AI in Cybersecurity Education- A Systematic Literature Review of Studies on Cybersecurity MOOCs

Samuli Laato
Dept. of Future Technologies
University of Turku
Turku, Finland
sadala@utu.fi

Ali Farooq
Dept. of Future Technologies
University of Turku
Turku, Finland
ali.farooq@utu.fi

Henri Tenhunen
Dept. of Future Technologies
University of Turku
Turku, Finland
henri.tenhunen@utu.fi

Tinja Pitkämäki
Dept. of Future Technologies
University of Turku
Turku, Finland
tinja.e.pitkamaki@utu.fi

Antti Hakkala
Dept. of Future Technologies
University of Turku
Turku, Finland
antti.hakkala@utu.fi

Antti Airola
Dept. of Future Technologies
University of Turku
Turku, Finland
antti.airola@utu.fi

***Abstract*—Machine learning (ML) techniques are changing both the offensive and defensive aspects of cybersecurity. The implications are especially strong for privacy, as ML approaches provide unprecedented opportunities to make use of collected data. Thus, education on cybersecurity and AI is needed. To investigate how AI and cybersecurity should be taught together, we look at previous studies on cybersecurity MOOCs by conducting a systematic literature review. The initial search resulted in 72 items and after screening for only peer-reviewed publications on cybersecurity online courses, 15 studies remained. Three of the studies concerned multiple cybersecurity MOOCs whereas 12 focused on individual courses. The number of published work evaluating specific cybersecurity MOOCs was found to be small compared to all available cybersecurity MOOCs. Analysis of the studies revealed that cybersecurity education is, in almost all cases, organised based on the topic instead of used tools, making it difficult for learners to find focused information on AI applications in cybersecurity. Furthermore, there is a gab in academic literature on how AI applications in cybersecurity should be taught in online courses.**

***Index Terms*—cybersecurity, MOOC, machine learning, AI, systematic literature review**

I. Introduction

Cybersecurity concerns everyone. Data management and privacy, malware protection on personal devices, secure banking services and democratic elections are part of ordinary peoples' lives and they are all impacted by the increase in *Big Data* collection and advances made in machine learning (ML) [1], which is a form of artificial intelligence (AI) as shown in Fig 1. In the current study, as we discuss the implications of the latest AI solutions for cybersecurity, we consider mostly ML.

AI has been embedded in information systems across all industries [2] and at the same time reports of AI failures and cybersecurity issues related to AI have increased. It has been projected that by 2021, 75% of all enterprise applications use some ML-based AI [3], increasing its relevance in cybersecurity. Scholars have pinpointed that AI failures in complex systems such as banking may have catastrophic consequences with no options for recovery [4]. These failures can be, for example, unintended behaviours which were not predicted by its creators such as *Sophia Android* proclaiming to destroy humans or the Russian *Promobot IR77* trying to escape its lab confounds [2], [5]. As AI failures have a potential impact on society as a whole, it is important to provide state-of-the-art training for cybersecurity professionals as well as basic education of AI in cybersecurity for citizens. Governments have also addressed the rising concerns about ML by introducing legislation such as the GDPR to restrict the data collection capabilities and privileges of companies [6].

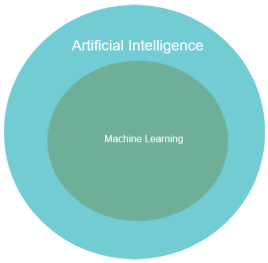

Fig. 1. The relationship between AI and ML

It is not only AI system failures that are troubling, as AI can be used offensively, as demonstrated by adaptive malware created using ML techniques [7]. ML is also capable of arming otherwise unusable data as a potential weapon [8]. On the other hand, AI has defensive potential such as being able to predict and fend off malicious activities [9]. Despite the

2161-377X/20/$31.00 ©2020 IEEE
DOI 10.1109/ICALT49669.2020.00009

increasing involvement of AI in cybersecurity, the current solutions do not make other cybersecurity measures obsolete.

Previous studies on information and cybersecurity domains have identified various key aspects and categories of cybersecurity. For example, Blanco *et al.* [10] proposed the following main concepts for cybersecurity: vulnerabilities, threats and attacks, controls and countermeasures, and, security protocols, mechanisms and policies. This list was further extended by Hakkala and Isoaho [11] by including the concepts of data security and information criticality. Other categorisations also exist, as Darraj et al. [2] categorise AI in the field of cybersecurity to 16 categories, some of which can still be divided further. Vähäkainu and Lehto, discuss 11 sub-categories of cybersecurity which have been impacted by ML -based applications: infrastructure security, endpoint security, application security, IoT security, web security, security operations and incident response, threat intelligence, mobile security, cloud security, identity and access management, network security and human security [12]. AI and ML have not been considered in these previous studies and categorisations explicitly, even though related concepts such as data security and criticality are examined.

Because of the increasing involvement of AI in cybersecurity, educators must consider in what ways and what kinds of AI applications should be involved in its teaching. Recently scholars have argued, that cybersecurity education should rely on the industry workforce to provide expertise in teaching, which would ensure students receive up to date information [13]. This idea has already been put into practice with several companies joining universities to create massive open online courses (MOOCs) on AI or building their standalone online courses. For example, Google offers a course on ML called *Machine Learning Crash Course*, which teaches to use their *TensorFlow* APIs, and a company called *Reaktor* has partnered up with The University of Helsinki to create a free course titled *Elements of AI*.

To investigate how AI and cybersecurity are currently taught together to the large public via MOOCs, we conduct a systematic literature review on empirical studies on cybersecurity MOOCs. As our primary concern is on how AI is currently being taught, we search the studies for actual courses, their description and design philosophy. Consequently, we formulate our research questions as follows:

- How do existing cybersecurity MOOCs describe and categorise their learning content. Does it involve AI?
- What lessons on how to implement AI teaching in cybersecurity MOOCs can be learnt from previous studies?

II. Research design

To answer the research questions, it is not enough to simply look at existing MOOCs from popular platforms such as *Coursera*, *edX* or *Future Learn*, as the courses do not necessarily include the academic analysis of the design philosophy and purpose of the course. Looking at studies on cybersecurity MOOCs on the other hand can include elaboration on the design principles and have the benefit of academic peer-reviewed commentary on what should be taught in the courses. The findings from studies can then be supplemented by looking at currently available courses.

A. Literature Search

The PRISMA guidelines for a systematic literature review were adopted [14] as the method has stood the test of time and is considered rigorous by scholars across disciplines. The Scopus database was chosen for the search, as it indexes most prominent computer science databases such as ACM, IEEE, Springer, and the DBLP Computer Science Bibliography [15]. Furthermore, it allows efficient objective search tools for researchers. We used the following search string:

"(TITLE-ABS-KEY ("mooc" OR "massive open online course" OR coursera OR edx OR udacity OR futurelearn OR "online course" OR "open learning") AND TITLE-ABS-KEY (cybersecurity OR "cyber security" OR "information security" OR "system security" OR "computer security" OR "network security" OR "IT security"))"*

to identify all potential papers discussing cybersecurity MOOCs. The search was carried out in January 2020 and resulted in 72 hits. All 72 papers were scanned and the following were excluded:

- Duplicate work
- Proceeding descriptions
- Unobtainable records (not even paid version available)
- Records in a language other than English

After this, 56 items remained. Following the initial scan, the abstracts of all 56 papers were read to identify which studies considered online cybersecurity courses. In case it was not clear after reading the abstract whether this was the case or not, the full text of the paper was scanned. Following this process revealed 15 peer-reviewed studies on cybersecurity online courses with the earliest one being published in 2003 and the latest in 2019. The entire literature search process is summarised and displayed in Fig 2. In addition to the paper - items, a scoping search was done on popular MOOC platforms as well as a regular search engine search identifying existing courses about AI applications in cybersecurity which were not covered in the found literature. 10 items were found, and even though they are analysed separately from the peer-reviewed studies, we have also included them in Fig 2.

B. Analysis and synthesis of results

Following the literature search, the resulting 15 papers were observed in detail. The papers were divided into two groups based on whether they considered a single cybersecurity courses or multiple. The primary information regarding the courses was obtained from the studies, as they were expected to contain meta-level information about the purpose of the course, its design philosophy and possible reasons for involving or not involving AI. Secondary information regarding the course was obtained from the actual course page when available, where the learning goals and topics of the course were retrieved. The aim was to look at how many of the courses involve AI, and those which do, in what way.

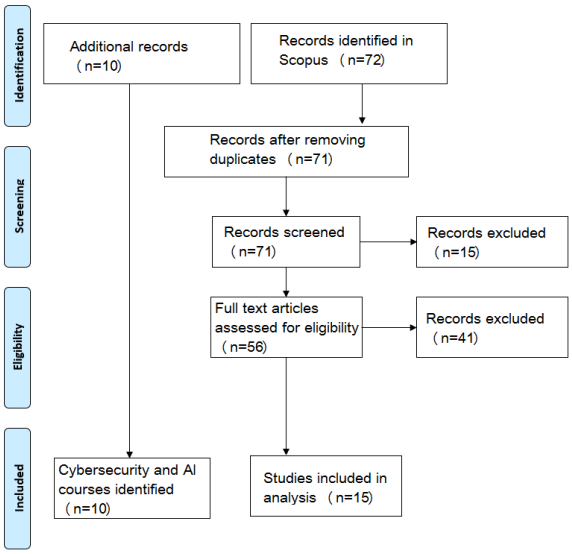

Fig. 2. The PRISMA method applied in the current research

III. Results

A. Studies on multiple courses

Three papers studying multiple cybersecurity MOOCs were identified and are displayed below in Table I. Gonzalez-Manzano and Jose de Fuentes used the NICE reference framework [16], which aims to describe cybersecurity work roles applied in any sector, to go through 35 currently available cybersecurity MOOCs [17]. The paper by Riabov discusses 24 online computer science courses of which especially one is a security course [18]. The paper by Poulova and Simonova briefly describe nine courses, some of which have a connection to cybersecurity, and invoke data from participants and subsequent analysis to evaluate the success of the courses [19].

TABLE I
Studies of multiple cybersecurity courses

Paper	N of courses	Analysis method
[17]	35	evaluation of content
[19]	9	analysis of participant data
[18]	24	evaluation of content

The studies looking at multiple cybersecurity courses were limited in scope and quantity. The only major work in the area was done by Gonzalez-Manzano and Jose de Fuentes in 2019 [17]. This particular work, however, is quite exhaustive, as it looks at 35 free cybersecurity MOOCs ranging from beginner level to advanced from the worlds currently most popular MOOC platforms. Even though the majority of the courses (25/35) were developed and produced by American universities, they are free MOOCs and hence offered to a global audience. What is interesting from the point of view of the current study is that despite these courses covering up to 33 speciality areas in cybersecurity, they do not explicitly involve the use of AI applications in cybersecurity. Many of the areas such as *Test and Evaluation*, *Network Services*, *Cyber Defense Analysis* and *Technology R & D* would naturally involve ML. It is possible that the courses have been updated with ML content, even if not explicitly mentioned.

B. Studies on individual courses

Altogether, twelve papers on individual cybersecurity MOOCs (or online courses) were found [20]–[31]. Whereas the studies observing multiple cybersecurity MOOCs were all published after 2013, the studies on individual courses are more separated, with first studies emerging in 2003 and a big chunk of new studies published later on in 2017 as shown in Table II. Out of the twelve observed studies, only eight could be traced back to an existing and operating course. However, even then, one course was in Japanese [28] and another in Russian [27] and therefore information of these could not be obtained beyond what was described in the paper. Thus, the studies which were linked to courses which we could explore and analyse were *Online Master of Science in Computer Information Systems Concentration in Security* [20], *Network Risk Assessment* [31], *Certified Ethical Hacker v.10* [26], *I Secure Agent* [25], *cybersecurity informatics* [29] and *Free 9-week online software security class* [24].

TABLE II
Studies on individual cybersecurity MOOCs

Paper	Publication year	Analysis Method
[23]	2003	design description
[22]	2005	hybrid course evaluation
[20]	2007	evaluation of content
[21]	2007	design description
[31]	2007	content and evaluation
[30]	2016	content evaluation
[24]	2017	content evaluation
[25]	2017	analysis of participants
[29]	2017	content evaluation
[28]	2017	content evaluation
[27]	2018	content evaluation
[26]	2019	content evaluation

The course topics varied from general security awareness courses [27], [28] to gamification of cybersecurity MOOCs [25]. The finding that the studies approach cybersecurity from multiple angles is understandable as there are tens of sub-categories and topics involved in cybersecurity [2], [17]. Table I showed that most studies focused on evaluating the content of their course in some method. One of the studies explicitly mention ML as part of the course objectives [29]. Chung describes *Applying machine learning and network science approaches to the prediction of cybersecurity phenomenon* as one of five primary course objectives [29].

C. Synthesis of Teaching AI techniques in Cybersecurity Courses

Unsurprisingly older studies on cybersecurity MOOCs did not contain evidence of AI being involved in the course even if they might have relevance in, for example, data security

and encryption, white-hat hacking and fuzzy penetration algorithms. The one study which gave design recommendation for cybersecurity MOOC designers did not consider the topic of the courses in detail or mention ML or AI, even if it might have been part of the course materials [17]. Only one of the screened studies explicitly discussed ML [29], however, 10 courses were discovered from popular MOOC platforms and by searching online which combined AI and cybersecurity. These courses are listed in Table III. It is unclear how well these courses have generated interest. For example, the Udemy course *"Cybersecurity data science"* currently has only 244 enrolled students, compared to another Udemy course *"The Complete Cyber Security Course: Network Security!"* which has 88211 enrolled students.

TABLE III
CYBERSECURITY MOOC FOCUSING ON AI APPLICATIONS

Course name	Offered by
Cybersecurity Informatics	UCF [29]
AI for Cybersecurity	Oxford University
ML in Cyber Security	Mario Fritz
Cybersecurity Data Science	Udemy
Elastic ML for Cybersecurity	Elastic
Applied DS for Cybersecurity	Center for CCT
Applied DS and ML for Cybersecurity	Blackhat
ML for Network Security	Tertiary courses
AI for Cybersecurity	Uni of Queensland
Training course ML in Cybersecurity	Peerlyst
CI in Cybersecurity	Uni of Jyväskylä

Current cybersecurity MOOCs sort content based on application areas, and not the method, meaning that AI is not explicitly mentioned in course descriptions of most courses, even if it would be present. This approach is logical from the point of view that AI is simply a tool to be used in the described application areas. Furthermore, looking at courses from a historical perspective, it is more natural to add the new upcoming content (AI) to existing structures than to change teaching suddenly to AI first. One of the benefits of making a course about the application of AI in cybersecurity would be that of highlighting the key areas where AI is revolutionising or has already revolutionised cybersecurity, as done in the courses displayed in Table III.

IV. DISCUSSION

A. Key Findings

We summarize the findings from the literature review in four points:

- There exists relatively few case studies on cybersecurity MOOCs compared to existing available MOOCs.
- Cybersecurity MOOCs organize educational content in most cases based on covered topics instead of the methods (such as AI).
- Even from the most recent MOOCs, only a few mention teaching ML techniques applied to cybersecurity.
- Domain-specific pedagogical studies on how to teach AI applied in cybersecurity, or which applications of AI should be covered, are missing.

B. Teaching Application of AI in Cybersecurity

Following the findings of the current study, a new question arises: *How should we teach about applications of AI in cybersecurity MOOCs?* Two strategies can be envisioned: (1) approach cybersecurity the traditional way and discuss AI when relevant; and (2) approach cybersecurity from the perspective of what new AI can offer existing solutions. Almost all areas of cybersecurity can find potential benefits from AI. On the other hand, there are some areas where advances in AI research have had more impact than in others. We argue that both approaches (1 and 2) are needed. The first one is useful for students who wish to obtain a holistic understanding of cybersecurity and learn to master some aspects of it. The latter is especially useful for lifelong learners and those who wish to obtain an understanding of the current state of cybersecurity quickly. As evident from Table III, some courses have been created focusing on ML techniques in cybersecurity, providing evidence that there is a need for such courses. In addition, some MOOC platforms are offering *overview* courses on the impact of technology on society which also discusses cybersecurity and AI generally [32]. Most these types of courses were designed for the beginner level.

C. Limitations

The systematic literature review method revealed that there are surprisingly few studies on cybersecurity MOOCs. Perhaps because of this, previous studies concerning cybersecurity MOOCs have resorted into searching existing courses instead of literature [17]. Furthermore, the literature search was carried out only in one meta-database, Scopus. Despite its versatility and coverage of most important publication venues, more studies could have been identified if additional research databases. Some literature search methods such as that proposed by Barbara Kitchenham [33] also involve *snowballing*, that is, searching through the bibliography of initially found studies until all possible leads have been exhausted. However, such method was not needed for the aim of the current study. The analysis of found studies and MOOCs was limited by the researchers linguistic capabilities as some courses were not in English [27], [28].

D. Future work

There are relatively few case studies on actual cybersecurity MOOCs. Designers looking into best practices or pedagogical strategies on how to teach certain topics would arguably benefit from such studies. Furthermore, following the result that there is a lack of MOOCs on applications of AI in cybersecurity, future work could involve creating more of such MOOCs focusing on key technologies and areas where ML is relevant. Undoubtedly such courses would be on the advanced level, narrowing down the number of participants looking for them, which might discourage some parties from committing resources into creating those MOOCs.

V. Conclusion

We investigated how the application of AI has been taught in cybersecurity MOOCs and what design philosophies exist by systematically reviewing existing peer-reviewed studies. The results showed that there are surprisingly few studies concerning cybersecurity MOOCs compared to the amount of courses currently offered. Furthermore, all courses, which were discussed in the papers, were organised based on their topic, and none based on the applied method (such as AI). This can be limiting for students looking to specifically learn about how AI is used in the domain of cybersecurity. Finally, only a couple of courses mentioned AI in their course content. These challenges have been addressed by previous work by suggesting that the industry would work together with academia to update course materials to include AI [13]. Altogether, the rapid increase in the popularity of ML applications does not yet show in studies on cybersecurity MOOCs. Updates on existing courses are required to ensure learners receive up to date information on the impact of AI on cybersecurity. New courses could look into organising content based on which applications of AI are most relevant for cybersecurity.

References

[1] J. Qiu, Q. Wu, G. Ding, Y. Xu, and S. Feng, "A survey of machine learning for big data processing," *EURASIP Journal on Advances in Signal Processing*, vol. 2016, no. 1, p. 67, 2016.

[2] E. Darraj, C. Sample, and C. Justice, "Artificial intelligence cybersecurity framework: Preparing for the here and now with ai," in *ECCWS 2019 18th European Conference on Cyber Warfare and Security*, p. 132, Academic Conferences and publishing limited, 2019.

[3] G. W. Romney, J. Guymon, M. D. Romney, and D. A. Carlson, "Curriculum for hands-on artificial intelligence cybersecurity," in *2019 18th International Conference on Information Technology Based Higher Education and Training (ITHET)*, pp. 1–8, IEEE, 2019.

[4] R. V. Yampolskiy and M. Spellchecker, "Artificial intelligence safety and cybersecurity: A timeline of ai failures," *arXiv preprint arXiv:1610.07997*, 2016.

[5] G. Bhattacharjee, "The curious cases of unruly robots," *Science Reporter*, 2018.

[6] P. Voigt and A. Von dem Bussche, "The eu general data protection regulation (gdpr)," *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, 2017.

[7] S. Mohanty and S. Vyas, "Cybersecurity and ai," in *How to Compete in the Age of Artificial Intelligence*, pp. 143–153, Springer, 2018.

[8] S. Wachter and B. Mittelstadt, "A right to reasonable inferences: rethinking data protection law in the age of big data and ai," *Columbia Business Law Review*, 2019.

[9] M. Taddeo and L. Floridi, "How ai can be a force for good," *Science*, vol. 361, no. 6404, pp. 751–752, 2018.

[10] C. Blanco, J. Lasheras, E. Fernández-Medina, R. Valencia-García, and A. Toval, "Basis for an integrated security ontology according to a systematic review of existing proposals," *Computer Standards & Interfaces*, vol. 33, no. 4, pp. 372–388, 2011.

[11] A. Hakkala and J. Isoaho, "Defining and measuring key expertise areas in information security for higher education students," in *Proceedings of the International Conference on Engineering Education ICEE 2015*, 2015.

[12] P. Vähäkainu and M. Lehto, "Artificial intelligence in the cyber security environment," in *ICCWS 2019 14th International Conference on Cyber Warfare and Security: ICCWS 2019*, p. 431, Academic Conferences and publishing limited, 2019.

[13] R. Heller., C. Toregas., and L. Hoffman., "Reach to teach: Preparing cybersecurity experts as adjunct community college faculty," in *Proceedings of the 11th International Conference on Computer Supported Education - Volume 1: CSEDU,*, pp. 338–343, INSTICC, SciTePress, 2019.

[14] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: the prisma statement," *Annals of internal medicine*, vol. 151, no. 4, pp. 264–269, 2009.

[15] B. Morschheuser, J. Hamari, J. Koivisto, and A. Maedche, "Gamified crowdsourcing: Conceptualization, literature review, and future agenda," *International Journal of Human-Computer Studies*, vol. 106, pp. 26–43, 2017.

[16] W. Newhouse, S. Keith, B. Scribner, and G. Witte, "National initiative for cybersecurity education (nice) cybersecurity workforce framework," *NIST Special Publication*, vol. 800, p. 181, 2017.

[17] L. González-Manzano and J. M. de Fuentes, "Design recommendations for online cybersecurity courses," *Computers & Security*, vol. 80, pp. 238–256, 2019.

[18] V. Riabov, "Tools and methodologies for teaching online computer-science courses in lms environment," *Learning Management Systems and Instructional Design: Best Practices in Online Education*, pp. 144–171, 2013.

[19] P. Poulova and I. Simonova, "Innovations in teaching computer networks subjects," *2016 IEEE Conference on e-Learning, e-Management and e-Services, IC3e 2016*, pp. 29–34, 2017.

[20] S. Kalathur, L. Chitkushev, S. Jacobs, T. Zlateva, and A. Temkin, "A course on computer and network security: Teaching online versus face-to-face," *IFIP International Federation for Information Processing*, vol. 237, pp. 57–64, 2007.

[21] C. York, D. Yang, and M. Dark, "Transitioning from face-to-face to online instruction: How to increase presence and cognitive/social interaction in an online information security risk assessment class," *International Journal of Information and Communication Technology Education (IJICTE)*, vol. 3, no. 2, pp. 41–50, 2007.

[22] S. Tabor, "Experiments with hybrid learning in a computer & network security course," *Association for Information Systems - 11th Americas Conference on Information Systems, AMCIS 2005: A Conference on a Human Scale*, vol. 4, pp. 1855–1858, 2005.

[23] N. Sinha, "Design and development of a distance learning course in computer data security," *Proceedings of the IASTED International Conference on Computers and Advanced Technology in Education*, pp. 391–394, 2003.

[24] C. Theisen, T. Zhu, K. Oliver, and L. Williams, "Teaching secure software development through an online course," *CEUR Workshop Proceedings*, vol. 1977, pp. 19–33, 2017.

[25] A. Antonaci, R. Klemke, C. Stracke, M. Specht, M. Spatafora, and K. Stefanova, "Gamification to empower information security education," *CEUR Workshop Proceedings*, vol. 1857, pp. 32–38, 2017.

[26] T. Nguyen, "Certified ethical hacker v.10 online course - a case study," *ACM International Conference Proceeding Series*, pp. 168–173, 2019.

[27] N. Aleksandrova, M. Khramova, and S. Kurkin, "Computer safety basics training for the older generation," *Proceedings of the 2018 International Conference "Quality Management, Transport and Information Security, Information Technologies", IT and QM and IS 2018*, pp. 542–544, 2018.

[28] H. Ueda and M. Nakamura, "Deployment of multilanguage security awareness education online course by federated moodle in japan," *Proceedings - International Computer Software and Applications Conference*, vol. 2, pp. 49–52, 2017.

[29] W. Chung, "Developing curricular modules for cybersecurity informatics: An active learning approach," *2017 IEEE International Conference on Intelligence and Security Informatics: Security and Big Data, ISI 2017*, pp. 164–166, 2017.

[30] C. Au, K. Lam, W. Fung, and X. Xu, "Using animation to develop a mooc on information security," *IEEE International Conference on Industrial Engineering and Engineering Management*, vol. 2016-December, pp. 365–369, 2016.

[31] Y. Bai, W. Summers, and E. Bosworth, "Teaching network risk assessment to online graduate students," *InfoSecCD'07: Proceedings of the 4th Annual Conference on Information Security Curriculum Development*, pp. 50–55, 2007.

[32] UCSSanDiego and B. Simon, "Teaching impacts of technology in k-12 education specialization," *Coursera, https://www.coursera.org/learn/teach-impacts-technology-data, checked 22th of January, 2020*, 2020.

[33] B. Kitchenham, "Procedures for performing systematic reviews," *Keele, UK, Keele University*, vol. 33, no. 2004, pp. 1–26, 2004.