Issues in Information Systems
Volume 23, Issue 2, pp. 231-241, 2022

DOI: https://doi.org/10.48009/2_iis_2022_119

Comparison of simulation and infrastructure-as-code training in IT courses

Natalya Bromall, *Robert Morris University, bromall@rmu.edu*

Abstract

With a substantial amount of hands-on training required in virtually every IT program and every IT career, instructors face difficulties in providing sufficient training resources. Simulation training has become one possible solution to this issue a while ago. Another possible solution that emerged quite recently is based on cloud-based Infrastructure-as-Code (IaC) training. This paper includes a comparison of the two methods on a test group that consists of the students enrolled in five different courses that used both types of training. Each training type was found to have a substantial impact on the student's perceived knowledge improvement; however, only a few significant differences were found between the students' perception of the effectiveness of simulated and IaC training in that area. While IaC training is found to be significantly more similar to the actual environment, it is relatively new and needs improvement in functionality. The paper concludes with several improvement recommendations to the IaC training, which would make it more satisfactory to the students.

Keywords: IT training, Infrastructure-as-Code, IaC, simulation, training environment

Introduction

The importance of hands-on training for IT professionals cannot be overstated. With the continuously growing demand for the IT professionals, the nature of the IT job market is shifting to more practical and less theoretical knowledge. This shift dictates certain changes in the IT education process. The IT students need a substantial amount of hands-on training, which is important for two reasons. First, it helps them become more competitive in the job market. Second, it helps them learn by making a clear connection between theoretical knowledge they received and its practical applications in the real-world (Kapoor and McCune, 2018). While setting up a real lab environment remains arguably the best option to provide hands-on training, it is not always a possibility, and in modern world is often unnecessary. For example, if the students are being trained to use data analytics tools, maintaining a physical lab and managing various data warehouse, data mining and other similar software becomes a time consuming and often a costly option. As a workaround, educators are looking for more manageable and less costly options, such as simulation training and cloud-based training.

In simulation training, the students are placed in the environment that mimics the tool that they are learning or its certain features. This method is especially popular in the courses where a substantial portion of training includes working with hardware. For example, in a networking class the students are introduced to routers, wiring, hubs and switches, and other physical devices. If setting up a physical lab with such hardware is not an option, the simulation labs are the next best choice, even though the students work with virtual hardware instead of the actual one.

Infrastructure-as-Code (IaC) is generally understood as "an approach to infrastructure automation based on practices from software development" (Carey, 2021). This approach allows automating an entire infrastructure of the cloud resources, such as as virtual machines (instances), data storage, network configurations, and other services. IaC labs present a recently emerged training model that allows provisioning certain virtual resources in the cloud environment by running a program (typically written in JSON or YAML) for a short period of time. Such training is currently available with Amazon Web Services (AWS), MS Azure, Google Cloud and several other top cloud service providers. For example, while training the students to program data connections with MySQL and PHP, the instructor can run an IaC lab with the code that provisions the entire virtual environment in the cloud: a MySQL server instance, an Apache server, and a virtual machine that can be used for PHP coding. Figure 1 shows the template code that creates a Redshift cluster in AWS for training the students to work with data warehouse tools. The resources are provisioned for the duration of the lab, and are terminated after the lab is completed, normally within a couple of hours. The IaC model provides a flexible, reusable and inexpensive means of training and exposes the students to the actual and not simulated tools.

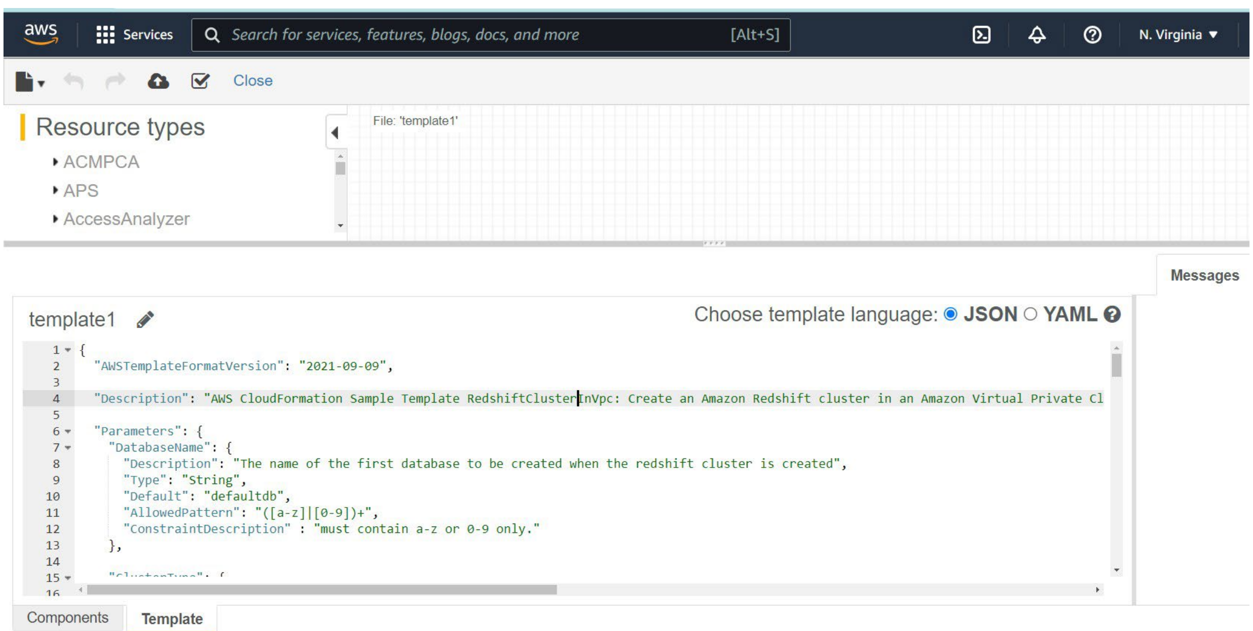

Figure 1: a sample JSON template in Amazon CloudFormation

The goal of this study is to determine if simulation and IaC training improve the students' understanding of the material, and further, if they are perceived as equally helpful in improving the students' understanding and in preparing them to their future career. The purpose of each learning tool is to (1) improve the learners' knowledge and performance in the course, (2) prepare them for their future career where they will utilize the knowledge, and (3) do it in a well-functioning and non-stressful learning environment. As the result, the following four research questions were formulated in an attempt to fulfil the goal of this research.

Research Questions and Hypotheses

RQ 1: Do simulation and IaC training improve the students' knowledge in IT courses? (The corresponding null hypothesis is $H0_1$= "Simulation and IaC training do not improve the students' knowledge in IT courses").

RQ 2: Do simulation and IaC training improve the students' readiness for their future career? ($H0_2$= "Simulation and IaC training do not improve the students' readiness for their future career")

RQ 3: Are current IaC tools more satisfactory to the students as a training option? ($H0_3$= "Current IaC tools and simulation tools are equally satisfactory to the students as a training option")

RQ 4: What are the limitations of simulated and IaC training?

Literature Review

This literature review includes three sections: first section discusses the studies that bring up the importance of hands-on experience in IT education in general. Second section includes an overview of the studies dedicated to simulation training. The third section includes an overview of the research dedicated specifically to the cloud-based education tools, or Infrastructure-as-Code (IaC) education platforms.

Hands-on Training in IT

The new trends in IT job markets and the shift of emphasis from theoretical to practical knowledge and specific skills are well reflected in literature. Wingreen and Blanton (2018) report a substantial number of research studies that emphasize the pragmatic relevance between the knowledge, skills and abilities of the IT professionals, and the organizational demands. At the same time, many employers prefer that their new IT hires have received the necessary training prior to starting their job, so that they do not require any substantial training (Taylor-Smith et al., 2019). One common path to recruiting qualified candidates is listing professional certifications as job requirements, and the importance of such certifications has grown substantially within the past years (Newstex, 2018). Since many employers request practical experience and professional certifications, more institutions include certification training programs and courses to compete with the "Certificate, then a Degree" paradigm (Gilbert & Horn, 2020). To pass a professional IT certification, the students need a substantial hands-on training, which is often problematic to the institutions due to the lack of resources/labs, lack of professional maintenance, and other issues. The studies mention two very effective methods to fill this training gap: simulations and cloud-based training with the use of virtual resources. While there are many studies of simulation-based training and a substantial number of studies dedicated to cloud-based training, there is a lack of comparison between these two training types in their effectiveness.

Simulation Training

Simulation training, or simulation environments are "such learning environments that model a real-life system or situation, computer software, a laboratory experiment setup, a business case, etc. to allow learners to practice skills and apply knowledge in a way similar to the real-life setting" (Benckendorff et al., 2015; Hallinger and Wang, 2020). Until recently, instructors used simulated labs as the most common training solution, and many studies showed this as an effective training tool. In simulated labs, the students are placed in a replica of the actual environment they are learning, and are offered to complete the specific tasks. In this type of training, the simulation offers some but not 100% of the functionality of the actual system. Simulation platforms substantially improve the education process by adding the hands-on component to the theoretical presentation (Sui et al, 2018). They prove to be especially effective when training the students in the use of hardware, such as power electronics (Chen and Lai, 2021).

The necessity of hands-on training is discussed in most research studies dedicated to IT training and education. Scholtz and Hughes, 2021 in their comprehensive review of research demonstrated, that with the correct educator facilitations, simulated training can benefit the students throughout multiple categories of skills in Bloom's taxonomy. An overwhelming number of researchers see simulation training as an excellent opportunity to provide this type of hands-on training. Ari et al. (2022) demonstrated that using simulation learning tools increased the students' higher perceived learning (assuming that they are satisfied with the simulated learning environment). Chernikova et al., 2020, found simulation-based training to be among the most effective types of training for complex skills across multiple domains. At the same time, researchers note certain issues associated with simulated learning environments, such as precision (the requirement to follow a certain sequence of steps) and inflexibility (Ari et al, 2022).

Infrastructure-as-Code Training

In the recent years, however, a new type of training emerged in education and in research gaining more and more popularity: the training based on cloud tools and "Everything-as-a-Service" concept (Qasem et al, 2019). Among other benefits, it allows the instructor to create an Infrastructure-as-Code learning environment. The instructors can utilize virtual environment with the actual tools that the students are required to learn, create temporary accounts, let the students work in the environment and terminate the account and resources after the time lapse. For example, if the students need to use a big data platform, it can be generated using the cloud tools and give the students an unaltered real-life experience (Wang, Wan and Liang, 2022).

Cloud-based training is a relatively new concept in IT education. While there are not many academic studies in that area, researchers do admit that partnerships between cloud providers and academia can solve many problems, such as the lack of physical labs and improved online programs (Doan, 2021). The first such partnership was AWS Academy platform, launched in 2015 and aimed at training the IT instructors deliver cloud courses to their students (AWS Academy, 2022). In addition to AWS Academy and AWS Educate, which are free platforms, there are many other proprietary platforms that offer cloud-based training, such as Cloud Academy and Qwiklabs. The top three cloud service providers, Amazon AWS, Microsoft Azure, and Google, are currently providing convenient educational platforms for faculty to offer training to their students utilizing cloud resources (Schwartz, 2020). As the IaC labs covered more and more content, research studies showed that along with the cloud skills, the cloud-based IaC platforms can help the students build hands-on skills in many other IT areas, such as software, operating systems, networking, security and data storage (Encalada & Castillo Sequera, 2017).

Methodology

The review of literature showed the following three directions in which hands-on training benefits the students (Table 1). The importance of pursuing these directions is stated by most employers looking for IT candidates (Hollister et al., 2017) across most IT areas: cybersecurity (Beuran et al., 2019), networking (Cui et al., 2012), bioinformatics (Revote et al., 2017) and many other.

Table 1. Summary of the benefits of hands-on training

Benefits of hands-on training
1. Develop technical knowledge and competencies
2. Improve career • Become better at the current job • Become a stronger candidate in the job market
3. Prepare for professional certification

The research instrument used in this study was designed as a survey, according to the three directions listed in Table 1. The survey included five groups of questions as listed below.

- Group 1: Knowledge and career questions. The questions in this group were formulated considering that many students did not work in their field yet, and could not provide adequate responses to the questions related to career improvement.
- Group 2: Certification questions
- Group 3: Respondents' current experience with the training tools, which impacts the effectiveness of these tools.
- Group 4: Demographic questions.

The survey was given to graduate and undergraduate students enrolled in five IT classes, all of which used either simulation or IaC cloud-based training (Networking Security, Data Analytics, Database Management, Cloud Computing, and IT Security Control & Assurance). The survey link was sent to 65 students, of which 48 students responded, and 46 response we complete and used in data analysis.

Descriptive statistics and frequency distributions were used to observe the overall tendencies in the students' responses. The scores were further analyzed individually and in groups. In determining the perceived effects of both trainings, a one-sample t-test was used to compare the scores with the "neutral" score of 3. Independent sample t-tests were done to compare the mean scores of the two groups (IaC and simulation) on each independent variable (knowledge improvement, certification, satisfaction, and individual test scores).

Findings

Demographic data and frequencies

The initial analysis of frequencies showed that the students were interested in completed the training and believed it was beneficial to them. In both groups (simulation labs and IaC labs) the majority of the students named gaining experience as their most important reason for completing the labs (38.1% of IaC group, 37.5% of simulation labs group, and 37.8% total). The most frequent unimportant reason to complete the labs was named passing the course (33.3% of IaC group, 50% of simulation labs group, and 42.2% total).

Most respondents were highly motivated students with high GPA and high performance in class (Table 2). A substantial group of students reported an intent to become certified (23% for IaC labs, 67% for simulation labs, and 45% total). This number was low for IaC labs because one of the classes that used them were not tied to any certification. In all groups, 43.5% of respondents reported that they were undecided on certification (54.5% for IaC group and 33% for simulation labs group). As expected, most respondents did not use the content of their labs prior to taking the training. Most respondents have minimal to none experience working in IT.

Table 2. Respondents' performance and intent to become certified

	IaC, %	Simulated, %	Total, %
High performance in class (A or B)	77	100	89
High GPA (3.5 or higher)	86	81	78
Intent to take cert	23	67	45
Less than 1 year in IT	65	83	74
Less than 1 year using the content	95	79	87

Knowledge improvement

The group of questions related to knowledge improvement included four statements, which the participants were asked to evaluate on a five-point Likert scale:

- IaC/Simulation labs help me master some specific components of a platform or a tool that I study
- IaC/Simulation labs help me become proficient with the entire platform or a tool that I study
- IaC/Simulation labs prepare me well for working with the actual platform or tool
- IaC/Simulation labs help preparing me for my future career in IT

The overall scores were compared to the "neutral" score of 3 using a t-test, which showed that the respondents believed the hands-on labs did help them to improve the knowledge and prepare to use it in practice (Table 3). The test was significant on .05 level for the total as well as each individual statement. It is worth noting that the average response was closer to "agree" (4 points) but not "strongly agree" (5 points), which can be explained with the fact that the respondents realized that they needed extra work to excel in using the tool in which they were trained. At the same time, so significant differences were found between the group that took simulation labs, and the IaC labs – both groups thought that the training them received helped them equally to build their knowledge in the future career. A weak significant (on the 0.1 level) difference was observed in how the respondents perceived their training helped preparing them for working with the actual platform: the group that used IaC cloud-based training rated it higher. As the result, the null hypothesis $H0_1$ was rejected.

Table 3. Total knowledge improvement scores (* = significant difference)

	t	p	Mean	St Dev	Mean Diff
Knowledge_improvement_GrandAvg	6.984	< .001	3.750	.73	.750*
Help master specific components	5.950	< .001	3.891	1.02	.891*
Helps master the whole system	4.317	< .001	3.696	1.09	.696*
Prepare for working with the actual platform/tool	5.783	< .001	3.674	.79	.674*
Prepare for future career	5.260	< .001	3.739	.95	.739*

In addition to evaluating the overall helpfulness of the hands-on training, participants were asked to specifically rank how helpful would the hands-on labs be in 11 common domains of IT; Table 4 lists the rating given to each domain.

Table 4: Helpfulness of the two types of training in different areas of IT (x^- = not helpful; x^+ = most helpful)

	IaC Group	Simulation Group
Cyber security	3.29	3.61
Development/web development	3.62	3.08^-
Database administration	3.95^+	3.54
Data analytics/science	3.76^+	3.37
Operating systems	3.52	3.92^+
Hardware	2.86^-	3.83^+
Networking	3.19	4.13^+
Human computer interaction	3.24	3.67
Cloud architecting	3.90^+	3.33
IT technical support	3.43	3.87^+
Specific application	3.82^+	3.55

The respondents in the first group ranked the IaC-based training the most helpful in learning database administration, cloud architecting, and data analytics/data science; they also admitted it can be helpful in training for a specific application. The same group ranked hardware as the only domain where IaC training is not helpful. The second group ranked the simulation straining as especially helpful in learning networking, operating systems, and IT tech support. Hardware was another category that received understandably high scores. According to the results, $H0_2$ was partially rejected for certain subject groups and IT domains.

Preparing for certification

No significant difference was found in the students' rating of helpfulness of the two types of training in preparing them for a certification exam. Due to the amount of material in the courses using IaC training, passing certification normally requires 2-6 months additional preparation after completing the course. With that in mind, it is not surprising that the majority of the respondents (70%) admitted that they were not ready to take the certification exam. In addition, data management and data analytics courses did not immediately aim at any specific certifications.

Level of satisfaction with training and limitations

Two important and statistically significant differences were shown between the IaC training and the simulation training (Table 5). As expected, cloud-based IaC training was ranked as significantly more similar to the actual platform or tool that the students were trained to use. At the same time, IaC labs were ranked significantly lower for functionality issues. The following section includes a few comments which clarify such difference. Most results did not show significant differences between IaC and simulation training; therefore, $H0_3$ was not rejected.

Table 5. Total satisfaction scores (* = significant difference; G1 – IaC group, G2 – sim group)

	t	p	Mean G1	St Dev G1	Mean G2	St Dev G2	Mean Diff
No differences between lab and actual platform	1.605	.05	3.23	.97	2.71	1.20	.52*
Includes all features of a real platform	.528	.30	3.41	.96	3.25	1.07	.16
Instructions are easy to follow	-.251	.40	3.82	.73	3.88	.80	-.06
Training system functions with no issues	-4.472	< .001	2.45	1.01	3.74	.92	-1.285*
Satisfied with grading process	-1.269	.11	3.55	.91	3.88	.85	-.33
Spend too much time completing (negative)	.181	.43	3.18	.91	3.13	1.19	.06
Satisfied with the feedback after completion	-1.200	.12	3.55	.91	3.88	.95	-.33
Easier that the actual platform (negative)	.312	.38	3.36	.58	3.29	.96	.07
Satisfaction grand average	-1.349	.09	2.9	.51	3.1	.48	-.2

Limitation of training and areas of improvement

Three issues with IaC labs and one with simulation labs were mentioned by the respondents as the most common and serious:

1. Sensitivity to network interruptions. When the lab resources are provisioned in the cloud, the system is very sensitive to the network outages, and if it detects there is no network, it may interpret it as log-out, and start terminating the resources.
2. Grading. In a simulation lab, the students are graded based on the steps they take, such as a menu option they click, or a command they type in a command line. Unlike that, the IaC labs are graded by the key benchmarks. For example, if a student is asked to create a database server with a specific DBMS and a default database, the grading system checks the presence of the server and the database, but does not care what sequence of steps the student took in creating them. If the system is not correctly programmed to grade a resource, the corresponding score is never recorded. At the same time, simulation labs lack flexibility by requiring the student to complete the task in a specific step-by-step manner.
3. Lack of provisioned compute capacity on the cloud provider's side. With the increased number of participating academic programs, the IaC labs require much more computing power, which does not seem to be matched on the cloud provider's side. As a result, the programs that provision lab resources sometime run slowly or incompletely, skipping or incorrectly creating some resources.

Conclusion

Improved knowledge in IT courses

The analysis of data showed that both types of training together were perceived by the students as helpful and improving their knowledge in the corresponding courses. At the same time, no significant difference was found between their perception of the knowledge improvement after IaC training and simulation training. The answer to the first research question is partially positive.

Improved career readiness

The answer to the second research question is similar to the first one: the study did not show any significant difference in the students' perception of how helpful the training was for the IaC group and the simulation group. Although both groups together reported that the training was significantly helpful in preparing them for certification, this result needs to be verified on larger groups with more students with intent to take certification tests. The numbers of the respondents with the intent to become certified was very uneven, with over 70% of the IaC group admitting they did not want certification, or were undecided (Table 6).

Table 6: Distribution of indent to become certified

"I intend to become certified"	IaC, %	Simulation, %
Strongly disagree	9.1	0
Disagree	13.6	0
Undecided	54.5	33.3
Agree	9.1	37.5
Strongly agree	13.6	29.2

Satisfaction with IaC training and its limitations

While the data analysis showed that the respondents were generally satisfied with either type of training, there were some significant differences between the IaC and the simulation groups' scores. The cloud-based IaC tools were admitted to be closer to the real environments; at the same time, simulation training showed less functionality issues. The recommendations to improve the IaC platforms include: (1) running the labs in the environment with stable network connections, (2) reviewing the grading algorithms to make sure they include all correct benchmarks, and (3) increasing the capacity of the temporary lab accounts to make sure all lab resources are provisioned in less time and with more accuracy.

Contribution

During the past decade, online education has shown to be the future of academia. Online learning has redefined education, especially in the post-pandemic world (Chellathurai, 2020). With the increased demands in IT employees with substantial practical knowledge and certifications, the IT programs need to provide rigorous hands-on training to the students, which is not possible without effective training tools. This study includes comparison two of such tools that are currently used: cloud-based IaC training and simulation training. It also includes incites on IaC issues and recommendations on how to improve this new but very promising technique.

The study needs will be repeated to involve a larger group of students; this will allow a multi-factor analysis that will show interaction between factors. For example, while there are no significant differences in knowledge improvement between the two groups (simulated vs. IaC), there may be a more significant difference for the interaction of the two factors (1) training type and (2) intent to become certified.

References

Ari, F., Arslan-Ari, I., Abaci, S. & Inan, F. A. (2022). Online simulation for information technology skills training in higher education. Journal of Computing in Higher Education, pp. 1-25.

AWS Academy: empowering higher education institutions to prepare students for industry-recognized certifications and careers in the cloud. Retrieved from https://aws.amazon.com/training/awsacademy/ on April 30, 2022.

Benckendorff, P., Lohmann, G., Pratt, M., Reynolds, P., Strickland, P., Whitelaw, P. (2015). Enhancing student learning outcomes with simulation-based pedagogies. Australasian society for computers in learning and tertiary education, pp. 618–620.

Beuran, R., Tang, D., Tan, Z., Hasegawa, S., Tan, Y. and Shinoda, Y. (2019). Supporting cybersecurity education and training via LMS integration: CyLMS. Educational and Information Technologies, 24, pp. 3619-3643.

Carey, S. (Dec 14, 2021). What is infrastructure as code? Automating your infrastructure builds. InfoWorld.com, retrieved from https://www.proquest.com/docview/2609864695 on May 4, 2022

Chellathurai, K. G. (2020). Future of education post pandemic COVID -19: online VS classroom learning - redefining education. The Researchers' International Research Journal, 6(2), pp. 23-30.

Chen, W.-H. and Lai, Y.-S. (2021). Module-based laboratory course for the industry-sponsored "Research and Design Master Program" in power electronics. IEEE Access, 9, pp. 100962-100971.

Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., & Fischer, F. (2020). Simulation-based learning in higher education: a meta-analysis. Review of Educational Research, 90(4), pp. 499-541.

Cui, L., Tso, F.P., Yao, D. and Jia, W. (2012). WeFiLab: a web-based WiFi laboratory platform for wireless networking education. IEEE Transactions on Learning Technologies, 5(4), pp. 291-303.

Doan, D.T. (2021). Academia-industry partnerships to develop cloud computing skills for students: a case study from AWS Academy. International Journal of Entrepreneurship, 25(1), pp. 1-10.

Encalada, W.L. and Castillo Sequera, J.L. (2017). Model to implement virtual computing labs via cloud computing services. Symmetry (Basel), 9(117), pp. 1-15.

Gilbert, C. G. and Horn, M. B. (2020). A certificate, then a degree. Education Next, 20(1), pp. 1-5.

Hallinger, P., & Wang, R. (2020). The evolution of simulation-based learning across the disciplines, 1965–2018: a science map of the literature. Simulation and Gaming, 51(1), pp. 9–32.

Hollister, J.M., Spears, L.I., Mardis, M.A., Lee, J., McClure, C.R. (2017). Employers' perspectives on new information technology technicians' employability in North Florida. Education & Training, London, 59(9), pp. 929-945.

Kapoor, A. and Gardner-McCune, C. (2018). Understanding Professional Identities and Goals of Computer Science Undergraduate Students, In proceedings of SIGCSE'18, February 21-24, 2018, Baltimore, MD, USA.

Qasem, Y.A.M., Abdullah, R., Jusoh, Y.Y., Atan, R. and Asadi, S. (2019). Cloud computing adoption in higher education institutions: a systematic review. IEEE Access, 7, pp. 63722-63744.

Revote, J., Watson-Haigh, N.S., Quenette, S., Bethwaite, B. McGrath, A. and Shang, C.A. (2017). Development of a cloud-based bioinformatics training platform. Briefings in bioinformatics, 18(3), pp. 537-544.

Scholtz, F., & Hughes, S. (2021). A systematic review of educator interventions in facilitating simulation-based learning. Journal of Applied Research in Higher Education, 13(5), pp. 1408-1435.

Schwartz, N. (2020). Why 4-year colleges are tapping amazon to help deliver cloud computing degrees. Education Dive, Washington. Retrieved from https://www.proquest.com/trade-journals/why-4-year-colleges-are-tapping-amazon-help/docview/2466124070/se-2 on May 3, 2022.

Sui, T., Liu, X., Lu, D., Shao, C. and Cheng, F. (2018). Research on the construction of teaching case library of the computer simulation technology. Multimedia Tools and Applications, 78, pp. 1183-1199.

Taylor-Smith, E., Berg, T., Smith, S., Meharg, D., Fabian, K. and Varey, A. (2019). Bridging the Digital Skills Gap, In Proceedings of ITiCSE '19, July 15–17, 2019, Aberdeen, Scotland UK, Association for Computing Machinery.

Wang, Z., Wan Y. and Liang, H. (2022). The impact of cloud computing-based big data platform on IE education. Wireless Communication and Mobile Computing, Article ID 9740407.

Weblog post (2018). Why has IT certification training become the need of the hour? (2018). Chatham: Newstex. Retrieved from https://www.proquest.com/docview/2102241423 on May 3, 2022.

Wingreen, S.C. and Blanton, J.E. (2018). IT professionals' person-organization fit with IT training and development priorities. Information Systems Journal, 28, pp. 294-317. https://doi.org/10.1007/s12528-021-09303-0.