular expression denial of service (ReDoS) still exits, due to automated code generated at the webserver level (not by the web application), which requires further investigation.

While both languages can produce secure websites, ASP.NET's built-in security features, such as Identity framework, and the support of and integration with Microsoft, makes it a clear choice for prioritising security and scalability. Microsoft also provides extensive security documentation for developers who are using ASP.NET, making it ideal for developers who need to build secure applications.

Finding and fixing these security vulnerabilities before the application is deployed, or even developed, allows a coding team to generate stronger, cleaner and more robust applications that will be less susceptible to vulnerabilities or breaches once deployed. This can be achieved by static testing and testing the code without execution.

The research is limited to a small web application developed with limited data. It will be interesting to see the impact when a similar approach is adopted for a larger web application. Further, the research needs more testing towards understanding the open problems that exist in a particular framework and how these can be addressed by moving to a newer version. At the same time, any update potentially gives rise to new vulnerabilities that need to be identified and fixed. This remains a significant concern.

REFERENCES

Bell, L. (2022). *Why we need to stop shifting cyber security left.* SafeStack. https://safestack.io/blog/resources/secure-development-stop-shifting-cyber-security-left

Pattison-Gordon, J. (2023). *New report: One-fifth of software has a severe security flaw*. Government Technology. https://www.govtech.com/security/new-report-one-fifth-of-software-has-a-severe-security-flaw

Veracode. (2023). *State of software security 2023*. https://info.veracode.com/report-state-of-software-security-2023.html

AUTHORS

Thomas Desmond Kerr-Smith is a student in the Bachelor of Digital Technologies at Manukau Institute of Technology Te Whare Takiura o Manukau. He likes to explore and evaluate coding practices to identify vulnerabilities at the development stage of SDLC.
kerr238@manukaumail.com

Dr Sreenivas Tirumala is a Senior Lecturer in the School of Digital Technologies at Manukau Institute of Technology Te Whare Takiura o Manukau. His research areas are privacy, security practices and implementing low-cost artificial intelligence solutions for cybersecurity.
Sreenivas.Tirumala@manukau.ac.nz

Michael Andrews is a Senior Lecturer in the School of Digital Technologies at Manukau Institute of Technology Te Whare Takiura o Manukau. His research interests are in educational practice and the effective adoption of technologies designed to enhance learning, engagement and content approach. He has recently been engaged in investigating best practises for developing software to avoid vulnerabilities in programming languages that are prone to attacks.
Michael.Andrews@manukau.ac.nz

Using Blended HTML-CSS-JS Semantic to Implement Web Accessibility Principles

Daniel Dang

https://doi.org/10.34074/proc.240110
Correspondence: ddang@eit.ac.nz

This publication may be cited as:

Dang, D (2024). Using Blended HTML-CSS-JS Semantic to Implement Web Accessibility Principles. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 70–77). ePress, Unitec. https://doi.org/10.34074/proc.240110

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga – New Zealand Institute of Skills and Technology

ABSTRACT

Web accessibility is the inclusive practice to overcome all barriers that prevent interaction with websites or web apps on the World Wide Web (WWW) by people with physical disabilities (cognitive/auditory/visual limitation), situational disabilities (small screen, bright sunlight), and socioeconomic restrictions on bandwidth and speed. Tim Berners-Lee, W3 Director and inventor of the WWW, stated that "the power of the Web is in its universality, access by everyone regardless of disability is an essential aspect" (World Wide Web Consortium, n.d.). Web accessibility is a crucial aspect of modern technology and is playing an increasingly critical role in shaping a more equitable and accessible future for all users. Teaching web accessibility as a topic to undergraduate students in the Bachelor of Computing Systems (BCS) at EIT is not only a technical necessity (meeting legal requirements and enhancing user experiences) but also promotes inclusivity and fosters social responsibility. As a result, students are prepared for successful and rewarding careers in the technology industry. However, this also comes with challenges. These challenges may stem from various factors, such as technical complexities, time constraints, resource constraints, and lack of awareness and understanding. This article discusses in more detail some challenges to teaching this topic to undergraduate students. It then suggests how to use HTML/CSS/Javascript semantics to implement four web accessibility principles – perceivable, operable, understandable and robust (POUR).

KEYWORDS

Web accessibility, web accessibility principles, HTML semantic, CSS semantic, Javascript semantic

INTRODUCTION

Web accessibility is the inclusive practice to overcome all barriers that prevent interaction with websites or web apps on the World Wide Web (WWW) by people with physical disabilities (cognitive/auditory/visual limitation), situational disabilities (small screen, bright sunlight), and socioeconomic restrictions on bandwidth and speed. Web accessibility ensures that digital content and services are available to everyone. Many countries have laws and regulations that mandate web accessibility for public and private organisations. By teaching web accessibility, we equip our students (future technology professionals) with the knowledge and skills needed to adhere to these legal requirements and uphold ethical standards in their work. As accessibility becomes an integral part of web-development practices, students who understand web accessibility will be better prepared for future job opportunities (Web Accessibility Initiative, n.d.b). Organisations increasingly prioritise accessible design, and candidates with accessibility skills will have a competitive advantage.

Teaching web accessibility promotes a sense of social responsibility among students, as they recognise the impact of their work on the lives of individuals with disabilities. By prioritising accessibility, they contribute positively to society and advocate for an inclusive online environment. Web accessibility challenges students to think creatively and innovatively to solve design and development problems. Overcoming accessibility barriers can lead to novel solutions and more robust web applications (Abuaddous et al., 2016).

As shown in Figure 1, teaching this topic comes with challenges. These challenges may stem from various factors, including technical complexities, attitudinal barriers and resource constraints. Educators must be aware of and address the challenges that can arise during the teaching process.

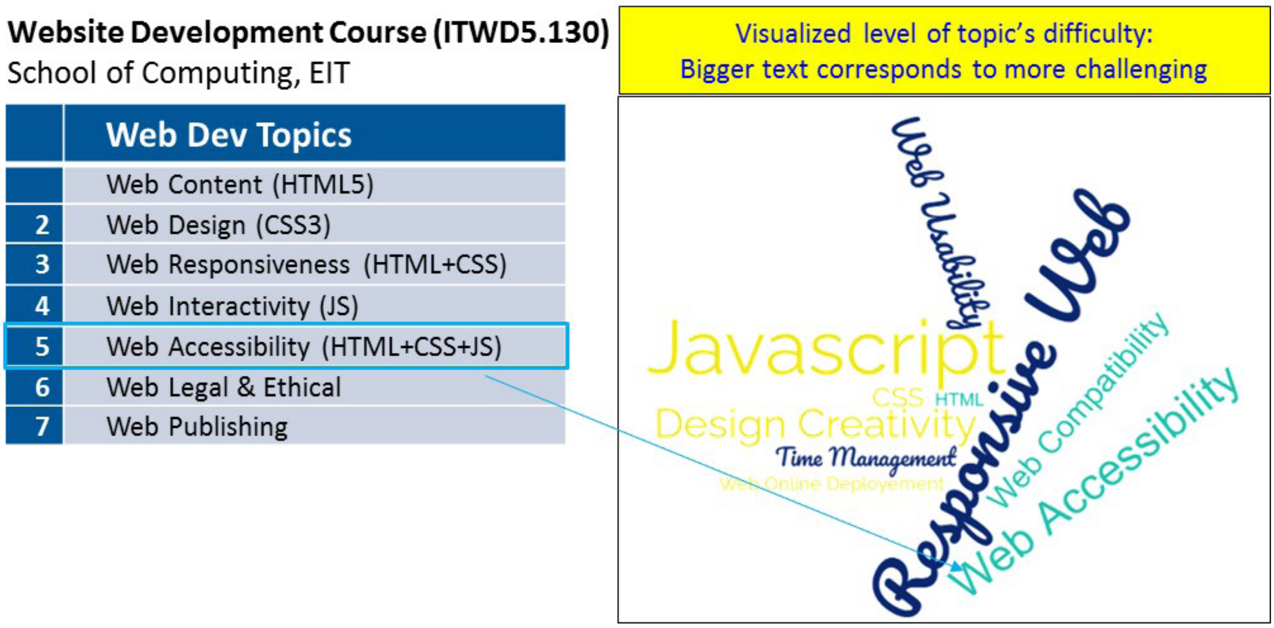


Figure 1. Visualised level of topic's difficulty: bigger text corresponds to more challenging (Dang, 2022).

AIMS AND RESEARCH DESIGN

People with disabilities represent a significant market segment with unique needs and preferences. Teaching web accessibility empowers students to tap into this underserved market and create products and services that cater to diverse audiences. As the population ages, the number of individuals with disabilities is expected to increase. We prepare students to address the needs of an ageing population and a more diverse user base by teaching web accessibility.

In an increasingly digital world, the internet plays a pivotal role in providing access to information, communication and services. However, without proper consideration of accessibility, many individuals with disabilities face barriers in accessing and utilising online resources. To address this issue, it is crucial to integrate web accessibility education into the curriculum of computer science and BCS students.

Research questions:

1. What are web accessibility principles?
2. What are the challenges to teaching web accessibility to undergraduate students?
3. How can we blend HTML/CSS/JS semantics to build an accessible website?

ANALYSIS AND FINDINGS

Web Accessibility Principles

Web accessibility principles are a set of guidelines and best practices designed to ensure that websites and their content are usable and accessible to individuals with disabilities. These principles aim to provide equal access and a positive user experience for all users, regardless of their abilities. The Web Content Accessibility Guidelines (WCAG), developed by the Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C) (Web Accessibility Initiative, n.d.a), are the most widely recognised and accepted set of principles for web accessibility. The WCAG guidelines are organised around four key principles, often referred to by the acronym 'POUR', as shown in Figure 2:

- **Perceivable:** Information and user interface (UI) components must be presented in ways that users can perceive. This involves providing alternatives for non-text content, such as images and multimedia, so that users with visual or cognitive disabilities can understand them. It also includes using clear and consistent navigation and layout structures.
- **Operable:** UI components and navigation must be operable by users with various disabilities. This includes ensuring that all functionality is available via keyboard navigation, as some users may not be able to use a mouse. Time-sensitive content or animations should also be adjustable or avoid causing confusion.
- **Understandable:** The content and operation of the UI must be understandable. This involves using clear and simple language, providing instructions and context, and designing consistent and predictable navigation.
- **Robust:** Content must be robust enough to be reliably interpreted by a wide variety of user agents, including assistive technologies. This means using valid code, adhering to web standards, and avoiding technologies that might create barriers for users.

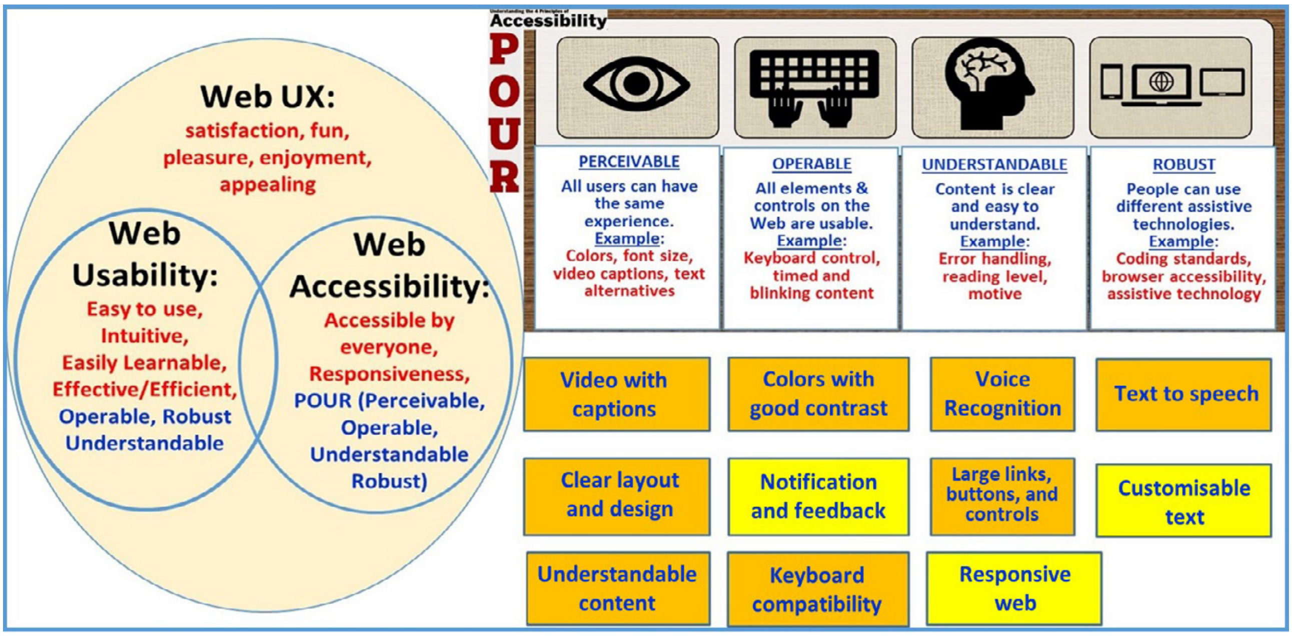


Figure 2. Four principles of web accessibility – POUR: Perceivable, Operable, Understandable and Robust (Web Accessibility Initiative, n.d.a).

Within these four principles, there are specific guidelines that provide more detailed instructions for making web content accessible, as shown in Figure 2. Some common practices include providing alternative text for images, ensuring sufficient colour contrast, providing captions for videos or voice recognition and text-to-speech support, clear layout and design, accessible large links/buttons/controls, and making sure interactive elements are easily navigable via keyboard.

Web accessibility is not only about accommodating individuals with disabilities; it also improves the overall user experience for everyone, including those using mobile devices or experiencing temporary impairments. Following these principles helps create a more inclusive and user-friendly online environment.

Teaching Web Accessibility: Challenges

Teaching web accessibility to undergraduate students is essential for fostering an inclusive digital landscape. However, educators must be aware of and address the challenges that can arise during the teaching process. Teaching web accessibility to undergraduate students faces challenges that stem from four main factors:

(1) technical complexities, (2) time constraints, (3) resource constraints, and (4) lack of awareness and understanding, as shown in Table 1.

By taking a proactive approach, staying up to date with evolving guidelines, providing hands-on experiences, and emphasising the real-world impact of accessible design, tutors can effectively equip students with the knowledge and skills needed to champion web accessibility in their future careers.

Table 1. Challenges when teaching web accessibility to undergraduate students.

Challenges	Explanation
Technical complexity	It is technically challenging to design and develop accessible websites that work seamlessly across different platforms, across browsers, and supporting assistive technologies. As web accessibility principles include all other fields, such as web responsiveness, web usability and advanced web technologies, implementing web accessibility features can be technically complex, especially when addressing diverse disabilities and various types of devices.
Time constraints	The bachelor programme has a packed curriculum with various courses to cover, and web accessibility is only one topic covered in one paper, Website Development, so allocating sufficient time to effectively teach this topic might be difficult, which leads to limited coverage or rushed teaching. As a result, students might not gain a comprehensive understanding of the topic. Students might have limited exposure to working on projects that prioritise web accessibility. Without practical experience, it can be challenging for them to fully grasp the complexities and benefits of accessible design.
Resource constraints	Accessing a range of assistive technologies (screen readers, text-to-speech, and braille computers) for testing and experimentation is challenging and limited for both educational institutions and individual students. Moreover, web accessibility guidelines and assistive technologies are continuously evolving to keep up with advancements in web development. This presents a challenge for tutors to stay updated with the latest standards and teach students the most current best practices.
Lack of awareness and understanding	At the beginning of the semester in 2021, we conducted a small survey by asking students a question: "Do you prioritise a responsive web design that provides a great mobile UX over web accessibility?" Surprisingly, many students said "Yes." It means that those students underestimate the significance of web accessibility, viewing it as an optional or secondary concern in web development. They might not be aware of the barriers faced by people with disabilities, or the legal and ethical aspects of web accessibility. As a result, tutors may need to invest time in raising awareness and building a foundational understanding of accessibility principles before delving into more technical aspects.

Combination HTML/CSS/JS semantics to implement web accessibility principles:

- Using HTML semantic elements appropriately is a fundamental aspect of improving web accessibility. Semantic HTML elements provide meaningful structure and context to web content, making it easier for assistive technologies and users with disabilities to understand and navigate the content.
- Using CSS semantic techniques improves web accessibility by enhancing the visual presentation and usability of web content.
- Using JavaScript semantic techniques enhances web accessibility by adding interactivity, dynamic behaviour, and improved user experiences for all users, including those with disabilities.

Table 2. Guidelines to using HTML/CSS/JS semantics to implement web accessibility principles (POUR).

Web Accessibility Principle	HTML Semantic Elements	CSS Semantic Techniques	JS Semantic Techniques
Perceivable This involves providing alternatives for non-text content, such as images and multimedia, so that users with visual or cognitive disabilities can understand them. It also includes using clear and consistent navigation and layout structures.	**Use semantic elements for multimedia:** + Employ <figure> and <figcaption> elements to associate captions with images and multimedia content. + Add <alt> attributes to images to provide alternative text descriptions for users who cannot see the visuals. **Structure content with sectioning elements:** + Utilise <header>, <main>, <section>, <article>, and <footer> elements to structure content logically. + Group related content within <section> and <article> elements, and include relevant headings for each section. **Offer text alternatives for audio and video:** + Use the <audio> and <video> elements for multimedia content and provide captions, transcripts, and audio descriptions for users who may have difficulties perceiving the content.	**Responsive design:** + Implement responsive CSS techniques to ensure that web content adapts to different screen sizes and devices, providing an accessible experience across various platforms. + <meta viewport=""> and CSS Grid or Flexbox for laying out webpage structure. **Use CSS Grid and Flexbox:** + Employ modern CSS layout techniques such as CSS Grid and Flexbox to create more accessible and flexible layouts that adapt to various screen sizes. **Optimise for print:** + Implement print CSS styles to ensure that printed versions of web pages are accessible, legible and well structured.	**Toggle buttons and ARIA checkbox/radio:** + Use JavaScript to create toggle buttons and custom checkboxes/radio buttons that convey their state to screen readers using ARIA attributes. **Progressive enhancement:** + Use JavaScript semantic techniques to progressively enhance the user experience, ensuring that the core functionality of the website remains accessible even without JavaScript. **Accessible dropdowns and menus:** + Use JavaScript to create accessible dropdown menus and expandable/collapsible sections using ARIA attributes and keyboard support.
Operable This includes ensuring that all functionality is available via keyboard navigation, as some users may not be able to use a mouse. Time-sensitive content or animations should also be adjustable or avoid causing confusion.	**Implement landmark roles with ARIA:** + Use ARIA landmark roles (role="banner", role="navigation", role="main", role="complementary", role="contentinfo", etc.) to indicate important sections of the page for assistive technology users. **Enhance link navigation:** + Include descriptive text within anchor tags using the title attribute or, preferably, the ARIA-label attribute. + Avoid using generic phrases such as 'click here' or 'read more' for links, as they may not provide enough context for screen-reader users.	**Focus states and keyboard navigation:** + Use CSS to style focus states for interactive elements, such as links and form fields. This makes it easier for keyboard users to identify their current position on the page. + Test keyboard navigation to ensure that users can access all interactive elements and content without the need for a mouse. **Avoid animation overload:** + Use CSS animations judiciously and provide options for users to control or disable animations if necessary, as they can be distracting for some users.	**Keyboard accessibility:** + Ensure that all interactive elements and components are fully accessible via keyboard navigation. + Add keyboard event listeners to elements to handle focus, key presses and actions. + Use appropriate ARIA roles and states to convey the purpose and behaviour of dynamic elements. **Control over animations:** + Provide users with the ability to control or disable animations using JavaScript, as rapid animations may cause issues for some users.
Understandable This involves using clear and simple language, providing instructions and context, and designing consistent and predictable navigation.	**Use semantic elements for data:** + For tabular data, use the <table>, <th>, <tr>, and <td> elements appropriately, and associate headers with data cells using the scope or headers attribute. + For data representation, use <dl> for definition lists and <dt> and <dd> for term-description pairs. **Use heading elements properly:** + Utilise <h1> for the main heading of the page and follow a hierarchical structure for subsequent headings (e.g., <h2>, <h3>, etc.). + Ensure that heading levels are used in sequential order and not skipped, as this helps screen-reader users navigate through the content more efficiently. **Properly label forms:** + Use <label> elements to associate labels with form fields explicitly. + Provide informative <legend> elements for <fieldset> elements in complex forms.	**Use logical and intuitive tab order:** + Order HTML elements in a logical sequence and use CSS to ensure that the tab order follows the visual layout. **Use clear and readable fonts:** + Choose fonts that are legible and easy to read, especially for body text. + Ensure that font sizes are adjustable by users without causing content overlap or loss of functionality. **Contrast and colour:** + Ensure sufficient colour contrast between text and background to make content easily readable for all users, including those with visual impairments. + Avoid using colour alone to convey important information; use text or symbols as alternatives.	**Focus states and styling:** + Implement custom focus styles for interactive elements using JavaScript, ensuring they are clearly visible for keyboard users. + Remove focus styles only when a custom focus indicator has been provided to maintain keyboard accessibility. **Navigation and skip links:** + Use JavaScript to create skip links that allow keyboard users to bypass repetitive content and jump directly to the main content. + Ensure that hidden navigation menus or off-screen content are accessible to screen readers.

Robust This means using valid code, adhering to web standards, and avoiding technologies that might create barriers for users.	**Create accessible buttons:** + Use <button> elements for interactive buttons rather than <div> or <span> elements with click event listeners. + Provide descriptive text within the button to convey its purpose.	**Hidden content and screen readers:** + Use CSS to visually hide non-essential content (e.g., skip links, off-screen menus) while making it available to screen readers using techniques like clip-path, position, or opacity. **Test with users with disabilities:** + Involve users with disabilities in usability testing to gather feedback and ensure the JavaScript enhancements are genuinely accessible.	**Optimise for assistive technologies:** + Test JavaScript-powered components with assistive technologies to ensure they are appropriately announced and interactable. **Error handling and validation:** + Implement real-time form validation using JavaScript to provide immediate feedback to users, including users with disabilities. + Ensure that error messages are accessible and associated with the corresponding form fields.

CONCLUSION AND RECOMMENDATIONS

Web accessibility is a crucial aspect of modern technology, ensuring that websites and web applications are usable and navigable by individuals with disabilities. As the internet continues to play an increasingly significant role in our daily lives, the need for an inclusive digital environment becomes more apparent.

Teaching web accessibility to undergraduate students is not just a technical necessity but a moral imperative. Tutors must be aware of and address the challenges that can arise during the teaching process. Teaching web accessibility to undergraduate students faces up to challenges that stem from four main factors: (1) technical complexities, (2) time constraints, (3) resource constraints, and (4) lack of awareness and understanding. By taking a proactive approach, staying up to date with evolving guidelines, providing hands-on experiences, and emphasising the real-world impact of accessible design, tutors can effectively equip students with the knowledge and skills needed to champion web accessibility in their future careers.

By using HTML semantic elements appropriately, we can create a more accessible and inclusive online experience for all users, including those with disabilities. Additionally, when combined with proper use of ARIA attributes, HTML semantic elements contribute significantly to making websites more user-friendly for individuals who rely on assistive technologies. By using CSS semantic techniques appropriately, we can create an inclusive and accessible design that enhances the user experience for all visitors, including those with disabilities. It is essential to test and verify the accessibility improvements using various tools and by involving users with disabilities in the testing process. By using JavaScript semantic techniques, we can create a more inclusive and accessible web experience that enhances the usability and functionality for all users, regardless of their abilities or assistive technology usage. Always combine JavaScript enhancements with appropriate ARIA roles and attributes to provide meaningful information to assistive technologies and ensure a seamless experience for everyone.

REFERENCES

Abuaddous, H. Y., Jali, M. Z., & Basir, N. (2016). Web accessibility challenges. *International Journal of Advanced Computer Science and Applications*, *7*(10), 172–182. https://doi.org/10.14569/IJACSA.2016.071023

Dang, D. (2022). Recommended topics and practical labs for the introductory Level-5 Website Development Course. In E. Erturk & B. Otinpong (Eds.), *Proceedings of the 13th Annual Conference of Computing and Information Technology Education and Research in New Zealand* (pp. 29–36). CITRENZ.

Web Accessibility Initiative (WAI). (n.d.a). *Accessibility principles*. https://www.w3.org/WAI/fundamentals/accessibility-principles/

Web Accessibility Initiative (WAI). (n.d.b). *Introduction to web accessibility*. https://www.w3.org/WAI/fundamentals/accessibility-intro/

World Wide Web Consortium. (n.d.) *Accessibility*. https://www.w3.org/mission/accessibility/

AUTHOR

Dr Daniel Dang is an ICT Lecturer in the School of Computing at Eastern Institute of Technology Te Aho a Māui. His research interests are sharing the best practices in teaching ICT subjects and using digital technologies to tackle social and environmental issues. ddang@eit.ac.nz

Predicting Learning Outcomes in an Online Learning Platform

Ying Bai, Michael Bosu and Diab Abuaiadah

https://doi.org/10.34074/proc.240111
Correspondence: Michael Bosu, michael.bosu@wintec.ac.nz

This publication may be cited as:

Bai, Y., Bosu, M., & Abuaiadah, D. (2024). Predicting Learning Outcomes in an Online Learning Platform. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 78–90). ePress, Unitec. https://doi.org/10.34074/proc.240111

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga –
New Zealand Institute of Skills and Technology

ABSTRACT

This paper carried out a comprehensive analysis of a dataset collected from the Open University's online learning platform, with the aim of understanding the relationships between various factors and the outcomes for learners. Data pre-processing and analysis showed that factors such as age, gender, different course modules, educational background, IMD band and total clicks positively correlate with learners' performance. The k-means clustering algorithm was utilised to identify distinct learning behaviours among learners by grouping them into three clusters. The random forest algorithm was then used to build machine learning models based on the identified learning behaviours, achieving a higher prediction accuracy of 86.1%. The findings emphasise the importance of targeted interventions and support tailored to the specific needs of different learner groups. The contribution of this paper is that it is the first to use the k-means clustering algorithm to divide the data into groups prior to using the random forest algorithm to predict the final outcomes for learners at the Open University. Furthermore, this is the first study to apply a random forest algorithm to the Open University's online learning platform dataset, with commendable results in predicting the outcomes for learners.

KEYWORDS

E-learning, machine learning, personalised content for online learners, k-means clustering, random forest algorithm

INTRODUCTION

In recent years, there has been a significant shift towards online education, particularly during the global Covid-19 pandemic, when it became the predominant access to education for many learners globally. Maatuk et al. (2021) conducted a study to investigate the perspectives of students and faculty tutors regarding the challenges and issues associated with remote learning during this period. In addition, their study found that students perceive online learning platforms as beneficial to their educational journey. This is primarily because e-learning platforms offer the advantage of accessing courses and classrooms at any time and from anywhere. On the other hand, it becomes crucial for providers of online training platforms to offer personalised learning pathways to accommodate learners' cognitive abilities, knowledge structures and study paces (Maatuk et al., 2021). Online education platforms are web-based systems that provide users with access to educational materials and enable interaction with instructors and fellow learners. These platforms have gained popularity over the past decade due to their flexibility and convenience for acquiring new skills and knowledge.

Various online learning platforms, such as massive open online courses (MOOCs), learning management systems (LMSs), adaptive learning platforms, language learning platforms and skills training platforms, have effectively leveraged data mining techniques to develop automatic grading and recommendation systems. By utilising intelligent algorithms, these platforms gather valuable user information, including the frequency of platform usage, accuracy in answering questions, and time spent engaging with learning materials (Despotović-Zrakić et al., 2012). Aher et al. (2013) proposed that machine learning (ML) methods, including collaborative filtering, content-based filtering, decision trees and artificial neural networks, can be employed to process and analyse the acquired information on e-learning platforms.

ML involves learning from data, identifying patterns and making predictions. The rise of data availability and advancements in cloud computing have contributed to the rapid growth of ML, enabling efficient analysis of complex data and mitigation of unforeseen risks (Aher et al., 2013). Although online education offers flexibility and affordability compared to traditional on-campus learning, it does present challenges due to reduced interaction between learners and instructors (Eom & Ashill, 2016). To address this, long-term log data from online platforms can be leveraged for learner and course assessment. ML algorithms can assist in predicting students at risk and estimating dropout rates through analysis of pre-processed log data (Essalmi et al., 2015).

The Felder and Silverman Learning Styles Model (FSLSM) is a well-recognised framework for identifying an individual's learning style. It assesses learners across four dimensions: information processing, input, understanding and perception (Nafea et al., 2019). In the context of online learning, Moodle, an online learning platform, utilises the FSLSM to offer adaptive course delivery (Essalmi et al., 2015). Despotović-Zrakić et al. (2012) conducted a study focusing on data mining techniques to classify students into clusters based on the FSLSM. Although the research primarily focused on learning styles, the results have been applied in Moodle LMSs to enhance personalised strategies.

ML techniques such as k-means clustering and the random forest algorithm can be employed to develop a model that leverages learners' profiles and online operation logs to refine personalisation strategies on a specific online learning platform. It is important to acknowledge that personalisation strategies can also be influenced by other factors, such as prior knowledge and expectations (Despotović-Zrakić et al., 2012).

The objectives of this paper are to:

- Identify distinct learning behaviours or learning styles among learners in an online learning platform using k-means clustering.
- Divide the learners into groups based on the identified learning behaviours.
- Predict learner outcomes using the random forest algorithm.

LITERATURE REVIEW

Various performance metrics are used to evaluate the effectiveness of online education platforms. These metrics include course completion rate, student engagement, satisfaction and learning outcomes (Essalmi et al., 2015). ML models can leverage these metrics and learner data to make predictions and provide personalised content and feedback to learners (Eom & Ashill, 2016). Eom and Ashill (2016) conducted an empirical investigation and identified six independent variables that influence student outcomes, including course structure, instructor feedback, self-motivation, learning style, interaction and instructor facilitation. Among these variables, instructor feedback and learning style were found to significantly impact learning outcomes, while user satisfaction could predict outcomes (Eom & Ashill, 2016). Instructor feedback encompasses cognitive, diagnostic and prescriptive feedback delivered through various channels (Eom & Ashill, 2016).

Learning style refers to an individual's preferred method of acquiring and processing information, influenced by physiological dimensions, cognitive dimensions and personality characteristics. ML algorithms can automatically detect learning styles based on learners' sequences extracted from e-learning system log files, enabling the provision of personalised content (El Aissaoui et al., 2019). Essalmi et al. (2015) found that a single personalisation strategy may not fit all courses and teachers, leading to the investigation of a generalised approach that considers learners' profiles and appropriate personalisation parameters (Essalmi et al., 2015). Additionally, Aher et al. (2013) demonstrated the practicality of a combined algorithm, using simple k-means clustering and apriori-association rule algorithm, for building a course recommendation system. Khanal et al. (2020) reviewed 101 recommendation-system-related papers and found that multiple algorithms are often used in a single system, making classification of ML algorithms challenging (Khanal et al., 2020).

Hyper-parameter optimisation is crucial for improving the performance of trained ML models. Different techniques exist for hyper-parameter tuning, and their selection depends on specific scenarios and models (Yang & Shami, 2020). Identifying key hyper-parameters requires data mining and analysis expertise, and choosing the appropriate tuning technique for a given model or dataset is important (Yang & Shami, 2020). Automating hyper-parameter tuning through optimal configuration improves reproducibility and model performance, and reduces human effort (Yang & Shami, 2020). The implementation of a hyper-parameter tuning approach for the ML models built in this research will be discussed in the sections below.

DATASET DESCRIPTION AND ANALYSIS

The dataset used in this study was sourced from the Open University's (OU) online learning platform. The OU is a British public university that has students who are mainly off campus and study online from everywhere in the United Kingdom (Kuzilek et al., 2017). Figure 1 shows the structure of the dataset, including course content in the courses table, assignment marks in the assessments table, student demographics (such as location, age group, disability, education level and gender) in the studentInfo table, forum discussion in the vle table, students' interactions with the virtual learning environment (VLE) in the studentVle table and students' registration information in the studentRegistration table.

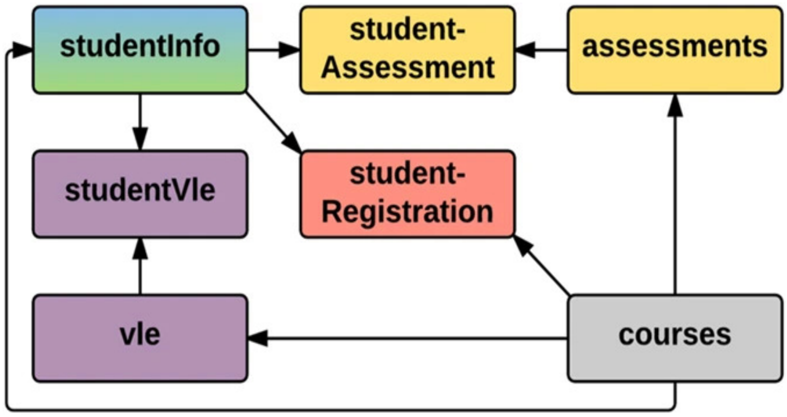


Figure 1. Data structure of Open University online learning platform dataset. Source: Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). https://creativecommons.org/licenses/by/4.0/

The final data used as an input for the ML models in the 'Model Development' section is obtained by combining data from studentInfo table, studentRegistration table, studentAssessment table, assessments table, courses table and studentVle table. The final data used in building the models comprises 24,845 records of 22,488 students and 22 course modules.

Figure 2 presents the relationship between parameters and students' final results. Parameters include age band, education background, the Index of Multiple Deprivations (IMD) band and total clicks, which refers to the cumulative number of clicks made by a student within the Open University (OU) online platform. The Index of Multiple Deprivations (IMD) is a measure used in the United Kingdom to assess relative deprivation at the small-area level. It combines various indicators across multiple domains, such as income, employment, education, health, crime and living environment, to provide an overall measure of deprivation (Noble et al., 2006).

Preliminary analysis of the data as demonstrated by Figure 2 indicates that older learners tend to perform better than younger learners. In fact, around 40% of learners older than 55 years received a distinction in the final results, as shown in Figure 2(a). Additionally, Figure 2(b) illustrates that learners with higher education qualifications have a higher chance of passing a course and achieving distinction performance when compared to learners with lower qualifications. Figure 2(c) indicates that learners from high IMD areas have a higher chance of achieving better performance than learners from low IMD areas. Lastly, Figure 2(d) indicates that frequent interaction with the OU online platform has a positive impact on final results. Learners with a total of over 1000 clicks tend to pass courses and perform better overall.

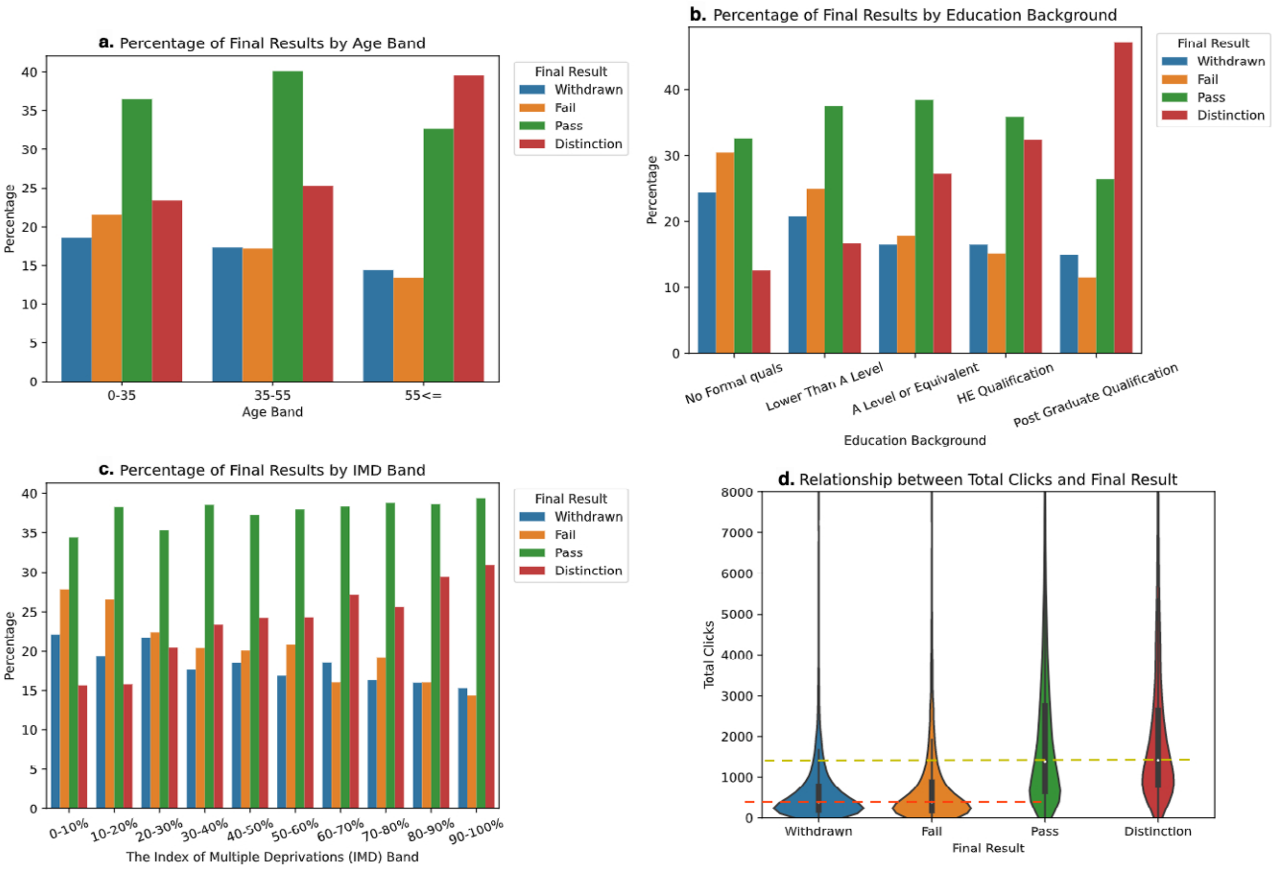


Figure 2. Preliminary results of data analysis.

Figure 3 shows a positive correlation between six factors and the final results, because their values in the following heat map are more significant than 0. These factors include IMD band (0.14), age band (0.04), education background (0.17), course modules (0.022), gender (0.049) and total clicks (0.34). The parameter with the strongest positive correlation with final results is total clicks, which has a value of 0.34, the highest among the other parameters.

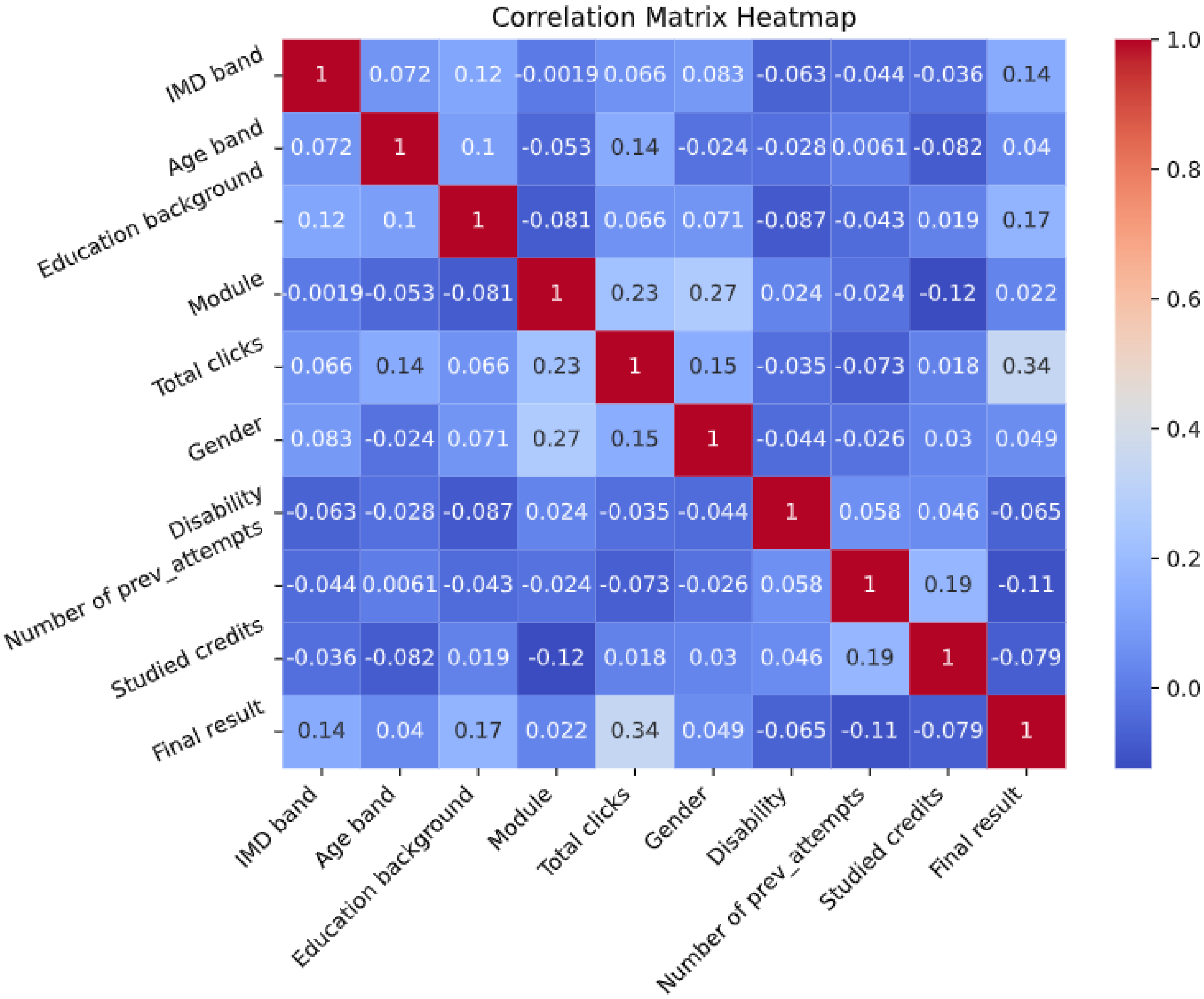


Figure 3. Correlation of age, education background and IMD.

The dataset analysis found that six factors are closely related to the learners' performance in the final results. These factors include age, educational background, IMD in the location, course, gender and frequency of interaction with the Open University online learning platform.

MODEL DEVELOPMENT

Based on the dataset analysis in the previous section, age (this refers to the "age_band" column in the combined dataset), educational background (this refers to the "highest_education" column in the combined dataset), IMD in the location (this refers to the "imd_band" column in the combined dataset), course modules (this refers to the "code_presentation" column in the combined dataset), a learner's gender (this refers to the "gender" column in the combined dataset) and frequency of interaction with the OU online learning platform (this refers to the "total_click" column in the combined dataset) were used as parameters to build ML models to predict learners' final results.

There are two fundamental types of ML algorithms: supervised and unsupervised. K-means clustering is an unsupervised algorithm to group learners based on shared characteristics or patterns. By identifying clusters of learners with similar attributes, k-means clustering helps to reveal the underlying structures within the data (Mahesh, 2020).

This section uses k-means clustering to identify distinct groups in the combined dataset. Learners with similar characteristics are grouped based on factors, including age, educational background, IMD in the location, course modules and frequency of interaction with the OU online learning platform. Then the random forest algorithm is utilised to build an ML model. The random forest offers a flexible and robust ensemble of decision trees, making it suitable for analysing complex datasets and capturing intricate relationships between variables (Mahesh, 2020). By leveraging the power of random forest, this model aims to provide predictions of learners' final results while considering a wider range of factors.

Utilising the K-Means Clustering Methodology to Classify Learners

To determine the appropriate number of clusters when utilising k-means clustering, it is essential to address the optimal value denoted as 'k'. The elbow method is a commonly employed means for identifying this optimal k value (Marutho et al., 2018). The elbow method is based on the observation that the within-cluster sum of squares (WCSS) tends to decrease as the number of clusters increases. The WCSS calculates the sum of squared distances between each data point and the centroid of its assigned cluster (Marutho et al., 2018). The quality evaluation of a clustering algorithm, particularly in k-means clustering, employs the sum of squared error (SSE) metric. The SSE quantifies the variation or dispersion within the clusters. By minimising the SSE, the algorithm aims to create compact and well-separated clusters (Bholowalia & Kumar, 2014).

By employing the elbow method to determine the optimal k value, a plot is generated with the number of clusters (k) on the x-axis and the corresponding SSE values on the y-axis. Typically, the plot exhibits a downward trend, where the SSE decreases as k increases. However, there is a certain point beyond which increasing k does not significantly reduce the SSE (Bholowalia & Kumar, 2014). The 'elbow' in the plot represents the point where the SSE starts to level off, forming a bend similar to an elbow. This bend indicates a diminishing return in SSE reduction beyond that point. The optimal k value is frequently selected at the elbow as it signifies a suitable trade-off between capturing meaningful patterns within the data and avoiding excessive complexity. Figure 4 illustrates SSE values on the y-axis and the number of clusters on the x-axis. It is evident that the optimal k value is 3, as indicated by the elbow point in the plot.

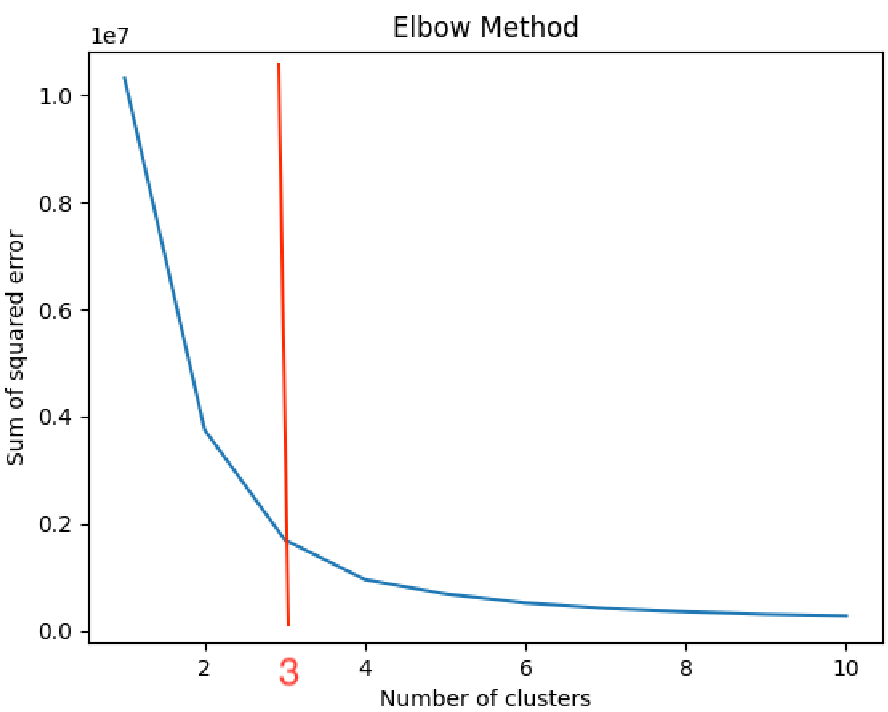


Figure 4. Using the elbow method to the find optimal k value.

Based on the implementation of the elbow method, it was determined that the optimal number of clusters (k) is 3. As a result, the learners in the combined dataset have been divided into three distinct groups, as shown in Figure 6. In Figure 5, for k = 3, in Group 1, the group size is 5409, and the overall performance is 'Pass'. The size of Group 2 is 1066 and the overall performance is also 'Pass' whilst the size of Group 3 is 18,370 and the performance comprises of 'Fail' or 'Withdrawn' or 'Pass'. Also, the group sizes and the overall performances for Groups 2 and 3 are 1066, 18,370 and 'Pass', and 'Fail or Withdrawn or Pass'.

a. Learner's performance index

Total Score	Learner's performance
below 4000	Fail or Withdrawn
4000-8000	Pass
8000 and above	Distinction

b. K=3

Group #	Group size	Overall learner's performance
1	5409	Pass
2	1066	Pass
3	18370	Fail or Withdrawn or Pass

Figure 5. Group sizes and types.

Figure 6 provides insights into the performance of learners in three different groups. It reveals that learners in Group 1 and Group 2 have a higher likelihood of successfully passing their courses, with Group 2 showing promising results. However, a significant proportion of learners in Group 3 are at a higher risk of withdrawing or failing their courses.

These findings highlight the existence of distinct learner profiles within the dataset, indicating the need for targeted interventions and support tailored to each group's specific needs and challenges. For example, to address the higher withdrawal and failure rate observed in Group 3, personalised support should be provided by staff of the OU online learning platform to assist and motivate learners in this group. By recognising and addressing the unique requirements of each group, educational institutions can enhance learner outcomes and improve overall success rates.

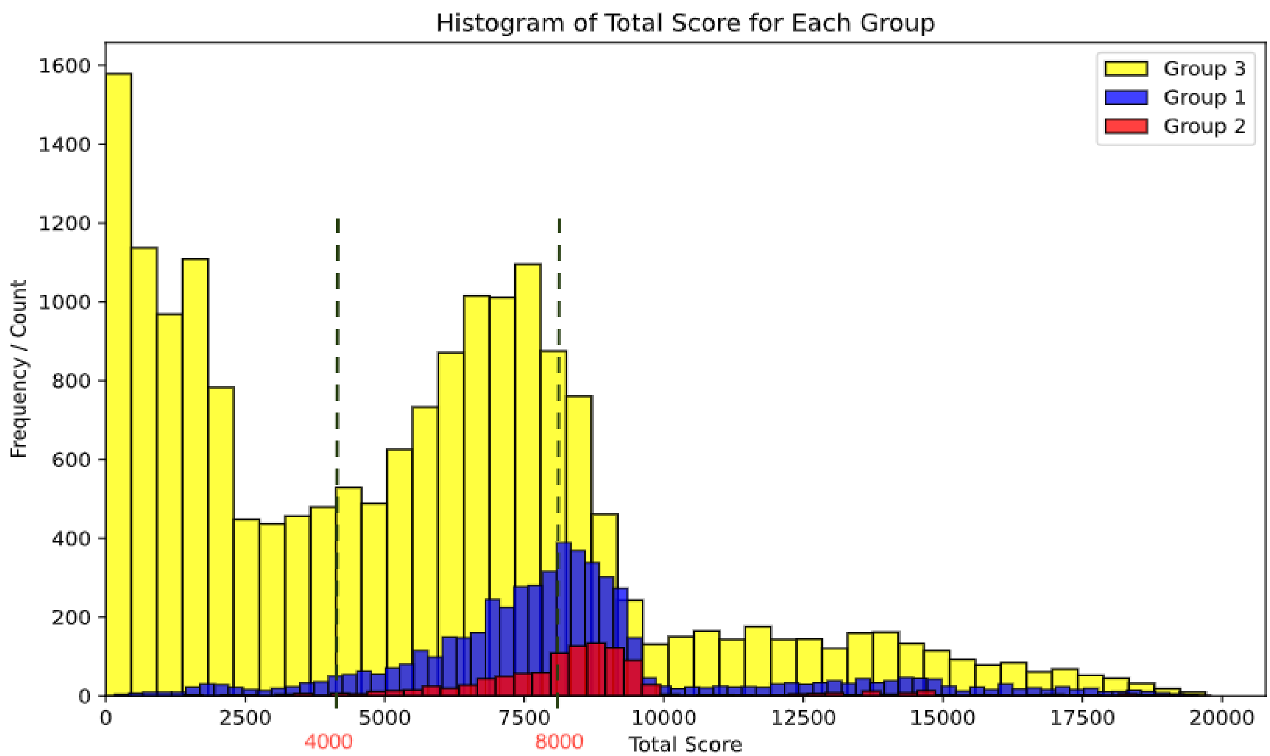


Figure 6. Dividing learners into three groups.

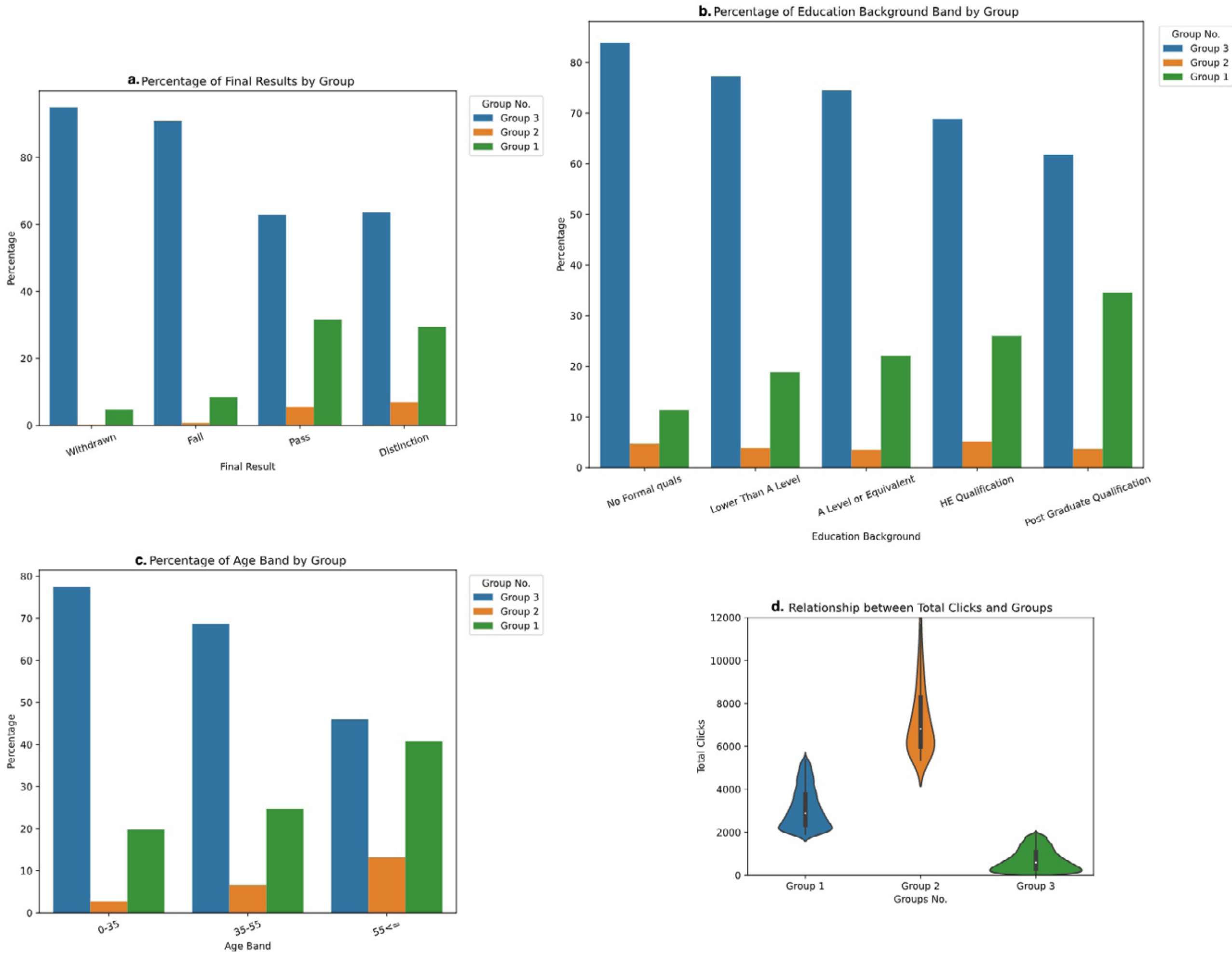


Figure 7. Comparison of three groups.

Figure 7 illustrates that almost all learners in Group 2 pass courses, and the learners in Group 1 have relatively high possibility of passing courses compared with learners in Group 3. Figures 7b, 7c and 7d show that most learners in Group 1 have A Level or higher education qualifications, are usually more than 55 years old, and have moderate interaction frequency with the online learning platform. In contrast, most learners in Group 3 have lower than A Level qualifications, are less than 35 years old, and have the lowest interaction frequency with the online learning platform.

Utilising Random Forest to Build a Machine Learning (ML) Model

In the previous subsection, learners were divided into three groups using k-means clustering, which means learners in the combined dataset have distinct characteristics, and by applying k-means clustering, learners with similar characteristics are grouped together. In this subsection, the random forest algorithm is utilised to build an ML model to predict learners' outcomes (final results) based on their learning characteristics identified through k-means clustering in the previous subsection.

The random forest algorithm is a classification and regression algorithm to build a forest containing many individual decision trees. It generates reasonable predictions across various parameters while requiring little configuration. Moreover, when processing extensive data, several missing pieces of data will not significantly affect its accuracy (Boulesteix et al., 2012).

The parameters used in random forest model training include two main parts. The first part is the relevant factors found in the section 'Dataset Description and Analysis': age, educational background, IMD in the location, course, gender, and frequency of interaction with the OU online learning platform. The second part is the factors that correlate less with learners' outcomes, including disability, number of previous attempts at a course, and learners' studied credits according to the heatmap in Figure 3.

The table of courses in the datasets contains seven different programmes. The domains of these programmes are social science and STEM (science, technology, engineering and mathematics). The assessment table shows three assessment types: tutor-marked assessment, computer-marked assessment and final examination (Kuzilek et al., 2017).

Figure 8 illustrates the visualisation of a decision tree generated by the random forest algorithm in the context of creating the ML model. The decision tree begins with a topmost node, which is split based on the most significant feature. In this particular tree, the topmost node represents the number of interactions a learner has had with the OU online learning platform (total_click). The internal nodes of the tree correspond to binary decisions made based on specific features. These decisions guide the data flow down the tree, ultimately leading to the terminal nodes. The terminal nodes represent the predicted outcomes or classes, such as "Pass" or "Fail", based on the given features and their associated paths within the decision tree. Implementation of the k-means clustering and the random forest was conducted in Jupyter Notebook in the Python 3. X programming environment, and the decision tree in Figure 8 is virtualised in Graphviz Online (https://dreampuf.github.io/GraphvizOnline).

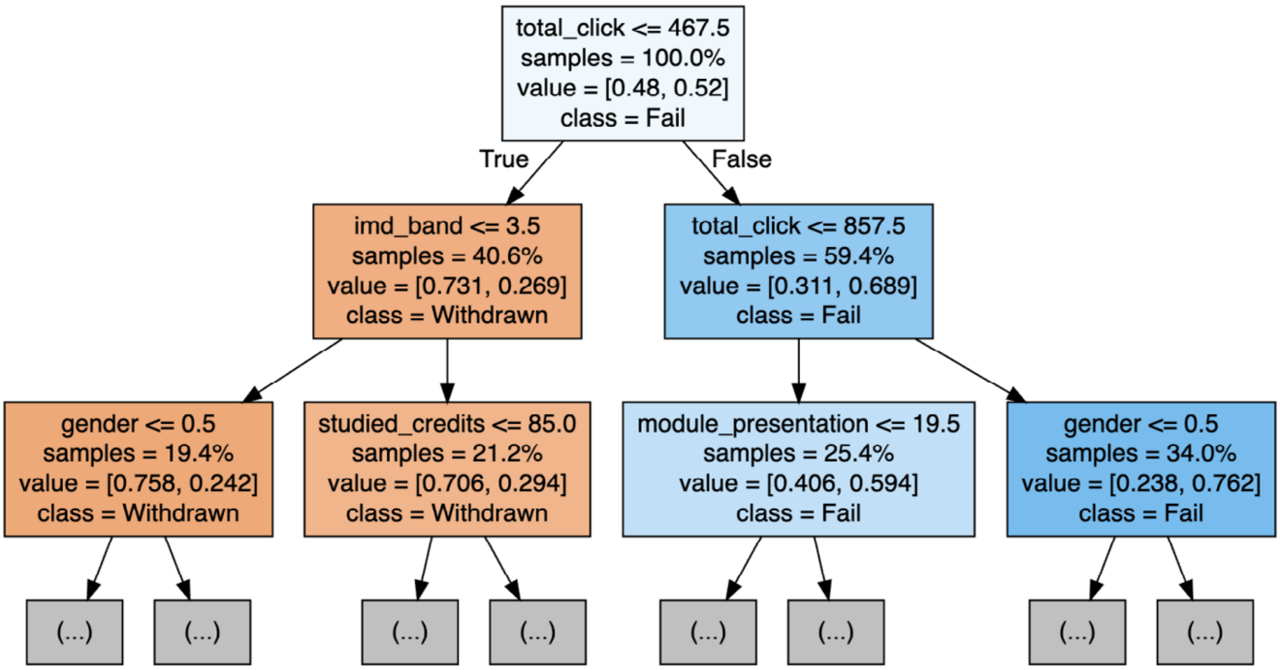


Figure 8. An example of decision trees.

Based on identifying distinct learners' characteristics through k-means clustering, as previously described, the original combined dataset was divided into three individual datasets. Using the random forest algorithm, these datasets were used to build three separate ML models (Model 1, Model 2, Model 3). Among the three models, Model 2, which is based on learners in Group 2, achieved the highest prediction accuracy of 93%, indicating that this group exhibits more consistent and distinct learning characteristics.

On the other hand, Model 3, which is based on learners in Group 3, demonstrated lower prediction accuracy than the other two models. The findings suggest that learners in Group 3 exhibit a more comprehensive range of learner characteristics, making it more challenging to accurately predict their outcomes.

Model 4 was built using the original combined dataset without applying k-means clustering. It achieved a prediction accuracy of 79%, indicating that the random forest algorithm can effectively classify datasets even without the explicit grouping provided by k-means clustering.

Model #	Data size (rows)	Train data size (rows) 80%	Test data size (rows) 20%	Prediction accuracy
1	5409	4327	1082	89%
2	1066	852	214	93%
3	18370	14696	3674	76.3%
4	24845	19876	4969	79%

Figure 9. Results of ML models built by the random forest algorithm.

K-means clustering and random forest can be used for dataset classification. However, k-means clustering is an unsupervised algorithm that does not require predefined labels or target variables. It aims to identify natural groupings or clusters within the data based on similarities in features or variables. While k-means clustering can be useful for exploratory analysis and understanding data patterns, it does not directly predict a target variable (Mahesh, 2020). On the other hand, random forest is a supervised algorithm that requires predefined labels and a designated target variable. It constructs an ensemble of decision trees, where each tree is trained on a subset of the data with random feature selection (Boulesteix et al., 2012). Random forest can predict the target variable based on the learned patterns and relationships in the training data, as shown in Figure 9.

Therefore, in this specific scenario, k-means clustering is more suitable for discovering underlying patterns and grouping similar learners. At the same time, random forest is a supervised algorithm that enables the prediction and classification of target variables (final results). Combining k-means clustering with the random forest algorithm can improve prediction accuracy. In Figure 9, the average accuracy of Models 1, 2 and 3 is 86.1%, which is higher than that of Model 4, 79%, built by the original combined data without applying k-means clustering. This indicates that combining k-means clustering with random forest can efficiently improve the ML models' accuracy.

CONCLUSION

This study focused on collecting, pre-processing and analysing Open University's online learning platform dataset. Applying ML algorithms showed that age, gender, educational background, IMD band, and frequency of interaction with the OU online learning platform (total clicks) positively correlated with learners' outcomes. The k-means clustering algorithm was employed to identify distinct learning behaviours among learners, forming them into three groups. The random forest algorithm was then used to build ML models based on learner groups with the identified learning behaviours. Comparisons between the models built with and without k-means clustering showed the effectiveness of clustering. The results show that the accuracy of the model without applying the k-means clustering algorithm is 79%, whereas the accuracy of the model with the application of the k-means clustering algorithm reached 86.1%. To the best of our knowledge, this study is the first to use the k-means clustering algorithm in pre-processing the Open University online learning platform dataset, and the random forest algorithm to predict the outcomes for learners.

FUTURE WORK

The next step is finding an approach to automatically set hyperparameters, learning from data instead of manually setting them; and executing hyperparameter optimisation techniques, such as Bayesian optimisation hyperband (BOHB), genetic algorithms (GA) and particle swarm optimisation (PSO), to improve the performance of the ML models built as described above. Therefore, future work could explore additional ML techniques, consider more features in the analysis, and evaluate the effectiveness of personalised interventions in improving learner outcomes.

REFERENCES

Aher, S. B., & Lobo, L. M. R. J. (2013). Combination of machine learning algorithms for recommendation of courses in e-learning system based on historical data. *Knowledge Based Systems*, *51*, 1–14. https://doi.org/10.1016/j.knosys.2013.04.015

Bholowalia, P., & Kumar, A. (2014). EBK-means: A clustering technique based on elbow method and k-means in WSN. *International Journal of Computer Applications*, *105*(9), 975–8887. https://www.ijcaonline.org/archives/volume105/number9/18405-9674/

Boulesteix, A-L., Janitza, S., Kruppa, J., & König, I. R. (2012). Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *2*(6), 493–507. https://doi.org/10.1002/widm.1072

Despotović-Zrakić, M., Markovic, A., Bogdanović, Z., Barać, D., & Krco, S. (2012). Providing adaptivity in Moodle LMS courses. *Educational Technology & Society*, *15*(1), 326–338. https://www.researchgate.net/publication/267846725_Providing_Adaptivity_in_Moodle_LMS_Courses

El Aissaoui, O., El Madani El Alami, Y., Oughdir, L., & El Allioui, Y. (2019). A hybrid machine learning approach to predict learning styles in adaptive E-learning system. In *Advanced Intelligent Systems for Sustainable Development (AI2SD'2018) Volume 5: Advanced Intelligent Systems for Computing Sciences* (pp. 772–786). Springer International Publishing.

Eom, S. B., & Ashill, N. (2016). The determinants of students' perceived learning outcomes and satisfaction in university online education: An update*. *Decision Sciences Journal of Innovative Education*, *14*(2), 185–215. https://doi.org/10.1111/dsji.12097

Essalmi, F., Ayed, L. J. B., Jemni, M., Graf, S., & Kinshuk (2015). Generalized metrics for the analysis of E-learning personalization strategies. *Computers in Human Behavior*, *48*, 310–322. https://doi.org/10.1016/j.chb.2014.12.050

Khanal, S. S., Prasad, P. W. C., Alsadoon, A., & Maag, A. (2020). A systematic review: Machine learning based recommendation systems for e-learning. *Education and Information Technologies*, *25*(4), 2635–2664. https://doi.org/10.1007/s10639-019-10063-9

Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open University Learning Analytics dataset. *Scientific Data*, *4*, 170171. https://doi.org/10.1038/sdata.2017.171

Maatuk, A. M., Elberkawi, E. K., Aljawarneh, S., Rashaideh, H., & Alharbi, H. (2021). The COVID-19 pandemic and e-learning: Challenges and opportunities from the perspective of students and instructors. *Journal of Computing in Higher Education*, *34*(1). https://doi.org/10.1007/s12528-021-09274-2

Mahesh, B. (2020). Machine learning algorithms: A review. *International Journal of Science and Research*, *9*(1), 381–386. https://www.ijsr.net/archive/v9i1/ART20203995.pdf

Marutho, D., Handaka, S. H., & Wijaya, E. (2018). The determination of cluster number at k-mean using elbow method and purity evaluation on headline news. In *2018 International Seminar on Application for Technology of Information and Communication*, 533–538. IEEE. https://doi.org/10.1109/isemantic.2018.8549751

Nafea, S. M., Siewe, F., & He, Y. (2019). On recommendation of learning objects using Felder-Silverman Learning Style Model. *IEEE Access*, *7*. https://doi.org/10.1109/access.2019.2935417

Noble, M., Wright, G., Smith, G., & Dibben, C. (2006). Measuring multiple deprivation at the small-area level. *Environment and Planning A*, *38*(1), 169–185. https://doi.org/10.1068/a37168

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, *415*, 295–316. https://doi.org/10.1016/j.neucom.2020.07.061

AUTHORS

Ying Bai is a Master of Applied Information Technology student at the Centre for Information Technology, Waikato Institute of Technology. Her main research interests are machine learning and cloud computing. yinbai10@student.wintec.ac.nz

Dr Michael Bosu is a Principal Lecturer at the Centre for Information Technology, Waikato Institute of Technology. His main research interests include data quality in empirical software engineering, software effort estimation, machine learning and mining of application reviews. michael.bosu@wintec.ac.nz

Dr Diab Abuaiadah is a Principal Lecturer at the Centre for Information Technology, Waikato Institute of Technology. His main research interests include development of clustering and classification algorithms, machine learning and software testing. diab.abuaiadah@wintec.ac.nz

Embedding Mātauranga Māori in Computing Courses: A Case Study

Abha Chitalia, Simon Dacey, Gerard Lovell, Lei Song and Bashar Barmada

https://doi.org/10.34074/proc.240112
Correspondence: achitalia@unitec.ac.nz

This publication may be cited as:

Chitalia, A., Dacey, S., Lovell, G., Song, L., & Barmada, B. (2024). Embedding Mātauranga Māori in Computing Courses: A Case Study. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 91–97). ePress, Unitec. https://doi.org/10.34074/proc.240112

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga –
New Zealand Institute of Skills and Technology

ABSTRACT

This paper discusses how to embed mātauranga Māori in computing courses to increase the awareness of ākonga (students) regarding Māori beliefs, language and practices. The study has been driven by the low scoring of the statement regarding embedding Māori beliefs, language and practices in the course survey conducted by Unitec for the School of Computing, Electrical and Applied Technology.

The paper considers different methods provided by Unitec to embed mātauranga Māori in teaching and learning, such Te Noho Kotahitanga, I See Me initiatives, and others. The paper then looks at how to apply these methods in computing through a case study of one of the courses in the New Zealand Diploma in Information Systems (Level 5) (NZDIS). The case study discusses the approaches that are currently used, then it moves to how these approaches can be improved, as well as applying new ones. The aim is to apply these proposed approaches through Semester 2 2023, to get the results in the course survey by the end of that semester. The expectation is that the applied approaches will increase the awareness of ākonga of how they feel that Māori beliefs, language and practices are embedded throughout their computing courses. The proposed approaches here can be considered as a reference to embed mātauranga Māori in other IT courses.

KEYWORDS

Mātauranga Māori, Māori beliefs and practices, computing courses

INTRODUCTION

Unitec takes its obligations under Te Tiriti o Waitangi seriously. In 2001 it developed a partnership document called Te Noho Kotahitanga (Unitec | Te Pūkenga, 2023c), which itemises shared values and encourages all kaimahi (staff) to follow these both with fellow kaimahi and with ākonga (students). It aims to see more mātauranga Māori embedded into courses and teaching and learning. Kaiako (teachers) on the New Zealand Diploma in Information Systems (Level 5) (NZDIS) were therefore concerned when, in the institutional student course survey, ākonga rated the statement "I felt that Māori beliefs, language and practices were embedded throughout my course" the lowest among other statements. In this paper, we will look at methods that are provided by Unitec to assist kaiako to understand how Unitec embeds mātauranga Māori in different aspects of teaching and learning. These include Te Noho Kotahitanga, kaihautū (Māori Academic Advisor), I See Me initiatives, and others. We will then consider one course in the NZDIS, namely, HTCS5607 IS Application Project, and look at what we are doing and how that can be improved to embed more mātauranga Māori. Our aim is to implement the improvements in this course in Semester 2 2023 and compare the rating. If the rating has improved, then the proposed work on the HTCS5607 IS Application Project course can be used as a reference to embed mātauranga Māori in other IT courses.

Unitec conducts a survey of each course each semester. In the Course Content section of the survey, there are seven statements and ākonga need to rate each as "Don't know" or on a scale of 1 (poor) to 10 (excellent). For the NZDIS programme in 2022, the average rating for the Course Content section was 7.8 out of 10. The lowest rating was 6.5 for the statement "I felt that Māori beliefs, language and practices were embedded throughout my course" (Unitec | Te Pūkenga, 2023a). A total of 35% of the responders chose "Don't know". This compares with the Unitec average of 7.6 out of 10 for this statement. For the HTCS5607 IS Application Project course, the referred statement scored 7.1

out of 10, while the average was 8.2, and 36% of the responders chose "Don't know". These statistics indicate that many of our ākonga are not aware of how Māori beliefs, language and practices are embedded in teaching and learning for NZDIS courses.

UNITEC METHODS TO EMBED MĀTAURANGA MĀORI

It is a challenge to apply Māori language in IT courses in general (Te Momo, 2022). However, Unitec has a strong desire to embed mātauranga Māori in teaching and learning, and to provide a culturally safe learning environment for all ākonga. Unitec has introduced many methods in support of this desire, which are discussed below.

Te Noho Kotahitanga (TNK) Framework

At Unitec, the teaching and learning principles are guided by Te Noho Kotahitanga framework (Unitec New Zealand, 2020). Through this framework, Unitec expresses its commitment to te Tiriti o Waitangi (Unitec | Te Pūkenga, 2023d). This framework consists of five values, which are as follows:

- **Rangatiratanga:** The authority and responsibility for all teaching and learning related to Māori dimensions of knowledge lie with Māori.
- **Wakaritenga:** Each stakeholder has a legitimate right to be here, to speak freely in their language and put its resources into use that will benefit all.
- **Kaitiakitanga:** Unitec takes responsibility as a critical guardian of knowledge.
- **Mahi Kotahitanga:** The spirit of generosity and co-operation will guide all the actions.
- **Ngākau Māhaki:** The heritage and customs, current needs and future aspirations of each stakeholder are valued and respected.

All kaimahi are expected to incorporate these values into their roles. Schools at Unitec have the assistance of a kaihautū, assigned to them to integrate these values into programme philosophy, content, teaching, learning and assessment.

I See Me

I See Me are Unitec's initiatives (Unitec | Te Pūkenga, 2023b) to make all ākonga feel welcome, supported and part of the Unitec whānau. The focus areas of these initiatives are in line with Te Noho Kotahitanga to improve Māori success rate.

Focus Area One places emphasis on the concept of whakawhanaungatanga – the process of establishing meaningful relationships. Ākonga are invited to attend pōwhiri when they are new to Unitec and when they return for second semester. Whānau evenings and tuākana–tēina mentoring support (Peer Assisted Study Session – PASS) are other important strategies suggested under this initiative.

Focus Area Three has two aspects. The first is embedding elements of Māori content into the courses so that ākonga see their culture reflected in the curriculum. Courses may include Māori equivalent words in the course descriptors and in the teaching materials. A glossary of digital technology terms and their equivalent in te reo Māori is provided by Maclennan (2023), while a glossary of general Māori words can be found in *Te Ara* (2023). Another way to embed Māori content is through introducing topics that are directly related to Māori culture and beliefs. For example, Māori data sovereignty, application of te Tiriti o Waitangi in the IT sector, Code of Ethics and professional practices, legal and privacy requirements in IT projects, and others.

The second aspect of Focus Area Three concerns the use of some teaching and learning practices that are culturally familiar to Māori ākonga. For example, mihi ki te ākonga (greetings), pronouncing names correctly, using Māori concepts in the classroom, and noho marae as a learning tool (using marae for educational purpose); teaching methods that appeal to different learning styles; keeping track of ākonga learning progress (Tomoana, 2012). In addition, ākonga may be exposed directly to many current, real-world Māori case studies through the assessments and class activities. They also have the option to submit their assessment work in te reo Māori, and their work will be evaluated in collaboration with kaihautū and Unitec's Maia Centre.

Focus Area Four is about increasing staff capability in embedding mātauranga Māori into their day-to-day work through internal professional development (digital credentials). Living Te Noho Kotahitanga is a digital credential that aims to support kaimahi in developing an understanding of the Unitec partnership document of Te Noho Kotahitanga. It also supports kaimahi with integrating the values into their day-to-day work (Unitec | Te Pūkenga, 2023e). Another digital credential is Te Tīpare: Embedding Mātauranga Māori, which aims to support kaimahi in embedding mātauranga Māori in their role. Kaimahi can use Te Tīpare framework to critically reflect on their professional practice to ensure the success of ākonga. They can also impact and influence the success and holistic wellbeing of all Unitec ākonga, especially Māori, through the framework (Unitec | Te Pūkenga, 2023e).

Te Reo Māori me ngā Tikanga

Many resources are available to learn about Māori culture and practices. *Te Ara: The Encyclopedia of New Zealand* is an online government resource that briefly describes Māori culture. Many academic providers, including Unitec, offer free classes to learn Māori language and culture (Unitec | Te Pūkenga, 2023c), from beginner to advanced levels. Libraries, whether they are council owned or belong to education providers, offer many resources and guidance about mātauranga Māori.

CASE STUDY – IS APPLICATION PROJECT COURSE

The HTCS5607 IS Application Project is designed to provide ākonga with the opportunity to apply their skills and knowledge in information systems development to an integrated project set within an unfamiliar context. This course is worth 30 of 120 credits in the NZDIS and enables ākonga to apply knowledge and skills gained in the previous six courses. Ākonga are assigned case studies of companies, some of which have Māori backgrounds. The case studies consist of a set of detailed user requirements, and ākonga are informed that the requirements are the product of the information-gathering phase (the process of identifying and defining what a software system needs to do to meet the needs of its users) of the system-development lifecycle. They are required to collaborate in groups (of usually three to four), which are allocated by the kaiako, to complete the projects. Due to the length of the project being restricted to eight weeks, the waterfall model is chosen as the system-development lifecycle to follow. The structure of the course is: project initiation and planning (Week 1), analysis of requirements (Week 2), design (Weeks 3 and 4), coding and testing (Weeks 5 and 6), technical report and training materials (Week 7), and demonstration and project close (Week 8).

Current Approach to Embedding Mātauranga Māori

We are already applying Te Noho Kotahitanga values in the course. The internal professional development (i.e., digital credential) on Living Te Noho Kotahitanga has helped kaiako to understand the values and how to embed them in teaching and learning. Ongoing support is provided by kaihautū to further integrate the values throughout the lifespan of the project. We have incorporated strategies such as role playing, case studies and group reflections (Smith, 2015). In each project, ākonga assume different characters that reflect real-life roles and responsibilities, which is in line with the value of Rangatiratanga. These roles include the project manager, business analyst, quality-assurance manager, developer and tester, with the kaiako acting as project sponsors. Kaiako are actively involved in providing project-management guidance and technical support to the ākonga, guiding them in building a high-quality product from the ground up. This is an example of how the value of Kaitiakitanga is applied. Ākonga in the groups are from diverse backgrounds, so in order to establish a sense of Wakaritenga (one of Te Noho Kotahitanga values) in the classroom, we have implemented clear rules that apply to all ākonga, such as attending weekly meetings, no use of inappropriate language and no absence without a legitimate reason. Having the rules also helps in ensuring that ākonga demonstrate the value of Ngākau Māhaki towards each other. The project gives ākonga an opportunity to learn from each other's perspectives and apply Mahi Kotahitanga in their work, towards achieving the final goal.

The concept of extended whānau is used throughout the project lifecycle. The project is structured and delivered in such a way that each ākonga is supported by a whānau consisting of fellow ākonga, kaiako, academic programme manager and guest speakers (Te Momo, 2022). For example, in Week 1 all groups focus on project planning and initiation. This approach can help detect issues early in the project, which can be resolved with whānau support. We organise a Whānau Day at the end of project. We invite students' families, friends and others to celebrate the ākonga and the project's completion with whakawhanaungatanga and shared kai.

The Future Approach in Embedding Mātauranga Māori

In future we would like to continue the practices mentioned above and emphasise them to support ākonga. We will meet each group individually once a week, and during the meeting we will get each ākonga to pick a Te Noho Kotahitanga value and ask them to talk about how it is working in their group. By using role play, we want to prepare ākonga for real-world scenarios. The group work with allocated roles gives them a safe place to practise Te Noho Kotahitanga values. This strategy will help them in developing better awareness of the application of the values in IT projects.

We will invite the school kaihautū to speak to our project ākonga in the first week of the course. We would like kaihautū to speak about mātauranga Māori and ways to implement the knowledge into their project. This session and the group-forming session can he held in the marae, providing a formal or informal atmosphere, as well as providing the right environment for Māori culture. We will also invite Māori experts to talk about Māori data sovereignty and te Tiriti o Waitangi.

We will use the tuakana–teina model (with PASS leaders as tuākana) as additional support to manage the progress of the projects. PASS leaders are ākonga who have been through the same process in an earlier semester. They can attend the class once a week and spend 15 to 20 minutes with each group, sharing their experiences and providing help and guidance to the ākonga in completing the project.

CONCLUSION

The paper discusses different methods introduced by Unitec to embed mātauranga Māori in teaching and learning, such as Te Noho Kotahitanga values and I See Me initiatives. In addition, we have also looked at some reo Māori me ngā tikanga (Māori language and culture) to build the capability of kaimahi and ākonga in Māori culture and practices.

A case study of the Level 5 IS Application Project is considered to show the current approaches of embedding mātauranga Māori. These approaches have been evaluated by ākonga through Unitec students' course surveys. The statement related to mātauranga Māori was rated the lowest among other statements. The case study also discusses the future approaches for how to improve the awareness of ākonga regarding embedded mātauranga Māori in their IT courses. Some of these future approaches are: putting more emphasis on Te Noho Kotahitanga values; inviting kaihautū and Māori experts to address the class; and to use tuakana–teina relationships during the class. The plan is to implement these approaches in Semester 2 2023, and compare the course survey results with previous ones by the end of the semester.

If the proposed approaches are successful in improving the awareness of ākonga regarding Māori beliefs, language and practices, then the work here will be used as a reference for other IT courses in the future.

ACKNOWLEDGMENT

The team would like to express their gratitude to the kaihautū of the School of Computing, Electrical and Applied Technology, Deacon Andrews, for revising this work from a Māori perspective and providing valuable inputs.

REFERENCES

MacLennan, V. (2023, November). *Te Reo Māori digital technology terms*. IT Professionals New Zealand. https://itp.nz/techblog/3156-Te-Reo-Maori-digital-technology-terms-

Smith, G. (2015). The dialectic relation of theory and practice in the development of kaupapa Māori praxis. In L. Pihama, S-J. Tiakiwai, & K. Southey (Eds.), *Kaupapa rangahau: A reader* (pp. 17–27). The University of Waikato.

Te Ara: The Encyclopedia of New Zealand. (2023, July). https://teara.govt.nz/en

Te Momo, H. (2022). Building Indigenous knowledge: Exploring the pedagogy of Māori knowledge in the digital computing information technology tertiary sector of New Zealand. *Journal of Contemporary Issues in Education*, *17(*2), 92–110. https://doi.org/10.20355/jcie29499

Tomoana, R. (2012). *Sharing successful teaching and learning strategies for Māori, Pacific, and Youth learners*. Whitireia New Zealand.

Unitec New Zealand. (2020). *The Unitec Learning and Teaching Strategy 2020–2022* [Unpublished document].

Unitec | Te Pūkenga. (2023a). *Course content statements, institutional reports* [Unpublished report].

Unitec | Te Pūkenga. (2023b, July). *"I See Me" initiatives*. https://moodle.unitec.ac.nz/course/view.php?id=4367

Unitec | Te Pūkenga. (2023c, July). *Māori night classes*. https://www.unitec.ac.nz/career-and-study-options/m%C4%81ori-night-classes/m%C4%81ori-night-classes

Unitec | Te Pūkenga. (2023d, July). *Our partnership*. https://www.unitec.ac.nz/maori/who-we-are/our-partnership

Unitec | Te Pūkenga. (2023e, July). *Te Rito Workshops 2023*. https://unitec.wufoo.com/forms/w13sqxvh0jq5svq/

AUTHORS

Abha Chitalia is an Academic Advisor – Teaching and Learning in Te Puna Ako at Unitec. Her main research areas are curriculum development, teaching and learning, and assessment design. achitalia@unitec.ac.nz

Simon Dacey is a Senior Lecturer in the School of Computing, Electrical and Applied Technology at Unitec. His main research areas are remote sensing, gamification, and teaching and learning. sdacey@unitec.ac.nz

Gerard Lovell is a Senior Lecturer in the School of Computing, Electrical and Applied Technology at Unitec. glovell@unitec.ac.nz

Dr Lei Song is a Senior Lecturer in the School of Computing, Electrical and Applied Technology at Unitec. His main research areas are machine learning and environmental data analysis. lsong@unitec.ac.nz

Dr Bashar Barmada is a Senior Lecturer and Academic Programme Manager in the School of Computing, Electrical and Applied Technology at Unitec. His main research areas are in IoT applications and network security. bbarmada@unitec.ac.nz

A W-WRAITEC Model for Elevating the Academic Writing Skills of International Postgraduate Students

Minjie Hu, Tony Assadi and Marta Vos

https://doi.org/10.34074/proc.240113
Correspondence: Minjie Hu, Minjie.Hu@weltec.ac.nz

This publication may be cited as:

Hu, M., Assadi, T., & Vos, M. (2024). A W-WRAITEC Model for Elevating the Academic Writing Skills of International Postgraduate Students. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 98–103). ePress, Unitec. https://doi.org/10.34074/proc.240113

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga – New Zealand Institute of Skills and Technology

ABSTRACT

In recent years, there have been more international students enrolling in information technology (IT) courses at postgraduate level than ever before. Although relevant courses about research such as research methods are provided, some students still struggle in writing their theses and research reports. This work-in-progress study proposes to build up a model of weekly writing retreats with AI tools embedded in a course (W-WRAITEC), in which students could be trained to apply both academic writing and language-related skills to their assignments in a postgraduate-level IT course. The evaluation will include the analysis of current students' perspectives and performance in different assignments, as well as the comparison of their performance between 2022 and 2023. We believe that our model, which includes explicit guidance on academic writing and AI tools in individual IT courses, will extend the opportunity to enhance postgraduate students' abilities and confidence in their thesis writing.

KEYWORDS

Academic writing, artificial intelligence, Bloom's Taxonomy, information technology, research

INTRODUCTION

In recent years, there have been more international students enrolling in IT courses at the postgraduate level than ever before. For admission to postgraduate study, minimum performance levels of English-language proficiency for the polytechnic sector are based on the International English Language Testing System (IELTS – Academic, 6.5 overall with no sub-score less than 6.0); or TOEFL Internet-Based Test (IBT Score of 79 with a writing score of 21). Our experience shows that international students with a bachelor's degree or even a master's qualification from their own country can still have poor research and academic writing skills. Consequently, our school provides compulsory research methods courses that include literature review and research method skills, as well as providing additional support through the institute's Learner Support Services (LSS). However, despite these efforts, some postgraduate students still struggle with writing their theses and research reports.

There is a common consensus that international students face challenges and difficulties in academic writing at postgraduate level (Bair & Mader, 2013; Kotamjani & Hussin, 2017; Kotamjani et al., 2018; Schulze & Lemmer, 2017; Singh, 2016, 2017). To investigate the challenges faced by Iranian postgraduate students in academic writing in English, a survey was conducted on 52 Iranian postgraduate students from various disciplines in 16 faculties at University Putra Malaysia (Kotamjani & Hussin, 2017). The results found that these students faced difficulty in both academic writing and English-language-related skills. In terms of academic writing skills, the most challenging skill was literature review, especially in critically reviewing and determining the research gap. Additionally, students experienced difficulty in writing introductions, discussing findings and making conclusions. Regarding English-language-related skills, it was very challenging for them to use appropriate academic vocabulary to construct sentences and paragraphs and to ensure the paragraphs' coherence.

Kotamjani et al. (2018) extended their study on academic writing to 128 international postgraduate students in different fields of study at both University Putra Malaysia and University Technology Malaysia. The survey findings revealed that although international postgraduate students believed academic writing skills (e.g., literature review) were less difficult than language-related abilities (e.g., writing coherent paragraphs), they considered the most problematic areas of academic writing to be critical writing, followed by writing introductions, addressing the research gap, conclusions and discussions. Once again, the most challenging skills in language-related problems were consistent with their previous findings, such as how to build sentences and paragraphs as well as write paragraphs coherently. Finally, the study advocated to "provide students with appropriate writing resources to foster academic writing skills and encourage students to become more critical in their writing process" (Kotamjani et al., 2018, p. 194).

Recently, a report by Tremblay-Wragg et al. (2021) described a Thèsez-vous writing retreat model that was designed by and for postgraduate students, including a literature review and consultation. The model was designed to address a lack of academic writing ability in Quebec and elsewhere in Canada. The three-day retreat paid attention to individual writing to meet the SMART (Specific, Measurable, Attainable, Relevant and Timely) goals. It also provided chances for participants to "readily implement newly acquired hints and tips in the actual writing context of the retreat ... and encourage students to sustain good writing practices after the retreat" (p. 6). The retreat had run more than 40 times for over 1500 graduate students and was reported to significantly improve their writing capabilities and productivity.

In summary, various writing resources or equivalent guidelines for improving academic writing have been widely utilised over the years. Similar issues exist for most international postgraduate students. We argue that academic writing is not only about remembering, understanding, and applying these resources and guidelines to the thesis at the lower-order thinking-skills level, but also about the need to analyse, synthesise and evaluate the collected literature and data into valuable information at the higher-order thinking-skills level in Bloom's Taxonomy Pyramid (Thompson et al., 2008).

Teaching higher-order thinking skills to postgraduate students is significantly different to the lower-order thinking skills that we teach undergraduate students. Moreover, the three-day writing retreat is a plausible strategy for improving academic writing for postgraduate students. However, this paid short-period service provides a limited chance for students to learn and apply these resources and guidelines to exercises before they start writing their research proposals and theses. Therefore, we would propose to build up a model of weekly writing retreats with the AI tools embedded in a course (W-WRAITEC). This model would allow students to receive training in applying both academic writing and language-related skills to their assignments in a postgraduate-level IT course.

AIMS AND RESEARCH DESIGN

Research question: What is the effectiveness of integrating a model of weekly writing retreats with AI tools into postgraduate-level courses to improve the academic writing skills of international students?

In our postgraduate course on Business Intelligence and Data Mining, we aim to provide students with a framework for identifying and aligning computer-based data-analysis strategies with organisational goals to create and sustain competitive advantage. The assessment for this course contains three assignments: a literature review, a project development of a data warehouse and data mining, and a research report.

We propose to set up the W-WRAITEC model to offer these students training in both academic writing and English-language-related skills before or while writing their theses or research reports. Consequently, students have more opportunities to learn higher-order thinking skills in terms of analysis, synthesis, and evaluation. The W-WRAITEC model contains three parts to support international students' learning in the above postgraduate course:

1. Initiating academic writing with a literature review based on a provided business case.
    a. Offering guidance on how to write a literature review that emphasises analysing, synthesising and critically evaluating research papers rather than just summarising them.
    b. Utilising video clips to support learning the literature review process step-by-step.

c. Incorporating artificial intelligence (AI) tools such as Paper Digest, Elicit and LitMaps to assist them in summarising the contents and visualising the relationship of each paper.
d. Supplying problem-solving guidance such as divide and conquer (Floyd, 1979) so that students practise ways of analysing, synthesising and evaluating each paper.
e. Explicitly providing guidance in academic writing, spanning from academic vocabulary to sentence, paragraph and literature writing skills.

2. Presenting findings and discussions from the controlled research scenario based on their practice in data warehouse and data mining.
   a. Offering prerequisite IT knowledge such as E-R modelling and SQL for students from non-IT backgrounds.
   b. Creating visual documents and video resources for understanding professional text-based instructions in practice.
   c. Providing explicit guidance to students to describe data visualisation in various charts and tables, such as describing a table, and the process of a diagram/figure; comparing numbers, charts and diagrams; describing trends from charts and tables, etc.
3. Crafting a formal research report under a designed template.
   a. Furnishing guidance throughout the entire research process.
   b. Offering step-by-step guidance on how to write a research report while ensuring students adhere to scientific writing guidelines.
   c. Creating a research report template for students to incorporate their previous work from the above two assignments and integrate them together to achieve a standard benchmark.

ANALYSIS AND FINDINGS

In this work-in-progress research, we will evaluate our W-WRAITEC model based on the analysis of the following information from students' perspectives and performance.

- We will collect a declaration form in respect of the ethical use of AI tools in students' reports to assess the impact of AI tools in their writing.
- We will collect and analyse data on students' access to the above learning-support materials in academic writing and general language skills, correlating it with their learning performance.
- To gauge the effect of the model, we will employ a paired sample t-test to analyse the impact of every student's performance in two written assignments: literature review and research report.
- The student t-test will compare current-year students' performance to that from the previous year in both assignments.
- Finally, a survey will be conducted before the end of the trimester to gather insights into students' perspectives on research and report writing, as well as the use of AI tools for assisting in their research report writing.

CONCLUSION AND RECOMMENDATIONS

Teaching an IT course at postgraduate level is different to that at the undergraduate level. In the postgraduate course of Business Intelligence and Data Mining, the approach involves not only training students in lab practices at the knowledge, comprehension and application levels, but also placing emphasis on teaching them research methods and reporting skills. This equips them to handle their practical results at analysis, synthesis and evaluation levels.

Our W-WRAITEC model is like a swimming pool for training international students in a designed and controlled environment, instead of letting them find their own way of swimming in a deep sea through lack of training. We strongly recommend that our colleagues who teach at the postgraduate level consider using our model, which includes explicit guidance, AI tools, academic-writing support, and pastoral care in their own courses. This approach will provide students with more opportunities to practise both academic writing skills and general language skills before they begin writing their theses and research papers. Ultimately, we anticipate that writing their theses and research papers will become 'a piece of cake' for international students.

REFERENCES

Bair, M., & Mader, C. (2013). Academic writing at the graduate level: Improving the curriculum through faculty collaboration. *Journal of University Teaching and Learning Practice*, *10*(1). http://ro.uow.edu.au/jutlp/vol10/iss1/4

Floyd, R. (1979). The paradigms of programming. Turing Award Lecture. *Communications of the ACM*, *22*(8), 455–460.

Kotamjani, S., & Hussin, H. (2017). Perceptions of challenges in writing academically: Iranian postgraduate students' perspectives. *International Journal of Applied Linguistics & English Literature*, *6*(4), 1–4. https://journals.aiac.org.au/index.php/IJALEL/article/view/3111

Kotamjani, S., Samad, A., & Fahimirad, M. (2018). International postgraduate students' perception of challenges in academic writing in Malaysian public universities. *International Journal of Applied Linguistics & English Literature*, *7*(7), 191–195. https://doi.org/10.7575/aiac.ijalel.v.7n.7p.191

Schulze, S., & Lemmer, E. (2017). Supporting the development of postgraduate academic writing skills in South African universities. *Per Linguam*, *33*(1), 54–66. https://doi.org/10.5785/33-1-702

Singh, M. (2016). An emic perspective on academic writing difficulties among international graduate students in Malaysia. *Journal of Language Studies*, *16*(3), 83–97. https://doi.org/10.17576/gema-2016-1603-06

Singh, M. J. M. (2017). International EFL/ESL master students' adaptation strategies for academic writing practices at tertiary level. *Journal of International Students*, *7*(3), 620–643. https://doi.org/10.5281/zenodo.570025

Thompson, E., Luxton-Reilly, A., Whalley, J., Hu, M., & Robbins, P. (2008). Bloom's Taxonomy for CS assessment. *Proceedings of the Tenth Conference on Australasian Computing Education*, 78, 155–161. https://dl.acm.org/doi/10.5555/1379249.1379265

Tremblay-Wragg, É., Chartier, M. S., Labonté-Lemoyne, É., Déri, C., & Gadbois, M. E. (2021). Writing more, better, together: How writing retreats support graduate students through their journey. *Journal of Further and Higher Education*, *45*(1), 95–106. https://doi.org/10.1080/0309877X.2020.1736272

AUTHORS

Dr Minjie (Min) Hu is a Senior Lecturer in IT at the School of Innovation, Design and Technology, Whitireia-WelTec. Having published extensively at domestic and international conferences in the last two decades, Min's research interests are teaching novices programming, e-learning, data analytics in business and financial information, software testing and postgraduate education. Minjie.Hu@weltec.ac.nz

Tony Assadi is a Programme Manager at the School of Innovation, Design and Technology, Whitireia-WelTec. He is responsible for the undergraduate programmes at Whitireia and the postgraduate IT programmes at the Wellington campus. Tony's research interests are in work-integrated learning, computational thinking in tertiary education, and digital divide research. Tony.Assadi@wandw.ac.nz

Dr Marta Vos is currently the Postgraduate Co-ordinator at School of Innovation, Design and Technology, Whitireia-WelTec, as well as being an external PhD supervisor. Her primary research interests lie in understanding the nature of the IT artefact, particularly from the cultural/Indigenous perspective, as well as in the implementation and management of data and IoT systems. Marta.Vos@whitireia.ac.nz

Teacher Capability Development at Unitec: A Case Study

Abha Chitalia and Diana Ayling

https://doi.org/10.34074/proc.240114
Correspondence: achitalia@unitec.ac.nz

This publication may be cited as:

Chitalia, A., & Ayling, D. (2024). Teacher Capability Development at Unitec: A Case Study. In H. Sharifzadeh (Ed.), ***Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September*** **(pp. 104–114). ePress, Unitec. https://doi.org/10.34074/proc.240114**

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga –
New Zealand Institute of Skills and Technology

ABSTRACT

This paper explores the Teacher Capability Development (TCD) project at Unitec, which aims to improve kaiako (teacher) professional development and ākonga (student) outcomes through a competency-based approach, applied practice, and the use of digital credentials (badges). The project was evaluated using the Kirkpatrick model, assessing participants' reactions, learning and behaviour, and organisational performance impact. Initially resistant, participants eventually embraced the project, demonstrating its success in addressing their concerns. Participants' learning was evident in their badge submissions, demonstrating their skills applied in teaching. The project's adaptability was highlighted during the Covid-19 pandemic, when the online environment facilitated the use of technology tools. Badges as digital credentials provided a tangible way to recognise achievements. The evaluation extended to the project's impact on behaviour and organisational performance. Organisational evaluations revealed increased adoption of course changes due to TCD professional development, fostering continuous improvement. Ākonga reported high net promoter scores, indicating improved satisfaction and outcomes.

Implications of the TCD project include designing competency-based projects, integrating applied practice for continuous improvement, leveraging technology for development, and recognising achievements through digital credentials. The findings emphasise the importance of investing in ongoing teacher development to enhance ākonga success.

KEYWORDS

Professional development, applied practice, digital credentials

INTRODUCTION

This paper explores the implementation and evaluation of the Teacher Capability Development (TCD) project at Unitec. The TCD project aimed to enhance kaiako (teacher) professional development and ākonga (student) outcomes through its competency-based approach, applied practice, integration into performance development plans, and the use of digital credentials (badges) to acknowledge achievements. Over four years, the project underwent evaluation using the Kirkpatrick model, which assessed participants' reactions, learning and behaviour, and the impact on organisational performance.

BACKGROUND

In 2017–18, Unitec Institute of Technology reviewed the current kaiako professional development programme. A new project, TCD, resulted from the review. The TCD project's design includes an integrated approach to kaiako professional learning and focuses on academic quality at each point of the programme and course design process. TCD was part of a more extensive body of work to ensure ākonga success. Unitec had developed a Student Success Programme to operate in parallel with the professional development project for kaiako. Student success programmes (SSPs) provide support services for all ākonga (Kahu & Nelson, 2018). The programmes specifically support ākonga who are Māori, Pacific, International, first generation in tertiary study, low income and/or ākonga with disabilities (Unitec Institute of Technology, 2020). SSPs focus on both academic and personal development of ākonga throughout the learning journey.

TCD was aligned with a range of teaching competencies developed by Unitec (Unitec Institute of Technology, 2019a). The competencies were drawn from a review of current domestic and international standards in higher-education teaching (Unitec Institute of Technology, 2018c), and are as follows (Unitec Institute of Technology, 2019a):

- Create learner-centred environments
- Design for effective learning
- Facilitate learning
- Assess and give feedback on learning
- Review learning and teaching
- Show discipline/industry expertise
- Engage in continuous professional learning about teaching and learning
- Contribute to programme's operation and academic success

See Appendix 1 for a list of digital credentials that are aligned to competencies.

Implemented from 2018, this innovative approach has the following features (Unitec Institute of Technology, 2018c):

1. A competency-based professional development approach that would clearly model best-practice teaching to kaiako.
2. Concepts of applied practice (i.e., kaiako completing professional development in conjunction with their actual teaching practice) and learning-in-work to support a culture of continuous improvement for all kaiako.
3. A flexible approach to professional development that includes recognition of formal and informal learning, comprehensive self-guided online resources, face-to-face support (individual or group) on request, three development modes (supported, independent and evidence only).
4. Adaptation of Pohatu's Mauri Model (Pohatu, 2011) to evaluate and acknowledge the progressive stages of competence development. The model includes three levels: Mauri Oho (Emerging), Mauri Tū (Demonstrating) and Mauri Ora (Modelling).
5. Recognition of achievement using digital credentials (popularly referred to at Unitec as 'badges'), which also facilitate the gathering of evidence of individual competencies. Each badge achieved by a kaiako represents a level of capability, based on evidence provided and validated by a reliable assessment process.
6. Integration into the Performance Development Plan and internal performance partnering process.

TCD is part of Unitec's demonstration of how academic practice supports the New Zealand Qualification Authority's (NZQA) Key Evaluative Questions (New Zealand Qualifications Authority, 2023). The NZQA's evaluative questions refer to the Key Evaluative Questions framework developed by NZQA. These questions serve as a guide for evaluating the quality and effectiveness of educational programmes and institutions in Aotearoa New Zealand. The framework consists of six key questions:

1. What is the educational purpose of the programme or institution?
2. How effectively does the programme or institution achieve its educational purpose?
3. How well does the programme or institution address learner needs?
4. How effective is the teaching, learning, and assessment within the programme or institution?
5. How well does the programme or institution manage its educational performance?
6. How effectively does the programme or institution promote valued outcomes for learners?

These evaluative questions are important because they provide a comprehensive and systematic approach to assess the quality and impact of educational programmes and institutions. They help in evaluating whether the programme or institution is aligned with its intended purpose, is meeting the needs of learners, and is achieving desired outcomes. By considering these questions, educational stakeholders can identify strengths and areas for improvement, and make informed decisions to enhance educational quality.

The evaluative questions also promote accountability and transparency within the education system. They provide a standardised framework for evaluating programmes and institutions, ensuring consistent and fair assessments. The questions encourage a continuous-improvement mindset, prompting institutions to reflect on their practices, identify areas of development, and implement strategies to enhance the quality of education (New Zealand Qualifications Authority, 2023).

Overall, the NZQA's evaluative questions have played a crucial role in evaluating and improving educational programmes at Unitec. They provide a robust framework to assess educational quality, inform decision-making, and promote ongoing improvement in teaching and learning practices.

TCD EVALUATION METHODOLOGY

Each year, the TCD project was evaluated using the Kirkpatrick model (Kirkpatrick, 1994). This is the best-known model for analysing and evaluating the results of training and educational programmes. It is based on four levels:

Level 1: Reaction – measures how participants react to the training (e.g., satisfaction).

Level 2: Learning – analyses whether participants truly understood the training (e.g., increase in knowledge, skills or experience).

Level 3: Behaviour – explores whether participants are using what they learned (e.g., change in behaviours).

Level 4: Results or Organisational Performance – determines whether the training has had a positive impact on the organisation.

During 2018–22, the TCD project was evaluated five times. Data was gathered from various sources to evaluate the project from different perspectives, including participants, learning effectiveness, participants' use of learning, and organisational impact.

Participants' reaction to the new TCD project was initially negative. One comment in the 2019 TCD evaluation identifies resistance to change:

> *"Over time the approach to PD changes, this imposes a quantum of adaptive load and stress for staff."* (Kaiako 1; Unitec Institute of Technology, 2019b)

Another comment highlights resistance to the applied practice approach:

> *"I dislike the whole new approach and find it patronising and demeaning, especially for experienced teaching staff."* (Kaiako 2; Unitec Institute of Technology, 2019b)

Over time, reactions moved to acceptance and, finally, to a positive endorsement from those that participated. Those participating in 2022 rated their TCD learning experience 8.88 out of 10 (Chitalia et al., 2022).

Course evaluation and planning (CEP) completed by kaiako for each course provided insight into the number of kaiako implementing changes to their courses because of their engagement with TCD professional development.

In the 2022 CEP process, two questions were put to teachers relating to professional development:

1. List recent continuous professional development or engagement with industry that influenced the delivery of the course.
2. In what ways did this professional development impact on the teaching and assessment of this course?

The authors reviewed individual comments (N = 868) in response to these questions at a broad level of keywords and frequencies. They noted that in relation to the first question the majority of professional development was badge based and there was a strong engagement in discipline professional development.

In relation to the text-based data analysed from the CEP question two, the authors found broad themes suggesting positive outcomes for course delivery of kaiako completing professional development. Kaiako reported improvement in their practice in three key TCD areas: Create Learner-Centred Environments, Design for Effective Learning, and Facilitate Learning.

Specifically, kaiako reported on the integration of Unitec's Te Noho Kohitanga values; cultural understanding and learning from the Teaching Pacific Learners digital credentials; learning more about Moodle and other delivery tools; design and delivery of courses; and integration of new discipline knowledge from professional development activity into teaching. Thirty-five kaiako reported improvement in their practice of assessment and feedback.

Unitec collected data from ākonga in each semester of study and asked whether they could recommend their programme of study. Net promoter scores remained strong throughout the implementation of TCD.[1] This data indicates a positive impact on the organisation, as a substantial number of ākonga recommended their programme of study (Chitalia et al., 2022).

The focus of Level 4 of the Kirkpatrick model is to assess whether professional development has had a positive impact on the organisation. The authors considered that data available from the net promoter score collected each semester from Unitec ākonga is relevant to organisational performance in respect of teaching capability. We may infer that high student satisfaction is related to good levels of teaching capability. However, we cannot assume a correlation or causal link. Further research would be required to fully substantiate a relationship.

DISCUSSION

The Value of Teaching Competencies

Kaiako have specific knowledge and skills to acquire in order to attain the position of a practitioner in vocational education (Handley et al., 2006). As kaiako develop their 'know-how' of teaching, they identify and form linkages between theory, previous research, current experience and their ākonga needs.

Early-career kaiako may not have the knowledge, skills and emotional strength to take risks in front of their ākonga. As Kreiner (2001, p. 71) explains, if kaiako do not take risks they "will lose the sense of excitement that comes from trying new things." Teachers who lack the confidence to take risks are likely to "have students who lack excitement" (Kreiner, 2001, p. 71).

Support that fosters deep reflection that involves scrutiny and clarification of their own educational beliefs, values and mission may assist early-career kaiako in moving from novice to practitioner (Handley et al., 2006; Korthagen & Vasalos, 2005; Loughran, 2014; Loughran & Hamilton, 2016).

The teaching competencies assist early-career kaiako to identify the key knowledge, skills and behaviours of vocational teaching that complement and reinforce their professional expertise. The TCD project facilitated

1 The net promoter score (NPS) is calculated by the % of students who give a high recommendation rating (promoters, who rate 9–10) minus the % of student who give a low recommendation rating (detractors, who rate 0–6).

participants' learning through engaging with various badges, encouraging the application of acquired knowledge and skills in their teaching practice. Reflecting on their teaching practice before and after the completion of digital credentials, participants demonstrated new insights gained and plans for further development.

As Ping et al. (2018) explain, there is a range of reasons why kaiako engage in professional learning. One reason is personal ambition ('inside reasons'), and external forces such as government policy and institutional policy are described as 'outside reasons'. The 2018 TCD Pilot Report (Unitec Institute of Technology, 2018a) identified that kaiako engagement was significantly higher when heads of faculty were actively involved in co-learning – an 'outside reason'.

Motivation for learning was a key element to kaiako participation. Motivation came from a range of sources, including personal-development plans, faculty capability-development plans and institutional requirements measured through performance management processes.

While business organisations have a long tradition of commitment to continuous improvement programmes for products and services, vocational education and training organisations do not have the same working history of quality improvement in teaching and learning. In the vocational education sector, quality improvement has been an administrative function closely related to NZQA quality assurance and compliance (Maurice-Takerei & Anderson, 2022). Kaiako often contribute to, but are not the critical elements of, a quality-improvement programme.

The TCD project provides kaiako with a valuable opportunity to embrace continuous improvement in their professional development and career progression. Aligned with the stated aims of Te Pūkenga, New Zealand Institute of Skills and Technology (NZIST), as defined in the Education and Training Act 2020 (the Act), the project aims to enhance consistency, quality and equity within the tertiary education system.

An essential aspect of achieving these goals is the integration of continuous-improvement principles through TCD. This connection between professional learning and ākonga learning becomes evident and meaningful to kaiako as they plan and develop their educational practices.

Leveraging Technology for Professional Learning in Higher Education

The original Teacher Capability Development project goals included the ideas of digital credentials (Unitec Institute of Technology, 2018b). The plan included use of the Unitec learner management system (Moodle). Over time, this changed due to the Covid-19 pandemic and the need for online learning. We adopted a new simple template for Moodle, we reduced the activities for kaiako to ten or fewer steps, we embedded the use of Zoom.

Bruggeman et al. (2021) perceived seven teacher attributes as key for the adoption of blended learning in higher education:

1. Putting teaching and education at the centre.
2. Holding a student-centred pedagogical belief.
3. Realising a pedagogical need for change.
4. Daring to experiment (and fail).
5. Sharing needs and concerns.
6. Being able to critically self-reflect as a teacher.
7. Being able to connect technologies to learning processes.

The TCD project aligns with this understanding of preparing kaiako to succeed in a highly blended teaching and learning environment. There is more work to be done in this area to support kaiako to succeed in the online learning environment. The feedback from participants in the project was strongly endorsing of digital credentials that supported their learning and development in blended learning (Chitalia et al., 2021).

On-campus and group workshops of the TCD digital credentials moved to online during the Covid-19 pandemic. Although the move to online workshops was under emergency, it allowed continuity of the TCD project. The Academic Advisors – Teaching and Learning had to quickly upskill themselves in online teaching practice and model this practice to the participants. The adaptability and integration of technology (i.e., Zoom and Echo360) demonstrates the project's responsiveness to changing educational contexts and the evolving needs of kaiako and ākonga.

Flexibility in offerings has become a normal practice and not an exception for 'new normal' education (Cadorna et al., 2022). Due to mix of offerings, kaiako are showing more interest and willingness to attend/participate in both on-campus and online workshops offered for the TCD project (Chitalia et al., 2021). This also allows kaiako to participate at a time and place that is convenient for their work schedules and personal needs.

Two additional digital credentials were developed so kaiako can gain recognition of the capability they have developed while rapidly adapting their teaching practice and revising assessments during the 2020–22 Covid-19 lockdowns (Chitalia et al., 2021; 2022).

Integrating Applied Learning into Work

In Aotearoa New Zealand there is a shift to prioritising student-centred learning in vocational education. The Education and Training Act 2020 (the Act) legislated for one large polytechnic body as Te Pūkenga, New Zealand Institute of Skills and Technology (NZIST) to provide vocational education and training, and to be supportive, flexible and relevant to the workplace.

As explained by Maurice-Takerei & Anderson (2022), when kaiako work in a range of learning environments, including online, in the classroom, lecture theatre, studio or workplace, a transition to student-centred learning, as required by the Act, is a complex change for many in the vocational education system.

Kaiako will need significant support to move to a student-centred teaching environment. Professional development, which both models and supports good practice, is a strength of the TCD project. Furthermore, teachers apply their learning in their context of a course and cohort of ākonga. Effectively, TCD fosters action research to improve understanding of personal and professional development (Banegas & Villacañas de Castro, 2019).

Recognising and Showcasing Achievements

As noted by Pathways For Employ (2017, p. 12):

> When a learner is aware that their learning path can be both documented and appreciated in some way it provides an important stimulus to learning. A motivated learner will identify new skills and paths they need to explore as they learn.

This approach acknowledges participants' accomplishments and provides a tangible way to demonstrate their learning to others. The use of digital credentials aligns with the growing trend of alternative forms of credentialing and provides a valuable means of publicly displaying professional development achievements (Carey & Stefaniak, 2018).

IMPLICATIONS

The TCD project at Unitec offers several implications for teachers' professional development and educational institutions.

Designing competency-based programmes: The TCD project's competency-based approach serves as a model for other institutions to align their professional development programmes with recognised teaching competencies.

By clearly defining and addressing the knowledge and skills required for effective teaching, institutions can ensure that professional development projects are relevant and impactful.

Integrating applied practice and learning-in-work: The TCD project's emphasis on applying learning in teaching practice fosters a culture of continuous improvement. This integration allows kaiako to immediately apply their acquired knowledge and skills, leading to more meaningful and effective professional development experiences.

Leveraging technology for professional development: The successful integration of technology tools, such as Moodle, Zoom and Echo360, demonstrates the importance of leveraging digital platforms to enhance kaiako professional development. Educational institutions should invest in technology-enabled learning environments that facilitate online collaboration, resource sharing and ongoing support for teachers.

Recognising and showcasing achievements: The use of digital credentials (badges) provides a valuable means of supporting kaiako career development and ensuring recognition of learning and commitment to continuous development.

CONCLUSION

In conclusion, the TCD project implemented by Unitec offers valuable insights into enhancing kaiako professional development and improving ākonga outcomes. The project's competency-based approach, applied practice, integration into performance development plans, and use of digital credentials (badges) provide a framework for designing effective professional development aligned with teaching competencies.

The evaluation of the TCD programme using the Kirkpatrick model revealed a positive shift in participants' reactions, demonstrating the project's success in addressing concerns and meeting expectations. Participants' learning was evident through their submissions for each badge, showcasing the project's effectiveness in facilitating the application of knowledge and skills in teaching practice. The project's adaptability to online environments, particularly during the Covid-19 pandemic, highlighted its responsiveness to changing educational contexts and the evolving needs of kaiako and ākonga.

The integration of technology tools and the recognition of achievements through digital credentials (badges) further contribute to the project's effectiveness. By leveraging digital platforms, educational institutions can enhance kaiako professional development, foster collaboration and provide ongoing support. The use of badges as tangible representations of achievements promotes continuous development and career advancement for kaiako.

The implications of the TCD project extend beyond Unitec, offering valuable insights for other educational institutions. Designing competency-based programmes, integrating applied practice and learning-in-work, leveraging technology and recognising achievements are key considerations for effective kaiako professional development. By investing in ongoing kaiako development, institutions can create a conducive learning environment and ultimately enhance ākonga success.

Overall, the TCD project at Unitec demonstrates the importance of aligning professional development with kaiako needs, utilising innovative approaches, and continuously evaluating and improving programmes. Through these efforts, institutions can empower kaiako, enhance their instructional practices, and positively impact ākonga outcomes. The findings of this study provide valuable guidance for future endeavours in kaiako professional development, ensuring the continuous improvement of education systems.

REFERENCES

Banegas, D. L., & Villacañas de Castro, L. S. (2019). A look at ethical issues in action research in education. *Argentinian Journal of Applied Linguistics*, *3*(1), 58–67. https://wrap.warwick.ac.uk/69491/1/WRAP_BanegasVillaca%C3%B1asAJALVol3%25281%2529_.pdf

Bruggeman, B., Tondeur, J., Struyven, K., Pynoo, B., Garone, A., & Vanslambrouck, S. (2021). Experts speaking: Crucial teacher attributes for implementing blended learning in higher education. *The Internet and Higher Education*, *48*, 100772. https://doi.org/10.1016/j.iheduc.2020.100772

Cadorna, E. A., Cadorna, E. F., Jaramilla, A. S., Rosette, K., & Cadorna, A. (2022). Readiness of teachers on flexible learning: Basis for a capability-building program. *International Journal of Evaluation & Research in Education*, *11*(4), 1969–1979. http://doi.org/10.11591/ijere.v11i4.22862

Carey, K. L., & Stefaniak, J. E. (2018). An exploration of the utility of digital badging in higher education settings. *Educational Technology Research and Development*, *66*, 1211–1229. https://doi.org/10.1007/s11423-018-9602-1

Chitalia, A., Ayling, D., & Simperingham, L. (2021). *Report on Unitec's Teacher Capability Development* [Unpublished report]. Unitec New Zealand.

Chitalia, A., Ayling, D., & Simperingham, L. (2022). *2022 Teacher Capability Development Evaluation Report* [Unpublished report]. Unitec |Te Pūkenga.

Education and Training Act. (2020). https://www.legislation.govt.nz/act/public/2020/0038/latest/LMS170676.html

Handley, K., Sturdy, A., Fincham, R., & Clark, T. (2006). Within and beyond communities of practice: Making sense of learning through participation, identity and practice. *Journal of Management Studies*, *43*(3), 641–653. https://doi.org/10.1111/j.1467-6486.2006.00605.x

Kahu, E. R., & Nelson, K. (2018). Student engagement in the educational interface: Understanding the mechanisms of student success. *Higher Education Research & Development*, *37*(1), 58–71. https://doi.org/10.1080/07294360.2017.1344197

Kirkpatrick, D. L. (1994). *Evaluating training programs: The four levels*. Berrett-Koehler.

Korthagen, F., & Vasalos, A. (2005). Levels in reflection: Core reflection as a means to enhance professional growth. *Teachers and Teaching*, *11*(1), 47–71. https://doi.org/10.1080/1354060042000337093

Kreiner, D. S. (2001). Taking risks as a teacher. In W. Buskist, V. Hevern, & G. W. Hill IV (Eds.), *Essays from excellence in teaching: 2000–2002* (vol. 1) (pp. 71–73). Society for the Teaching of Psychology.

Loughran, J. (2014). Professionally developing as a teacher educator. *Journal of Teacher Education*, *65*(4), 271–283. https://doi.org/10.1177/0022487114533386

Loughran, J., & Hamilton, M. L. (2016). Developing an understanding of teacher education. In J. Loughran & M. L. Hamilton (Eds.), *International Handbook of Teacher Education: Volume 1* (pp. 3–22). Springer. https://doi.org/10.1007/978-981-10-0366-0_1

Maurice-Takerei, L., & Anderson, H. (2022). Vocational education and training reform in Aotearoa New Zealand: The value of educators and education in a new VET environment. *New Zealand Annual Review of Education*, *27*, 116–130. https://doi.org/10.26686/nzaroe.v27.8034

New Zealand Qualifications Authority. (2023, June 29). *Key evaluation questions*. https://www2.nzqa.govt.nz/tertiary/quality-assurance/self-assessment/key-evaluation-questions/

Pathways For Employ. (2017). Research on good practices establishing Open Badges eco systems. *Pathways For Employ*. https://www.voced.edu.au/content/ngv%3A83297

Ping, C., Schellings, G., & Beijaard, D. (2018). Teacher educators' professional learning: A literature review. *Teaching and Teacher Education*, *75*, 93-104. https://doi.org/10.1016/j.tate.2018.06.003

Pohatu, T. W. (2011). Mauri: Rethinking human wellbeing. *Mai Review*, *3*. https://www.journal.mai.ac.nz/system/files/maireview/380-3362-1-PB.pdf

Unitec Institute of Technology. (2018a). *TCD Pilot Report* [Unpublished report].

Unitec Institute of Technology. (2018b). *Teacher Capability Development Plan* [Unpublished document].

Unitec Institute of Technology. (2018c). *Teacher Capability Development – Evaluation Report* [Unpublished report].

Unitec Institute of Technology. (2019a). *Course: TCD – Teaching Competencies home page*. https://moodle.unitec.ac.nz/course/view.php?id=4591

Unitec Institute of Technology. (2019b). *Teacher Capability Development – Evaluation Report* [Unpublished report].
Unitec Institute of Technology. (2020). *Unitec student success strategy 2020–2022* [Unpublished document]. https://thenest.unitec.ac.nz/TheNestWP/wp-content/uploads/2020/09/Unitec-Student-Success-Strategy-2020-2022-Final.pdf

AUTHORS

Abha Chitalia is an Academic Advisor – Teaching and Learning in Te Puna Ako at Unitec. Her research areas are mainly in curriculum development, teaching and learning, and assessment design. achitalia@unitec.ac.nz

Diana Ayling is an Academic Advisor – Teaching and Learning in Te Puna Ako at Unitec. Her research areas are mainly in work-based learning and digital communities of practice. dayling@unitec.ac.nz

APPENDIX 1

Unitec Teaching Competencies and Related Badges

The table below shows the eight Unitec Teaching Competencies and related badges that are currently available. More will be added over time (Chitalia et al., 2022).

Unitec Teaching Competencies and Related Badges

Competency	Badges
Create learner-centred environments	Working with International Learners Teaching Pacific Learners
Design for effective learning	Design for Collaborative Learning Design Group Assessment Moodle Standards Introduction to Moodle Improve Summative Assessment Writing Learning Outcomes
Facilitate learning	Active Learning Echo360 – Classroom Engagement Interactive Presentation Literacy Learning Strategies: Writing Literacy Learning Strategies: Reading
Assess and give feedback on learning	Intro to Summative Assessment Pre-Moderate Assessment Materials Post-Moderate Assessor Decisions
Review learning and teaching	Peer Teacher Coaching Changing Assessment for Remote Online Delivery Teaching Under-25 Learners
Show discipline/industry expertise	There are currently no courses available for this competency
Engage in continuous professional learning about teaching and learning	There are currently no courses available for this competency
Contribute to programme's operation and academic success	Academic Integrity Working with Graduate Profiles Pacific Learners: Success and Retention

Study of Conference Presentations by Novice and Expert Presenters in Computing

Farnaz Rajabi Mehr

https://doi.org/10.34074/proc.240115
Correspondence: fmehr@unitec.ac.nz

This publication may be cited as:

Mehr, F. R. (2024). Study of Conference Presentations by Novice and Expert Presenters in Computing. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 115–119). ePress, Unitec. https://doi.org/10.34074/proc.240115

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga –
New Zealand Institute of Skills and Technology

ABSTRACT

This study investigates the organisational structure of academic conference presentations in the field of computing. It examines possible differences between novice (postgraduate students) and expert (faculty) presenters. A self-built corpus of 15 presentations was analysed using NVivo 12. Findings reveal 14 moves in the presentations, with variations based on presenter expertise, particularly in audience orientation and content presentation. The study has pedagogical implications for teaching conference-presentation skills to postgraduate students.

KEYWORDS

Computing, conference presentation, expert presenter, novice presenter, move analysis

INTRODUCTION

The academic conference presentation is a specialised form of public speaking and a challenging task due to its real-time delivery to an expert audience (Heino et al., 2002). According to Seliman (1996), academic conference presentations hold great significance as they offer more up-to-date information than written versions. They are considered an essential component of the research cycle and constitute a key academic genre (Hyland, 2009).

Some scholars may believe there is no difference between writing a paper for publication and presenting a paper at a conference, ignoring the fact that the organisation and language use of the two texts are different. Academic conference presentations are particularly vital for postgraduate students, as novice researchers and new members of the academic community, as these help them to be accepted within the target academic community. Therefore, the presenter should be aware of the rhetorical organisation of the presentation, ensuring it is well organised, clear, and effectively conveys the information.

To date, the extensive and rapidly growing body of literature on the language of research reporting has mainly concentrated on written forms (e.g., research articles, theses, essays or proposals). Spoken forms of research reporting, such as academic conference presentations, have received less attention.

The pioneering study on conference presentation was conducted by Dubois (1980) on biomedical conference presentations from a rhetorical perspective. Since then, there have been valuable contributions to studying the academic conference-presentation genre. Thompson (1997) and Webber (2002) compared interactive features used by presenters in delivering presentations. Heino et al. (2002) focused on interpersonal and text-organising aspects of conference presentations. Hood and Forey (2005) examined how speakers shape their talks to interact with the audience. Wulff et al. (2009) also compared the presentation session and discussion session in terms of phraseological patterns, discourse management aspects and chairs' utterances.

Furthermore, some studies have paid attention to the multimodal nature of conference presentations, especially in technical, medical and scientific fields. For example, Dubois (1982) stressed the structuring role of visuals in conference presentations. However, some scholars, including Selimen (1996), Webber (1997), Shalom (2001), Ventola (2002), Carter-Thomas and Rowley-Jolivet (2003), Rowley-Jolivet and Carter-Thomas (2005), and Querol-Julián (2011), investigated the structural organisation of conference presentations in various academic fields such as applied linguistics, engineering, health science and medical science.

The available literature mainly focuses on expert presenters who have been in the academic community for years, while neglecting the new members of the academic community. Additionally, none of the studies so far has focused on the organisational structure (i.e., moves and steps) of conference presentations in the field of computing. As a result, this under-researched academic genre still requires the attention of scholars and more in-depth research.

AIMS AND RESEARCH DESIGN

The main purpose of the study is to contribute to the description of computing conference presentations by investigating the organisational structure (moves and steps). According to Swales & Freak (2004), a move is a rhetorical unit that performs a communicative function in a discourse. In other words, a move is a functional, not a grammatical unit. A step is a sub-unit of a move, and its function is to achieve the purpose of the move to which it belongs (Cortes, 2013). Each move can comprise several steps.

This study aims to identify possible differences in the rhetorical structure of conference presentations between novices in the academic community (postgraduate students) compared to experts (faculty members). The study addresses the following research question:

Does the structure of the conference presentation vary based on the presenter's expertise (novice or expert)?

To conduct this research, a self-built corpus was created, comprising 15 conference presentations in the field of computing. These presentations consist of texts from two sources: fresh data collected from PST 2016: 14th International Conference on Privacy, Security and Trust, held in Tāmaki Makaurau Auckland, and recorded data from a public-domain website called Video Lectures (http://videolectures.net). Among the 15 presentations, eight were delivered by novices, while seven were delivered by experts. Each presentation was transcribed, and the data was analysed using NVivo 12.

The selected conference presentations were all research presentations delivered at parallel conference sessions. The novice presenters of the study were either master's students or first/second-year PhD students in the computing field. The expert presenters were faculty members teaching at different universities around the world.

To identify the rhetorical moves and steps in conference presentations, the study employed Seliman's (1996) move model. Seliman's (1996) model is grounded in her examination of engineering conference presentations and incorporates Swales's (2004) definition of a move, which regards it as a functional unit identifiable by its communicative function within the text.

The move analysis of the conference presentations has been achieved through three phases. Phase One is the move model development, which includes the initial coding of text segments based on their communicative function according to Seliman's (1996) model. In Phase Two, three of the conference presentations were coded using NVivo 12. At this phase, a specialist advisor in the computing field was consulted during the coding procedure on discipline-specific technical concepts that needed clarification. The data was coded twice, using the identified coding scheme, to minimise human errors during the coding procedure and ensure consistency in the coding. There were four-week gaps between each coding round to avoid any previous biases in the new coding round. Finally, in Phase Three, the inter-coder reliability test was conducted to assess the reliability of the coding.

ANALYSIS AND FINDINGS

In total, 14 moves were identified in computing conference presentations: three in the introduction section, seven in the body section, and four in the conclusion section. However, the findings reveal differences in the structural organisation (i.e., moves and steps) of various presentation sections based on the presenter's level of expertise (novice or expert). The study demonstrates that a presenter's expertise also influences the sequence of moves and steps, particularly in orienting the audience and presenting the content in both the introduction and body sections of conference presentations.

CONCLUSION AND RECOMMENDATIONS

The study has significant pedagogical implications for teaching the structure, organisation and delivery of conference presentations to postgraduate students in computing as novice members of the academic community.

The available books on conference presentations, such as Silyn-Roberts (2012), Jalongo and Machado (2015), and Freiermuth (2022), offer very general guidance and are not informed by comprehensive fieldwork or corpus-driven analyses, such as the approach taken in this study. Previous research demonstrates that learners' understanding of genre knowledge positively influences their production of that genre (Kaufhold, 2017). Therefore, the genre-knowledge framework developed in this study can be utilised in designing and implementing courses for postgraduate students in computing.

Despite its contributions, the study has several limitations. The first limitation is the small sample size, which restricts the generalisability of the results. To enhance the reliability of the findings, future research should consider conducting similar studies with a larger sample size. Additionally, this study did not distinguish between conference presentations delivered by native-speaking and non-native-speaking presenters, as its primary focus was on exploring differences between experts and novices in delivering presentations. Hence, future research could investigate the distinctions in the move and step structure among presenters from diverse cultural backgrounds.

REFERENCES

Carter-Thomas, S., & Rowley-Jolivet, E. (2003). Analysing the scientific conference presentation (CP): A methodological overview of a multimodal genre. *ASP. La Revue du GERAS* (39–40), 59–72. https://doi.org/10.4000/asp.1295

Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes*, *12*(1), 33–43. https://doi.org/10.1016/j.jeap.2012.11.002

Dubois, B. L. (1980) Genre and structure of biomedical speeches. *Forum Linguisticum*, 5, 140–68.

Dubois, B. L. (1982). "And the last slide, please" – Regulatory language function at biomedical meetings. *World Language English*, *1*(4), 263–268.

Freiermuth, M. R. (2022). *Academic conference presentations: A step-by-step guide*. Springer International Publishing.

Heino, A., Tervonen, E., & Tommola, J. (2002). Metadiscourse in academic conference presentations. In E. Ventola, C. Shalom & S. Thompson (Eds.), *The language of conferencing* (pp. 127–146). Peter Lang.

Hood, S., & Forey, G. (2005). Introducing a conference paper: Getting interpersonal with your audience. *Journal of English for Academic Purposes*, *4*(4), 291–306. https://doi.org/10.1016/j.jeap.2005.07.003

Hyland, K. (2009). *Academic discourse: English in a global context*. Continuum.

Jalongo, M. R., & Machado, C. (2015). *Making effective presentations at professional conferences: A guide for teachers, graduate students and professors*. Springer.

Kaufhold, K. (2017). Interdisciplinary postgraduate writing: Developing genre knowledge. *Writing & Pedagogy*, *9*(2), 251–274. https://doi.org/10.1558/wap.30568

Querol-Julián, M. (2011). *Discussion session in specialized conference paper presentation*. Universitat Jaume.

Rowley-Jolivet, E., & Carter-Thomas, S. (2005). The rhetoric of conference presentation introductions: Context, argument and interaction. *International Journal of Applied Linguistics*, *15*, 45–70. https://shs.hal.science/hal-00275963/file/RJ_CT_Revised.pdf

Seliman, S. (1996). *The genre and the genre expectations of engineering oral presentations related to academic and professional contexts* [Doctoral thesis, The University of Stirling]. STORRE: Stirling Online Research Repository. https://dspace.stir.ac.uk/handle/1893/1778

Shalom, C. (2001). *The academic conference: An analysis of the research process presentation genres of contrasting academic discourse communities* [PhD thesis, The University of Liverpool]. Livrepository. https://livrepository.liverpool.ac.uk/3175297/1/368678.pdf

Silyn-Roberts, H. (2012). *Writing for science and engineering: Papers, presentations and reports*. Elsevier.

Swales, J. M., & Feak, C. B. (2004). *Academic writing for graduate students: Essential tasks and skills* (2nd ed.). University of Michigan Press.

Thompson, S. E. (1997). *Presenting research: A study of interaction in academic monologue* [PhD thesis, The University of Liverpool]. Livrepository. https://livrepository.liverpool.ac.uk/3175129/1/DX216427.pdf

Ventola, E. (2002). Why and what kind of focus on conference presentations. In E. Ventola, C. Shalom, & S. Thompson (Eds.), *The language of conferencing* (pp. 15–50). Peter Lang.

Webber, P. (2002). The paper is now open for discussion. In E. Ventola, C. Shalom, & S Thompson (Eds.), *The language of conferencing* (pp. 227–253). Peter Lang.

Wulff, S., Swales, J. M., & Keller, K. (2009). "We have about seven minutes for questions": The discussion sessions from a specialized conference. *English for Specific Purposes*, *28*(2), 79–92. https://doi.org/10.1016/j.esp.2008.11.002

AUTHOR

Farnaz Rajabi Mehr is a Lecturer in the School of Computing at Unitec. She is completing her PhD at the University of Auckland. Her research interests are teaching and learning, technology integration in education, and assessment methods. fmehr@unitec.ac.nz

Identifying Challenges with Unifying Information Technology Qualifications Within New Zealand Institutes of Technology and Polytechnics: Project Proposal

Dipendra Ghimire, David Weir and Bernard Otinpong

https://doi.org/10.34074/proc.240116
Correspondence: Dipendra.Ghimire@ara.ac.nz

This publication may be cited as:

Ghimire, D., Weir, D., & Otinpong, B. (2024). Identifying Challenges with Unifying Information Technology Qualifications within New Zealand Institutes of Technology and Polytechnics: Project Proposal. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 120–127). ePress, Unitec. https://doi.org/10.34074/proc.240116

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga – New Zealand Institute of Skills and Technology

ABSTRACT

Unifying qualifications is a complex process. The unification process aims to establish a comprehensive portfolio of programmes that support learners, employers and industry while ensuring sustainable ongoing delivery. This paper presents a literature review that highlights challenges faced in unifying qualifications in other countries. This paper also explores the benefits and challenges of a unified qualification. Specifically focusing on IT qualifications, the study aims to identify the challenges that exist in unifying IT qualifications in Aotearoa New Zealand, and what impact these challenges have on the qualifications. This paper outlines the proposed methodology that will be used for the study and discusses the significance of understanding these challenges and their impact. This research also aims to provide insights and propose a model to address the challenges faced in unifying IT qualifications. This research will contribute to the development of an effective and streamlined qualification system within Te Pūkenga.

KEYWORDS

IT qualifications, unifying qualifications, ITP sector, Te Pūkenga, unification challenges

INTRODUCTION

In addressing the problems in having different qualifications in Institutes of Technology and Polytechnics (ITPs), the government in Aotearoa New Zealand has proposed to create unification across the programmes offered. This process will happen through the formation of one institution, named Te Pūkenga. Within Te Pūkenga, network partners (ITPs and transitional Industry Training Organisations [ITOs]), there are many qualifications that currently have multiple programmes structured and delivered in various ways. Te Pūkenga is taking two approaches to the unification of programmes of study. The first approach is one in which most programmes of study will transition to a current existing programme of study with some minor updates. The second approach is a full transformational redesign and development of programmes of study.

BACKGROUND

In April 2020, Te Pūkenga (New Zealand Institute of Skills and Technology) was established. The objective was to create a wide range of programmes that benefit learners, employers and business (Te Pūkenga, n.d).

Te Pūkenga, proposing strategic direction for innovation in academic delivery, seeks to create a wide range of programmes that assist learners, employers and industry in making knowledgeable decisions about their future education and employment while assuring sustainable continued delivery. To accomplish this, Te Pūkenga Programme Development Excellence Framework, Whiria Te Pūkenga (Mātauranga Māori Framework), and Te Pūkenga Ako Framework are all aligned with the unification process's collaborative approach to programme renewal.

The main goal of unification is to collaborate to create a set of unified programmes that work for all modes of delivery within Te Pūkenga network. Most of the time, current programmes will receive modest adjustments to make sure they are suitable for continuous delivery in a variety of locations and delivery methods. The emphasis is on creating a set of programmes that guarantee efficient and streamlined learner routes are maintained or improved.

The first step in unification is switching to an accepted existing programme, which will then go through additional collaborative redevelopment to make it better. After an initial evaluation, the redevelopment and review process will last for five years. Programme transformation development, which entails the joint reconstruction of a new unifying programme for the network that functions for all delivery modes, may occasionally be initiated.

Programme transformation projects might be needed for a number of reasons: a qualification undergoing a significant change; modifications made by standards-setting bodies; the unification of multiple degree-level qualifications into a single qualification and programme of study; or the pressing need to redevelop a programme due to success and equity issues with the current offerings. The co-operative redevelopment strategy makes sure that the programme design satisfies the needs of all stakeholders and is in line with Te Pūkenga's strategic objectives.

LITERATURE REVIEW

In a review carried on the current state of vocational qualifications in England during 2013, Whitehead (2013) identified several weaknesses, including over-prescriptive design features, a narrow model of occupational competence, poor labour-market returns, and limited employer involvement and ownership. This was based on the review of vocational education for 14–19-year-olds, apprenticeships, and adult vocational teaching and learning. The review notes that some critics believe prescribed qualifications are the main problem. However, others argue that the issue is with the content of the course, its concept of learning outcomes, assessment model and bureaucratic procedures. Whitehead (2013) also notes that vocational qualifications at Level 2 or below often do not attract labour-market return. There are too many vocational qualifications that do not have productive outcomes. Furthermore, the landscape of vocational qualifications is complex and confusing. The pathways from vocational qualifications into higher education tend to be restricted. The review concludes that the current system needs fixing and calls for urgent action to address the identified weaknesses (Whitehead, 2013).

A literature review conducted in Scotland discusses why it has been challenging to put unified qualifications into practice in that country (Raffe, 2007). The author notes that the qualifications have to be relevant to individuals and employers, affordable for individuals, rigorous and based on occupational standards. These qualifications need to be designed and assessed by the sector; qualifications developed must be worthy of the investment, giving a clear signal of the economically valuable skills, and must be useful for current and future occupations. The author suggests that more effective reforms may require changes in the way policies are made and implemented. There should be a change in governance and power distribution in education, training and skills. Overall, the review raises questions about the challenges of translating general principles into concrete actions in policymaking. The study makes the following recommendations:

- The Office of Qualifications and Examinations Regulation should require awarding organisations to engage actively with employers in sectors where they provide vocational qualifications; and awarding organisations should require training providers offering their vocational qualifications to engage with employers locally.
- Encourage more leading employers to work in partnership to develop recognised, rigorous and relevant vocational qualifications.
- The UK Commission for Employment and Skills should work with employers to agree on the future model for occupational standards; and employer-led collaborations (industrial partnerships) should ensure that future occupational standards articulate their ambition and aspiration for their workforce clearly and effectively.
- Awarding organisations and training providers should report on the impact of their vocational qualifications to their customers.
- Ofqual, with the SFA and UK Commission, should create a single point of access to the different qualifications databases.

Raffe's (2007) literature review discusses why it has taken Scotland some time to introduce a viable and effective National Qualifications Framework (NQF). The review identifies internal and external pressures for a review. Internal pressures for reform in Scotland were the strongest drivers for change, including pressures to increase control and co-ordination, arising from changes in the education system in the 1980s and 1990s. External democratic and economic pressures for an NQF were weak, and there was no vision of a unified outcomes-based NQF until the mid-1990s. However, the review notes that Scotland enjoyed advantages in developing an NQF. The review highlights that NQFs introduce a language and operational logic that can take a long time to learn and embed in educational practice. The Scottish experience suggests that other countries may make faster progress in introducing an NQF, but any process involving learning the language, changing institutional logics and developing trust is likely to take time. Raffe (2007) has identified some benefits and challenges with the unified qualification.

Benefits

Raffe (2007) summarises the benefits of the unified qualification system introduced in Scotland in 1999, called Higher Still. The reform was aimed at providing equal opportunities for all, improving collaboration between schools and colleges, and creating a common currency for qualifications. Overall, the following benefits could be achieved:

- Flexibility in implementing different curricular models.
- Quality education for learners at all levels, including 16-year-olds staying in school, individuals with learning difficulties, and those with special needs.
- The change strategy adopted a consensual and evolutionary approach, with priority given to academic school courses during implementation.
- Higher Still was designed and implemented in a way that did not pose a significant threat to those already advantaged by the existing system.

Challenges

The Higher Still reform did not fully replace college-based vocational programmes. Furthermore, this reform did not expand to include work-based programmes as had been initially planned. To address this, the Scottish Credit and Qualifications Framework (SCQF) was developed (Raffe, 2007).

Higher Still was designed to be a flexible model of a unified system, and schools and colleges took advantage of this flexibility by implementing different curricular models. However, this flexibility reduced the tightness of the qualifications framework and created tension between the framework's tightness and coverage (Raffe, 2007).

Higher Still's main aim was to provide 'opportunity for all' by offering well-regarded provisions for learners at all levels of prior accomplishment. While it achieved this, it did not establish an even gradient of progression. Raffe (2007) noted some specific design issues, practicalities of delivery, perceived unsuitability of external assessment for some college students, and low priority for vocational courses in implementation. The reform faced political barriers.

The Irish qualification framework is regarded as a descriptive framework and serves as a comprehensive endeavour covering all the educational and training qualifications in Ireland (Allais, 2017). Its primary objective is to enhance the coherence of, and interconnect qualifications within, the educational landscape. This framework has gained a certain level of achievement; however, the implementation of this framework has come through some challenges. These are particularly in the transformation of institutional structures, regulations and organisation governing the framework itself. These alterations have introduced a degree of instability and uncertainty into the system (Allais, 2017).

IT QUALIFICATIONS IN AOTEAROA NEW ZEALAND ITPS

Te Pūkenga (New Zealand Institute of Skills and Technology) offers a range of IT qualifications. Currently, different ITPs offer different version of IT qualifications. These qualifications range from Level 1 to Level 10, as defined by the New Zealand Qualifications Authority (NZQA). Table 1 presents the NZQA level, and some examples of the IT qualifications currently delivered.

Table 1. NZQF structure – levels and qualifications.

Level	Qualification Types	Qualifications
10	Doctoral degrees	Doctor of Computing Doctor of Professional Practice
9	Master's degrees	Master of Applied Technologies – Computing Master of Information Technology
8	Postgraduate diplomas Postgraduate certificates Bachelor honours degrees	Postgraduate Certificate of Applied Technologies – Computing Postgraduate Diploma of Applied Technologies – Computing Postgraduate Diploma of Applied Technologies Postgraduate Diploma in Information Technology Postgraduate Certificate in Information Technology
7	Bachelor's degrees Graduate diplomas Graduate certificates	Bachelor of Computing Systems Bachelor of Information and Communication Technologies Bachelor of Information Technology Bachelor of Digital Technologies Graduate Diploma in Computing Graduate Diploma in Information and Communication Technologies Graduate Diploma in Information Technology Graduate Diploma in Data Analytics (Level 7) Graduate Diploma in Networking (Level 7) Graduate Diploma in Software and Web Development (Level 7) Graduate Certificate in Information Technology (Level 7)
6	Diplomas	New Zealand Diploma in Cybersecurity (Level 6) New Zealand Diploma in Systems Administration New Zealand Diploma in Software Development (Level 6) New Zealand Diploma in Information Systems (Level 5) New Zealand Diploma in Information Technology Technical Support (Level 5) New Zealand Diploma in Web Development and Design (Level 5) New Zealand Certificate in Information Technology (Level 5)
5		
4 3 2 1	Certificates	New Zealand Certificate in Information Technology (Essentials) (Level 4) New Zealand Certificate in Study and Career Preparation (Level 3) New Zealand Certificate in Computing (Intermediate User)

The aim of Te Pūkenga is to create unified qualifications (including IT) across the different subsidiaries. Creating unified IT qualifications is a major undertaking. The review of literature suggests that there are challenges in unifying the qualifications.

Therefore, the proposed research question is:

> To what extent do challenges exist in unifying the IT qualifications in Aotearoa New Zealand, and what impact do these challenges have on the qualifications?

The purpose of this paper is to give an overview of a proposed study investigating challenges arising in unifying IT qualifications and the impacts on qualifications. This investigation is important in order to understand the impact of the challenges. The study will examine the qualifications that have been unified or are in the process of unification. In particular, this study will focus on IT qualifications. With the identification of challenges, a model will be proposed to address those challenges. The methods used to conduct the research and to analyse the findings are discussed in following the sections.

METHOD

To answer the research question, an exploration of the views and experiences of participants working in Institutes of Technology and Polytechnics is required. Therefore, it is necessary to both explore and explain, rather than only explore or only explain, the experiences of participants. In order to obtain detailed and in-depth answers for the research question, a mixed-methods approach is proposed. In this approach both quantitative and qualitative data are collected, based on the research question (Creswell, 2011).

A mixed-methods approach is used as a strategy whereby the strength of a questionnaire can be enhanced by understanding the information in detail through semi-structured interviews (Creswell, 2011). This study will consist of a questionnaire that aims to gain an understanding of the unification process used, and the challenges that occurred during the unifying of the qualification, combined with a semi-structured interview. A questionnaire was chosen as it provides a quick, inexpensive and accurate means of accessing information (Groves et al., 2011). Mixed methods can be used to improve the power of the analysis by combining qualitative and quantitative techniques (Johnson & Onwuegbuzie, 2004). For example, mixed methods help in developing context-specific instruments and provide a complete and broad understanding of the research problems (Johnson & Onwuegbuzie, 2004)

Questionnaire

The initial data collection will be undertaken by the sampling of different roles (managers, teaching staff, administrative staff) in Te Pūkenga using a survey questionnaire. The questionnaire will include demographic and experience questions that explore the extent to which participants have experienced challenges, and the details of the challenges they have experienced.

Questionnaires allow researchers to evaluate variables that may not be directly observable, such as attitude or perceptions of the participants. Examples of questions to be included are:

- What is your current job title?
- How long have you been involved in teaching IT courses?
- In your opinion, what are the biggest challenges that might be faced in implementing a unified IT qualification in New Zealand?
- How do you think unified qualifications will impact students' learning experience and outcomes?
- How will unified qualifications affect the alignment between educational programmes and industry needs?

Interviews

Interviews, which involve in-depth exchange between researchers and researched participants, are often presented as the 'gold standard' of qualitative research. It is proposed that for this study a representative sample of participants from all roles from the Ako Delivery ICT Division in Te Pūkenga who have completed the questionnaire are interviewed. Interview questions will expand on the questions asked in the questionnaire, specifically those related to the qualification.

Analysis Process

The quantitative data will be analysed using a combination of descriptive methods. Statistics such as minimum, maximum, mean and standard deviation will be used. Data from the discrete scale questions will be analysed using appropriate statistical techniques, such as a regression analysis. Qualitative data collected from interviews will be transcribed verbatim and analysed using thematic analysis. Thematic analysis is a way of identifying, analysing and reporting distinguishable and repeated patterns often found in interview data. Themes are the patterns across the entire set of data that can be used to answer research questions and address hypotheses (Braun & Clarke, 2006). To determine themes, data from the interviews will be transcribed and stored in Microsoft Word. The responses from the respondents will be read through to gain an understanding of the information. Possible codes will be identified through reading the responses. After the initial coding of the data, a complete list of codes will be created and restructured to group similar codes. The codes will be revisited, and new themes will be identified. These similar codes will then be compared, reviewed and combined with the initial codes to form comprehensive themes.

In quantitative research, a researcher can apply statistical methods to establish validation (Corbin & Strauss, 1990). However, in qualitative research, methodological strategies are designed to ensure the 'trustworthiness' of the findings (Noble & Smith, 2015). A researcher will adopt different validity options that are considered to be appropriate, using terms such as quality, rigour and trustworthiness (Davies & Dodd, 2002). To maintain the validity of the data, a researcher needs to remain true to the respondents by ensuring the information given by them is well presented. For this reason, a number of strategies will be used to ensure rigour in this study. These strategies include:

- Interviews will be recorded and transcribed verbatim.
- Audio recordings will be replayed to check emerging themes and remain true to the respondents, and the coding process used for this study will be described.

FURTHER STUDY

Further study will follow the initial study and confirmation of challenges in unifying IT qualifications. This will use co-ordination theory, which is based on organisational theory, and has been applied in fields such as economics and computer science (Malone & Crowston, 1990). Co-ordination theory is constructed on the principle that activities can be co-ordinated between actors (Malone & Crowston, 1990), where actors may be the stakeholders, managers or staff working under the management. This is a descriptive-based theory that is used to understand the particular activities in organisational settings.

REFERENCES

Allais, S. (2017). Labour market outcomes of national qualifications frameworks in six countries. *Journal of Education and Work*, *30*(5), 457–470. https://doi.org/10.1080/13639080.2016.1243232

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://www.tandfonline.com/doi/abs/10.1191/1478088706QP063OA

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, *13*(1), 3–21. https://link.springer.com/article/10.1007/bf00988593

Creswell, J. W. (2011). Controversies in mixed methods research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (pp. 269–284). Sage Publications.

Davies, D., & Dodd, J. (2002). Qualitative research and the question of rigor. *Qualitative Health Research*, *12*(2), 279–289. https://doi.org/10.1177/104973230201200211

Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology*. John Wiley & Sons.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14–26. https://doi.org/10.3102/0013189X033007014

Malone, T. W., & Crowston, K. (1990). What is coordination theory and how can it help design cooperative work systems? In *Proceedings of the 1990 ACM Conference on Computer-Supported Cooperative Work* (pp. 357–370). Association for Computing Machinery. https://doi.org/10.1145/99332.99367

Noble, H., & Smith, J. (2015). Issues of validity and reliability in qualitative research. *Evidence-Based Nursing*, *18*(2), 34–35. https://doi.org/10.1136/eb-2015-102054

Raffe, D. (2007). Making haste slowly: The evolution of a unified qualifications framework in Scotland. *European Journal of Education*, *42*(4), 485–502. https://doi.org/10.1111/j.1465-3435.2007.00322.x

Te Pūkenga. (n.d.). *Mō Te Pūkenga. About us*. https://www.xn--tepkenga-szb.ac.nz/our-work/about/

Whitehead, N. (2013). *Review of adult vocational qualifications in England*. UK Commission for Employment and Skills. https://assets.publishing.service.gov.uk/media/5a7dfe0f40f0b6230268877f/review-of-adult-vocational-qualifications-in-england-final.pdf

AUTHORS

Dr Dipendra Ghimire is a Lecturer in the ICT Portfolio at the Ara Institute of Canterbury. His research interests are in software engineering, educational technologies, and teaching and learning. dipendra.ghimire@ara.ac.nz

Dr David Weir is a Principal Lecturer in the ICT Portfolio at the Ara Institute of Canterbury. His research interests are in the use of technology to enhance learner engagement and the challenges in preparing ākonga for work integrated learning (WIL) in the IT sector. david.weir@ara.ac.nz

Dr Bernard Otinpong is a Senior Lecturer in the ICT Portfolio at the Ara Institute of Canterbury. His research interests are in data analytics, information visualisation, simulation and modelling. bernard.otinpong@ara.ac.nz

Professions in the Data Analytics Job Market – Australia and New Zealand

George Huang, Emre Erturk and Farhad Mehdipour

https://doi.org/10.34074/proc.240117
Correspondence: Emre Erturk, eerturk@eit.ac.nz

This publication may be cited as:

Huang, G., Erturk, E., & Mehdipour, F. (2024). Professions in the Data Analytics Job Market – Australia and New Zealand. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 128–137). ePress, Unitec. https://doi.org/10.34074/proc.240117

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga –
New Zealand Institute of Skills and Technology

ABSTRACT

This article investigates the data analytics job market in Australia and Aotearoa New Zealand with a focus on finding the major professions that exist and how they are distinguished one from another in technical competencies. The main purposes are to inform data analytics curriculum design and development to enhance market relevance and graduate employability, as well as to provide advice to assist quality assurance for the associated professional course accreditations. Currently, data analytics, either as a course or as a specialised programme, is offered by different schools and in different disciplines. The variances and discrepancies among them are large and concerning. This article is the first part and partial literature review of a larger study, which later includes text mining and examination (using R) of 540 job advertisements in five industries that demand data analytics skills the most. It is found that business analyst, data analyst and data scientist are three major professions in Australia and Aotearoa New Zealand. The larger study fills a gap in the data analytics literature by focusing on countries other than the United States. It also offers key specific competencies for universities and the professional accrediting bodies to focus on when developing and evaluating data analytic courses and programmes. The article will benefit students, practitioners, employers and other parties interested in achieving better outcomes for professional development, employment and economics in capitalising values, capabilities and potentials pertaining to data analytics.

KEYWORDS

Business analyst, data analyst, data scientist, technical competencies, text mining

INTRODUCTION

A competency-based curriculum is promoted by global IT professional bodies such as ACM, AIS and IEEE, and universities that provide IT and IS programmes (The Joint ACM & AIS IS2020 Task Force, 2021; CC2020 Task Force, 2021). A curriculum with competency focus may bridge the gap between education and the professional employment market, create market-relevant courses and enhance graduate employability (Topi, 2019). To achieve a curriculum as such, knowledge of what professions exist in the job market and how they are distinguished or substantiated in technical competencies must be known.

In the case of data analytics, large variances and discrepancies exist in the curricula of different universities and in different disciplines (Delen & Zolbanin, 2018; Weathers & Aragón, 2019; Sarkar et al., 2021). The situation is concerning, as it may confuse educators and developers in course and programme design, undermine the standard of curriculum quality-assurance for the related professional course accreditations, and disorient students, professional workers, employers and other parties (Clayton & Clopton, 2019; Radovilsky et al., 2020).

Data analytics has long been hosted in the IT discipline, and a dominant share of data analytics jobs are taken by the IT industry (De Mauro et al., 2018; Halwani et al., 2021). The hosting status of IT is being challenged by other disciplines, as they are offering data analytics in various courses and programmes. The data analytics curriculum has to cover a wide range of topics, content and competencies to accommodate the multi-disciplinary or interdisciplinary needs in the job market. Given that a course or a programme has a constrained timeframe, limited resources and facilities, it is not practically possible to cover all that is needed (Clayton & Clopton, 2019). For a course or a programme to enhance graduate employability and be market relevant, it must target a profession. Therefore, a clear understanding of what the major professions are in the job market and how each profession is distinguished from another one is important.

In the literature, there is a large body of studies reflecting professions and competencies. Most of them are conducted in the US, while less is known about other countries such as Australia and Aotearoa New Zealand. Studies are required so that data analytics courses and programmes can be offered to accommodate the local employment market and economy.

This study will endeavour to answer three questions:

1. What are the professions that exist in the data analytics job market in Australia and Aotearoa New Zealand?
2. What are the important technical competencies that identify each profession?
3. How are the professions distinguished from one another in technical competencies?

The reason for Australia and Aotearoa New Zealand being studied together is that the two countries share the same employment market. Professional workers move between the countries freely so they can work in either country with no legal restrictions. Many large companies have their headquarters in one country and have branches and subsidiaries in both countries.

LITERATURE REVIEW

This section presents an analytical review of the literature towards discovering professions and technical competencies in relation to data analytics.

Data, Data Analytics and Competency

Data, data analytics and competency are three concepts that are instrumental to this study. They are particularly used to inform data collection, analysis and interpretation. Data is thought of as having started from generic forms in statistics, mainly numerical and categorical, and in computer science. As information technology evolves, data grows in all Seven Vs (volume, velocity, variety, variability, veracity, visualisation and value) (Aasheim et al., 2015). Data analytics grows from three typical analyses (descriptive, predicative and prescriptive) to a range of mechanisms and capabilities, including artificial intelligence, machine learning, data mining, automation, and streaming analytics in information and cyber-physical systems (Chang et al., 2018; Wang et al., 2018; Radovilsky et al., 2020; Sun & Huo, 2021).

The definition of competency also varies. Among various conceptual frameworks, the common ones are KSD (knowledge, skills and dispositions), KSA (knowledge, skills and abilities) and CFS (cognitive, functional and social) in the fields of IT and data analytics research (ACM Task Group on Information Technology Curricula, 2017; Dong & Triche, 2020). Competency is captured in two categories: technical skills and soft skills (Radovilsky et al., 2020). Soft skills, such as communication, teamwork and collaboration, are required for all professions today and are unlikely to distinguish one profession from another (Radovilsky et al., 2020; Verma et al., 2019). For entry-level positions that apply to most university graduates, technical skills (what a graduate can do) are more important for employers (Kurtzke & Setkute, 2021).

Home of Data Analytics

Data analytics is demanded by many professions and its courses and programmes are offered in many disciplines. This raises a question as to which discipline it belongs or whether it should be hosted by IT. If hosting, IT should educate in the basics of data analytics and cover the various analytics needs. A review of the historical literature reveals that data analytics has been a major domain in the IT curriculum since the 1990s (Kang et al., 2018). The view that IT is the host discipline of data analytics is also due to IT placing a higher emphasis on the professional curriculum or accreditation guidelines. For example, in IT 2107 (ACM Task Group on Information Technology Curricula, 2017), data analytics is a primary competency, while in the 2021 Accounting Professional Accreditation Guidelines (CPA & CAANZ, 2021) it is treated as an item for enrichment.

The scope of data analytics is expanding and becoming more multi-disciplinary or interdisciplinary. This is caused by the growing demands from various industries but, mainly, in the broad business category (Mamonov et al., 2015). The demands are so huge that they cannot be satisfied only by the expansion of IT. Many disciplines have started offering their own data analytics courses in various forms, such as marketing analytics, financial informatics and business analytics, and this trend is continuing (Radovilsky et al., 2020). In the data analytics job market, professions are not in a binary division (IT versus non-IT) anymore, but have now possibly adopted a transformed structure. In the literature that studies professions in association with data analytics, professions are varied and are identified in different ways. For example, De Mauro et al. (2018) identify the professions as business analyst, data scientist, developer and engineer, while Verma et al. (2019) identify them as business analyst, business intelligence analyst, data analyst and data scientist. These variations confuse educators, developers and evaluators when they are trying to design and develop a data analytics curriculum that targets the professional market. Therefore, finding the major professions in the job market and identifying the important competencies that distinguish each profession are crucial.

Business analyst, data analyst and data scientist appear to be three professions that are commonly found in the previous studies (Wilder & Ozgur, 2015; Mamonov et al., 2015; Halwani et al., 2021). These studies share a view that business analyst professionals graduate from the traditional business schools, possessing competencies such as consumer behaviour analysis, sales forecasting and financial fraud detection; data analyst professionals grow from the traditional IT schools (more specifically, in the IS specialisation), featuring skills such as database management, programming and application of infrastructures (distributed, cloud, etc.); and, lastly, data scientist professionals expand from the traditional applied mathematics or statistics schools, demonstrating capabilities such as data description, statistical analysis, and algorithmic and intelligence modelling. The professionals' titles are sometimes used interchangeably. For example, data scientist is used interchangeably with data analyst by De Mauro et al. (2018), but it is identified as an umbrella term, including data analyst, in Dong and Triche (2020). In Aasheim et al.'s opinion (2015), both data analyst and data scientist are developed from IT schools. The composite set of competencies, skills and capabilities that are contributed from all the three professions largely remains the same across the various studies (De Mauro et al., 2018; Halwani et al., 2021).

Professions in Further Examination

There are a number of studies that reveal how professions are formed and divided in the data analytics job market. Table 1 shows the professions, and how they are substantiated in competency focuses and emphasised skills that are reported in this brief literature review.

Table 1. Professions, competency focus, and emphasised skills identified in previous studies.

Author (year)	Data	Study context	Division of professions	Competency focus	Emphasised skills
Wilder & Ozgur, 2015	University course descriptions	U.S.	Data scientist	Quantitative expertise	Development of models and communication of results
			Data specialist	Data management expertise	Accessing and displaying data
Aasheim et al., 2015	University course descriptions	U.S.	Data analyst	Apply data warehousing and programming in business	Application of data mining
De Mauo et al., 2017	Job advertisements	U.S.	Data scientist	Apply data analytics and methods to insights	Data methods, understanding of data warehouse queries
			Developer	Design, develop and modify data-reliant application software	Coding, cloud computing and distributed technology
			Engineer	Build and maintain full technology infrastructure	Data architecture and enterprise eco-systems
			Business analyst	Transform insights to business impact	Manage projects and create business impact
Radovilsky et al., 2020	Job advertisements	U.S.	Data science	Apply algorithms and programming language	Machine learning, programming, and big data technology
			Business analytics	Database design and big data technology application	SQL, python, statistics in design, processing and analysis
Verma et al., 2019	Job advertisements	U.S.	Data scientist	Apply computer science to analyse big data sets	Statistical and programming skills
			Data analyst	Apply statistical analysis and business acumen to big data	Communication and statistical skills
			Business intelligence analyst	Use computer-based tools to solve business problems	Applying structured data management skills and statistical knowledge
			Business analyst	Disciplinary analysis in business domain	Using domain knowledge and data management skills
Dong & Triche, 2020	Job advertisements	U.S.	Data scientist	Find values in data and create data products	Data hacking, analysis, communication and advice
			Data analyst	Clean, transform and model data for decision-making	Recognition of data problems and solving them
			Business intelligence analyst	A range of applications, technologies, and architectures to store, access and analyse data	Using fact-based support systems
Halwani et al., 2021	Job advertisements	U.S.	Data scientist	Mine and analyse data for business, and design architecture for relational databases	Discovering, extracting and analysing data for decision-making and prediction
			Data analyst	Initialise to develop tools, data collection processes and data management systems	Using analytical and technical skills (esp. relational database and data modelling)
			Big data analytics professional	Apply semantic correlation, ontology and text analytics techniques	Developing datasets, automation and data validation processes

There are obvious discrepancies in the competency focuses and emphasised skills for each profession. Data analyst, for instance, as identified by Verma et al. (2019), has expertise in both statistics and business. But, according to Dong and Triche (2020), it specialises in data cleaning, transformation and modelling. In the key skills that are identified for business analyst, De Mauro et al. (2018) emphasise managing projects and making business impacts, while Halwani et al. (2021) highlight creating data-driven strategies and performing cross-functional (visualised) communications.

Competencies in Deeper Examination

Data-analytics related competencies include coding, database management, project management, distributed computing, analytics and business impact; and after adding visualisation, as informed by Chang et al. (2018) and Sun and Huo (2020), a comprehensive set of dimensions is achieved, as seen in Table 2.

Table 2. Data analytics competencies in the data scientist profession.

Author (year)	Coding	Database management	Project management	Distributed Computing	Statistics and algorithms	Business impact	Visualisation
Aasheim et al., 2015	DS	DS			DS		DS
Wilder & Ozgur, 2015					DS		DS
De Mauro et al., 2018		DS	DS		DS	DS	
Radovilsky et atl, 2020	DS		DS	DS	DS		
Halwani et al., 2021					DS	DS	

We examine data scientist (DS) as an example. Coding is claimed for data scientist professionals by Aasheim et al. (2015), Radovilsky et al. (2020) and Verma et al. (2019), while it is not recognised by Wilder and Ozgur (2015). Some skills, such as Microsoft Access, SAP and Congos (for database development and systems management), have declined, while Python, R and Tableau (for data analytics and visualisation) are on the rise (Dong & Triche, 2020).

FINDINGS AND DISCUSSION

Business analyst, data analyst and data scientist are the major professions in the data analytics job market. As shown in Figure 1, the business analyst profession shows the range of technical competencies developed in traditional business schools; the data analyst profession, competencies from IT schools; and the data scientist profession, skills from applied mathematics. In the technical dimension, the business analyst profession takes the lower range, and the data scientist profession takes the higher range. The data analyst profession takes a wider range, from both the business analyst and data scientist areas.

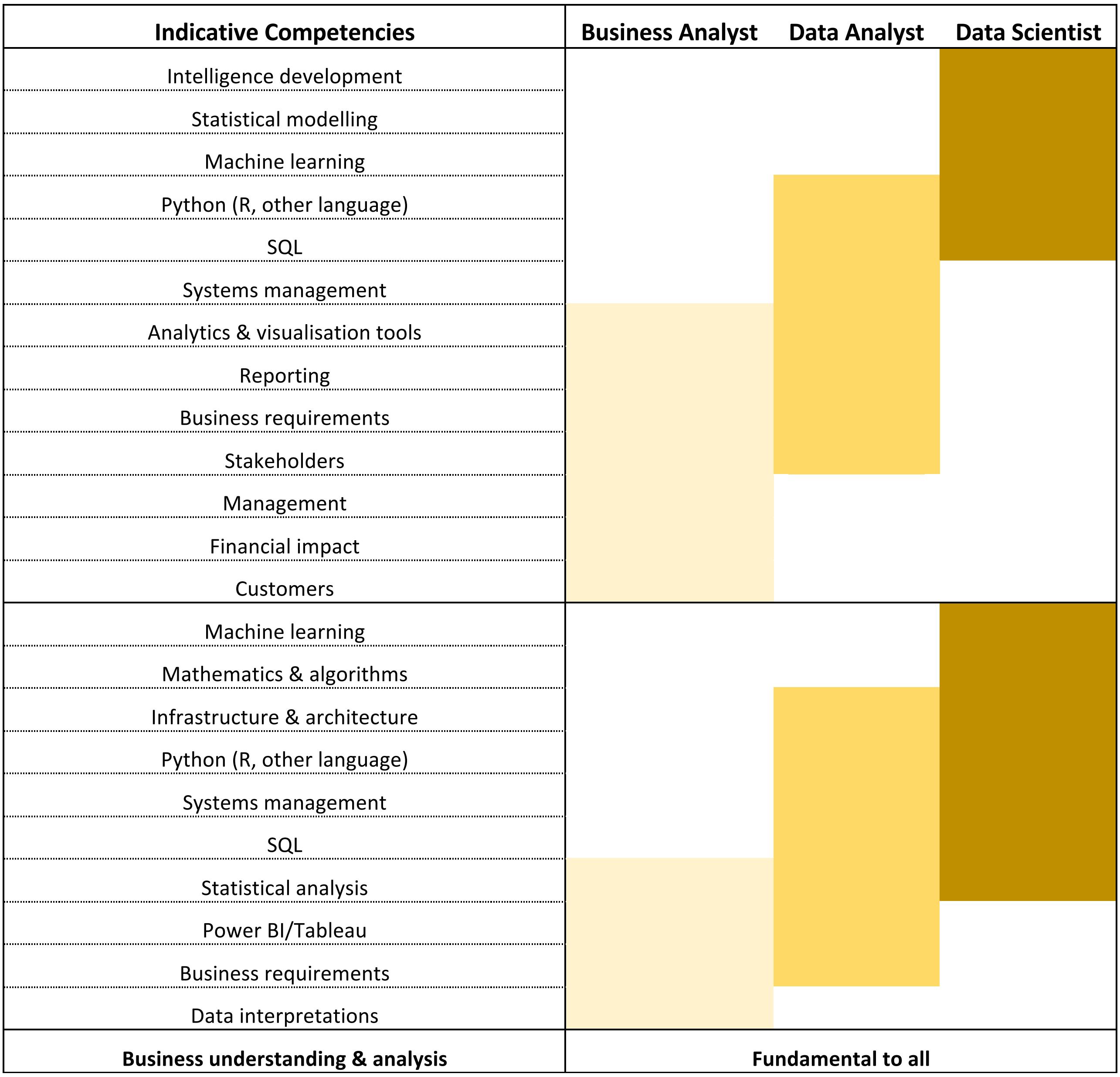


Figure 1. Competencies for the three professions in the data analytics job market.

Data scientist professionals extend their applied mathematical and statistical skills through competencies in programming, algorithmic and intelligence modelling, and machine learning. Apart from roles in database management, software development and infrastructure development, data analysts share some competencies from the business analyst and data scientist areas. They are involved with visualisation, application of systems and tools, programming, and using machine learning algorithms. However, an IT school curriculum should also take in modules from the business and science domains.

In Australia and Aotearoa New Zealand, the data analytics job market is largely distributed in IT and business industries in the order of marketing, accounting, banking and government administration, from large to small, as shown in Figure 2.

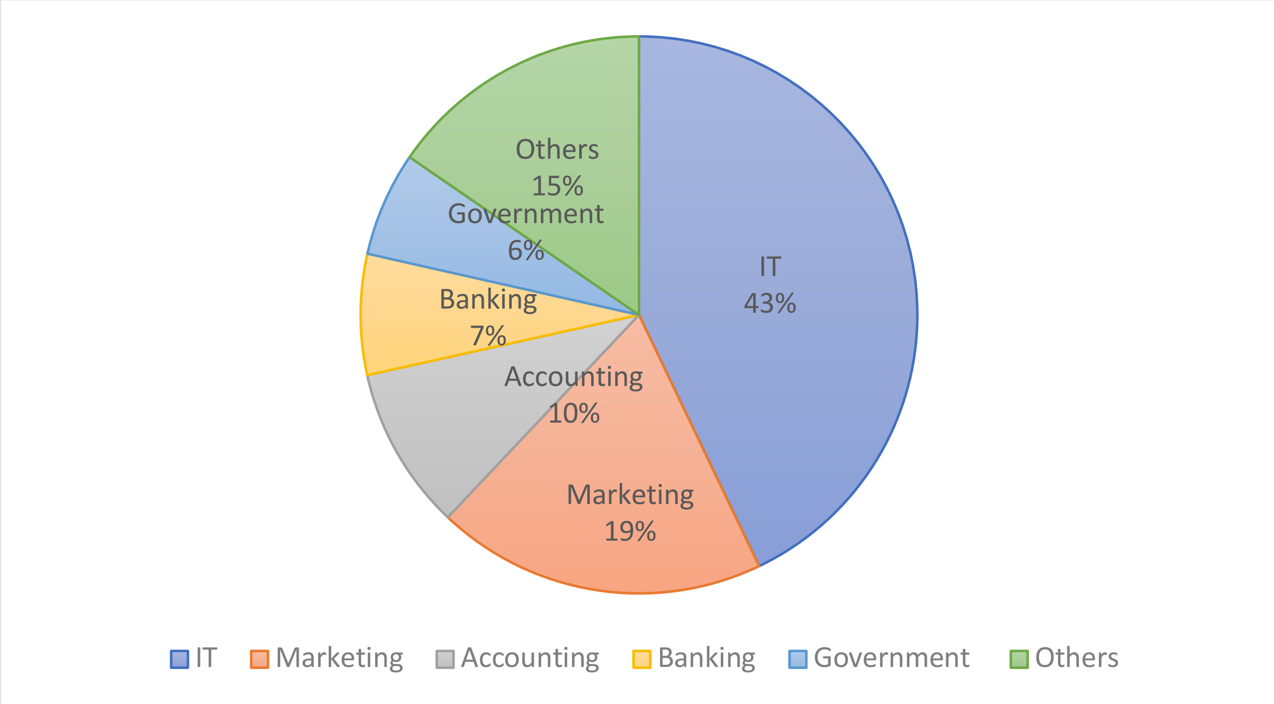


Figure 2. Five industries that demand data analytics, in percentage of the total job number.

Surprisingly, those industries such as healthcare, logistics and mining that are often targeted by many IT and business schools are not in the mainstream job market.

IMPLICATIONS FOR SCHOOLS AND PROFESSIONAL ACCREDITATION BODIES

There are several implications that are derived from this study:

- When developing a programme or a course in data analytics, a school should follow its disciplinary tradition. For example, if it is an IT school, then it should grow the curriculum from its existing IT skillset. In other words, do not forgo the existent repertoire but extend or enrich it.
- Certain competencies, as shown in Figure 1, will still need to be retained to maintain the discipline's core, the curriculum coherence, and to target the job-market niche that is well established.
- No matter what a discipline is, the list of competencies as shown in Figure 1 must be covered (more or less). This is potentially important for a professional accreditation body when accrediting a programme or a course aiming to have a data analytics focus or as a major learning outcome. The range of the competencies that each profession takes should also be referred to by the corresponding professional accreditation bodies.

LIMITATIONS AND DIRECTIONS FOR FUTURE STUDIES

There are certain limitations associated with this study. The most salient ones are:

- Data collected for this study was from a certain period. It may need updating to recognise that competencies for data analytics (as a young professional domain) are still being shaped and reshaped in the job market.
- Methods in this study are based on the literature and the job descriptions, which probably do not fully and correctly reflect the data analytics job reality.

It is suggested the following approaches should be attempted for studies in the future:

- Apply data mining of a larger set of data (data analytics job descriptions) in the job market to refine the distinctive competencies for the different professions.
- Interview employers, educators and practitioners in data analytics. This may obtain more insights about the professions and associated technical competencies.

REFERENCES

Aasheim, C. L., Williams, S., Rutner, P., & Gardiner, A. (2015). Data analytics vs. data science: A study of similarities and differences in undergraduate programs based on course descriptions. *Journal of Information Systems Education*, *26*(2), 103. https://jise.org/Volume26/n2/JISEv26n2p103.pdf

ACM Task Group on Information Technology Curricula. (2017, December 10). *Information technology curricula 2017: Curriculum guidelines for baccalaureate degree programs in information technology.* Association for Computing Machinery and IEEE Computer Society. https://www.acm.org/binaries/content/assets/education/curricula-recommendations/it2017.pdf

CC2020 Task Force. (2021). *Computing curricula 2020*. Association for Computing Machinery. https://doi.org/10.1145/3467967

Chang, H-C., Wang, C-Y., & Hawamdeh, S. (2018). Emerging trends in data analytics and knowledge management job market: Extending KSA framework. *Journal of Knowledge Management, 23*(4). https://www.emerald.com/insight/content/doi/10.1108/JKM-02-2018-0088/full/html

Clayton, P. R., & Clopton, J. (2019). Business curriculum redesign: Integrating data analytics. *Journal of Education for Business*, *94(*1), 57–63. https://www.tandfonline.com/doi/full/10.1080/08832323.2018.1502142

CPA and CAANZ. (2021). *Professional accreditation guidelines*. https://www.cpaaustralia.com.au/become-a-cpa/academic-institution-support/professional-accreditation-guidelines

Delen, D., & Zolbanin, H. M. (2018). The analytics paradigm in business research. *Journal of Business Research*, *90*, 186–195 https://doi.org/10.1016/j.jbusres.2018.05.013

De Mauro, A., Greco, M., Grimaldi, M., & Ritala, P. (2018). Human resources for Big Data professions: A systematic classification of job roles and required skill sets. *Information Processing & Management*, *54*(5), 807–817. https://doi.org/10.1016/j.ipm.2017.05.004

Dong, T., & Triche, J. (2020). A longitudinal analysis of job skills for entry-level data analysts. *Journal of Information Systems Education*, *31*(4), 312. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1862&context=jise

Halwani, M. A., Amirkiaee, S. Y., Evangelopoulos, N., & Prybutok, V. (2021). Job qualifications study for data science and big data professions. *Information Technology & People, 35*(2). https://doi.org/10.1108/ITP-04-2020-0201

Kang, J. W., Yu, Q., Golen, E., & Holden, E. P. (2018). IT curriculum: Coping with technology trends and industry demands. In Ayers, T., & Zilora, S. (Eds.), *SIGITE 18: Proceedings of the 19th Annual SIG Conference on Information Technology Education* (pp. 44–49). Association for Computing Machinery.

Kurtzke, S., & Setkute, J. (2021). Analytics capability in marketing education: A practice-informed model. *Journal of Marketing Education*, *43*(3), 298–316. https://journals.sagepub.com/doi/10.1177/02734753211042404

Mamonov, S., Misra, R., & Jain, R. 2015. Business analytics in practice and in education: A competency-based perspective. *Information Systems Education Journal*, *13*(1), 4–13. https://files.eric.ed.gov/fulltext/EJ1137326.pdf

Radovilsky, Z., Hegde, V., & Damle, L. H. (2020). Data mining in business education: Exploratory analysis of course data and job market requirements. *Journal of Supply Chain and Operations Management*, *18*(1), 119. https://www.csupom.com/uploads/1/1/4/8/114895679/jscom18n1p07_formatted.pdf

Sarkar, S., Boss, S. R., & Gray, J. (2021). Pedagogical practices of accounting departments addressing AACSB technology requirements. *Issues in Accounting Education*, *36*(4), 59–85. https://doi.org/10.2308/ISSUES-19-082

Sun, Z., & Huo, Y. (2021). The spectrum of big data analytics. *Journal of Computer Information Systems*, *61*(2), 154–162. https://doi.org/10.1080/08874417.2019.1571456

The Joint ACM & AIS IS2020 Task Force. (2021). *IS2020: A competency model for undergraduate programs in information systems*. Association for Computing Machinery. https://doi.org/10.1145/3460863

Topi, H. (2019). Reflections on the current state and future of information systems education. *Journal of Information Systems Education*, *30*(1), 1. https://aisel.aisnet.org/jise/vol30/iss1/1/

Verma, A., Yurov, K. M., Lane, P. L., & Yurova, Y. V. (2019). An investigation of skill requirements for business and data analytics positions: A content analysis of job advertisements. *Journal of Education for Business*, *94*(4), 243–250. https://doi.org/10.1080/08832323.2018.1520685

Wang, Y., Chen, Q., Hong, T., & Kang, C. (2018). Review of smart meter data analytics: Applications, methodologies, and challenges. *IEEE Transactions on Smart Grid*, *10*(3), 3125–3148. https://ieeexplore.ieee.org/document/8322199

Weathers, D., & Aragón, O. (2019). Integrating analytics into marketing curricula: Challenges and effective practices for developing six critical competencies. *Marketing Education Review*, *29*(4), 266–282. https://doi.org/10.1080/10528008.2019.1673664

Wilder, C. R., & Ozgur, C. O. (2015). Business analytics curriculum for undergraduate majors. *INFORMS Transactions on Education*, *15*(2), 180–187. https://doi.org/10.1287/ited.2014.0134

AUTHORS

Dr George Huang is currently working at the Australian Technical and Management College. George has great experience in teaching, and in curriculum development and review across many universities and institutes. guozhen.huang@atmc.edu.au

Associate Professor Dr Emre Erturk is from the School of Computing at the Eastern Institute of Technology Te Aho a Māui, Aotearoa New Zealand. Emre is experienced in industry-based research and interested in the continuous improvement of IT education. eerturk@eit.ac.nz

Dr Farhad Mehdipour is a Principal Lecturer of Information Technology at Otago Polytechnic Te Kura Matatini ki Otago, Aotearoa New Zealand. Farhad is an experienced data scientist and teaches data analytics at various institutes, including at postgraduate level. farhad.mehdipour@op.ac.nz

Discovering the Most Needed Technologies and Skills for IT Jobs on LinkedIn

Athavaloshan Thirulingam and Sarita Pais

https://doi.org/10.34074/proc.240118
Correspondence: Athavaloshan Thirulingam, qbz6506@autuni.ac.nz

This publication may be cited as:

Thirulingam, A., & Pais, S. (2024). Discovering the Most Needed Technologies and Skills for IT Jobs on LinkedIn. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 138–149). ePress, Unitec. https://doi.org/10.34074/proc.240118

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga –
New Zealand Institute of Skills and Technology

ABSTRACT

With the rapid pace of technological growth, many information technology (IT) jobs are created every year, but, as IT is a vast field with enormous sub-disciplines, not enough people are trained to match the demand. Educational conferences, research papers and question-and-answer platforms have been studied in the past to map the leading technologies. IT technologies are ever changing and the most recent research is required to identify the trending technologies. This research answers the question of the most sought-after technologies and skills in the current IT job market by analysing job postings on LinkedIn with latent Dirichlet allocation (LDA).

KEYWORDS

Web scraping, latent Dirichlet allocation, topic modelling, automation

INTRODUCTION

Information technology is a dynamic field, and both computer science and information systems are considered under the umbrella term IT in this paper. We are experiencing rapid development in recent times, and, as a result, many sub-disciplines are continuing to emerge in the territory while some become obsolete. With the phase of growth, computer and information technology occupations are created in large numbers and predicted to grow further despite the recent mass layoffs in the IT industry. As a result, the number of students choosing IT-related studies and institutions offering science, technology, engineering and mathematics (STEM) qualifications has been increasing.

IT jobs defined as those that deal with computer programs, hardware, software, networking and maintaining systems are among the best-paying categories of careers. According to the US Bureau of Labor Statistics (2023), IT executives and managers are among the highest-paid employees, and the second-most high-paying jobs are computer and information research scientists and database architects, followed by software developers in US firms.

Further, considering the unemployment rate in the US, while the monthly unemployment rates were between 3.4% and 3.8% in 2023 (U.S. Bureau of Labor Statistics, 2024), the unemployment rate in the tech sector has been at 2.3% in 2023 (Kolakowski, 2024). These statistics indicate that the probability of being unemployed with an IT background is less than the overall average. Furthermore, New Zealand Immigration reports that 4–5000 new digital technology professionals are required per year into the near future (New Zealand Immigration, 2023). With all this evidence we can safely conclude that IT jobs are expected to grow regardless of the recent job cuts in big IT firms such as Amazon, Twitter and Meta.

It is no surprise that students wish to study in a field that promises attractive salaries and career growth. Information technology studies – computer science, information systems and other information technology studies – are popular choices among students. The statistics about education from the New Zealand Government show that the number of students, including international and domestic, enrolling in IT-related studies has increased year by year since 2012 (Education Counts, n.d.). Figure 1 shows the number of students enrolled in IT-related studies from 2012 to 2021.

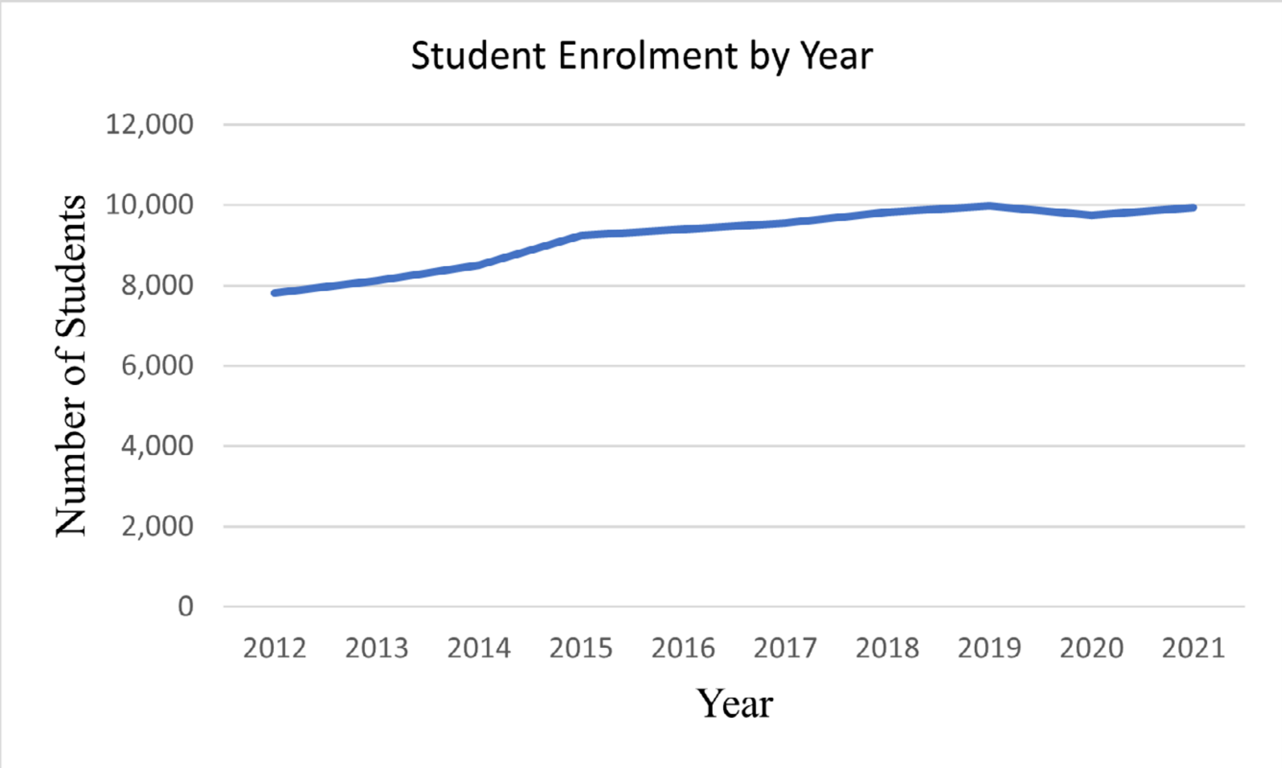


Figure 1. Student enrolment by year.

Though many qualifications are dedicated to a specific technology, such as cybersecurity, networking, data science and software engineering, it is difficult to know which one to choose to maximise the benefits, because they come in many flavours. For instance, cloud computing has three major platforms: Amazon Web Services (AWS), Microsoft Azure and Google Cloud Platform (GCP). One cannot study all the platforms, but must dedicate resources to one or two. Furthermore, software engineering has a diverse mixture of databases, programming languages and frameworks. SQL, PostgreSQL and MongoDB are popular database technologies, while programming languages have plenty of options, such as Python, C++ and Java. Frameworks such as Solid, Bootstrap, React, Ember and Angular are common in product development.

On the other hand, educational organisations also struggle to cope with the rapid growth of technology and diversity. Knowing the prominent technologies will help organisations to better align their curricula. For instance, a university offering blockchain courses might include the following modules: blockchain architecture, cryptography, data structures, smart contracts, programming languages and distributed systems. Choices for a programming language are huge, but knowing what is expected in the job market simplifies the issue.

Furthermore, technology adoption is one of the essential factors in the survival of a business. With the continuous surge of the diverse array of technologies, there is a pressing need to adopt new technologies. One solid piece of evidence is the migration to cloud platforms such as AWS, Azure and GCP by many small and medium businesses.

The job market reflects the demand for skills that employers are seeking. It gives a hint as to what technologies are in demand or being used widely by businesses. Many sites offer the service of matching job seekers with potential employers. Initially, SEEK, a leading employment marketplace, was considered for this study, but the number of vacancies published on this site was lower than on others. LinkedIn job posts are a rich source of information to identify current technologies, and were analysed using a topic modelling algorithm in this research work.

Text mining is a process of automatically extracting useful information from unstructured text corpus such as written resources. Though there are many text mining techniques, topic modelling was the most suited technique for our research. Topic modelling is an unsupervised learning technique to extract latent topics from a collection of documents (Rubin et al., 2012). The unsupervised nature removes the extra effort of labelling documents, which is a time-consuming process of manually analysing documents and assigning topics to each one.

AIMS AND RESEARCH DESIGN

Research question: What are currently the most popular technologies and skills for IT jobs on LinkedIn?

Employment-focused social media platforms connect employers with potential candidates. LinkedIn is a popular employment platform where thousands of jobs are published every day. This rich source of information can be studied to identify in-demand knowledge that can help both students and educational institutions. With the output of the research, students can be more focused on what is required in the job market, while universities and educational institutions can optimise their curricula with prominent technologies. It can even benefit content creators writing IT articles on platforms such as Medium and creating educational videos on streaming platforms such as YouTube, helping them to choose their subjects so they can reach a wider audience.

Most of the relevant research work identifying leading technologies and skills is based on research publications and question-and-answer platforms with text mining techniques, namely semantic analysis and topic modelling, in addition to capturing trends by considering data over a period.

Research publications are a great source of knowledge about technologies that cover a very diverse array of topics while allowing researchers to see the developments over a period. Research (Viet et al., 2021) surveys methods for performing technology trend analysis and forecast (TTAF) using data mining techniques. Scientific publications from SCOPUS over a ten-year timespan were reviewed to discover new and promising technologies using the best method from the results of the survey. Though the research focused on connecting impactful technologies over some time, it did not capture the commercial viability of the research work.

Semantic analysis (Yang et al., 2015) is used in technology trend analysis. Technology information is extracted from science, technology and innovation (ST&I) records to build a Subject-Action-Object structure that emphasises the role of verbs in the study of technology documents. The research can build a map of nodes representing technological factors within a specific technology, enabling the visualisation of the growth of various topics. However, it fails to provide a wide view of the growth and impact of various technologies, as the work only focuses on a specific technology.

Topic modelling-based analysis on question-and-answer websites (Johri & Bansal, 2018) was conducted to find the topics most searched by developers. The research employed a latent Dirichlet allocation (LDA) topic modelling algorithm to extract topics on Stack Overflow. LDA is an unsupervised learning and generative probabilistic model widely used in the field of information retrieval. The input data is treated as a collection of independent words, and it randomly allocates each word with a topic. After several iterations, the topics are finalised. However, the results are mainly bound to programming languages, and related tools and frameworks, and do not cover a wide range of technologies. Further, the topic modelling algorithm creates tags from the given data, which limits the level of granularity in topics. For example, JavaScript, Jquery, HTML and CSS can be grouped to form the topic web development. To resolve the issue, pre-defined labels are used in this work with a text classifier.

Our research took a similar approach to that of Alsmadi and Omar (2022), and Fortino and You (2022), whose data source was job posts on a job board. In Fortino and You's study (2022), 30 job descriptions were scraped from Indeed.com for 30 job titles. The job titles were limited to frequently occurring job titles from the US Bureau of Labor Statistics O*NET occupation database, filtered for STEM IT occupations. Term frequency count was obtained, and standard technologies from O'Reilly's taxonomy were used to categorise the terms in the job descriptions. The output produced a bar chart with technologies outlined in O'Reilly's taxonomy. The technologies analysed were confined to O'Reilly's taxonomy, but it only captured a small proportion of the technical skills. Though some skills and technologies were outlined in the taxonomy they could be lost in the output, as the number of job descriptions considered were low (Fortino & You, 2022). Table 1 shows a summary of the past relevant research work.

Table 1. Literature review.

Title	Type of Data	Data Source	Dataset Size	Methodology	Results
Identifying trends in technologies and programming languages using topic modelling (Johri & Bansal, 2018)	Question and answer website	Stack Overflow	11,097,716	LDA	Topic popularity, topic trends over time
Semantic-based technology trend analysis (Yang et al., 2015)	Science, technology and innovation (ST&I) records	Derwent Innovations Index	7413	Subject-action-object-based semantic analysis	Technology road-mapping
Tracking technology trends using text data mining (Fortino & You, 2022)	Job postings	Indeed.com	900	Term frequency	Technology frequency grouped by IT categories
Analysing the needs of ICT job market in Jordan using a text mining approach (Alsmadi & Omar, 2022)	Job postings	wzfne.com, kalamntina.com, akhtaboot.com, bayt.com	500	LSA	Skills and knowledge
Data mining methods for analysis and forecast of an emerging technology trend (Viet et al., 2021)	Research publications	SCOPUS papers	50	Clustering	Future technology tendencies

ANALYSIS AND FINDINGS

Data Extraction

Data collection is a time-consuming and costly process in many studies, but web scraping has recently become a popular technique to collect required data from web sources. Web scraping is a proven way of collecting data, in which a program accesses the internet and visits web pages to gather the required information (Ruchitaa Raj et al., 2023). A hybrid approach was taken to collect the job postings in our study. An automated Python script was used for extracting job-posting information. Selenium and Beautiful Soup Python libraries were employed for this purpose. Selenium is an excellent choice to control Chrome to get access to web pages. Beautiful Soup serves the purpose of removing HTML tags from the downloaded HTML documents containing information about the job advertisement. As web content is dynamic, a consistent scraping scheme based on tags and attributes is difficult to achieve. Therefore, the task was divided into four steps.

Each job post is associated with a job ID, a numeric value. The first step was manually collecting all the job IDs. When the LinkedIn job page is searched with some keywords and filters, it lists the job titles and the details of the company posting the vacancy in the left-hand panel. At the same time, a job description associated with a job post is also displayed in the main window. Each page may contain 25 job IDs, but they may differ. When we inspected the source code, which is possible in most of the browsers – for example, Google Chrome has the developer tools option, which will display the source code of the webpage – job IDs of all the job posts on a page could be copied at once and stored in a text file to be used in the second stage. To increase the scope of results, keywords were changed after every 250 job posts. The keywords were selected in a way that better represented the IT sub-fields such as networking, blockchain and cloud computing.

The second step was downloading HTML files. Selenium requires a URL to get a web page in which a job vacancy is published. A URL consists of several parameters, such as job ID, keyword, geoid and location. The URL for each job posting was constructed dynamically by changing the parameters.

https://www.linkedin.com/jobs/search/?currentJobId="str(id)"&geoId=91000015&keywords="subfield" &location=Australia%20and%20New%20Zealand&refresh=true

In the above URL, the subfield and ID are variables; currentJobId and keywords are kept dynamic, where the parameter currentJobId is a numeric value, and keywords is a string. Location and geoid are constant, as the research focused on jobs available exclusively in Australia and New Zealand.

The keyword field was filled with various software engineering disciplines so that the data collected represented a wide range of topics. For example, when searching for robotics jobs the keywords field was filled with robotics and currentJobId was filled with an ID collected in Step 1. Step 1 collected 250 IDs for each subfield. The resulting URL looks like this:

https://www.linkedin.com/jobs/search/currentJobId=3595395342&geoId=91000015&keywords=robotics &location=Australia%20and%20New%20Zealand&refresh=true}

The raw HTML file was dumped into a file for future use. Step 3 was intermediate processing to extract job descriptions. The HTML documents downloaded from Step 2 were noisy, with much unwanted information such as metadata and headers. In this stage, the code made use of getting elements using HTML tags and attributes. To identify which tags and attributes encapsulated the job description and job title, a manual examination of the HTML document was required.

It was found that an article tag with the class attribute jobs-description_container contained the job description information, and the t-24.t-bold.jobs-unified-top-card_job-title attribute had the job title, but some of these tags changed over time, as web pages frequently go through modifications. When the tags or attributes changed the results were mostly empty. Inspection was required to ensure the correct extraction of data. A Python code checked the result, and if it was found empty the job ID and subfields were dumped into a file called "errors.txt". This helped to analyse the problem and find new correct tags and attributes.

Stage 4 was removing tags to get a clear text, as a few HTML tags remained after Stage 3. Stage 4 employed Beautiful Soup to get rid of remaining HTML tags with simple regular expressions. Figure 2 illustrates the overall processing pipeline.

A total of 2264 job posts were extracted, which is lower than expected, as some fields such as edge computing and metaverse did not have enough job vacancies in Australia and New Zealand.

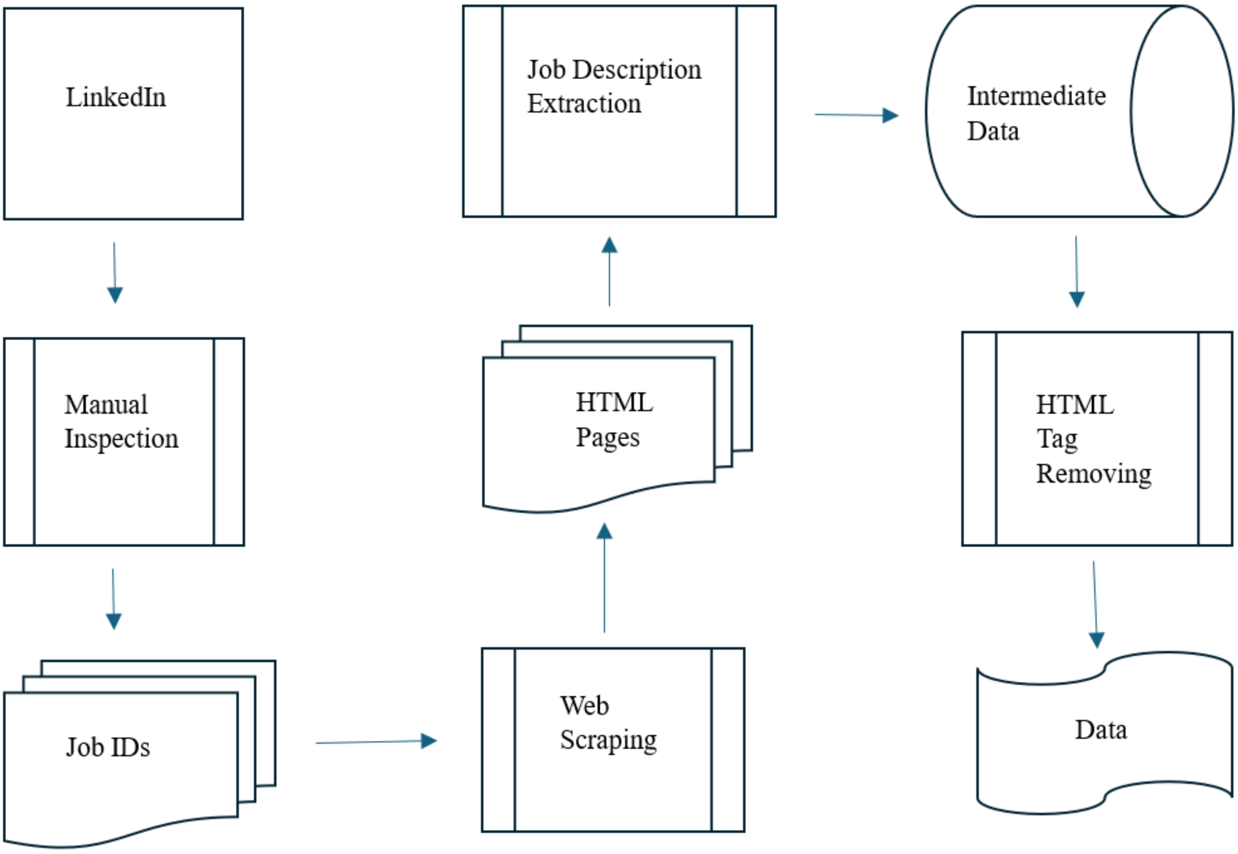


Figure 2. Data-collection process.

Data Cleaning

Data cleaning plays a major role in the accuracy of models as it helps to reduce the feature map dimension. Data visualisation can be used to explore data, and Wordnet is a visualisation technique extensively used in text processing (Sendhilkumar et al., 2017). Since LDA is vulnerable to high-frequency terms, it is better to remove stop words. Removing stop words, numbers, and punctuation with regular expressions is a fast method. In this study, Python regular expression (RE) library was used for this purpose.

Classification and Topic Extraction

The process of classification and topic extraction is divided into three parts: pre-processing, vectorisation and topic modelling. Pre-processing involves cleaning the data, as it contains special characters and emojis.

Vectorisation is an inevitable part of the process, as the LDA model can only understand numeric data. The bag-of-words model, term frequency and term frequency-inverse document frequency (TF-IDF) can be used to vectorise text. Bag-of-words is a very simple model that keeps the count of each term that occurs in the document collection. Term frequency is more complicated than bag-of-words, because it divides the term frequency with the total number of terms in the document so that common terms in the documents are assigned with a small weight. However, it does not capture the term rarity across the entire dataset. TF-IDF was proposed as a solution for the issue in this study. Therefore, the TF-IDF vectoriser, available in scikit-learn, was chosen for vectorisation.

The vectoriser offers many parameters. The parameters max_df and min_df can be fine-tuned to improve the output. The max_df parameter serves as an upper bound for document frequency, while min_df acts as a lower bound, which ignores terms with less value of document frequency than the min_df when building the vocabulary.

A vast amount of unlabelled data can be analysed with topic modelling to unearth the underlying topics. Latent semantic analysis (LSA), probabilistic latent semantic analysis (PLSA) and LDA are some relevant techniques (Liu & Tang, 2018). LSA is a fundamental technique in which the collection of documents is transformed into vectors using TF-IDF, and the resultant sparse matrix is processed with truncated singular value decomposition (SVD) to reduce the dimension. Subsequently, applying a similarity calculation such as cosine similarity will lead to dominant topics in the documents. But the LSA technique has some limitations, due to the spare matrix and subsequent dimension reduction step forcing symmetry in the similarity of words and imposition of the triangle inequality. PLSA is a variant of LSA, but it takes a probabilistic approach instead of SVD.

LDA is a widely used model for topic modelling, which can produce more accurate results than previously discussed methods, and it is possible to tune the number of topics and the number of top words. Gensim and scikit-learn offer LDA implementation, but Gensim has poor documentation that limits experimenting with various parameters to achieve desired accuracy. Scikit-learn has good documentation with many parameters for tuning – hence the LatentDirichletAllocation function from scikit-learn was chosen for this study.

When the whole data was fed into the model, the topics were inaccurate because the feature words did not represent the data. Therefore, the data was split into 20 batches and each batch was treated as an independent data collection. Ten topics with 20 top words were used and the output was manually labelled from the pre-defined label set. The label set was: cloud computing, edge computing, IoT, big data and analytics, machine learning, cybersecurity, metaverse, digital twin, quantum computing, high-performance computing, computer vision and pattern recognition, computational biology and bioinformatics, robotics, networking, blockchain. Figure 4 and Figure 5 were obtained by plotting the topic scores against the topics. Topic scores are the output of the LDA function.

The dataset consists of 2247 job posts collected from LinkedIn. Figure 3 shows the major sub-disciplines in IT. Cloud computing and networking are the major sources of employment in IT. On the other hand, metaverse, digital twin and cloud computing have a smaller number of opportunities.

Among the ten topics, the top 20 words were considered to find leading technologies and skills in the specific sub-discipline. Figure 4 shows the important subjects in networking.

Figure 5 shows the leading technologies and skills in cloud computing. AWS is the leading platform for cloud computing, followed by Azure, and knowledge about AWS services is also desired, especially KMS, SQS, CDK and SNS.

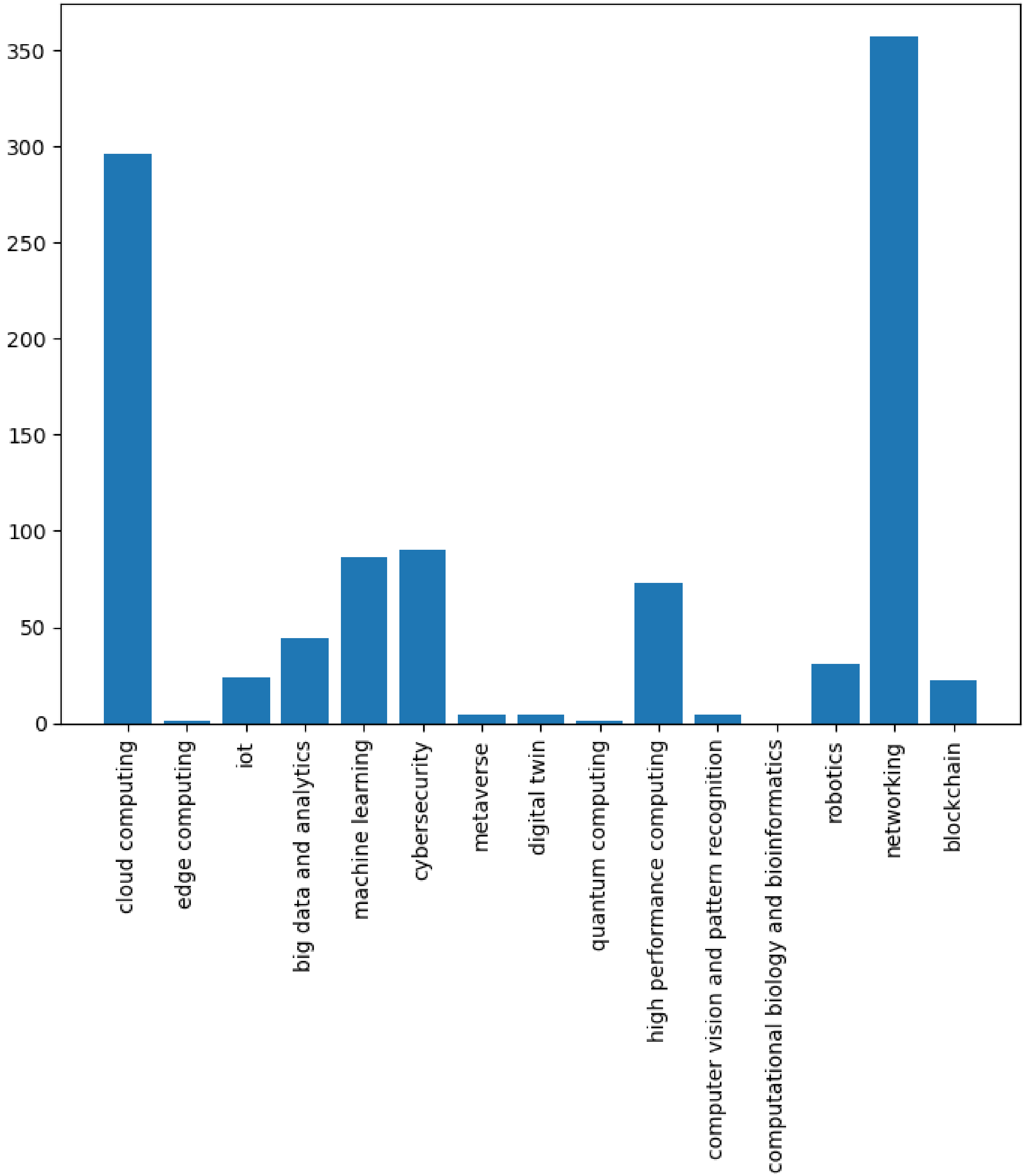


Figure 3. Number of jobs available in each field.

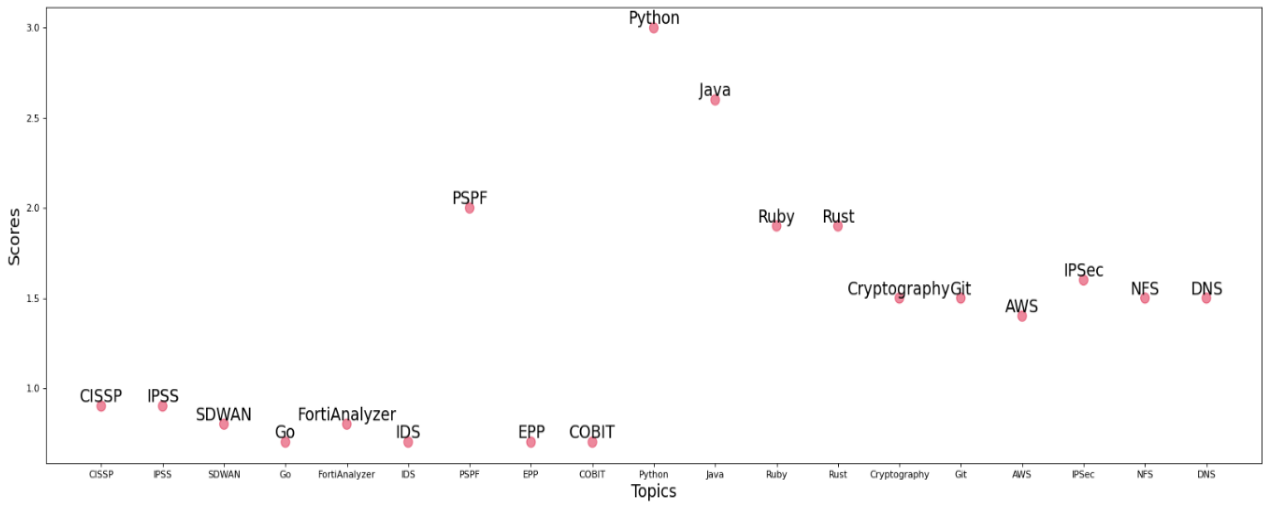


Figure 4. Dominant topics in networking.

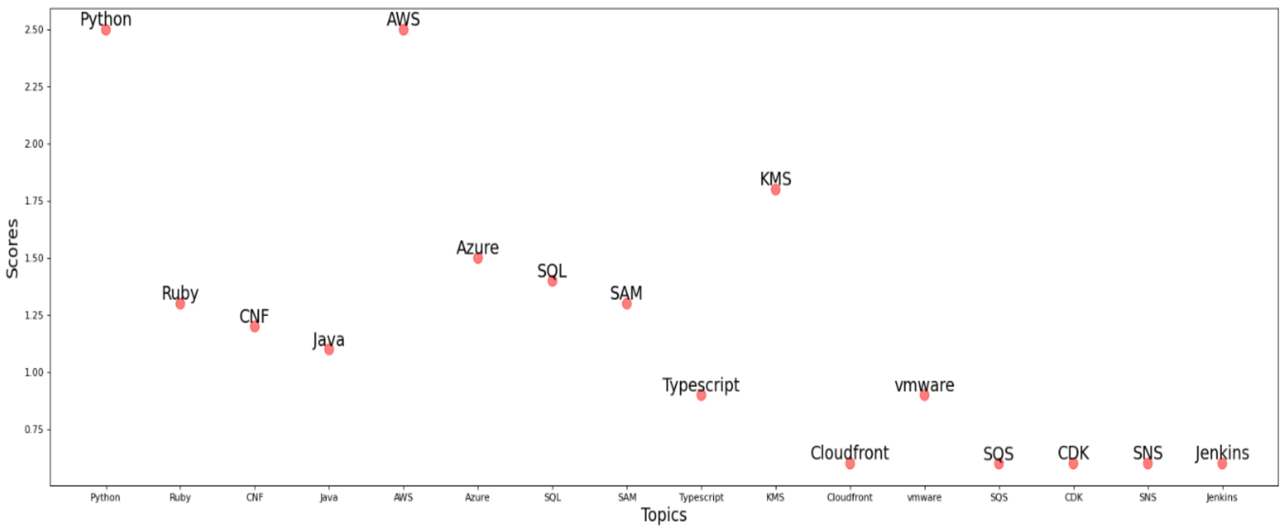


Figure 5. Dominant topics in cloud computing.

CONCLUSION AND RECOMMENDATIONS

In conclusion, our approach has proved to be efficient in analysing jobs posted on LinkedIn in 2023 to determine the main technologies and skills required in the job market, in addition to automated data collection. LDA was used to unearth latent topics from job postings, and web scraping with Python libraries was demonstrated to be an effective way to harvest data on the internet.

Our analysis primarily depended on topic extraction, but some parts of job posts shadow the important technical details. Removing some sections from job posts, such as job role, company name and description about the company, reduced the feature dimension and improved the accuracy of the LDA. The number of jobs posted for some categories was far fewer than for others, which created an inherited bias when extracting topics.

Though some fields such as edge computing and blockchain have fewer vacancies, there may be fewer qualified professionals who can fill these positions, compared to other fields where the supply of engineers may surpass the demand. Capturing this gap requires additional information. Adding this information to our current work will produce more valuable insights.

Developing a user-friendly package that can scrape university websites to automatically recommend courses according to current job-market requirements so that students do not need to manually visit each university website to get information about the courses will add more value to the current work.

REFERENCES

Alsmadi, D., & Omar, K. (2022, February). Analyzing the needs of ICT job market in Jordan using a text mining approach. In *2022 International Conference on Business Analytics for Technology and Security (ICBATS)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICBATS54253.2022.9759009

Education Counts. (n.d.). *Tertiary participation*. https://www.educationcounts.govt.nz/statistics/tertiary-participation

Fortino, A., & You, Y. (2022, March). Tracking technology trends using text data mining. In *2022 IEEE Integrated STEM Education Conference (ISEC)* (pp. 415–421). IEEE. https://doi.org/10.1109/ISEC54952.2022.10025069

Johri, V., & Bansal, S. (2018, January). Identifying trends in technologies and programming languages using topic modeling. In *IEEE 12th International Conference on Semantic Computing (ICSC)* (pp. 391–396). IEEE. https://ieeexplore.ieee.org/document/8334502

Kolakowski, N. (2024). *Tech unemployment stayed low at end of 2023*. Dice. https://www.dice.com/career-advice/tech-unemployment-stayed-low-at-end-of-2023

Liu, L., & Tang, L. (2018, June). A survey of statistical topic model for multi-label classification. In *26th International Conference on Geoinformatics* (pp. 1–5). IEEE. https://doi.org/10.1109/GEOINFORMATICS.2018.8557113

New Zealand Immigration. (2023, November 23). *Information technology*. https://www.live-work.immigration.govt.nz/work-in-new-zealand/job-market-key-industries/information-technology

Rubin, T. N., Chambers, A., Smyth, P., & Steyvers, M. (2012). Statistical topic models for multi-label document classification. *Machine Learning*, *88*, 157–208. https://link.springer.com/article/10.1007/s10994-011-5272-5

Ruchitaa Raj, N. R., Nandakumar Raj, S., & Vijayalakshmi, M. (2023, February). Web scrapping tools and techniques: A brief survey. In *4th International Conference on Innovative Trends in Information Technology (ICITIIT)* (pp. 1–4). IEEE. https://ieeexplore.ieee.org/document/10068666

Sendhilkumar, S., Srivani, M., & Mahalakshmi, G. S. (2017, February). Generation of word clouds using document topic models. In *Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM)* (pp. 306–308). IEEE. https://ieeexplore.ieee.org/document/8057554

U.S. Bureau of Labor Statistics. (2023). *Occupational employment and wage statistics*. https://www.bls.gov/oes/current/oes_nat.htm

U.S. Bureau of Labor Statistics. (2024). *Labor force statistics from the current population survey*. https://data.bls.gov/timeseries/LNS14000000

Viet, N. T., Kravets, A., & Duong Quoc Hoang, T. (2021, October 11–16). Data mining methods for analysis and forecast of an emerging technology trend: A systematic mapping study from Scopus papers. In S. M. Kovalev, S. O. Kuznetsov, & A. I. Panov (Eds.), *Artificial Intelligence: 19th Russian Conference, RCAI 2021* (pp. 81–101). Springer International Publishing. https://doi.org/10.1007/978-3-030-86855-0_7

Yang, C., Zhu, D., & Zhang, G. (2015, November). Semantic-based technology trend analysis. In *10th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)* (pp. 222–228). IEEE. https://ieeexplore.ieee.org/document/7383052

AUTHORS

Athavaloshan Thirulingam is a postgraduate student at Auckland University of Technology. His research interests are parallel computing, cloud computing and machine learning. qbz6506@autuni.ac.nz

Dr Sarita Pais is a Lecturer at the School of Engineering, Computer and Mathematical Science, Auckland University of Technology, with research interests in health informatics and data analytics. sarita.pais@aut.ac.nz

Recreating Goto in Python: How and Why

Carl Cerecke

https://doi.org/10.34074/proc.240119
Correspondence: Carl.Cerecke@nmit.ac.nz

This publication may be cited as:

Cerecke, C. (2024). Recreating Goto in Python: How and Why. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 150–159). ePress, Unitec. https://doi.org/10.34074/proc.240119

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga – New Zealand Institute of Skills and Technology

ABSTRACT

The Python language has no unrestricted goto statement, but because of Python's deep introspection capabilities, along with its dynamic nature, an unrestricted goto can be added to the language by rewriting a function's bytecode at runtime. Two primary use cases for goto in Python are discussed: porting existing historic code to Python while maintaining the existing code's flavour, and implementing state machines. While theoretically unnecessary, goto can make the first case easier, and the second case faster. An implementation is described and performance tests on an example state machine show using goto is a very fast method of implementing state machines in Python, and existing efforts to port historic code are already using the goto implementation.

KEYWORDS

Python, goto

INTRODUCTION

Dijkstra's well-known 1968 letter "Go To Considered Harmful" (Dijkstra, 1968, p. 147) popularised the problems that can be caused by careless use of the unrestricted jump that goto provides. The core of Dijkstra's complaint is that "unbridled use of the go to statement" makes reasoning about the behaviour of a running program difficult when presented with the static program code. In the programmer's struggle to constrain complexity, the goto statement provides far more opportunities for increasing complexity than it does to reduce it. Knuth (1974) brought some balance to Dijkstra's opinion and found that even with the significant then-recent advance of structured programming in academia, there were both cases where goto should be eliminated, and cases where goto was justified or even preferred. It was about two decades later that mainstream programming languages emerged (such as Java, Python, and then later JavaScript) with no native goto statement available. These languages instead have two restricted forms of the goto statement: *break* and *continue*. These restricted jump instructions cover many of the use cases where goto was used traditionally, but without the same potential for spaghettification of code.

While *break* and *continue* cover the majority of 'sensible' use cases for the goto statement, there are two main areas where goto still proves to be a useful construct for control flow: (1) porting existing software that uses goto statements; and (2) complex state machines.

Ceccato et al. (2008) compare different automatic goto-elimination techniques in porting a large (8 million lines of code) banking system to Java from a "BASIC-like language" which (initially, at least) had very limited control flow statements. Out of the 8 million lines, around 500,000 were goto statements. While most goto statements were automatically eliminated in the generated Java code, a minority could not be translated without a significant decrease in the understandability of the generated code. In these cases, placeholder `jlabel()` and `jgoto()` calls were introduced to the code, and a post-compilation step replaced these the `jgoto()` calls in the java byte code with a JVM goto instruction to the appropriate label. This post-compilation step is essentially a 'hack' to add goto to Java; the underlying bytecode interpreter has a goto, and the post-compilation step rewrites the bytecode to allow direct access to it. It is similar to the approach used in this paper, though with Java a separate post-compilation step is required, and in Python this can be done at runtime.

At the other end of the program size scale there is historical interest in early programs in languages that required the use of goto for control constructs, such as BASIC. There were many books and magazines with BASIC code listings in the 1970s and 1980s. These programs tended to be quite small, as not only did they have to be typed in from scratch, but the programs were intended for computers with very limited RAM: the Apple II, Commodore PET, and Tandy TRS-80 Model 1 were all available in 1977 with 4Kb of RAM (Reimer, 2005). A typical example is the program Hammurabi (Willaert, 2019), an economics simulator. It was already nearly a decade old when it was included in the early compendium *101 BASIC computer games* (Ahl, 1973). The code, to our modern eyes, looks very primitive and is littered with goto statements (Figure 1).

```
320 PRINT "HOW MANY ACRES DO YOU WISH TO BUY";
321 INPUT Q:IF Q<0 THEN 850
322 IF Y*Q<=S THEN 330
323 GOSUB 710
324 GOTO 320
330 IF Q=0 THEN 340
331 LET A=A+Q:LET S=S-Y*Q:LET C=0
334 GOTO 400
340 PRINT "HOW MANY ACRES DO YOU WISH TO SELL";
341 INPUT Q:IF Q<0 THEN 850
342 IF Q<A THEN 350
343 GOSUB 720
344 GOTO 340
350 LET A=A-Q:LET S=S+Y*Q:LET C=0
```

Figure 1. A scanned extract from the book *101 BASIC computer games*. Fourteen lines with 8 gotos!

In order to faithfully reproduce the code in a modern language while also retaining the flavour of the original code (Massey, 2014), some form of goto is required. As an example, here are the first 6 lines from the original code in Figure 1 translated to Python in Massey's port:

```
label .line320
Q=get_number("HOW MANY ACRES DO YOU WISH TO BUY")
if Q<0:
    goto .line850
if Y*Q<=S:
    goto .line330
sub710()
goto .line320
label .line330
if Q==0:
    goto .line340
```

A second common use of goto is where the control constructs of the programming language are not powerful or convenient enough to express the control flow that the programmer has in mind. One such example, "Centralized exiting of functions", is explicitly condoned in the Linux kernel coding style guide (The Linux kernel documentation, n.d.). In Python a centralised exit of a function might be idiomatically constructed using a try-except-else-finally block.

Another example where a goto is convenient is when coding a state machine. For example, a typical state machine implementing a shift-reduce parser for a modern programming language may have many hundreds of states in the state machine. Non-goto approaches to large state machines might use tables (e.g., YACC, Bison) or a loop with a switch statement. In either case, a core loop is required to iterate for each state transition.

Figure 2 shows an example of a table-style finite state machine for an LR parser with 12 states that implements a simple mathematical expression parser, including the four basic operators and parentheses. While typically such state machines are represented as tables, an early attempt at directly representing a parser's state machine using goto statements (Pennello, 1986) resulted in speedups of 6–10 times over a table-driven approach, though this approach directly generated assembly code. A later approach generated the parser state machine in c code and strived for compatibility with the popular early parser generator YACC, and yielded speedups of 2–6 times (Bhamidipaty & Proebsting, 1998).

Even with modern structured programming control flow constructs, the goto statement still has some potential niche applications where it is the best tool for the job.

LR(1) Parsing Table

	id	+	×	(	)	$	E	T	F
0	shift 5			shift 4			1	2	3
1		shift 6				accept			
2		$E \rightarrow T$	shift 7		$E \rightarrow T$	$E \rightarrow T$			
3		$T \rightarrow F$	$T \rightarrow F$		$T \rightarrow F$	$T \rightarrow F$			
4	shift 5			shift 4			8	2	3
5		$F \rightarrow id$	$F \rightarrow id$		$F \rightarrow id$	$F \rightarrow id$			
6	shift 5			shift 4				9	3
7	shift 5			shift 4					10
8		shift 6			shift 11				
9		$E \rightarrow E + T$	shift 7		$E \rightarrow E + T$	$E \rightarrow E + T$			
10		$T \rightarrow T \times F$	$T \rightarrow T \times F$		$T \rightarrow T \times F$	$T \rightarrow T \times F$			
11		$F \rightarrow (E)$	$F \rightarrow (E)$		$F \rightarrow (E)$	$F \rightarrow (E)$			

Figure 2. Table representation of a state machine for parsing simple mathematical expressions.

ADDING A GOTO STATEMENT TO PYTHON

The first attempt to add goto to Python (Hindle, n.d.), a 2004 April Fool's joke, used the **`sys.settrace`** function of Python, which registers a function that is then called before executing every line of code. The **`sys.settrace`** function was intended for implementing a debugger and greatly slows a running program, but it is powerful. Hindle's version also included the humorous **`comefrom`** statement (the nefarious opposite of the goto statement).

A more efficient way to add goto to Python is to take advantage of the dynamic nature of the Python runtime environment to rewrite the Python bytecode for a function. This is the approach used by the code implemented by the author (Cerecke, n.d.) and discussed in the remainder of this section. The source code is available at https://github.com/cdjc/goto. Only the reference implementation of Python (https://python.org) is compatible with the goto implementation discussed in this paper, as other implementations (such as Jython, IronPython and PyPy) take a different runtime approach.

Though we often think of Python as interpreted, rather than compiled like other languages that use a bytecode interpreter (or "VM") such as Java (the JVM) or C# (the .NET CLR), this is technically incorrect. In Python the compilation step is hidden – there is no separate compile-then-run steps in Python, like there are in Java or C#.

These are the high-level requirements before goto statements can be added to Python using bytecode rewriting:

1. Access a function's bytecode at runtime.
2. Set a function's bytecode at runtime.
3. Mark a function as requiring a rewrite of its bytecode.
4. Specify labels and gotos within the bounds of current Python syntax.

Requirement 1 is met in Python, as each function has a corresponding code object (accessed via the function's `__code__` attribute), and each code object has an attribute to get the bytecode string for the code object. The helpful dis module in the standard library provides a function that can produce a disassembly of another function.

Requirement 2 is met by the `CodeType` constructor in the `types` standard library module, through which a new code object can be created. A recent addition to Python (version 3.11) makes this more convenient, with the `replace` function of a code object. Furthermore, this new code object can then be assigned to a function's code object at runtime.

Requirement 3 is met by using a Python function decorator – a convenient language feature for creating and using higher-order functions.

A convenient way requirement 4 can be met is by misusing object attribute access. For example, the line foo.bar in Python means 'access the attribute bar from the object foo'. Although this is an expression, expressions are also valid Python statements, and replacing the object with the word 'goto' (or 'label') will allow specification of goto (and label) statements using Python syntax. This does not result in an error during compilation, as Python's dynamic typing does not attempt to resolve variables during the compilation step. Without rewriting the function's bytecode, running the function would result in the exception: `NameError: name 'goto' is not defined`

Figure 3 shows a trivial example of all four requirements in the Python interpreter. Line 3 imports the `dis` module for viewing the disassembly of a function. Line 4 creates a simple function named fn, where the first line (line 5 of the listing) is specifying a label as if it was a Python attribute (requirement 4). The statement on line 8 prints a disassembly of the function. The `label.start` from the function results in two instructions in the bytecode, one for the global object `label` (lines 11–12) and one for `start` attribute access of that object. If the function were to be called at this point, Python would raise an exception because it cannot find any globals with the name `label`.

Line 17 extracts the bytecode of the function as an immutable Python bytes object (requirement 1); not all of the bytecode is visible in the Figure due to its truncation at the edge of the Figure. The many zeros in the bytecode (represented by `\x00`) are a new Python 3.11 feature that adds cache space within the bytecode to be used during runtime optimisations. In order to change the bytecodes, a list is created from the bytes, and the bytes representing the `label .start` line of the function are deleted (line 20).

Line 23 creates a new code object from the function's existing code object, with the bytecodes replaced, and assigns the new code object to the function code object (requirement 2). The `replace` function is a new addition to Python 3.11; prior versions required creating a new code object from scratch.

Finally, the function is called (line 24) and the function's new bytecode, now containing only code for the return 1 statement, is executed and the resulting value (1) is returned.

```
Python 3.11.2 (tags/v3.11.2:878ead1, Feb  7 2023, 16:38:35) [MSC v.193
Type "help", "copyright", "credits" or "license" for more information.
>>> import dis
>>> def fn():
...     label .start
...     return 1
...
>>> dis.dis(fn)
  1           0 RESUME                   0

  2           2 LOAD_GLOBAL              0 (label)
             14 LOAD_ATTR                1 (start)
             24 POP_TOP

  3          26 LOAD_CONST               1 (1)
             28 RETURN_VALUE
>>> fn.__code__.co_code
b'\x97\x00t\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00j\x01\x00\x00\x
>>> bytecode = list(fn.__code__.co_code)
>>> del bytecode[2:-4]
>>> bytes(bytecode)
b'\x97\x00d\x01S\x00'
>>> fn.__code__ = fn.__code__.replace(co_code = bytes(bytecode))
>>> fn()
1
```

Figure 3. Trivial example showing the function bytecode rewriting mechanism required to implement goto in Python.

While Figure 3 illustrates the main mechanism by which a goto statement can be added to Python, there are still a few tasks remaining to create a working function decorator. At a high level these are:

1. Identify all the labels and gotos in the function. Check whether any gotos are missing labels, and labels are unique, and no illegal jumps are attempted.
2. Overwrite the instructions in the bytecode for each label and goto (and the following CACHE instructions) to NOP instructions.
3. For each position in the bytecode where the goto statement was, insert a JUMP_FORWARD or a JUMP_BACKWARD to the corresponding label. Take care to stop any in-progress iterators if we are jumping out of an iterator loop.

There are a few cautions that must be exercised:
No jumping into any part of function that is inside a block that has some special initialiser or meaning. For example, **`for`** blocks, **`with`** blocks, and **`try`** blocks. It's not clear what the semantics should be in those cases. Jumping into a **`while`** block is perfectly fine, though. Jumping out of a **`try`** block is also forbidden. There should be no way to bypass the **`finally`** block.

A new feature of the recent Python 3.11 is the removal of unreachable code during bytecode compilation. This means that code that is only reachable by goto may be in danger of being optimised away. This makes the following code fail:

```
for ls in matrix:
    for num in ls:
        if num == 1234:
            goto .foundIt

# didn't find it
return False

label .foundIt  # "unreachable" code removed by Python compiler
return True
```

To work around the unreachable code elimination, the `return False` would need to have an always-true 'guard', such as `if __name__: return False` The Python special variable `__name__` is normally always set to some value, either the name of the current module, or `"__main__"` (*The Python language reference*, n.d.) so the `if`-statement will always be true, barring any explicit meddling with the module's `__name__` attribute.

There is a maximum depth of iterators (`for` loops) that can be accommodated in the bytecode before lengthening the bytecode is required – lengthening the bytecode is tricky, as any existing jumps in the bytecode may also have their destinations recalculated. Each iterator that must be stopped before jumping must be removed from the top of the stack with a POP_TOP bytecode. By reusing the CACHE instructions where the goto statements appear in the bytecode, up to 12 instructions can be overwritten before the bytecode list itself has to be extended. This allows up to 11 POP_TOP instructions before the jump instruction. The current implementation limits the nested iterators to 10, which seems more than adequate for a reasonable function. The remaining 2 bytes allow for an extra instruction in case the jump offset is greater than what can be represented in one byte, and an EXTENDED_ARG opcode is required.

Each instruction in Python bytecode uses 2 bytes: one for the opcode, and one for the argument. For an opcode requiring a value larger than one byte, up to three preceding EXTENDED_ARG opcodes are used to add as many higher-order bytes as required. If a jump over 255 instructions is required, then an extra EXTENDED_ARG instruction will need to be inserted. This must be done with care, as inserting an EXTENDED_ARG instruction also changes the length of the jump by +1 (for a backward jump) or -1 (for a forward jump). The code assumes that jumps will never be more than what would fit in 2 bytes (65536 would be a very long jump!) and so will only ever need a single EXTENDED_ARG bytecode instruction.

A full example, showing the import, the decorator, a goto, a label, and two calls of the goto-decorated function, is shown in Figure 4.

```
>>> from goto import goto
>>> @goto
... def matrix_find(matrix, n):
...     for square in matrix:
...         for line in square:
...             for value in line:
...                 if value == n:
...                     rval = "found"
...                     goto.found
...     rval = "not found"
...     label.found
...     # ... other code here ...
...     return rval
...
>>> matrix_find([[[1,2]],[[3,4]]], 3)
'found'
>>> matrix_find([[[1,2]],[[3,4]]], 5)
'not found'
```

Figure 4. An example of using the goto decorator in Python.

RESULTS

Implementing state machines is one of the main uses of a goto statement. To gain an approximate measure of the speed of using gotos for state machines compared to other methods, an example state machine was implemented using four methods:

1. The popular python-statemachine library (Macedo, n.d.).
2. Python **`for`**-loop around the Python **`match`** statement. Each case in the **`match`** statement matches an input character, and the next state is set based on the current state using **`if`** statements.
3. Using the equivalent regular expression in Python. The regular expression matching (and building) code is implemented in C inside the Python interpreter. The regular expression corresponding to the state machine was compiled in Python using the following regular expression string:
   `r'\d+(\.\d+)?(e[+-]\d+)?(\+\d+(\.\d+)?(e[+-]\d+)?)*\$'`
4. Using the goto library described in this paper and using labels for states, and goto for transitions.

The state machine used for the test (Figure 5) recognises valid expressions consisting of a sum of floating point numbers. Each character in the input string triggers a transition to another state if that character is part of a valid string. The state machine is non-trivial enough for a reasonable test, yet small enough to be easily and quickly implemented (and debugged!) using a variety of methods. The implementation details are in the **`goto_test_speed.py`** file of the associated GitHub project (https://github.com/cdjc/goto).

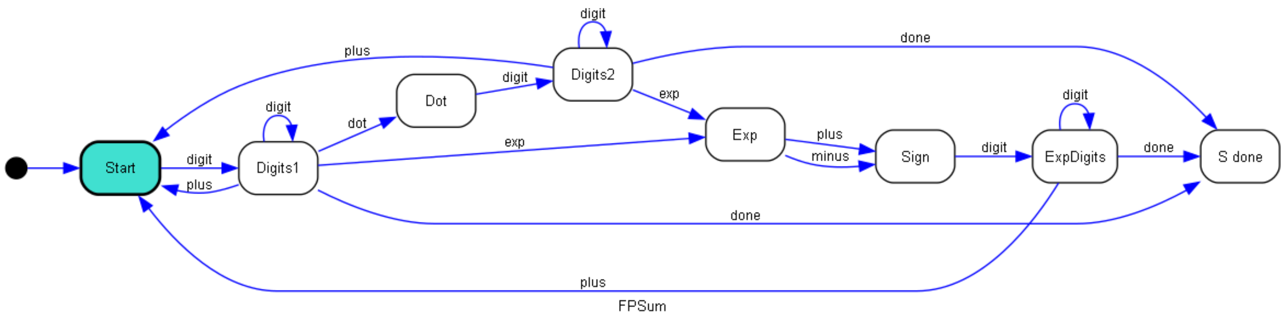


Figure 5. Finite state machine for recognising an expression of a sum of floating point numbers.

For the speed comparison between the methods, a 30,000 character randomly generated string exercising all non-final transitions of the finite state machine was generated. The Python standard library timeit module was used to time each method. Performance was measured on an i7-7820HQ laptop with 32GB RAM running Windows 11.

Table 1. Speed of different finite state machine implementations.

	Time to process 30,000 characters	**Slowdown compared to goto method**
Python-statemachine	504ms	180×
For-loop with match	21ms	7.5×
Regular expression	3.1ms	1.1×
Github.com/cdjc/goto	2.8ms	1×

Table 1 shows that the goto implementation presented in this paper is far superior in speed to the two other 'pure' Python implementations for programming state machines. The regular expression speed was comparable, which is perhaps surprisingly slow considering it does not execute any Python bytecode when matching the input (other than the regular expression `match` method call), but rather calls in to the regular expression module written in C inside core Python.

It can be argued that the goto implementation's much faster speed is somewhat offset by its lower readability (the regular expression suffers from this as well) and by being non-standard. Knuth's aphorism "premature optimization is the root of all evil" (Knuth, 1974, p. 268), in a paper about structured programming with gotos, is perhaps appropriate here.

CONCLUSION AND RECOMMENDATIONS

The dynamic nature of Python, along with its object introspection capabilities, provides a way to implement goto in Python at runtime. This technique could be used for other situations where directly manipulating bytecode is required. A major disadvantage is the instability of the Python bytecode specification; it is subject to significant changes between language versions with no effort to maintain backward compatibility. Also, the technique as implemented is only applicable to the standard reference implementation Python (CPython).

Wanting to use the goto statement in Python is a niche use case mostly restricted to porting some old code while retaining the flavour of the original code. State machines are another possible use case, and the goto statement provides a very fast alternative to other pure-Python state machine implementations. This speed does come at the cost of both reliability (because the underlying bytecode is subject to incompatible changes between versions) and readability, as the non-standard use of familiar Python constructs will be unfamiliar to many.

REFERENCES

Ahl, D. (1973). *101 BASIC computer games*. Ditigal Equipment Corporation. http://www.bitsavers.org/pdf/dec/_Books/101_BASIC_Computer_Games_Mar75.pdf

Bhamidipaty, A., & Proebsting, T. A. (1998). Very fast YACC-compatible parsers (for very little effort). *Software: Practice and Experience*, *28*(2), 181–190. https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-024X(199802)28:2%3C181::AID-SPE139%3E3.0.CO;2-4

Ceccato, M., Tonella, P., & Matteotti, C. (2008). Goto elimination strategies in the migration of legacy code to Java. In *2008 12th European Conference on Software Maintenance and Reengineering* (pp. 53–62). IEEE. https://doi.org/10.1109/CSMR.2008.4493300

Cerecke, C. (n.d.). *Goto in Python* [Computer software]. Retrieved July 22, 2023, from https://github.com/cdjc/goto

Dijkstra, E. W. (1968). Letters to the editor: Go to statement considered harmful. *Communications of the ACM*, *11*(3), 147–148. https://doi.org/10.1145/362929.362947

Hindle, R. (n.d.). *Goto for Python*. Retrieved July 18, 2023, from http://entrian.com/goto/

Knuth, D. E. (1974). Structured programming with *go to* statements. *ACM Computing Surveys*, *6*(4), 261–301. https://doi.org/10.1145/356635.356640

Macedo, F. (n.d.). *Python StateMachine*. Retrieved July 22, 2023, from https://python-statemachine.readthedocs.io/en/latest/readme.html

Massey, B. (2014). *HAMURABI.BAS in Python 3* [Computer software]. https://github.com/BartMassey/hamurabi

Pennello, T. J. (1986). Very fast LR parsing. *ACM SIGPLAN Notices*, *21*(7), 145–151. https://doi.org/10.1145/13310.13326

Reimer, J. (2005, December 15). Total share: 30 years of personal computer market share figures. *Ars Technica*. https://arstechnica.com/features/2005/12/total-share/

The Linux kernel documentation. (n.d.). *Linux kernel coding style*. Retrieved July 17, 2023, from https://docs.kernel.org/process/coding-style.html

The Python language reference. (n.d.). The import system. Retrieved July 22, 2023, from https://docs.python.org/3/reference/import.html

Willaert, "Critical Kate." (2019, September 9). The Sumerian game: The most important video game you've never heard of. *A Critical Hit!* https://www.acriticalhit.com/sumerian-game-most-important-video-game-youve-never-heard/

AUTHOR

Dr Carl Cerecke is an IT Tutor at Nelson Marlborough Institute of Technology Te Whare Wānanga o te Tau Imu. He has a wide array of research interests, some of which are useful.
carl.cerecke@nmit.ac.nz

The Challenges of TinyML Implementation: A Literature Review

Riya Adlakha and Eltahir Kabbar

https://doi.org/10.34074/proc.240120
Correspondence: Eltahir Kabbar, ekabbar@unitec.ac.nz

This publication may be cited as:

Adlakha, R., & Kabbar, E. (2024). The Challenges of TinyML Implementation: A Literature Review. In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 160–169). ePress, Unitec. https://doi.org/10.34074/proc.240120

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga –
New Zealand Institute of Skills and Technology

ABSTRACT

This study aims to sensitise and summarise the tiny machine learning (TinyML) implementation literature. TinyML is a subset of machine learning (ML) that focuses on implementing ML models on resource-constrained devices such as microcontrollers, embedded systems, and internet of things (IoT) devices. A systematic literature review is performed on the works published in this field in the last decade. The key focus of this article is to understand the critical challenges faced by this emerging technology. We present five significant challenges of TinyML, namely, limited and dynamic resources, heterogeneity, network management, security and privacy, and model design. This article will be of interest to researchers and practitioners who are interested in the fields of ML, IoT and edge computing.

KEYWORDS

TinyML, edge computing, machine learning, IoT, microcontrollers

BACKGROUND

With the evolution in hardware, integrating enhanced machine learning (ML) models running on low-power embedded devices such as microcontrollers has become possible. This convergence of ML and embedded systems is called tiny machine learning (TinyML). TinyML enables assembling ML to hardware segments without processing the data in an external location. Processing and execution occur at the edge, making the applications close to the data source. This concept is called edge computing. In recent times, edge computing has become a shining example of an emerging technology. The technology has numerous applications in healthcare, finance, smart cities and transportation. TinyML delivers promising results when deployed on small edge devices that facilitate fast processing and data analysis without needing a server response. Figure 1 showcases an architectural framework of TinyML to compute and analyse data from multiple IoT devices into an edge device (microcontroller unit [MCU], for example), eliminating the requirement for processing at cloud servers.

The three basic tool-sets required to enable TinyML for processing and predicting results are hardware, software and libraries (Janapa Reddi et al., 2022). Various hardware platforms are TinyML aware, such as Arduino Nano 33 BLE Sense, Apollo3, Nicla Sense ME, ST IoT Discovery and Nordic Semi nRF52840 DK, to name a few. Most hardware operates at a flash memory of less than 1 MB and SRAM of less than 1 MB. Most of these are battery operated and have Li-Po on top of a DC power source. ARM Cortex M4 stands out as the most popular and widely used microprocessor. TensorFlow Lite (TFL) and uTensor are open-source software designed to run ML on MCUs and rapidly deploy IoT devices. Edge Impulse, another software, is a cloud service that enables TinyML by running ML models for edge devices. It supports AutoML for platforms at the edge. The capability of local execution is also supported with Python, C++ and SDK, along with numerous platforms such as smartphones endowed with Edge Impulse, to build training models.

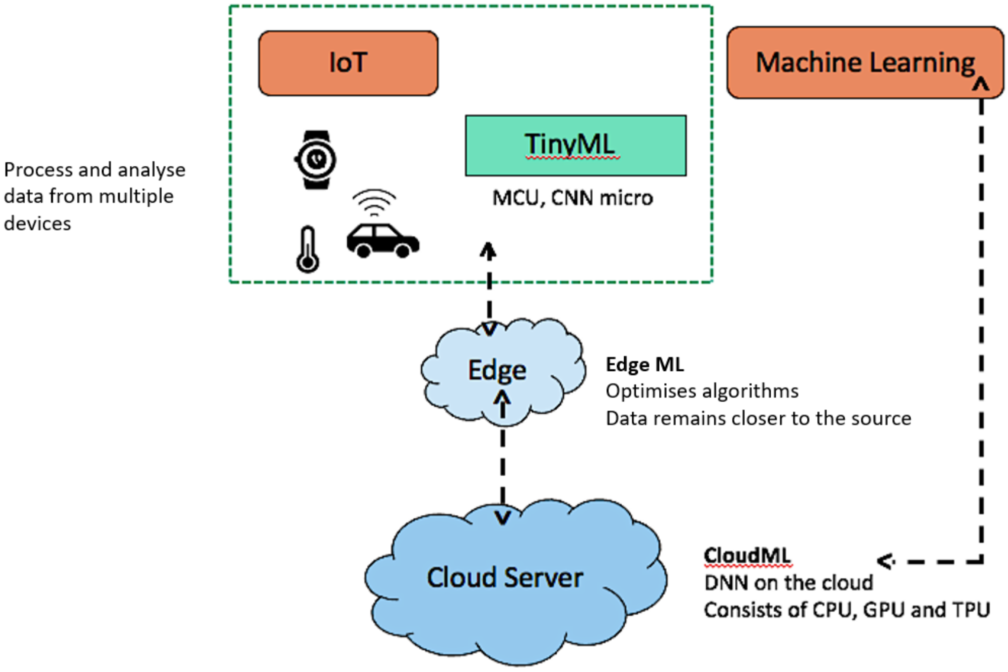


Figure 1. Enabling IoT devices with cloud computing, edge computing, and TinyML.

TinyML has a range of opportunities, given that the microcontrollers consume relatively less energy than the graphic processing units (GPUs), which enables the IoT devices that work with MCUs to be placed anywhere without plugging in (Sanchez-Iborra & Skarmeta, 2020). Given their low energy requirement, they can also be paired with devices driven by large batteries, allowing them to emerge as connected smart devices, such as smart watches and doorbells. In addition, MCUs are more cost effective compared to high-end processors, making the TinyML systems cost effective. With the TinyML model running on edge, the data is stored, processed and analysed internally rather than at an external server or cloud. These advantages increase data privacy, reducing the risk of compromising sensitive information. With complete independence, embedded devices increase autonomy by eradicating interference from outside sources. Thus, decisions and information can only be accessed and shared within the system (Soro, 2021).

One recent study that informs this research is Han and Siebert (2022), who presented a systematic literature review of TinyML focusing on five aspects: hardware, framework, datasets, use cases and algorithms. This study adds to their work by focusing on understanding the challenges associated with TinyML implementation in detail. We also explore the history, benefits and implementation of TinyML in the following sections.

HISTORY OF TINYML

TinyML emerged from the internet of things (IoT). It is a game changer, emphasising that big is not always better. Traditional methods included putting complex ML models into hardware for developing applications and products. The computing was performed on the cloud, which introduced the hurdle of latency and depended entirely on connectivity. All these factors made computing not only slower but also expensive and inefficient. The development of TinyML by Pete Warden bridged the gap between intelligence and embedded systems. More and more companies across the globe are moving towards TinyML. Azip, a leading company in AIoT (artificial intelligence for IoT) ("Artificial intelligence of things," 2023), is making TinyML models available for high-performance and intelligent product solutions.

KEY ENABLER: TINYML AS-A-SERVICE

Conventional ML processing is powered by cloud providers that run on efficient CPUs, GPUs and tensor processing units (TPUs). However, embedded systems face constraints because of these devices' limited computation and processing powers to run complex ML models. TinyML as-a-Service (TinyMLaaS) is modelled to solve this fundamental problem for compact devices. Using an ML compiler, it aims to transform such models to fit the target device's resource size. It uses techniques to squeeze the model size; for example, quantising with fewer computing bits, pruning less critical parameters, and fusing multiple computational operators into one (Doyu et al., 2020). The framework uses ML compilers to generate optimised low runtime for popular ML models. TinyMLaaS also includes specialised ML models for embedded hardware accelerators, which are chip-manufacturer dependent.

In a typical TinyMLaaS ecosystem, an appropriate ML model is generated using lightweight machine-to-machine (LwM2M) software with an 'on-the-fly' model inferencing module. TinyMLaaS is a demand-driven cloud service that resolves privacy issues by keeping the data on-premise.

RESEARCH METHODS

To identify the critical challenges of TinyML implementation, a detailed literature review, guided by Brereton et al. (2007), was conducted using relevant keywords. The researchers then identified the articles that discussed TinyML implementation, eliminating any study out of the scope. Once the final list was devised, the researchers analysed and categorised the literature into five themes. The taxonomy of the literature categorisation is discussed in the following section.

RESULTS AND DISCUSSION

This section outlines the five key challenges and problems encountered in research and deployment pertaining to the numerous TinyML applications: limited and dynamic resources, heterogeneity, network management, security and privacy, and model design. We review the five main challenges and propose solutions in the following section. Figure 3 illustrates the taxonomy of obstacles TinyML faces.

Limited and Dynamic Resources

TinyML devices often have constrained energy, memory, and computing capabilities. These constraints present a barrier when deploying sophisticated machine-learning models that demand a lot of computational power.

Limited power: Maintaining accuracy throughout the range of TinyML devices is challenging because their power consumption can vary greatly. As articulated by R. Kallimani et al. (2023), the performance of the algorithms is severely hampered by the lack of power at the edge devices. Secondly, establishing what comes inside the scope of the power measurement is problematic when data pathways and preprocessing procedures differ dramatically between devices. TinyML frameworks may encounter energy mismanagement because sensors and other accessories are frequently attached to edge devices (Yelchuri & R, 2022). As a result, developing a power-efficient TinyML system remains a significant challenge.

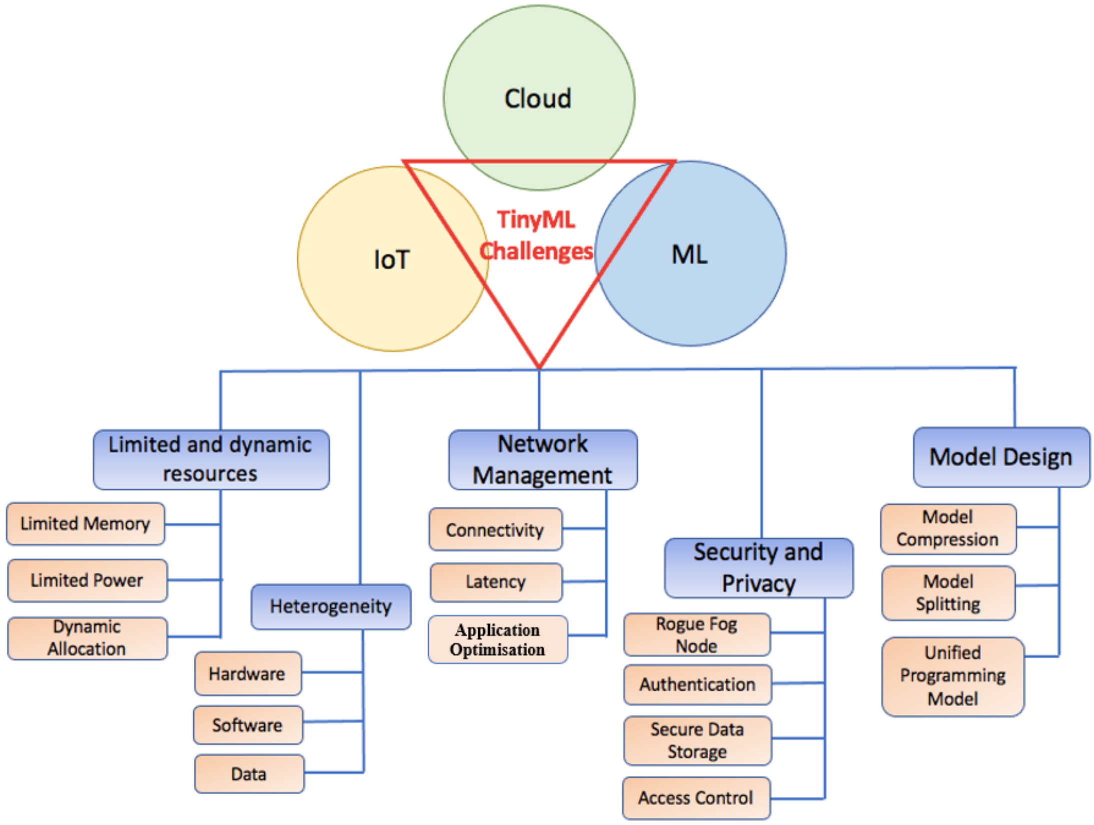


Figure 3: Taxonomy of TinyML challenges.

Limited memory: Due to the resource-constrained nature of tiny devices, one of the primary issues with TinyML is limited memory. Inference models used in traditional ML benchmarks have significantly greater peak memory requirements (in the range of GBs) than TinyML devices can offer (Banbury et al., 2020). Common ML systems frequently don't have resource use restrictions while running on workstations or the cloud. As a subset of edge devices, frugal gadgets, however, are much more limited. Most MCUs come with memory measured in kilobytes rather than megabytes and processors that run at megahertz rather than gigahertz in the most recent mobile phones. The restricted memory on TinyML devices leads to two key issues:

Catastrophic forgetting: One common implication of limited memory in TinyML is catastrophic forgetting. This is a phenomenon in which the model forgets previously learned information while continuing to learn new information. Catastrophic forgetting can be a significant worry in the setting of TinyML, because memory and computing resources are constrained (Rajapakse et al., 2023). TinyML models are frequently used on hardware with limited memory, making it challenging to retain vast amounts of data for the model's retraining on earlier jobs.

Volatility of SRAM: Due to resource limits, MCUs' primary memory (SRAM) spans from a few to a few hundred kilobytes. Neural networks stored in a flash as a C/C++ array are handled as frozen graphs, which means that any amendments to this graph are not permitted. As a result, many systems train the existing model entirely or partially in SRAM without putting it in flash memory. Because SRAM is volatile, any progress in training a model is lost when the MCU is reset or powered off.

Dynamic resource allocation: TinyML's dynamic resource allocation is a significant challenge, especially when dealing with resource-constrained devices with low memory, processing power and energy. Dynamic resource allocation entails efficiently managing and distributing the ever-dynamic resources to various TinyML system activities or components. Currently, the edge platform suffers from a problem with dynamic resource allocation, necessitating the development of techniques/algorithms for analysing dynamic data. New methods that consider the varying computational power and specifications of processing entities, as well as the requirement for consistent portability among heterogeneous devices in the event of local and parallel processing, are required (Doyu et al., 2021).

Heterogeneity

Due to the considerable variation of hardware, software, and data characteristics across different tiny devices and platforms, heterogeneity can appear in various ways in TinyML. The models must be compatible with the target device's hardware, software and data architectures. It can be challenging to ensure compatibility across multiple microcontrollers and embedded systems, because each platform may have its limits and optimisations. Below are the three significant types of heterogeneity observed in TinyML models.

Hardware heterogeneity: TinyML models are implemented on various hardware architectures, from microcontrollers to embedded systems and IoT devices. These devices may feature different processor types (for example, ARM or RISC-V), clock speeds, memory sizes and specialised hardware accelerators. Each hardware platform may have its own set of constraints and optimisations, making it challenging to create models that perform well across many devices. Furthermore, a significant challenge is the issue of standardising performance findings across different implementations (Banbury et al., 2020).

Software heterogeneity: TinyML's software heterogeneity refers to the variety of software components, frameworks, libraries and runtime environments utilised for deploying ML models on resource-limited devices. This variation is caused by changes in operating systems, ML frameworks and other software dependencies among small devices. The constraints of software heterogeneity include maintaining TinyML application consistency and performance, compatibility issues and program capability restrictions. Tiny devices frequently use several operating systems or RTOS, each with its own memory management and language support, which influences application behaviour. Furthermore, the availability and compatibility of ML frameworks may differ, resulting in model development and deployment inconsistencies. Software libraries, inference engines and compiler variations further complicate TinyML development and deployment.

Data heterogeneity: Handling data heterogeneity is difficult, because it requires careful consideration of data pretreatment, augmentation, and adaption approaches to maintain model resilience and generalisation. TinyML models are trained using data from various sources, such as sensors or edge devices, resulting in multiple data types and formats. Model accuracy depends on managing data quality and dealing with noisy or missing data (Kallimani et al., 2023). Variations in data distribution between devices can also impact model performance during deployment. TinyML application portability (Lakshman & Eisty, 2022) is affected by data heterogeneity, because models must be flexible to multiple data distributions, necessitating preprocessing, transfer learning, and robust generalisation techniques to maintain consistent performance across different devices and contexts.

Network Management

Network management is essential for dependable and effective communication among resource-constrained devices, edge nodes and cloud servers. Three challenges related to maintaining a consistent network for TinyML are discussed in this section.

Connectivity: Gateways in the architecture must be connected to the internet. Provisioning such connectivity necessitates both capital expenditure (CAPEX) and operating expenditure (OPEX), as well as administrative overheads associated with infrastructure maintenance (Zaidi et al., 2022). Local inference capabilities lessen reliance on connectivity, allowing for providing services in places where internet connectivity is sporadic or non-existent. Managing the network connectivity of small devices is critical for data transmission, model updates and contact with cloud services. Tiny devices may have different communication capabilities, such as wi-fi, Bluetooth, Zigbee or cellular connectivity, each with limitations and trade-offs.

Latency: Real-time inference is critical for many applications, including robotics and voice-based assistive technology. For example, real-time inference is essential in responding to user commands and operating in a voice-controlled home automation system.

Sending data to the cloud for inference or training may result in network delays, rendering it unsuitable for time-sensitive, interactive applications. For example, offloading sensor data to a cloud server for processing in a real-time robotics application would result in considerable end-to-end latency (Chen & Ran, 2019). Applications with low-latency requirements can accomplish real-time inference and provide smooth user experiences by leveraging TinyML and edge computing, which makes them particularly successful in dynamic and time-critical contexts.

Application optimisation: Because of the nature of the computing platforms, applications deployed in edge-cloud computing settings confront a slew of challenges, including restricted bandwidth, unreliability and heterogeneity of wireless connections, and computation offloading. Changing workloads across different components has an impact on application performance. Maintaining quality of service necessitates elasticity and remedial skills. Thang Le Duc et al. have discussed techniques such as load balancing, application scaling and migration to achieve these needs for reliable resource supply in edge-cloud computing (Le Duc et al., 2019).

Security and Privacy

The leakage of private information, such as data, location, or usage, is a crucial challenge for end users using services such as cloud computing, wireless network, or IoT. In addition to encryption of private data, secure proxies required for rendezvous, communication and access control are currently not popular in TinyML implementation. Yi et al. (2015) have extensively presented the challenges of security and privacy problems related to TinyML implementation. This section discusses these challenges.

Security: As discussed in the work presented by Lopez et al. (2015), the coexistence of malicious and trusted nodes in distributed edge-based overlays is considered by edge-centric computing. This will necessitate using secure routing, redundant routing, trust topologies and past peer-to-peer research on this novel set. Below are the problems faced in obtaining a secure TinyML framework (Yi et al., 2015):

> **Rogue fog node:** A rogue fog node is a fog device or instance that lures end users into connecting by pretending to be genuine. This can lead to the threat of cyber-attacks, as the system's security will be compromised if connected to a fake node. Due to the dynamic instance creation and deletion in TinyML, it is difficult to block or blacklist the false nodes, making them prone to such attacks.
>
> **Authentication:** Authentication is a prime focus for TinyML, as many end users are connected to its devices and applications. Each node requires authentication, which poses a challenge to TinyML. In contrast to the conventional public key infrastructure authentication, face, fingerprint and touch-based authentication would be beneficial to transform the security standards used in TinyML.
>
> **Secure data storage:** Edge computing involves outsourcing user data and transferring user control over data to the edge, which has the same security risks as cloud computing. It is challenging to guarantee data integrity, because the outsourced data may be deleted or erroneously manipulated. Furthermore, unauthorised parties might exploit the uploaded data to threaten the TinyML systems.

Privacy: Data transmission to the cloud raises privacy issues for users who own that data or whose actions are being traced in the data. This leads to the mental barrier of exposing sensitive and confidential information to the cloud without learning how the data will be consumed. Data, usage and location privacy are often the most challenging aspects of TinyML. Edge nodes fetch sensitive information produced by IoT devices and usually lack privacy-preserving methods. These methods cannot be deployed directly because edge computing has no reliable third party. Access control has been shown to be a reliable solution for maintaining user privacy while ensuring system security. While access control in cloud computing is typically performed cryptographically for outsourced data, building access control that spans client, edge and cloud in fog computing will be challenging while continuously meeting the design and resource limitations.

Model Design

Machine learning researchers frequently concentrate on developing models with a smaller number of parameters in the deep neural network (DNN) model in order to lower memory usage and latency while achieving good accuracy when designing DNN models for resource-constrained devices. Major challenges that arise in the implementation of TinyML are model compression, model splitting and a unified programming model.

Model compression: To enable running ML models such as DNNs on the edge, model compression is crucial to compress the existing models while preserving high accuracy at the same time. Compression techniques such as parameter quantisation and pruning are useful in doing so (Chen & Ran, 2019). By changing the precision of DNN parameters from floating-point numbers to low-bit numbers, parameter quantisation reduces the computational load by avoiding costly floating-point multiplications. Parameter pruning, on the other hand, reduces the size and efficiency of the DNN model while maintaining acceptable performance by removing the least significant parameters, which are frequently the ones close to zero.

Model splitting: A DNN model can be processed using the model splitting technique across many computer devices or resources. This method is very helpful in context with limited resources, such as edge devices and microcontrollers. The method entails breaking up the initial DNN into more manageable sub-models or layers. These sub-models can then be installed and run on distinct edge devices with constrained memory and processing power. The goal of model splitting is to maximise inference efficiency and resource utilisation without adversely affecting the performance of the model as a whole.

Unified programming model: TinyML's unified programming model addresses the heterogeneity by offering a consistent and standardised framework for constructing machine learning models that can operate quickly on a diverse variety of edge devices. Automatic model optimisation and quantisation techniques are used to reduce huge models into compact forms suitable for edge devices. This approach emphasises low-latency and real-time inference, ensuring that TinyML models can accomplish tasks with little delay, making them suited for robotics, IoT, and other real-time use cases. TinyML intends to lower the barrier of entry for developers, speed the creation of TinyML applications, and encourage the use of machine learning in resource-constrained situations by providing a consistent programming model.

CONCLUSION

TinyML can transform multiple technologies such as cloud computing, IoT and ML. The open challenges identified in this study make TinyML unfeasible for some industries, such as finance and healthcare, due to the high demand for security and privacy. The results of our study indicate that limited resources, security and privacy are the most significant difficulties TinyML implementation faces. While the heterogeneity of hardware, software and data stands out as a challenge, there is a huge potential for this technology to shape the future of machine learning.

REFERENCES

Artificial intelligence of things. (2023, September 18). In *Wikipedia*. https://en.wikipedia.org/wiki/Artificial_intelligence_of_things

Banbury, C. R., Reddi, V. J., Lam, M., Fu, W., Fazel, A., Holleman, J., Huang, X., Hurtado, R., Kanter, D., Lokhmotov, A., Patterson, D., Pau, D., Seo, J., Sieracki, J., Thakker, U., Verhelst, M., & Yadav, P. (2020). *Benchmarking TinyML systems: Challenges and direction*. http://arxiv.org/abs/2003.04821

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software*, *80*(4), 571–583. https://doi.org/10.1016/j.jss.2006.07.009

Chen, J., & Ran, X. (2019). Deep learning with edge computing: A review. *Proceedings of the IEEE*, *107*(8). https://doi.org/10.1109/JPROC.2019.2921977

Doyu, H., Morabito, R., & Brachmann, M. (2021). A TinyMLaaS ecosystem for machine learning in IoT: Overview and research challenges. *2021 International Symposium on VLSI Design, Automation and Test (VLSI-DAT)*. https://doi.org/10.1109/VLSI-DAT52063.2021.9427352

Doyu, H., Morabito, R., & Höller, J. (2020). *Bringing machine learning to the deepest IoT edge with TinyML as-a-Service**. https://www.researchgate.net/publication/342916900

Han, H., & Siebert, J. (2022). TinyML: A systematic review and synthesis of existing research. *4th International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 269–274. https://doi.org/10.1109/ICAIIC54071.2022.9722636

Janapa Reddi, V., Plancher, B., Kennedy, S., Moroney, L., Warden, P., Suzuki, L., Agarwal, A., Banbury, C., Banzi, M., Bennett, M., Brown, B., Chitlangia, S., Ghosal, R., Grafman, S., Jaeger, R., Krishnan, S., Lam, M., Leiker, D., Mann, C., … Tingley, D. (2022). Widening access to applied machine learning with TinyML. *Harvard Data Science Review*, *4*(1). https://doi.org/10.1162/99608f92.762d171a

Kallimani, R., Pai, K., Raghuwanshi, P., Iyer, S., & L'opez, O. (2023). TinyML: Tools, applications, challenges, and future research directions. *arXiv*. https://doi.org/10.48550/ARXIV.2303.13569

Lakshman, S. B., & Eisty, N. U. (2022). Software engineering approaches for TinyML based IoT embedded vision: A systematic literature review. *4th International Workshop on Software Engineering Research and Practice for the IoT (SERP4IoT '22)*, 33–40. https://doi.org/10.1145/3528227.3528569

Le Duc, T., Leiva, R. G., Casari, P., & Östberg, P. O. (2019). Machine learning methods for reliable resource provisioning in edge-cloud computing. *ACM Computing Surveys (CSUR)*, *52*(5). https://doi.org/10.1145/3341145

Lopez, P. G., Montresor, A., Epema, D. H., Datta, A., Higashino, T., Iamnitchi, A., Barcellos, M., Felber, P., & Riviere, E. (2015). Edge-centric computing: Vision and challenges. *Computer Communication Review (CCR)*, *45*(5), 37–42. https://iris.unitn.it/retrieve/handle/11572/114780/429278/ccr15.pdf

Rajapakse, V., Karunanayake, I., & Ahmed, M. (2023). Intelligence at the extreme edge: A survey on reformable TinyML. *ACM Computing Surveys*, *55*(13s), 1–30. https://doi.org/10.1145/3583683

Sanchez-Iborra, R., & Skarmeta, A. F. (2020). TinyML-enabled frugal smart objects: Challenges and opportunities. *IEEE Circuits and Systems Magazine*, *20*(3), 4–18. https://doi.org/10.1109/MCAS.2020.3005467

Soro, S. (2021). TinyML for ubiquitous edge AI. *arXiv*. https://doi.org/10.48550/arXiv.2102.01255

Yelchuri, H., & R, R. (2022). A review of TinyML. *arXiv*. https://doi.org/10.48550/arXiv.2211.04448

Yi, S., Qin, Z., & Li, Q. (2015). Security and privacy issues of fog computing: A survey. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9204, 685–695. https://doi.org/10.1007/978-3-319-21837-3_67

Zaidi, S. A. R., Hayajneh, A. M., Hafeez, M., & Ahmed, Q. Z. (2022). Unlocking edge intelligence through tiny machine learning (TinyML). *IEEE Access*, *10*, 100867–100877. https://doi.org/10.1109/ACCESS.2022.3207200

AUTHORS

Riya Adlakha is an aspiring researcher currently pursuing a Postgraduate Diploma in Applied Technology at Unitec, with a keen interest in diverse fields such as machine learning, data sciences, data and business analytics, and IoT. Her career as a Business Analyst at Amazon, India, served as the starting point for her foray into data-driven storytelling. Riya's work is an invitation to other enthusiasts to join them in their pursuit of knowledge and discovery. She is passionate about examining the transforming potential of research and is dedicated to making the difficult understandable, and works to close the knowledge gap in all of her endeavours.
riya.work04@gmail.com

Dr Eltahir Kabbar is a Senior Lecturer at the School of Computing Electrical and Applied Technology, Unitec. His research interests include diffusion of innovation (DOI), technology adoption, and e-government systems. His publications include a book chapter, journal articles and international conferences. ekabbar@unitec.ac.nz

Can hDAS Produce Tailored Instructional Design Methods, for the Design of Technology-Based VET Interventions?

Todd Cochrane

https://doi.org/10.34074/proc.240121
Correspondence: Todd.Cochrane@nmit.ac.nz

This publication may be cited as:

Cochrane, T. (2024). Can hDAS Produce Tailored Instructional Design Methods, for the Design of Technology-Based VET Interventions? In H. Sharifzadeh (Ed.), *Proceedings: CITRENZ 2023 Conference, Auckland, 27–29 September* (pp. 170–178). ePress, Unitec. https://doi.org/10.34074/proc.240121

Contact:
epress@unitec.ac.nz
www.unitec.ac.nz/epress/
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISBN: 978-1-99-118344-6

Unitec is a business division of Te Pūkenga –
New Zealand Institute of Skills and Technology

ABSTRACT

Hybrid educational design-based agile software development (hDAS) for producing tailored instructional design has been applied in the development of multi-user virtual environments. hDAS is being developed and applied in three contexts that provide new feed into the design and implementation of hDAS processes. The first two contexts are the design and implementation of a short course that introduces AI using a pretrained AI model to implement gesture-based game control, and vine-pruning worker training for the Regional Skills Employer programme. The third context applies hDAS in the design of computationally theoretic graphs to model learning systems. Undertaking research on hDAS systems on relatively different projects triangulates the application of the designed systems and enhances hDAS. This paper presents a summary of those projects and presents a brief reflection on the ongoing development of hDAS.

KEYWORDS

Instructional design, design-based research, agile software development, vocational education and training

INTRODUCTION

A study on how to design multi-user virtual environments (MUVEs) for vocational education and training (VET) discovered hybrid educational design-based agile software development (hDAS) for producing tailored instructional design methods for MUVE-based interventions in VET (Cochrane, 2020). Hence, hDAS is being reified in its application to instructional design to determine characteristics of its application in other educational contexts.

hDAS emerged in response to instructional design as undertaken for MUVEs. These had not been designed using strict instructional design (ISD) models; they were designed using fit-for-purpose approaches. For example, investigating instructional design using MUVEs, Soto (2013) sought to answer the question: Which ISD models are educators using to design virtual-world instruction? In his survey of 62 experienced designers, developers and published researchers of virtual world (VW), Soto found that about 67% used an ISD model, with 75% of those using the 'phased' Analysis, Design, Development, Implementation and Evaluation (ADDIE) model (Allen, 2006), and 30% using the text *The systematic design of instruction* (Dick et al., 2009). About 75% of respondents adjusted or adapted their ISD model to suit the situation. Soto (2013) concluded that ISD models are outdated, lack flexibility, do not incorporate game-design theory, and the choice of an ISD model depends on the project; that "participants in this study believed that merging various ISD models could help create unique processes to address specific learning needs." This perspective is supported by Larson and Lockee (2019) who have identified that various ISD models may not be as required for the specific situation, agreeing

> with others in the field of Instructional Design and Technology (IDT) who've called for a more responsive and rapid process for designing learning experiences and environments. … some resources on instructional design are geared for a specific career environment (higher education, K-12, business and industry, health care, non-profit, and/or government/military) and fail to provide globally applicable guidelines for the practice of ID. Others take a global approach without addressing the obvious differences in practice that exist across career environments. (p. xi)

Larson and Lockee (2019) revived ADDIE for ISD, by introducing an ADDIE iterative design model that includes continual improvement in an "infinitely" adaptive system (see Figure 1.2, Larson & Lockee, 2019, p. 9). They move ADDIE beyond Allen's (2006) assertion that ADDIE was created for developing ISD tools to meet the "challenge of the low level of training expertise being required of those functioning as the training organisation" (p. 435). Nonetheless, ongoing discussion of the position and derivation of ADDIE processes is the subject of research; for example, Kapp and O'Driscoll (2009) argue that in the development phase of ADDIE, there must be close collaboration between the instructional designer, content developer and programmer to realise the vision of the previous step. Furthermore, scripting by the programmer cannot start until the content developer has produced the environment. By treating ADDIE as a process that undergoes continual improvement to meet new situations Larson and Lockee (2019) bring ADDIE closer to agile software development practices, for example continuous development and continuous integration (Beck et al., 2001, Eck et al., 2014). In response, hDAS applies and integrates agile methodology, by inducing methods tailored for each vocation and intervention. Hence, undertaking hDAS leads to a potentially dynamic methodological process in each project (Cockburn, 1999).

This paper presents intermediate results from the application of a reified version of hDAS, which is being enacted for ISD in three different VET contexts: an introduction to programming with AI, Recognised Seasonal Employer (RSE) viticulture worker training using 3D technology, and the application of computational theoretic graphs to model VET learner journeys and systems. hDAS, as applied in each case, is considered through early reflective notes on this work in progress.

AIMS AND RESEARCH DESIGN

In the current research, hDAS methodological frames are adjusted and applied in design research for VET contexts other than ISD using MUVEs. The current research applied a minor reification in the potential hDAS methodological frames to produce structures for more general cases. The following considers generic aspects of design-based research (DBR), which were applied in the development of MUVEs for VET, that are integrated into hDAS. These are then described in terms of a reified hDAS that is to be applied in general technology-based ISD for VET.

According to Reeves (2000), in contrast to empirical research, educational design-based research starts with an analysis of practical problems by researchers and practitioners. Solutions that are developed are underpinned by theoretical frameworks and are evaluated and tested in practice. Design principles are then produced from an analysis of documentation and reflection. DBR originates from design experiments (Brown,1992; Collins, 1992). Brown (1992) describes "classroom life" (p. 141) as:

> [S]ynergestic: Aspects of it that are often treated independently, such as teacher training, curriculum selection, testing, and so forth actually form part of a systemic whole. Just as it is impossible to change one aspect of the system without creating perturbations in others, so too it is difficult to study any one aspect independently from the whole operating system. (pp. 141–142)

Brown (1992) adopted design experiments in her research in complex classroom situations. According to Cobb et al. (2003), design experiments on learning processes reach beyond the classroom. They describe the purpose of design experiments to develop theories for the broader context of learning processes:

> The purpose of design experimentation is to develop a class of theories about both the process of learning and the means that are designed to support that learning, be it the learning of individual students, of a classroom community, of a professional teaching community, or of a school or school district viewed as an organization. (p. 10)

Sandoval and Bell (2004) cite Collins (1992) as putting "forth a notion of educational research as a 'design science,' like aerospace engineering, that required a methodology to systematically test design variants for effectiveness" (p. 2).

DBR is guided by design and developmental research goals. At each stage, the problems, solutions, evaluation, and design principles are refined. The derived design principles can feed back into the other process stages in further iterations. In educational DBR, educational theories are enacted in practice and what emerges from that practice is evidence, which can be evaluated. Plomp (2007) identifies three phases in educational DBR to improve the theory base of education; that is, in the design of educational interventions with their accompanying technology. Preliminary research develops a theoretic conceptual frame as identified through need and content analysis, and literature review. In iterative design prototyping, each iteration is a research "micro-cycle" (Plomp, 2007, p. 17) from which formative evaluation produces research outcomes for refined improvement through the next cycle. An assessment phase is a semi-summative evaluation of the intervention. According to Plomp (2007), throughout DBR the researcher undertakes systematic reflection and documentation to produced theories or design principles – these are the research outcomes. Plomp (2007) identifies that educational theory is a key aspect of DBR and is used to inform the design of educational products and processes, and the evaluation of them. Improvement to educational theory is the most important output or finding. Easterday et al. (2017) consider nested scientific processes for testing theory DBR:

> DBR … does not just produce an educational intervention but makes use of nested scientific processes to produce theory in the form of novel design models of how people learn that do not just promote learning but expand our capacity to promote learning. (pp. 24–25)

However, according to Dede et al. (2004) the focus of DBR is on how the strengths and limits of a design inform theories of learning (educational theory) rather than the focus presented earlier by Easterday et al. (2014), in which the theoretical focus is in developing "theory that is useful for guiding design [of a learning situation]" (p. 317).

Collins et al. (2004) identify problems with design experiments, such as "resistance to experimental control" (p. 2), and comparison of designs due to the complexity and unpredictability of real-world situations that tend to produce large volumes of data with combinations of methodologies.

hDAS addresses these potential problems as follows. It treats data as a valuable resource that can be revisited; by undertaking hDAS 'Swim' narrative analysis the resource can be revisited, providing further insights and investigation into the structure of processes. Undertaking systematic agile software development processes provides a trail of the development of the technology and interventions; hybrid agile software development enables research into those processes.

The hDAS methodology takes a paradigmatic shift in perspective that focuses on tailored fit-for-purpose methods to produce a technology-based intervention for a vocational context and, in the process, test educational theory in the effective design of interventions in VET. The hDAS methodology deploys professionals from at least three disciplines to develop an instructional design: software engineering, education and the selected vocational discipline. Figure 1 depicts the relationships of key components of hDAS methodology. The hDAS methodology in phases determines the tailored hDAS method. Enacting the hDAS method then leads to discoveries for and from method and methodology, which leads to further tailoring of the method. The hDAS methodology is enacted in three phases. Phase Two is depicted at centre stage because the goal is to implement an intervention for VET. Phase Three builds on the other two phases with ongoing feedback and evaluation of design in research practices. Phase One is depicted between Phase Three and Phase Two to represent the origins of design and development between research-design evaluation and running an intervention.

Adjusting hDAS to produce tailored instructional design methods, for the more general design of VET teaching and learning interventions, required minor changes in its structure. Reification of hDAS frames is achieved by replacing MUVE with "for ākonga", see Figure 1. hDAS is also reified by focusing on continuously integrating learner experience into learning interventions, towards the development of a set of skills, as depicted in Figure 2, and described below.

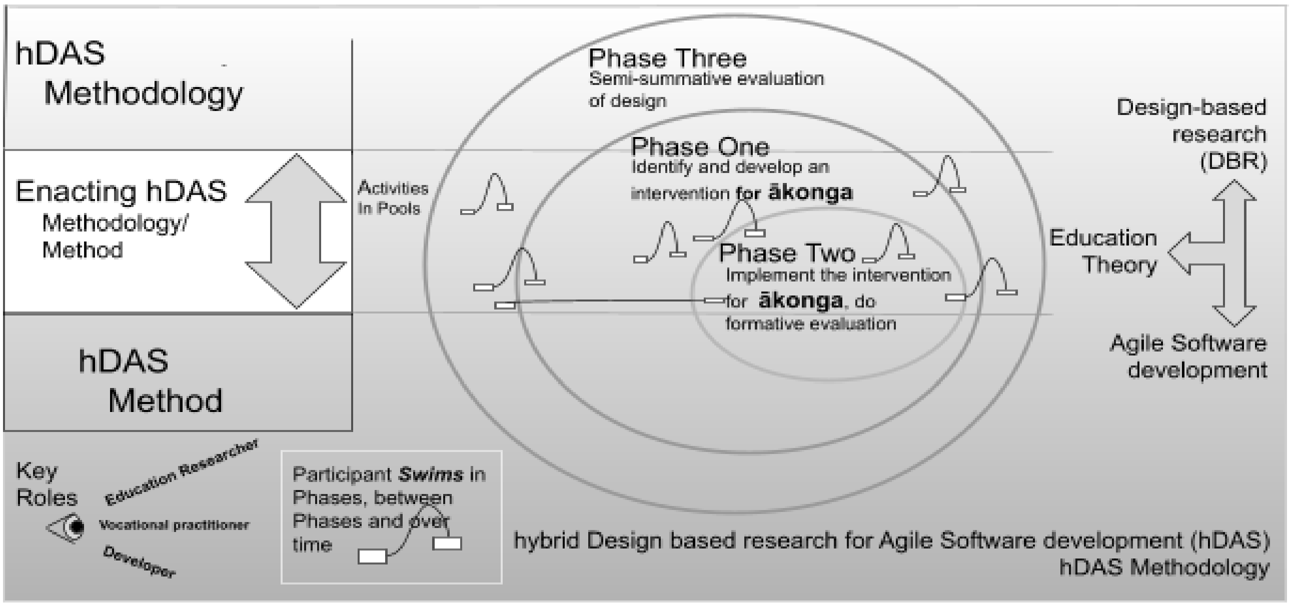


Figure 1. Reification of the hDAS methodological frame.

After removing MUVE from the methodological frame, participants in the reified frame supply the same roles as provided for the MUVE-based hDAS, see Table 1.

Table 1. Key hDAS roles by discipline and phase – reified.

Role	Discipline	Brief description	Phases
Developer	Software engineering, computer graphics, design information, technology support	A developer designs and develops digital artefacts, i.e., a software developer, or a graphics designer. A technician provides technical support for the development and during the run of an intervention.	One and Two
Educational researcher	Education	The educational researcher directs intervention development according to the research design, collects and analyses the data in research on theory in practice. At least one educational researcher is required.	One and Three
Vocational practitioner, and kaiako Vocational trainees	Selected vocation	A vocational practitioner is an expert member of the selected vocation. A teaching practitioner is preferred – kaiako. Ākonga in the vocation participate as trainees in practices of the vocation.	One and Two

Figure 2 applies the depiction of agile software development in process, to depict ākonga (learner) experiences as an ongoing process with lessons as interventions on the pathway to gaining vocational skills. As ākonga work towards gaining vocational skills they move up the diagram; interventions are provided to ensure progress towards achieving skills in their chosen vocation. Vocational practitioners identify and develop interventions for ākonga, for learning as well as for situations that impede their progress. The interventions are developed in a continuously implemented hDAS method, which applies and tests inculcated educational theory. The hDas method is tailored in phases that may run at the same time. Hence, for example, kaiako undertake ‘pathfinding' Phase One activities of the ākonga, during a lesson; at a higher level, kaiako are designing a lesson and a course for achieving skills. The derived method develops through each phase, which can be applied synchronously or interleaved over time.

The reified hDAS methodology imposes a slight meta-logical abstraction, which mildly affects the originally discovered methodological semantics, while retaining the same overall form. The adjusted approach focuses on the ākonga, while continuing to include vocationally centred design.

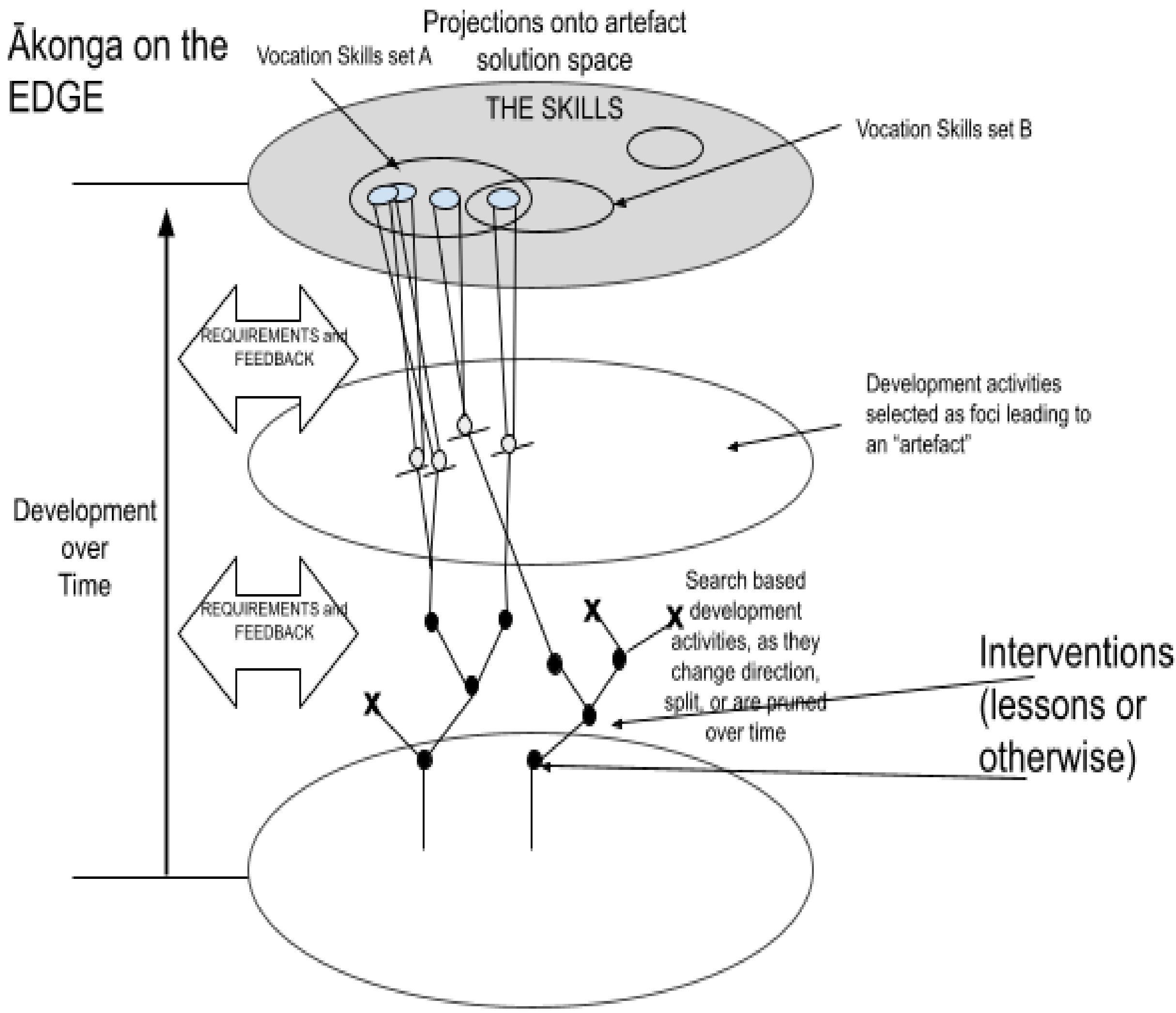


Figure 2. Agile software development, with continuous integration and development. Ākonga experiences are on the edge.

ANALYSIS AND FINDINGS

Reified hDAS is being applied to the design and implementation of interventions in three contexts: an introduction to programming that embeds artificial intelligence (AI), RSE viticulture worker training using 3D technology, and the application of computational theoretic graphs to model VET learner journeys and systems. Table 2 identifies each project and presents a summary of work undertaken.

Table 2. Three hDAS research projects.

Project	Motivation – project overview	Summary of work undertaken – with reflections
1. An introduction to programming with AI. hDAS in the design and delivery of an ML/AI workshop "You are the controller: PoseNet ML/AI game controller for flappy bird."	The goal of this research is to assess the ability to apply hDAS in an introduction to ML/AI context, through design of the workshop "You are the controller: PoseNet ML/AI game controller for flappy bird", a two-hour 'taster experience' workshop on applied software development of machine learning/artificial intelligence (ML/AI). The research tests the application of hDAS in the development of a method using legitimate peripheral participant theory and technological pedagogical content knowledge (TPACK). It will determine if the adapted software design is appropriate, and the workshop lesson plans reach the educational milestones. Data is collected from three types of participant: students – survey and focus group; software developer – software trial; focus group made up of AI experts – expert survey for design and focus group. At this point the form of the focus group is a semi-structured survey to be determined when the workshop has been designed. Expert survey is a semi-structured survey to be refined once ethical approval is given; that is based on previous theory-inculcated semi-structured interviews.	Two vocational practitioners have been recruited. One is working in the AI industry, the other is a post-doctoral academic with a teaching role. They have been interviewed using a theory-inculcated semi-structured interview, based on the original designed for the MUVE version of hDAS. A "flappy bird" application has been implemented and the workshop designed and run several times. Ākonga participants are to be recruited.
2. RSE worker vineyard training 3D tech research	The goal of this research is to assess the ability to apply hDAS in the design and application of 3D technology for RSE worker preparation. This project applies the reified hDAS method, with the same participant roles as Project One. This research considers how to design 3D digital vocational training interventions, so workers can prepare to work before they arrive. Previous studies on how to design digital technology for MUVE-based VET have been undertaken with specific disciplines – surveyors, civil engineers and maritime personnel – and a current study is underway on design of training for ML/AI. In this project, mobile telephones and a haptic pruning-shear device are to be used as technological targets in the design of learning interventions for RSE viticulture workers. The theoretic underpinning is to include educational choreography tools that are extended to team-based situations in the fields. The project includes workers, employers and kaiako practitioners.	Just received funding to undertake this research. Next stage is to interview RSE workers, employers and kaiako. Have communicated with wine researchers – who have run a short trial of a VR headset-based system. We have an agreement that allows exchange of software technology and commercially sensitive IP. A kaiako researcher is a participant as a vocational expert.
3. Te whakataurite i ngā huarahi ako mā te whakamahi kauwhata o te ako a ngā ākonga	The goal of this research is to assess the ability to apply hDAS in the design. Note, hDAS is applied in phases to determine a tailored method for the study of the selection and application of computational theoretic graphs to model learner journeys and to discover, through graph-based modelling and analysis, potential ākonga trajectories. Graph-based theoretic models of learner trajectories and their application is the subject of the design; the artefact in this design-based research. A graph-based model of vocational ākonga trajectories is to be produced that intentionally attempts to understand and reduce situational constraints. In an iterative process, ideal models are derived from an educational organisation's operating constraints and ākonga situations. These are to be compared with models of a currently operating educational organisation, and a proposed educational organisation. The design process evaluates a model's representation of learning trajectory constraints in an operating organisation and in the proposed organisation. As ākonga move towards becoming part of a discipline, they develop skills that, through practice, enhance their identity as members of that discipline. The realisation of those skills is assessed according to the discipline's criteria. Ākonga develop their skills from their situation in life. That situation identifies prerequisite skills or the need for equity in their journey. The process can be thought of as a learning path or trajectory that moves the learner further into becoming a fully participant member of the discipline. How to design graph-based models from constraints on ākonga as they learn is studied using educational DBR.	This project focuses on the application of reified hDAS for the design and testing of graph-based models of ākonga experience – the major outcome from this research at this point is depicted in Figure 2. The research enhances hDAS by reflection on its application in ISD, founded on research on learner experience.

By undertaking three research projects that apply reified hDAS, the intention is to provide data that can be used to triangulate outcomes. Data collection has started in Project One, with records of interview transcripts and the development of supporting software system. For Project Two, analysis of the outcomes from the application of a 3D VR grapevine-pruning environment for RSE workers is to be used to determine characteristics for the design of new 3D technology that supports educational choreographic design tools (Cassola et al., 2022). Project Three is substantially different from the other ongoing projects, because the artefact to be designed does not relate directly to the design of a learner intervention; it is concerned with learner experience fundamentally in the design of interventions using graph (or network) based approaches. hDAS as an approach to research in this context requires identification of vocational participants as well as ākonga. An outcome from Project Three has had the most effect on hDAS – by integrating agile software tracing notations, which are used for depicting ākonga experience trajectories, to reflect dynamic adaptive ISD.

CONCLUSION

Applied educational design-based research (DBR) has paradigmatically changed the approach to designing MUVEs for interventions in VET. DBR tests theory in practice to produce useful artefacts, and design research is a foundation through which to develop an understanding of theory in practice while producing vocationally suited MUVEs that are VWs of the vocation. The current research is ongoing, with several outcomes to be discovered. The reified hDAS is being applied in other educational technology contexts in a way that may be successful. This research needs to be pursued further to draw a clear conclusion regarding the application of reified hDAS for the design of VET interventions using technologies other than MUVE.

REFERENCES

Allen, W. C. (2006). Overview and evolution of the ADDIE training system. *Advances in Developing Human Resources*, *8*(4), 430–441. https://doi.org/10.1177/1523422306292942

Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R., C., Schwaber, K., Sutherland, J., & Thomas, D. (2001). *Manifesto for agile software development*. http://agilemanifesto.org

Brown, A. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, *2*(2), 141–178. https://doi.org/10.1207/s15327809jls0202_2

Cassola, F., Mendes, D., Pinto, M., Morgado, L., Costa, S., Anjos, L., Marques, D., Rosa, F., Maia, A., Tavares, H., Coelho, A., & Paredes, H. (2022). Design and evaluation of a choreography-based virtual reality authoring tool for experiential learning in industrial training. *IEEE Transactions on Learning Technologies*, *15*(5), 526–539. https://doi.org/10.1109/TLT.2022.3157065

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9–13. https://doi.org/10.3102/0013189X032001009

Cochrane, T. (2020). Hybrid design-based research for agile software development (hDAS) in ISD contexts: A discovery from studying how to design MUVEs for VET. *Proceedings of the Sixth International Conference on Higher Education Advances*. Polytechnic University of Valencia Congress. https://doi.org/10.4995/HEAd20.2020.11167

Cockburn, A. (1999). *Methodology per project (Humans and Technology Technical Report HaT TR 1999)*. http://alistair.cockburn.us/Methodology+per+project

Collins, A. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology*. NATO ASI Series, vol. 96. Springer. https://doi.org/10.1007/978-3-642-77750-9_2

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, *13*(1), 15–22. https://doi.org/10.1207/s15327809jls1301_2

Dick, W., Carey, L., & Carey, J. O. (2009). *The systematic design of instruction* (7th ed.). Pearson.

Dede, C., Nelson, B., Ketelhut, D. J., Clarke, J., & Bowman, C. (2004). Design-based research strategies for studying situated learning in a multi-user virtual environment. In *Proceedings of the 6th International Conference on Learning Sciences* (pp.158–165). Association for Computing Machinery. http://dl.acm.org/citation.cfm?id=1149144

Easterday, M., Lewis, D. R., & Gerber, E. (2014). Design-based research process: Problems, phases, and applications. *Proceedings of the International Conference of Learning Sciences*, *1*, 317–324. https://www.scholars.northwestern.edu/en/publications/design-based-research-process-problems-phases-and-applications

Easterday, M. W., Rees Lewis, D. G., & Gerber, E. M. (2017). The logic of design research. *Learning: Research and Practice*, 1–30. https://doi.org/10.1080/23735082.2017.1286367

Eck, A., Uebernickel, F., & Brenner, W. (2014, August 7–9). Fit for continuous integration: How organisations assimilate an agile practice. *Twentieth Americas Conference on Information Systems*, *4*, 362. https://www.proceedings.com/22766.html

Kapp, K. M., & O'Driscoll, T. (2009). *Learning in 3D: Adding a new dimension to enterprise learning and collaboration*. Center for Creative Leadership.

Larson, M. B., & Lockee, B. B. (2019). *Streamlined ID: A practical guide to instructional design*. Routledge. https://doi.org/10.4324/9781351258722

Plomp, T. (2007). Educational design research: An Introduction. In T. Plomp & N. Nieveen (Eds.), *Educational design research* (3rd ed.) (pp. 9–36). SLO. Netherlands Institute for Curriculum Development. https://ris.utwente.nl/ws/portalfiles/portal/14472302/Introduction_20to_20education_20design_20research.pdf

Reeves, T. C. (2000). Enhancing the worth of instructional technology research through "design experiments" and other development research strategies. *International Perspectives on Instructional Technology Research for the 21st Century*, *27*, 1–15. https://www.researchgate.net/publication/228467769_Enhancing_the_worth_of_instructional_technology_research_through_'Design_Experiments'_and_other_development_research_strategies_online

Sandoval, W. A., & Bell, P. (2004). Design-based research methods for studying learning in context: Introduction. *Educational Psychologist*, *39*(4), 199–201. https://doi.org/10.1207/s15326985ep3904_1

Soto, V. (2013). Which instructional design models are educators using to design virtual world instruction? *Journal of Online Learning and Teaching*, *9*(3). http://jolt.merlot.org/vol9no3/soto_0913.htm

AUTHOR

Dr Todd Cochrane is a Senior Academic Staff Member at Nelson Marlborough Institute of Technology Te Whare Wānanga o te Tauihi o te Waka o Māui, who has worked as a science programmer in epidemiology, and as a dairy-farm management software developer. He has worked with local AI companies and came to NMIT in 2013 from WelTec. He is constantly curious and is developing various systems to understand how to transfer technology to learners. He is a member of the ACM, IEEE, IEEE Education Society, and a member of the Technical Committee on Immersive Learning Environments. Todd.Cochrane@nmit.ac.nz