E-exams in Norwegian higher education: Vendors and managers views on requirements in a digital ecosystem perspective

Aparna Chirumamilla [*], Guttorm Sindre

Department of Computer Science, Norwegian University of Science and Technology, 7034, Trondheim, Norway

ARTICLE INFO

ABSTRACT

E-assessment has been supported in Learning Management Systems for decades. More recently, dedicated e-exam systems have emerged on the market, more specifically supporting the workflow and security needs surrounding high stakes exams. For instance, in Norway, LMS's Canvas and Blackboard are only used for ungraded assessment tasks, while e-exam systems like WISEflow and Inspera Assessment are used for graded ones. Since the systems are mass-market software, vendors must satisfy the needs of several customers, and needs that are specific to only one or a few customers will receive low priority, perhaps forcing teachers to adapt their assessments to what the tool supports, rather than having the tool adapt to the preferred pedagogy. So far, there has been considerable research on views of students and teachers on e-exam systems, much less on the views of vendors and managers. In this paper, we investigate what these stakeholder groups consider to be the key features of e-exam systems, and by what process they are determined. An exploratory case study was conducted, based on interviews with 12 participants belonging to three different groups: vendors, process manager and system managers in Norwegian universities. Our findings indicate much agreement among these groups about key features of e-exam systems, though observing that not all functionality requested by end-users will be prioritized. Also, there was much agreement that a movement towards standardization, open interfaces and digital ecosystems would allow smoother integration with other information systems in the higher education sector, and easier addition of plug-ins for specific functionality – but that there still is a way to go to reach the ambitions of a flexible ecosystem. Currently, vendors give more priority for adding functional features in e-exam systems rather than better interoperability, and integration with third-party tools remains a challenge.

1. Introduction

The higher education sector (HE sector) is currently going through a massive digital transformation (Llamas-Nistal, Fernández-Iglesias, González-Tato, & Mikic-Fonte, 2013; Sandkuhl & Lehmann, 2017). One exciting trend is the digitisation of assessment. While e-assessment has been with us for decades (Buzzetto-More & Alade, 2006; Frosini, Lazzerini, & Marcelloni, 1998), we are now witnessing a much more consistent shift. Many countries have ambitions to abandon traditional pen and paper exams within five years (Butler-Henderson & Crawford, 2020; Fluck, 2019).

The general trend for e-assessment products is towards mass-market software, with competing vendors targeting the same needs.

Many universities around the world use Blackboard or Canvas as their learning management system (LMS). For high-stakes exams, they may supplement the LMS with software to safeguard against cheating, such as the Respondus Lockdown Browser (Cluskey Jr, Ehlen, & Raiborn, 2011). Norwegian universities also have Blackboard or Canvas but use dedicated e-exam systems such as Inspera or WISEflow for graded tests. Mass-market software allows development cost to be divided on a large number of customers. On the other hand, there are also challenges with mass-market software:

- Functionality will often be a compromise between needs of various customers, often imperfectly understood by vendors (Naous, Giessmann, & Legner, 2020). Less common needs (e.g., of specific courses or innovative assessment practices) will tend to get low priority.
- For high-stakes e-assessment, the need to mitigate cheating makes security requirements particularly important (Dawson, 2016). Cheating can be defined as breaking the rules of an exam to gain unfair advantage (Cizek, 1999) and is a many-faceted issue. Partly, it is about ensuring the correct identity of the candidate and correct authorship of the delivered answer (Mellar, Peytcheva-Forsyth, Kocdar, Karadeniz, & Yovkova, 2018), but also about ensuring that candidates do not use forbidden aids or illegitimate collaboration during the exam (Dick et al., 2002).
- Mass-market software products may be hard to integrate with other systems that the university already has. Integration requirements are often hard to capture in the acquisition stage (Lauesen, 2006). The modern approach to interoperability is to move away from proprietary systems and instead use open standards and governance frameworks, so that many different software products can collaborate smoothly in a digital ecosystem (Kerssens & Dijck, 2021, pp. 1–14) – as also proposed within e-learning (Uden, Wangsa, & Damiani, 2007) and e-assessment (Llorens, Molina, Compañ, & Satorre, 2014; Luo & Lin, 2013). In spite of improvements towards this vision in recent years, interoperability remains a major challenge in the e-learning domain (Chituc & Rittberger, 2019). Poor interoperability may cause a lot of double work, e.g., re-entering of data, meaning that the administrative simplifications and cost savings that one hoped to achieve from e-exams, may not materialise (or be smaller than expected).

Both limited functionality and trade-offs between functionality, cheating prevention and interoperability mean that teachers may end up having to adapt their assessment practices to what the exam tool allows, rather than having the tool adapt to the assessment practices that are ideal from a pedagogical point of view.

Our main research question is: *What are the key features for e-exam software, according to vendors, process managers, and system managers, and how are such features identified and agreed upon?*

In connection with this main question, we have four sub-questions:

1. What process is followed by vendors, process managers, and system managers to identify features and agree upon requirements?
2. What do vendors, process managers and system managers see as key features for functionality and security in e-exam systems?
3. What are the goals and challenges concerning integrations between the e-exam system and other exam supporting systems?
4. To what extent do the stakeholders envision a move towards a more open digital ecosystem for e-exams, and would this be assumed to impact security?

Notably, our main research question omits two of the most important user groups for e-exam systems, namely students and teachers. Instead, we focus on stakeholders involved in developing e-exam systems (i.e., vendors), and in acquiring and operating e-exam systems for the university sector (i.e., system managers involved in acquisition and operation of e-exam systems, and process managers involved in negotiating requirements). This selection of informants does not mean that the views of students and teachers are considered unimportant. However, the views of students and teachers on e-exam systems have been studied in several other publications (Fluck, 2019; Kuikka, Kitola, & Laakso, 2014; Wibowo, Grandhi, Chugh, & Sawir, 2016), while less has been published on the views of vendors, system and process managers. Moreover, with research questions focusing on rather technical issues like security, interoperability, and digital ecosystems, students and teachers except for the most technologically competent might be weak informants. Obviously, the requirements of students and teachers should be essential when a university is acquiring e-exam technology. However, the requirements of teachers would likely differ a lot from person to person, e.g., depending on the discipline taught and the assessment practice followed. Among students, there will also be much variation, e.g., based on personal preferences and special needs. The advantage of system managers (i.e., license administrators, project managers, team leaders, advisors, engineers) at a university is that they receive feedback from many students and employees, e.g., concerning dissatisfaction with the system or requests for new functionality, thus able to present a more aggregate idea of requirements. Similarly, process managers employed by the Ministry of Education and Research run joint procurement processes on behalf of all public universities and vendors (i.e., development managers, head of product development, product managers) who sell to several universities should also possess such aggregate views on needs. Hence, these three groups of people should be knowledgeable about general trends concerning requirements for e-exam software.

2. Literature review

There have been several case studies concerning usage of e-exam systems, e.g., (Fluck, 2019; Fluck, Adebayo, & Abdulhamid, 2017; Fluck, Pullen, & Harper, 2009; Kuikka et al., 2014; Wibowo et al., 2016), typically eliciting viewpoints of students and teachers to inform the further development and operation of such systems. They have reported various advantages for e-exam systems, such as improved exam logistics, support for auto-marking and question authoring – and for harvesting data that can be used in learning analytics (Fitzharris & Kent, 2020). Especially, Bring Your Own Device (BYOD) e-exams, which are cloud services letting students use

their own laptops during the exam – have become increasingly popular, offering reduced costs and increased scalability for universities (Fluck & Hillier, 2017; Hillier & Fluck, 2013), and better accessibility and usability for students (Fitzharris & Kent, 2020) by using the equipment they are already familiar with.

Necessary security features for high stakes e-exams have been discussed by several authors. Huszti and Petho (2010) focussed on authentication, anonymisation of candidates, and confidentiality of questions. Mitigation of cheating is considered especially challenging for BYOD e-exams since the student controls the hardware (Dawson, 2016; Frankl, Schartner, & Zebedin, 2012). The two most prominent approaches to secure such exams – requiring booting from a memory stick or accessing the exam through a lock-down browser – both have their pros and cons. Security features for BYOD exams using Moodle LMS are provided through "Secure Exam Environment" in (Frankl et al., 2012) that mainly use wired LAN and boot from USB or DVD. Kaiiali, Ozkaya, Altun, Haddad, and Alier (2016) proposed a "Secure Exam Management System" for the security of BYOD exams run on Moodle LMS through WIFI. With remote exams, it becomes even more challenging to mitigate cheating through impersonation (another person taking the exam) or collaboration, answer sharing and plagiarism among students (D'Souza & Siegfeldt, 2017), which has been observed with many remote exams during the Covid lockdown period (Bilen & Matros, 2021). For such situations, authorship verification combined with remote proctoring technology utilizing biometry and candidate monitoring can be helpful (Okada, Noguera, Alexieva, Rozeva, Kocdar, Brouns et al., 2019). However, our study took place before the Covid lockdown, so the context for our interviews was the then-typical usage of e-exam systems in Norwegian universities, namely for on-campus exams, with face-to-face human invigilators checking the identity of students. Hence, remote proctoring technology is not much addressed in this paper.

As for design research, Adebayo and Abdulhamid (2014), Brink and Lautenbach (2011), Ferdiana and Hoseanto (2018), and (Fluck, Pálsson, et al., 2017) all report on various design and implementation experiences with e-exam systems, with combinations of in-house development and already available software. When it comes to more general frameworks, Striewe (2019, pp. 220–228) based on a literature review, proposes components and design alternatives for e-assessment systems, for each component pointing out also the features that the component would be supposed to cover. Isaias, Miranda, and Pífano (2019) made a framework of eight evaluation criteria for e-assessment systems: variety of design options, scalability, security, access and usability, feedback features, personalisation, cost and interoperability. The framework was validated by a questionnaire gaining responses from academic staff across 37 countries. Among the criteria, the highest level of agreement was for variety of question types and feedback features, and for interoperability. The EU project TeSLA is a collaboration between a number of universities across Europe, implementing a system for secure online exams, as reported for instance in (Baró-Solé, Guerrero-Roldan, Prieto-Blázquez, Rozeva, Marinov, Kiennert et al., 2018; Mellar et al., 2018; Okada, Noguera, et al., 2019; Okada, Whitelock, Holmes, & Edwards, 2019). The project has focussed on security against cheating especially for remote online exams, combining biometry and authorship verification, and also on how technology can facilitate pedagogical improvement in assessment. TeSLA has also had a strong focus on accessibility for students with special needs, and interoperability with other e-learning tools (Ladonlahti, Laamanen, & Uotinen, 2020, pp. 213–238).

E-exam systems need to interoperate with other software products performing complementary functions in the organisation, such as the Student Information System, LMS, and single sign-on system for authentication. Such interoperability challenges are somewhat discussed by (Dagger, Connor, Lawless, Walsh, & Wade, 2007; Jakimoski, 2016). Also, e-exam systems may need to interoperate with other e-exam systems, to allow universities to move and exchange information and collaborate on joint question bases for various disciplines (Sclater, Low, & Barr, 2002). Major obstacles to interoperability could be that systems use different interfaces and data formats. Standardization can reduce this problem. Specifically for e-learning tools, the IMS Learning Tools Interoperability (LTI) standard allows external tools to be launched within an application (Queirós, Leal, & Paiva, 2016, pp. 381–386; Severance, Hanss, & Hardin, 2010). The Question and Test Interoperability (QTI) allows tests and questions exported from the e-exam system of one university to be imported to the e-exam system of another university (Wills, Davis, Gilbert, Hare, Howard, Jeyes et al., 2009). However, Piotrowski (2011) considered the QTI specification too complex, ambiguous, and challenging to implement, and Sclater (2007) argued that interoperability testing must be included at acquisition to ensure that the vendors really deliver on this. In addition to adhering to standards, products with an open and well-documented Application Programming Interfaces (API's) will more easily allow for plug-ins so that universities may customise the e-exam system according to their needs (Chirumamilla & Sindre, 2019).

All in all, the publications reviewed above do present required features for e-exam systems in various ways, but none of them directly respond to our main research question of investigating these features as seen by vendors, system and process managers. Closely related to our study is one by Foss-Pedersen and Begnum (2017), targeting the same stakeholder groups that we have interviewed, using a questionnaire survey plus interviews. However, their focus was on the e-exam systems' support for universal access, not towards functional features in general, nor for security or interoperability.

3. Research approach

3.1. Case context

Previously, each university or college in Norway might run separate procurement processes for e-learning products, though some might also have collaboration and joint procurement. The last big procurement done by a single university was NTNU's choice of Blackboard as its new Learning Management System (LMS) in 2017. Nowadays, the national organization Unit[1] (Directorate for ICT

---

[1] https://www.unit.no/.

and Joint Services in Higher Education and Research, Merger of CERES, BIBSYS and parts of UNINETT)) works as process manager and runs joint procurement processes on behalf of all Norwegian higher education institutions, which could both increase bargaining power and save resources, as procurement processes are labour intensive. Hence, when most other Norwegian universities got Canvas in 2016, this was a result of such a joint process. Similarly, Unit currently manage Norwegian HE institutions' dialogue with e-exam system vendors and development of requirements. They also have responsibility for the development and maintenance of custom software, such as FS (Felles Studentsystem) which is a Student Information System (SIS) in use by almost all higher education institutions in Norway. The architecture diagram in Fig. 1 shows various systems involved, with links indicating information exchange. FS (second from left) contains authoritative information about students (e.g., personal information, enrolment, course registration, exams scheduled, grades received, etc.), courses, teachers, etc. StudentWeb (left) is a front-end to FS where students can register or withdraw from courses/exams, view and appeal grades, etc. Blackboard and Canvas are typical LMS's, which in Norway are used to handle communication within courses – except for exams and graded coursework, which will be delivered through the e-exam system. Two mass-market software products are used for e-exams in Norway, and some universities use Inspera Assessment (hereafter IA), others WISEflow (hereafter WF). Both are proprietary software products, run as cloud services using lock-down browsers (top and bottom) to mitigate cheating. Further to the right are some other systems involved, the document archival system, the single-sign-on authentication (used with several systems, but we only show links to the e-exam systems to avoid messing up the diagram), and the plagiarism checking tool, where Norwegian HE currently use Urkund.

3.2. Research design

The research questions for this study are exploratory, hence a qualitative approach would be most appropriate, as indicated by Robson (2002). A quantitative survey might have been used for this study if the purpose was to compare stakeholder views on features of e-exam tools in terms of ranked preferences. However, the research questions are not about the relative importance of the features in numerical terms, but rather what the key features are and why they are important. As argued by Yin (2002), such research questions are best suitable for case studies. The exploratory nature of the research questions points towards inductive reasoning (Braun & Clarke, 2006; Twining, Heller, Nussbaum, & Tsai, 2017), using the exploratory case study approach (Yin, 2017) to develop theory from the case.

In the literature, case studies have been mainly designed to be as either single case study or multiple-case study (Baxter & Jack, 2008; Gustafsson, 2017; Weishaupl, Yasasin, & Schryen, 2018). Although our study spans several universities and two different system vendors, it is most appropriate to define it as a single case study, rather than multiple cases since the two e-exam systems have been acquired in a joint process run by Unit, for use in various universities in Norway, with both vendors responding to the same set of requirements. The participants for this study were people involved in the larger procurement and development process coordinated by Unit. The responses from interviews did not provide much variation depending on whether a university was using Inspera Assessment and WISEflow – if there were more variation, a multiple case study would have been more appropriate (Yin, 2002), While a single case study does not achieve the same breadth as multiple cases, the single case gives more time for investigating that one case, thus potentially promoting deeper understanding of the subject, as suggested by Dyer & Wilkins, 1991. The two system vendors also have customers in other countries, but since only Norwegian HE institutions were included, the case is "e-exam tool support for higher education in Norway".

3.3. Data collection and analysis

Fig. 2 illustrates the process of data collection and analysis used in this study. Based on research questions, suitable participants were identified. We used a combination of key informant sampling (contacting persons known to be in central roles, with expertise on the topic) and snowball sampling (getting suggestions from initial participants about other potential participants), with the aim of covering both vendors IA and WF, and universities using each system. All in all, we interviewed n = 12 participants from 7 different organizations. Participants (cf. Section 1 & Appendix B) were vendors, process managers, system managers at universities who were primarily involved in the execution of digital exams, and all consented to have their interview data researched and published in anonymized form.

Semi-structured interviews (Kallio, Pietila, Johnson, & Kangasniemi, 2016) were most appropriate for our purpose, keeping some structure for comparability between participants while at the same time allowing for participants to bring forth issues we might not have thought about. An interview guide (cf. Appendix C) was prepared and distributed to participants before interviews. All interviews were done by the first author, during April 2019–Aug 2020, some face-to-face (4) and some (8) via Skype and Zoom video calls. Each interview lasted approx. 40 mins.

The interviews were transcribed line by line by the first author and coded in NVivo 12, using the constant comparative method, which has the advantage of making the analysis more explicitly theoretical (Urquhart, Lehmann, & Myers, 2010). It was first proposed by Glaser, Strauss, and Strutzel (1968) in their grounded theory methodology, and further practically explained by others (Charmaz, 2006). We used it for data analysis in the same way as (Fluck, 2019). The purpose of our study is to mainly focus on answering the research questions rather than identifying emerging theory (Boeije, 2002).

We had a set of predefined categories from our research questions. Hence, our analysis was focused on finding the relation between the concepts that emerged from the analysis, and on grouping those concepts under our predefined categories. First, data analysis involved open coding of the responses using 'in vivo' in NVivo 12. This open coding abstracts concepts from the data. Second, axial coding was performed for grouping of the codes from open coding and further categorising the codes. Lastly, selective coding was
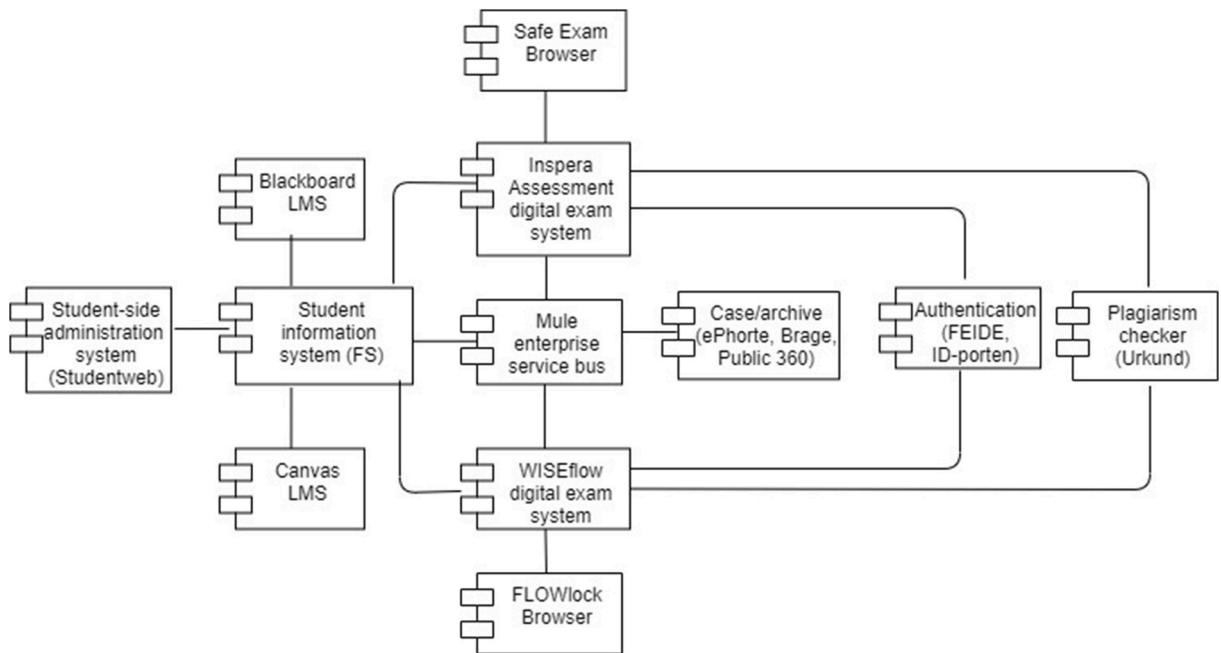
Fig. 1. Exam solutions interfaces [Adapted from (Melve & Smilden, 2015, p. 102)].
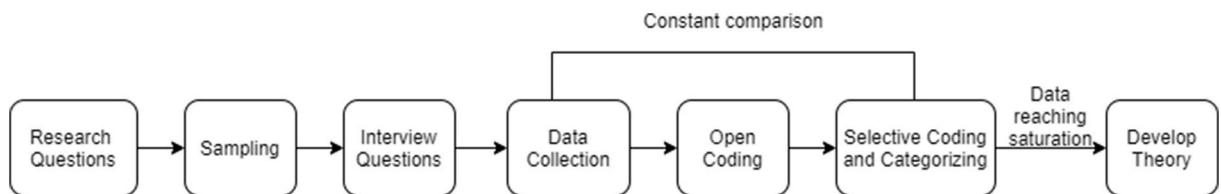

Fig. 2. Figure illustrating case study design.

conducted to interpret the relation between codes and categories from axial coding. The data collection, coding, and analysis were done together to enrich the existing category. During the coding process, the constant comparison of data and codes was made to compare responses and decide what data will be gathered next until the data get saturated. The coding was done by the first author but discussed with the second author along the way.

4. Results

This section presents the results emerging from the interviews, structured according to the research questions. Table 1 shows the most prominent concepts identified in data analysis.

4.1. Requirements engineering process

There were findings (cf. Table 2) related to several aspects of the Requirements Engineering (RE) process (also called procurement process by participants), e.g., elicitation, analysis, negotiation, prioritisation, and specification of requirements.

Elicitation of requirements was facilitated by UNIT, the directorate overseeing joint procurement as mentioned in the previous section. They collect requirements from system managers in universities who again get input from end-users. UNIT align requirements with government strategy and evaluate e-exam systems based on offers from vendors. Pilots and feedback from universities were given high weight in the comparison of products, and some universities had local procurement processes before UNIT took the coordinating role. However, system managers still asserted that there was no structured RE process followed during procurements. Vendors and system managers used different tools for specifying the requirements, e.g. Confluence as their collaboration software program. Prodpad and Confluence were used by vendors to document customer input, and system managers are using OneNote and Visio for designing of the processes.

After procurement, requirements for future releases have been analysed and prioritized through seminars conducted by UNIT and vendors, with participation by universities. If universities have critical requirements that were not included in the requirements

Table 1

Most prominent concepts identified from data analysis.

Categories	Grouped concepts (or codes)
Requirements engineering process	procurement process, requirements elicitation, requirements negotiation, requirements prioritisation, requirements specification
Key features	key functional features, key security features
Security	cheating prevention, cheating detection, cheating threats, security in BYOD exams vs exams in university-owned PCs, technical issues, vulnerabilities in tools, mitigations
Integration and interoperability	integration architecture, integration between e-exam system and LMS, integration between e-exam system and student information system, integration between e-exam system and lockdown browser, security challenges during integrations
Digital ecosystems	content sharing (e.g., sharing questions), monolithic vs digital ecosystem, open API requirements, third-party tools integration with e-exam system, impact of digital ecosystems on security

Table 2

Findings on the RE process.

RE activity	Stakeholder	Statements of Stakeholders
Elicitation	• Vendors	• "Unit come up with requirements from universities, based on those we provided offer, we also try to have a direct dialogue with end-users." (IA1)
	• System managers	• "We have had pilots on IA and WF, and in total, there were only four vendors [to choose from] at the start. So we [gathered] experiences of other schools, and then we tested it ourselves, but a lot has changed since [2015]" (NTNU2).
		• "We don't have a structured RE process yet, that's something we would like to do. We collaborate with [… some Norwegian universities] and Unit and write user stories. Then we [discuss] with vendor." (NTNU1)
	• Process managers	• "We ask universities about their requirements and align them with government strategy. We will have bi-weekly meetings with IA and WF together, to agree on workflow and integration for everyone so that we don't have to make a separate integration for every university." (Unit)
Analysis	• Vendors	• "We have webinars, user groups, annual seminar with customers for analysing, negotiating and prioritising which features to develop further in which order." (IA1)
		• "We negotiate in multiple ways. We have to negotiate to change contracts because […. When] the contract is a year or two old [… customer] needs might have changed." (IA1)
		• "Requirements prioritisation is always tricky. Because there might be functionality that's important to one segment that isn't important to another customer segment." (IA1)
		• "We prioritise […] based on votes and polls." (IA2)
		• "We don't build stuff for only one customer, we build it for everyone" (WF)
	• System managers	• "We had the local procurement in the beginning before national procurement. There, we gave points on requirement specifications, and WF ended up with the best score." (HVL).
	• Process managers	• "There haven't been that many conflicts. When something developed not in line with our requirements, we have dialogue with the vendors. We discuss on checkpoints. […]" (Unit)
Specification	• Vendors	• "We use many different tools, e.g., 'Prodpad' to get customer input and the final roadmap, different templates to define the requirements and acceptance criteria in confluence. For public tenders, we use 'UPO', which is a specialised tool for RE processes." (IA1)
	• System managers	• "Unit is more like a facilitator rather than specifying our requirements themselves" (NTNU1)
		• "We use OneNote and Visio. When we collaborate on the documents with other universities, we have one institution responsible for that area, and it is the editor, but always others can contribute. Then we move requirements to vendor's wiki and get a specification." (NTNU1)

specification, this may lead to contract changes. Requirements differ from country to country, and even between universities in the same country. Products aim to satisfy the generic needs of several customers. One of the challenges mentioned by vendors is to balance needs of various customers. Hence, they will consider the prioritized customer list or perform votes and polls about future features.

Regarding the similarities and differences in procurement process between countries, IA vendors and process managers mentioned that only Sweden has a similar setup as Norway with 'Sunet'. In contrast, UK has certain de facto standards concerning LMS systems identity federation (eduGAIN), but no central authority organizes procurement. Process managers especially mentioned that International organizations, Geant (Europe) and NORDUNET (Nordic countries) attempted to conduct joint procurements. Vendors and process managers thought that other countries could benefit from considering a procurement process similar to the Norwegian one:

"There are benefits from standardizing workflows and integrations, for reaching many customers with the same solution." (IA1)

"I think it requires a culture of sharing and collaborating rather than competing in addition to a government requirement to do so […] and depends on the Higher education sector attitudes, autonomy, and willingness to work together (even though in some cases it will result in development taking longer) rather than competing." (Unit)

4.2. Key features and qualities for e-exam systems

Functional features of e-exam systems face various user groups and address several process stages related to the exam. Table 3 shows functionality mentioned in interviews sorted by key user group (students, teachers/censors, administration) and process stage (before, during, after exam). Some of the cells are empty, as no such features were mentioned as important by the participants.

E-exams need several different questions types, ranging from multiple choice and other auto-scored formats (where WF's question types are based on Learnosity[2]) to free-text essays and file uploads. For the writing of lengthy text, some participants favoured a limited scope of support from the e-exam system. A mainstream word processor would have better functionality than a more limited text editor inside the e-exam system, hence it was perceived more user friendly to write the text outside the e-exam system and just upload the document: *"Most of our teachers except a few in engineering and health subjects, don't use that [the in-built editor]"* (HVL) *"Most of our exams are made outside WF system and are uploaded into WF as a pdf, and students write answers using a word processor in FLOWlock."* (UiT).

Some participants were familiar with both IA and WF from procurement elaborations, often comparing the two systems in their answers. Generally, participants tended to focus less on well-established functionality, more on recent and maybe challenging features. Especially, many talked about grade explanations and appeals, previously done outside the e-exam systems but recently included in the feature scope of e-exam systems. Examples of statements concerning functional features can be found in Table 4.

The key non-functional features required for e-exam systems were found to be scalability, usability, integrity and interoperability, security, and reliability (cf. Table 5). Scalability is important due to a huge load of exams in the peak period, some with large classes. Usability is essential as end-users are quite diverse, some with limited computer skills. We received more responses on integration and interoperability between systems, which we provide in detail in Section 4.3. As for integrations and interoperability, participants felt integration of exam systems with student information system and archival systems were the most important.

Security was seen as essential for the validity of the exam results: *"To make sure that data is authentic and secure both in terms of securing the data of the ones taking the exams when somebody is breaking in to tamper with exams and test-takers breaking out [of the lock-down] during exams."* (IA1). Security and reliability are also crucial due to the cost of redoing an exam: *"it has to be very secure because it's challenging to do it over again if something went wrong with exams and grades"*. (NTNU4). As can be seen in some of the quotes in Table 5, authentication of graders through single-sign-on was considered to have good security, while less secure ad hoc solutions like access links had to be used for foreign graders unable to use the national single-sign-on systems. Some participants considered WF's closed source FLOWlock browser more secure than the open source SEB browser used by IA: *"SEB being an open-source browser is however a problem we will have to deal with for as long as we still use the open source browser."* (NTNU2).

Participants were reluctant to reveal concrete ways of cheating but indicated that the system would not be 100% secure: *"We don't want to share […] what we know about how you can cheat. We […] caught one student who [utilized] one of the known issues that we have had in WF. That issue was fixed. […] However, I'm sure that a hacker student would probably be able to do something with the coding of any program installed."* (UiT).

Table 6 shows a number of quotes related to the mitigation of cheating during exams. One finding is that mitigation of cheating has a trade-off with scalability, e.g., it is risky to allow usage of third-party tools like MATLAB or Excel in BYOD exams, so these must rather be held in computer labs – but many institutions do not have sufficient capacity: *"University PCs might be secure, but we don't have labs, so we only use BYOD laptops."* (KU).

Some of the BYOD exam cheating practices mentioned during interviews include code injection using Macbook pro touch bar and Macro keyboard, speech to text, manipulating configuration file, modifying open source code, cheating via SEB configuration tool and third-party tools. Some of the mitigations mentioned include logging, penetration testing at vendors-site, using lockdown browsers, using ad-blocker and proxies for university computers, using virtual desktop infrastructure (VDI), in-built monitoring feature from the e-exam solution, and using passwords from invigilators. Participants mentioned that students also use cheating approaches which are totally outside the BYOD exam laptop, such as cheat notes, communicating through mobiles, wireless earpieces: *"We don't have many [discovered cheating] cases on invigilated onsite exams because it's difficult to find out cheating, e.g., with a note in their pocket and when reading it in the toilet, use their mobile phones when they are not allowed to use."* (HVL) and *"it happens outside of the computers like writing on the notes or hiding a phone in the bathroom. But that will happen if the exam is on the paper or if it is on their computer."* (KU).

Universities in Norway mainly depend on human invigilators for monitoring students during exams, typically retired elders who are hired part-time for the exam season. Biometric surveillance, though available in WF, is not used out of privacy concerns: *"We haven't used the facial recognition function of WF at HVL because we have the invigilators checking the ID of the student"* (HVL) and *"WF has face recognition functionality, but we don't have that because of regulation laws, and I think nobody in Norway uses that functionality."* (KU).

4.3. Digital ecosystems

Participants addressed different aspects of how e-exam systems might fit into a larger digital ecosystem, such as integrations between e-exam systems with other information systems that the university already has (cf. Fig. 1), support of third-party tools by e-exam systems, well-documented Application Programming Interfaces (API) to enable easy integration, support of third-party plug-ins, and usage of standard data formats to enable sharing of content between systems and universities.

---

[2] https://learnosity.com/.

Table 3
Functional features stressed by participants, sorted by user group and process stage.

User group \ Process stage	Before exam	During exam	After exam
Students		Answer questions Upload documents Deliver exam	Receive grades Seek, receive explanations Appeal grades
Teachers (and censors)	Authoring (of tests, questions) Upload documents		Grading Explanations Analytics
Admin	Logistics	Monitoring	

Table 4
Findings on key functional features.

Key features	Stakeholders	Statements of Stakeholders
Authoring	• System managers	• "In WF, you make a flow based on what kind of exam you're holding, e.g., FLOWlock, FLOWassign, whereas in IA, you make the exam and add features to exam based on what kind of exam you're holding." (UiT) • WF's authoring tool only supports FLOWmulti exam." (UiT)
Logistics system	• Vendors	• "Scheduling of exams, including things like the number of students and rough estimation of requirements for that room, e.g., power, special software, and grant access for students with disabilities." (IA2)
Question analytics	• Vendors	• "Ensure that exams that can be provided are fair, accurate, have relevant question types, help users providing good questions and give insight into the difficulty of questions and ability of students." (IA1)
Grading	• System managers	• "WF is more user-friendly than IA. it wasn't easy to see the whole grading process in one page in IA, but in WF it's easier to keep overlook of the whole." (KU)
Explanation of grades	• System managers	• "IA supports explanation for the test but not for the whole course since there is no synchronisation between FS and IA, e.g. if the project weighs 30% of the grades and final test weigh 70% of the grade." (NTNU2) • "Currently, there is an opportunity to add an explanation in IA, but it will not notify the student, so students have to log on to IA to check the explanation." (NTNU5) • "At the moment, censor may write a comment in WF, and share it with a students, soon they will see it when the grade is published. But it's not automatic yet." (HVL) • "At UiT, WF has been integrated with FS for grading explanations. So, an explanation written in WF can be exported to Studentweb now." (UiT)
Appeals and complaint grading	• Vendors	• "We have implemented appeal function in IA, which can manage requests for appeals, appeals grading, and invite graders into IA for appeal grading. It's based on integrating additional data from FS." (IA1)
	• System managers	• "Students request appeal from Studentweb then it goes to FS. When examiners get notified from FS, and when it is graded in WF, it goes back to FS. Students find the grade in Studentweb." (HVL)

There was no overarching standard or framework used by vendors for integrations. Instead, different standards, practices and a collection of APIs have been used as a framework for integrations:

We follow a lot of best practices and standards. Data are often presented via different industry standards, e.g. IMS QTI or OneRoster. We use REST API's and provide Swagger documentation which is a widely used way to document APIs that sort of serves us as the framework. It is not an ISO standard as such, but it is sort of one of the industrial practices for documenting. But well-established standards often ease the integrations with new systems. (IA1).

We build our own framework and then use some technology standards depending on the customer systems that we integrate with. (WF).

Our developers use MuleSoft for the integrations, which transfers the information from exam systems, FS, and other systems to the right place. (Unit).

Vendors mentioned that they use REST API's for in-house integrations, whereas process managers mainly used Mule soft Enterprise Service Bus (ESB) for API integrations of different systems at a time.

Several challenges with integrations were reported during interviews (cf. Table 7). First, some of the integrations were not standard based yet, e.g., SEB lockdown browser was integrated at the functional level of IA rather than full integration between these two systems. Integration with the FS administrative information system was complicated both by low data quality and synchronisation issues, many of these problems were on the side of FS due to lack of features and interfaces. System managers also thought that non-uniformity in choosing solutions by universities for integrations have led to slower implementations.

According to participants, neither WF nor IA have been integrated with the LMS (Canvas or Blackboard) although many teachers have expressed wishes for such an integration (e.g., wanting to use the same pool of questions both for formative and summative assessment). One key obstacle for the integration is incompatibility between their Question and Test Interoperability (QTI) specifications. As an alternative solution, these e-exam systems support interoperability through IMS Learning Tools Interoperability™ (LTI),

Table 5
Findings on key non-functional features.

Key features	Stakeholders	Statements of Interviewees
Scalability	• Vendors	• "Earlier IT has not involved with exams other than special needs accommodation, so our involvement has been different than before, e.g., instead of 100 people, we have 1700 at one time." (IA1)
	• System managers	• "E-exam system can allow as many students, e.g. 1500 students to take a different type of exams at the same time on the same day." (KU)
Usability	• System managers	• "It's easier for students to download, use, update, follow system requirements, and find the information they need in WF, but in IA the steps are difficult when you're not known to the system. In contrast, the dashboard is better in IA." (KU)
Integration and Interoperability	• Vendors	• "From the content point of view, questions and data in IA can be exported and can be moved to other exam systems and the same for submissions that they can be archived and exported in standardised formats." (IA1)
	• System managers	• "IA is integrated with 'Brage' [publication system for academics and research] and 'ePhorte' [archiving system for exam data], but most of the public sector uses public 360." (NTNU1)
		• "Integration between FS and WF transfers student and assessors' data, exam times, assessor deadlines, exams relevant data from FS to WF, and WF to FS such as grades." (UiT)
Security	• Vendors	• "Currently IA auto-save exam data in cloud storage by Amazon web services." (NTNU1) and "[...] same for WF" (WF)
		• "Security depends on the type of encryption and its length." (IA1)
		• "IA exam data is hosted within a virtual private cloud of Amazon web services. So the data has limited access, firewalls protection, encryption, logging, and data lies in redundancy in three different physical locations." (WF)
	• System managers	• "Feide authenticates students and internal censors, and we use ID-porten for the external censors. But sometimes when we have foreign censors, we have to send them a link through WF to access that specific exam." (HVL)
		• "We can also use access tokens, but we try not to use it unless external doesn't have a Norwegian social security number. But sometimes, we have to give the access token to faculty, e.g., when systems or routines are so slow." (KU)
		• "WF supports in-built monitoring feature called FLOWmonitor for monitoring exams." (HVL)
Reliability	• Vendors	• "We manage and execute the exams securely from end-end in a stable manner. Because when something happens, that obstructs the exam then reset of the exam is very cumbersome, and expensive for customers." (WF)
		• "There is an allocated time and place for the exam to happen. Otherwise, the exam might not be fair to everyone or able to be completed." (IA1)
		• "When a submission is handed in, it is being reflected as test taker intended, and evaluator can see what the student intended for them to see." (IA1)

to launch e-exam application as a tool from LMS. But this received low priority at universities:

Currently, this has not been highly prioritized in Norway. So, we have not started working on integration. Norway has somewhat different regulations, thereby mandating that assignments that count towards final grade should be conducted in a specific assessment system and not in the LMS and wanting to avoid confusion for students of what system is to be used for assessments and assignments. (Unit).

In this paper, interview responses are presented according to categories (themes) and user group (i.e., stakeholders) in Tables 2, 4–7, 11 since the responses are related to different categories associated to main themes mentioned in the captions of those tables. Whereas Tables 8–10 represent only statements according to user group since all the statements are directly associated to the main theme.

As for access to third-party tools from the two e-exam systems, as indicated by statements of our interviewees in Table 8, IA has the option to configure the third-party tools in its SEB browser. This functionality is similar in WF's FLOWlock exam. However, universities instead run open-book exams or use their computer labs if the test requires usage of third-party tools (e.g., MATLAB, SOLIDWORKS, GeoGebra) during BYOD exams. Alternatively, both IA and WF support whitelisting of websites and web tools (e.g., EXCEL), using this practice, teachers can allow students to access external material during lockdown exams. Some universities have already tried whitelisting Excel, but this created a security vulnerability so that students could access files and desktop. At the same time, constraints and increased cheating vulnerabilities around the use of this feature made system managers avoid it. According to participants, the need for using third-party tools depended much on the discipline. Those who had mainly essay exams (e.g. social sciences) or multiple-choice tests (e.g., written exams in medicine, to supplement clinical/oral examinations done outside the e-exam system) were largely satisfied with standard features of the e-exam systems. Whereas, STEM subjects were less satisfied, needing, e.g. design, programming and math tools to make exams more authentic. Regardless of discipline, there might also be students with special needs for which extra tools would be required. Some of the system managers found in-built Virtual Desktop Infrastructure (VDI) feature in e-exam systems would be the best solution for accessing third-party tools; however, they thought it would require extensive testing to assure the usability, stability and security of high-stakes exams with VDIs.

Vendors and system managers thought open APIs would enable universities to integrate third-party tools or develop plug-ins (cf.

Table 6
Findings on detection and prevention of cheating during onsite invigilated exams.

Findings	Stakeholders	Statements of Stakeholders
Prevention of cheating	• Vendors	• "The lockdown browsers close a lot of security holes, but there's no perfect solution for BYOD. Some organizations use computer labs to have a higher degree of security." (IA1)
	• System managers	• "If the student goes out of the FLOWlock, then invigilator has to help him with a password. So, we have not had any cheating cases that I know." (HVL)
		• "FLOWlock of WF works like SEB, but it is closed source. We experienced some problems with SEB that's why we changed to WF. We had more cases of cheating in SEB." (KU)
		• "If students are going to use MATLAB, then we will conduct exams in computer labs or VDIs." (NTNU2)
		• "We locked out our school computers with ad-blocker, proxies. We experienced fewer issues." (NTNU1)
		• "We have test computers for Mac OS, Windows, Linux. We tested if running exam from different virtual machines is possible, some of them worked, but we only test what is reported, more responsibility lies at Inspera and SEB." (NTNU2)
		• "It is always possible to hack something using internet, especially using third-party tools. So, vendors have had data protection agreements with third-party solutions." (NTNU1)
	• Process managers	• "Lockdown of the network can prevent students from sharing files when they use Excel in BYOD exams, but it is a challenging part of closed-book exams. It is better to change the type of exam to open-book exams to avoid cheating. (Unit)
Detection of cheating	• Vendors	• "We do case analysis, and flag if there is suspicion, e.g., copy-pasting or sudden increase in text. We log and store every response. There are not too many false positives, but we need to make sure that text is not copy-pasted from the question or the response." (IA1)
	• System managers	• "We have seen Macro keyboard and Mac book touch bar save something to the clipboard, but SEB wipes clipboard. SEB also detects if a large amount of text is dumped through macro, but if student use speech to text, it will not be able to tell the difference." (NTNU2)
		• "There is a monitoring feature in WF, but sometimes it doesn't indicate cheating, e.g., in WF, sometimes when students don't get the exam text up, they will copy-paste exam text into their text, and it gets the detection that students copy-pasted a lot of text." (KU)
		• "There's a built-in feature in IA that enables to get the list of active students when events triggered, e. g. when they go offline, screen share, copies a large amount of text. But when many events triggered, it's hard to differentiate what was the real issue." (NTNU2)
		• "A few students had manipulated certain files, and they got it to work in demo test but […] not when they had their exam. Because we had a few students that have gotten error messages for SEB config file and config tool was in the recently viewed steps." (NTNU2)

Table 9), for instance, using the Portable Custom Interactions (PCI) standard. However, the current APIs of IA and WF were felt not be open enough, and system managers of IA at NTNU thought currently available APIs only facilitated administrative workflows, and also thought it was not within their capacity to develop plug-ins.

As for using monolithic vs digital ecosystems (cf. Table 10), vendors mentioned an ecosystem is an ideal choice for e-exam toolsets, as a single tool cannot provide every needed functionality, but they feared more issues with availability when third party tools are integrated within the ecosystem compared to the monolithic system. System managers believed ecosystems would improve integration and benefit users and vendors. One of the reasons for not choosing monolithic system is that teachers want specific features from different systems, e.g., from LMS and e-exam solutions.

Currently, both IA and WF support the functionality of sharing questions within the university. But sharing outside university requires export/import questions through a QTI file. Participants thought that access control would be a significant problem if e-exam systems support sharing across universities (cf. Table 11). Moreover, vendors mentioned that the culture of the university was the main factor that encourages sharing. A system manager said that if, e.g., medical faculties create questions together for national exams, that would require questions to be made for different exam systems, requiring extra work. So, a shared database would enable easier collaboration within and across universities.

5. Discussion

5.1. Interpretation of findings

1. What process is followed by vendors, process managers, and system managers to identify features and agree upon requirements?

The interviews indicated that the identification and negotiation of features and requirements is an important process that needs a lot of attention. This is in accordance with literature in software engineering in general (Dick, Hull, & Jackson, 2017), and in digital ecosystems (Immonen, Ovaska, Kalaoja, & Pakkala, 2016), and e-learning ecosystems (Uden et al., 2007). Some e-exam systems requirements may differ between countries, institutions, disciplines, and assessment practices. Yet, the interviews revealed that the

Table 7
Findings on challenges of integrations between e-exam systems with exam supporting technologies.

Findings	Stakeholders	Statements from Stakeholders
Integration is not standard based	• Vendors	• "We control more of the ecosystem. Integration with SEB isn't standards-based, so in that sense, documentation is not as good as standards-based." (IA1)
Low data quality	• Vendors • System managers	• "There had been challenges with data quality." (IA1) • "The quality of the registration of data in FS is not uniform enough to do all the automation NTNU would like to do." (NTNU1) • "Duplicate users for both students and employees may happen." (NTNU2)
Non-uniformity in choosing solutions for integrations	• System managers	• "Biggest issue is that larger institutions that use IA have integrated their central user account with IA whereas NTNU integrated their central use account directly only to FS." (NTNU2) • "Since we have chosen different solutions […] for developing solutions for the same purpose, and it takes a long time to implement changes." (NTNU2)
Poor synchronisation between systems	• System managers	• "Sometimes the integrations won't export the right sources from FS to WF due to synchronisation issues, e.g., synchronisation doesn't do the changes when grading commission is changed." (KU) • "Slow process of integration between FS and IA due to loading of the large database." (NTNU4)
Incompatibility of QTI formats	• Vendors	• "Integration of IA and Blackboard is not supported due to incompatibility in their QTI formats." (IA1)
Integration of exam system with LMS is not prioritized	• Vendors • System managers	• "We did not receive requests from our clients." (WF) • "Integration with Blackboard is not something we prioritise because we're not going to send any results from Blackboard to FS or from Blackboard to Inspera. But a lot of teachers have said that why should I have to make questions twice, so that's why we're looking at." (NTNU2) • "There is no integration between WF and Canvas. I don't see any need for that rather what I do see a need for better integration between Canvas and FS if that would exist then all necessary integration between the LMS and exam system would go through FS." (UiT)

Table 8
Findings on third-party tools access during onsite invigilated e-exams.

Stakeholders	Statements of Stakeholders
Vendors	• "Universities have a certain number of licenses to use third-party software. So, students have to be on a university computer. We do not sell the third-party license, but we do integrate with some of these." (IA2)
System managers	• "Third-party tools are added through the config file, we use mainly PDF, and we can also use whitelisting, e.g., some X URL, but behind that URL there is a static list of rules we are going to use, so we don't use that function." (NTNU2) • "WF doesn't support third-party integration as of today, but they support whitelisting. We used Excel through whitelisting, but students were able to access the other files and desktop that we cannot allow during our exams." (KU) • "Students with special needs must access third-party tools. Currently, they deliver exam on Word, and we will upload the file, that's better in IA than WF." (KU) • "When we need third-party tools in onsite exams, we conduct open book exams. So teachers have to choose what is more important, is it to test the students what they learnt in the course or is it that it's 100% secure." (HVL) • "We don't use FLOWlock when we use third-party tools. But there is a project currently going on to access VDI in FLOWlock, giving you access to third-party tools. but the question is, are VDIs stable enough to handle many students at the same time? […] what happens if it breaks down? How about user experience? Another solution would be to make specific third-party integrations with specific programs, e.g., MATLAB, GeoGebra." (UiT)
Process managers	• "Vendors mostly do integration of tools, a lot of people want to integrate Excel in WF, the problem is that when students use excel sheet, they can share it" (Unit)

requirement process was somewhat ad hoc, and no standard requirements process seemed to be followed across various stakeholders. This is also in accordance with findings in the literature. For instance, Achimugu, Selamat, Ibrahim, and Mahrin (2014) observed that although several advanced methods for requirements prioritisation have been proposed in the literature, these have had limited uptake in general software development practice. The stakeholders who addressed requirements prioritisation seemed satisfied with their rather simple approach based on polling and web meetings.

A study from Norway by Foss-Pedersen and Begnum (2017) looked especially at universal access requirements for e-exam systems IA and WF, in a study involving some of the same organizations that we interviewed. Their findings indicate unclear division of responsibility between vendors and buyers for ensuring universal access and lack of quality assurance through user testing. Some of our participants did mention getting useful feedback from piloting and user testing at universities. This may indicate that the process has been somewhat improved after 2017, which is natural since more widespread use of the e-exam systems will have generated more user

Table 9
Findings on availability of open APIs for IA and WF.

Stakeholders	Statements of Stakeholders
Vendors	• "Currently, our API's are designed to enable data flow with the admin side of IA mostly, but additional ways of extending and customising are possible for universities through IMS PCI standard." (IA1)
	• "There are primary advantages by opening APIs, so we're working on adding support for APIs to more parts of the application, but we have to ensure that they're being used correctly." (IA1)
	• "Our APIs are tightly controlled, and we don't let people query things against our database. To do integrations, developers have to work with us. You get normal standard problems with our APIs that you get with regular APIs security." (IA2).
System managers	• "There is a national project going on with Unit, and it is looking into it to find a good solution through API. Every institution that uses WF gets the API through Unit, but we haven't done much with it yet." (KU)
	• "Security always depends on the level of documentation of open API. Currently, Inspera's REST API's help with the administrative workflow." (NTNU1)
	• "Working on APIs would be interesting to me. We have access to do that if we saw the need to tweak the API in some cases to fit some of our needs. but that's not something we've discussed at UiT." (UiT)
	• "Inspera is using PCI standard to create exercises, and they're open to enable others to do it as well, but we haven't explored that, and it is something our teachers are working on." (NTNU1)
Process managers	• "In general, having open API's allows people with competence to create applications, but specific considerations have to be kept in mind when opening them up to ensure that they are being used correctly." (Unit)
	• "Currently, API's are not open enough to pull out real data because of not having access to the access token." (Unit)

Table 10
Findings on opinions on monolithic vs digital ecosystem.

Stakeholders	Statements of Stakeholders
Vendors	• "It is an absurd idea to say one piece of software will do everything for the exam. The ecosystem is the only ideal choice for the universities because the idea of monolithic does not work" (IA2)
	• "From a high level, we have to focus on an ecosystem because there are so many different needs that we cannot possibly fulfil them all. The different systems are in use already a handful in specific tasks, so it makes sense to continue using those by integrating with the rest of the ecosystem." (IA1)
	• "When you have a combination of companies work together to achieve a goal […] it might be slower to implement because each of these companies has their own agenda. However, with a monolithic solution, they have the same agenda. In theory, they have many processes digitised, so when customers get an incident, they can immediately send it to the vendor, and they can immediately look at it." (WF)
System managers	• "We have talked a lot with teachers about what system they would like, and what's become very apparent is that a lot of them want completely different systems. Blackboard has the functionality to perform exams as well, and we could use IA for delivering assignments, but in the end, IA is designed as an exam system and Blackboard as an LMS." (NTNU2)
	• "We get requests when teachers see a certain functionality in Canvas that does not exist in WF. They wish to use Canvas for a certain part of their exams, but that system again lacks all other necessary functionality to conduct an exam." (UiT)
	• "People always want more features, but the more features you put into a system, the more difficult it will be to use one system. It will be simple to use multiple systems because there are so many ways to do things already in Blackboard and IA. If you combine them into one system, people will struggle with finding the right option for them." (NTNU2)
	• "The most important thing is providing a good tool for our students and our employees, having well-standardised integrations and being able to plug-in from another system into e-exam system." (NTNU1)
	• "Having ecosystem would be a better solution. […]. Even it would benefit vendors. I would be surprised if they would not be open to working with other parties" (UiT)

experience. To some extent, our findings indicate a clearer assignment of responsibilities between vendors, universities and system managers, thus approaching a clearer definition of roles as recommended in the literature (Immonen et al., 2016).

The arrangement to have one national organization – Unit – oversee the requirements and integration process, seemed to be satisfactory to the participants. No participant indicated a wish to go back to the previous situation when each institution would run separate procurement and integration projects. Unit maintains a requirements specification for e-exam systems for Norwegian institutions, negotiates with vendors on universities' behalf, and takes responsibility for developing integration software to make the e-exam systems fit into the common system infrastructure. Similar approach is used for Sweden and Netherlands higher education institutions (Boezerooy, Cordewener, & Liebrand, 2007). An interesting question is whether a similar collaboration could work even internationally. As indicated by vendors, some requirements will differ from country to country, such as those relating to the national IT infrastructure and those relating to laws, regulations, and student rights concerning exams and grade appeals. On the other hand, many features would be of interest regardless of country, such as support for specific exam types, question types, disciplines. E.g., a maths exam in country X would likely have more common needs with a maths exam in country Y, than with a history exam in country X. However, collaborating between several universities to establish requirements for e-exam systems as a basis for acquisition from commercial vendors is not the only viable path for international collaboration. Another option would be to have multiple universities collaborate on developing e-exam technology, as has been pursued, for instance, by the EU project TeSLA (Okada, Noguera, et al.,

Table 11
Findings on factors that influence sharing questions among universities.

Factor	Stakeholders	Statements of Stakeholders
Security	• Vendors	• "If we open up the ability for users to share through exam system, there is certainly a concern of users sharing the stuff they are not allowed to share, so a degree of access control should exist." (IA1)
	• System managers	• "I wouldn't recommend open sharing between the universities like opening it up to everyone who has authoring access to the system at any university. Because you would have to start trusting everyone else's security measures that would affect your security." (UiT)
Culture	• Vendors	• "It depends on the culture of the university as to whether sharing is important thing or not. In Norway, there is more share of information. Every year all the questions for the previous year's exams are available. If professors know each other […], they may talk about some of the old exam questions." (IA2)
	• System managers	• "Universities need to have a good collaboration outside the exam if they want to create the exams together. E.g., In medicine, they share questions across universities, and they had national tests where questions were imported into IA to conduct exams locally." (NTNU1)
Sharing between Cross-platforms	• System managers	• "We're involved in several national exams where universities use the same questions in different exam systems. They are made in a particular format by the national groups, but we have to do a lot of work after they've been imported to have them ready for the exam day. It will be a major development if it's possible to share questions between the different systems in a more dynamic and real-time way. (UiT)
		• "Having a shared database in cross platforms would be a major step for not only to cooperate and sharing questions, basically sharing the load." (UiT)

2019). However, this alternative was not suggested by any of the informants, likely because none of the Norwegian universities were involved in the TeSLA project.

2. What do vendors, process managers and system managers see as key features for functionality and security in e-exam systems?

Explaining all the needed features of an e-exam system would take a lot of interviewee time. It is thus reasonable to assume that participants focussed on significant features that they found most worthwhile to mention, based on their experience. It was also noted that they spent more time discussing recent and perhaps challenging features, than more straightforward features. It is a well-known phenomenon in the field of requirements engineering that assumedly obvious features will tend to be omitted by interviewees, especially if they are experts on the technology in question (Firesmith, 2003).

Of the features that participants did talk about, there was comparatively more mention of authoring, explanation of grades, and appeals grading, less about the logistic system, question analytics and ordinary grading. Of course, ordinary grading (for all students) is a much more used functionality than appeals grading (for the smaller fraction of students who dispute their grades), so when interview participants talked more about appeals grading, this again is likely an effect of omitting the obvious. Moreover, appeals grading had probably gained extra attention from many participants because it was initially unsupported by the e-exam systems, thus addressed by a cumbersome manual workaround, and had just recently been included as a feature.

Our results on key features are in line with previous findings (Kuikka et al., 2014; Striewe, 2019, pp. 220–228). The features highlighted by our participants relate to all four major components proposed in the framework by Striewe (2019, pp. 220–228): front-end components (user interface), educational components (underlying logic of tests and question types, pedagogical modules), knowledge representation and storing components (e.g., question pools, tests, exam answers and results), connector components (related to integration and interoperability, as discussed in our section 4.3 on digital ecosystems). Admittedly, the development of features for knowledge representation is currently much less advanced than suggested by Striewe. Kuikka et al. noted that none of their compared systems satisfied all the needs of teachers. Likewise in our study, participants indicated that features with less backing would end up not being prioritized. Also, teachers would sometimes have to make trade-offs, such as only being allowed to have open book exams if providing usage of third-party tools. The expert survey by (Isaias et al., 2019), found authoring, interoperability and feedback features to be important, this corresponds to our findings, but participants in our study felt feedback as the basic feature. Different from the situation in Norway, Kuikka et al. suggested that it is an advantage to have the same software product work both as LMS and e-exam system. This may be true, but some of our system manager participants believed otherwise, thinking that too many features in the same system would confuse some teachers.

Among the non-functional features, usability, integrity and interoperability, and security were considered as essential for e-exam systems. Vendors also felt scalability and reliability were important for their customers as the BYOD exams were mainly run on the internet. System managers considered the availability of technical support from vendors during exams to be vital for them.

3. What are the goals and challenges concerning integrations between the e-exam system and other exam supporting systems?

Several integration challenges are presented in this study. There was functional level integration between a few systems, e.g., IA exam system and SEB lockdown browser. Although there was no issue explicitly mentioned with this particular integration,

specification of this particular integration was mentioned as not as good as standards based. There had been challenges with data quality and synchronisation were reported due to universities in user group used different solutions for integrations. Hence, there should be a need for data management and governance before integrating e-exam systems with other systems.

System managers mentioned that teachers were concerned about content interoperability between LMS and e-exam systems. Both WF and IA e-exam solutions use IMS QTI specification for sharing and exchanging of the content. But question types do not follow a common standard, meaning a QTI exchange would be very complex and manual, and not fully supported across systems. The lack of common standard format for representation of questions was seen limiting the exchange of questions between LMS and e-exam solutions. This is in line with findings by several others (Boussakuk, Ghazi, Bouchboua, & Ouremchi, 2019; Tomberg, Savitski, Djundik, & Berzinsh, 2017, pp. 159–170). The lack of interoperability standards for LMSs to communicate with external applications due to their monolithic architecture has received more attention in recent research (Alier, Guerrero, Gonzalez, Penalvo, & Severance, 2010; Chirumamilla & Sindre, 2019; Dagger et al., 2007).

Moreover, our results show that low prioritisation of integration between LMS and e-exam systems and limited demand for such integrations from HE institutions seemed to be the reason for not having upgrades in QTI specifications of LMSs towards better interoperability with other systems. Such interoperability shortcomings of LMS's were pointed out already by Sclater (2007), arguing that the standards and specifications had not reached a critical mass of adoption due to the insufficient demand by users in higher education institutions. More than a decade later, although the uptake of such systems has increased enormously, the progress concerning interoperability is not too impressive. Literature shows that only in the Netherlands, their primary education school system has considered integration between LMS and e-exam systems as a mandatory requirement in procurement of e-assessment tools (Kerssens & Dijck, 2021, pp. 1–14). While it would also be useful, e.g., to be able to reuse test questions across several platforms (Chirumamilla & Sindre, 2019), participants indicated that so far, this had not been considered important *enough* to make the priority list, as there had always been other issues needing more urgent attention. Both IA and WF vendors also support the IMS Learning Tools Interoperability™ (LTI) standard for seamless integration and secure data communication between several platforms. But the lack of personnel resources and time for testing seemed to be the reason for system managers not to pursue increased use of that functionality.

4. To what extent do the stakeholders envision a move towards a more open digital ecosystem for e-exams, and would this be assumed to impact security?

Both IA and WF vendors are developing RESTful API's for flexible integrations of e-exam systems with other systems and allowing institutions to build plug-ins on their platforms. They followed effective API development and integration practices for e-exam systems adapting to institutions pedagogical requirements. However, e-exam systems and supporting systems use different APIs, so it is not easy to integrate multiple APIs of different systems. Moreover, API's are strictly controlled for security reasons, in literature this has been addressed as a strategy for limiting access to external users and protecting the system from malicious components (Striewe, 2019, pp. 220–228). So, integrations of e-exam systems with other systems were mainly performed at the process manager's site using a standardised third-party tool, Mule Soft ESB. MuleSoft serves as the framework for service organisation and orchestration, providing secure delivery of services (Menge, 2007). Still, integration of IA and WF with third-party tools seemed to be at the early stage. Both IA and WF support minimal external software during BYOD exams, requiring whitelisting or integration. Unit is running pilots, to use third-party software/tools in lockdown browser. So, currently, third-party tools can only be accessed at the cost of reduced security from the lock-down browser. Though the literature reports alternatives to lock-down browsers to allow third-party tools while maintaining good security, such as booting a dedicated operating system from memory stick (Fluck, Pálsson, et al., 2017; Frankl et al., 2012), this approach is not used with the e-exams used in Norway. One reason could be that the booting approach is more cumbersome, requiring additional actions at the start of every exam, plus preparations in copying a sufficient number of memory sticks. Participants instead suggested a VDI solution for allowing third-party tools without getting too high cheating vulnerabilities but admitted this would require extensive testing and implementation costs not yet undertaken.

There was clear collaboration reported between institutions, process managers and vendors in implementation, integrations, user testing. All participants preferred moving towards a digital ecosystem rather than a monolithic system, as this could provide more features and user convenience for teachers and students to organize learning and tests, better interoperability between systems, increased quality, reusability, and sharing of questions. This is in line with findings in (Kerssens & Dijck, 2021, pp. 1–14; Veiga, Campos, Braga, & David, 2016). Our results show that security, culture, sharing between cross-platforms seemed as factors that influence sharing questions among universities. On the other hand, system managers in our study suggested that more sharing within and between universities would be possible with the integration of question banks. This is in line with the previous findings (Chituc, Herrmann, Schiffner, & Rittberger, 2019, pp. 155–162; Laine, Sipilä, Anderson, & Sydänheimo, 2016). However, interoperability and IPR issues remain possible obstacles for the development of shared item banks (JISC, 2007).

5.2. Limitations of this study

5.2.1. Credibility (internal validity)

Several threats that are relevant to discuss.

Selection bias: The selection of whom to interview may have affected the results. Our approach to mitigate this threat was to interview a considerable number of persons, from several different organizations.

Participant bias: May the participants knowingly or unknowingly have given inaccurate information? As already mentioned in the discussion, some very obvious features of e-exam systems were hardly mentioned – omitting information that was taken for granted is a

well-known phenomenon. There may also be other reasons to reply inaccurately, such as memory, embarrassment (if something went wrong with the system or project) or secrecy (e.g., vendor representatives not wanting to reveal business secrets). Again, interviewing several persons will reduce this threat. Moreover, we have explicitly reported cases where participants were reluctant to answer about cheating vulnerabilities and concrete ways to utilize them.

Researcher bias: Researchers may tend to interpret data in ways that confirm their preconceived ideas. Various measures taken to mitigate this threat include avoiding leading questions, not pushing participants in any particular direction, following a well-defined protocol for analysing the data.

Participant checking is a key measure in mitigating researcher bias. Transcriptions were sent to participants before analysis to verify whether transcriptions indicate what participants intended to say. After analysis, a copy of this article has been sent to participants for comment before journal submission, to let participants point out any cases where their statements may have been misinterpreted. Their suggestions have been accommodated in the article.

Triangulation of analyses can also contribute to mitigating researcher bias. The first author was primarily involved in data collection and transcription of data, and data have been further analysed together with a co-author. Participant checking and triangulation of analyses correspond to the best practice guideline for implementing and reporting of qualitative research (Twining et al., 2017).

5.2.2. Transferability (external validity or generalizability)

To what extent can findings from this study be generalized to other settings? One notable limitation is that the study is focusing on the situation in only one country, Norway. All the system managers were from Norwegian universities, and the purchasing organization a Norwegian directorate. True, the two vendors have customers in several countries, and one of the companies (WF) is Danish, so the vendors take on required features for their products will have had a somewhat more international perspective, also exemplified by specific statements from participants that requirements would be different from country to country. Still, the context of the case is specifically the situation in Norway, and a study including universities from other countries might have come up with different findings. The two companies involved in the study were both vendors of dedicated e-exam software, hence catering to universities who use different products for high stakes e-exams than what they use for e-learning in general. As mentioned in the introduction, many universities around the world may be using the same system (e.g., Canvas, Blackboard, Moodle) both for e-learning and for high stakes tests, which may lead to differences in expectations towards the products. So, further work is needed to take a more international approach and to get findings covering a broader spectrum of educational software products. Nevertheless, challenges such as security and interoperability are key to e-exams in many countries – as indicated by related work – so findings are believed to be of interest also outside the specific Norwegian context.

5.2.3. Dependability (reliability) and confirmability (objectivity)

Often researchers find it challenging to synthesize results from the interpretation of qualitative data analysis (Twining et al., 2017). We followed guidelines by (Anney, 2014; Korstjens & Moser, 2018) to ensure dependability and confirmability of our findings.

Authors applied triangulation of analyses and participant checking to conduct audit trail of collected data, analysis documents. Authors were also in prolonged engagement with participants before submission to the journal to ensure whether the analysis was consistent with collected data rather than their own imaginations. Also, the analysis process has been constantly verified with the grounded theory approach to ensure whether the analysis is consistent with the approach. So, the research approach and instrument (i. e., interview questionnaire) can be used in a similar group of subjects and research settings.

6. Conclusion

This paper has reported on an interview study with 12 persons having central roles in the development, procurement, and operation of e-exam systems in Norwegian higher education, some being vendors, some system managers at various universities, and one being a process manager at the national organization Unit, coordinating the acquisitions and IT infrastructure development for all Norwegian public universities. Participants generally seemed satisfied with the requirements process, feeling that the coordinating role of Unit had improved the process compared to the previous situation where each university was running separate acquisitions and integration projects – though it was observed that parts of the requirements process were still somewhat ad hoc. There was much agreement between participants about both functional and non-functional features of e-exam systems, system managers much focussed on security and interoperability. As for security, while there might be some vulnerabilities for cheating, especially with the BYOD type of e-exam, stakeholders generally felt that systems were satisfactory and believed that more cheating was taking place outside the digital exam infrastructure – e.g., old fashioned cheat notes, concealed mobile phones. As for interoperability, stakeholders generally agreed that long-term ambition should be a move towards a digital ecosystem where open standards and APIs would allow for smooth integrations between various tools. However, at the current stage, many integrations were challenging – both considering the usage of third-party tools during the exam and exchanging information between e-exam systems and other tools, such as Learning Management Systems. Since our case study was restricted to Norway, some findings may be specific to the national setting. However, both e-exam system vendors have customers in several countries, in Europe and beyond. Hence, the challenges surrounding requirements specification, security and interoperability of e-exam systems are likely of much broader interest. Interesting avenues for further work could be to perform similar studies in other countries, or across several countries, and also to compare the views of the stakeholder groups interviewed here with those of students and teachers. It will also be exciting to see how the e-exam systems of these two vendors – and other competitors – develop over the next couple of years. While the move towards open digital ecosystems is a fine ideal that many agree about, there are also obvious business interests that would go in favour of maintaining a situation dominated by proprietary

software.

What advice should then be given to universities pursuing better e-exam systems in the future? Is it a good idea to continue relying on commercial vendors of learning management systems and e-assessment systems – as is currently done in Norway – or would it be better for universities to develop their own systems, as has been done, for instance, in the TeSLA project? If continuing with commercial vendors, our study indicates that it is a good idea that several universities collaborate in the requirements and acquisition process, to have more influence and better bargaining power versus the vendors. Likewise, with a development approach, a collaboration between many universities seems more likely to give successful results than each university making its own system. In any case, it is important to have a long-term view, focusing on the interoperability between the e-exam systems and other IT systems in the higher education sector, and to keep in mind that the ultimate goal is not the technology itself, but pedagogical improvement of assessment practices.

Credit author statement

Aparna Chirumamilla: Conceptualization, Methodology, Investigation, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. Guttorm Sindre: Conceptualization, Validation, Writing – review & editing, Visualization, Supervision.

Acknowledgement

Thanks to all participants for taking the time to participate in interviews.

Appendix

Appendix A. Participant Information Letter

Invitation
I would like to talk with you about your experiences with e-exams.
Purpose of this study
I hope to compare experiences of e-exam systems between several individuals, institutions and
    countries to publish in an academic journal.
What does participation involve for you?
The participant will be engaged in the interview. The interview will last approximately 45–60 min.
If you are agreeable, I would like to record our conversation. I will ask you to review the draft article to
    ensure your comments have been accurately reported.
Participation is voluntary
Participation in this study is voluntary. All information about you will then be made anonymous.
Your personal privacy – how we will store and use your personal data
We will process your personal data confidentially following data protection legislation (the General
    Data Protection Regulation and Personal Data Act). The data will not be handed to other
    researchers or professors at NTNU or outside of the institution.
What will happen to your personal data at the end of this study?
The project is scheduled to end 2020. The data will be stored in a university computer and will be kept
    until the end of the project after that data will be deleted.
Your rights
So long as you can be identified in the collected data, you have the right to:
- access the personal data that is being processed about you
- request that your personal data is deleted
- request that incorrect personal data about you is corrected/rectified
- receive a copy of your personal data (data portability), and
- send a complaint to the Data Protection Officer or The Norwegian Data Protection Authority
    regarding the processing of your personal data
What if I have questions about this study?
- Please contact at aparna.vegendla@ntnu.no or guttorm.sindre@ntnu.no
This study has been approved by the Norwegian Centre for Research Data AS.
This information letter is for you to keep. Your consent will be indicated by your agreement to speak
    with me at a mutually agreed appointment.

Appendix B. Demographic information of interview participants

Person & date	Position	Institution and site code	E-exam technology
IA1, 15Aug19	Development manager	Inspera AS, Norway, **IA**	Inspera Assessment, SEB, Blackboard, FS
			(*continued on next page*)

(*continued*)

Person & date	Position	Institution and site code	E-exam technology
IA2, 11Sep19	Product manager	Inspera AS, Norway, **IA**	Inspera Assessment, SEB
WF, 17Sep19	Head of product development	UNIWISE, Aarhus, Denmark, **WF**	WISEflow, FLOWlock, Canvas, FS
Unit, 29Aug19	Senior advisor	The Norwegian Directorate for ICT and joint services in higher education and research, Norway, **Unit**	Inspera Assessment, WISEflow
NTNU1, 13Aug19	Project manager, IT Digital Assessment	Norwegian University of Science and Technology, Norway, **NTNU**	Inspera Assessment, SEB, Blackboard, FS
NTNU2, 7July19	Team leader, IT Digital Assessment	Norwegian University of Science and Technology, Norway, **NTNU**	Inspera Assessment, SEB, Blackboard, FS
NTNU3, 14June19	Engineer, Digital Security Section	Norwegian University of Science and Technology, Norway, **NTNU**	Inspera Assessment, SEB
NTNU4, 14June19	Engineer, IT department	Norwegian University of Science and Technology, Norway, **NTNU**	Inspera Assessment, SEB, Blackboard, FS
NTNU5, 8Apr19	Senior consultant, IT Digital Assessment	Norwegian University of Science and Technology, Norway, **NTNU**	Inspera Assessment, SEB
HVL, 3July20	Senior advisor, IT Digital Assessment	Western Norway University of Applied Sciences, Norway, **HVL**	WISEflow, FLOWlock, Inspera Assessment, SEB, Canvas, FS
KU, 2July20	Senior advisor, IT Digital Assessment	Kristiania University College Norway, **KU**	WISEflow, FLOWlock, Inspera Assessment, SEB, Canvas, FS
UiT, 21Aug20	Advisor	University of Tromsø - The Arctic University of Norway, **UiT**	UiT, 21Aug20

Appendix C. Research questions and corresponding interview questions

Research Question	Interview Questions
What process is followed by vendors, process managers, and system managers to identify features and agree upon requirements?	• How do you conduct the RE process?
What do vendors, process managers and system managers see as key features for functionality and security in e-exam systems?	• What are the key functional requirements for digital exam system? • What are the key security requirements for digital exam system?
What are the goals and challenges concerning integrations between the e-exam system and other exam supporting systems?	• How do you conduct integrations? • What are the challenges of integrations between your organisation e-exam system with exam supporting systems, e.g., LMS, student information system, and lockdown browser?
To what extent do the stakeholders envision a move towards a more open digital ecosystem for e-exams, and would this be assumed to impact security?	Questions specific to vendors: • Would you like to provide monolithic or a system as a part an ecosystem? • How your e-exam system supports customisation? • How do you support different third-party plug-ins in your e-exam system? • What are the challenges of integrations between your e-exam system and third-party tools? Questions specific to process managers and system managers: • Which system would you like to experience between monolithic or an ecosystem? • Will the requirements to have an open API and allow for third-party and university-made plug-ins be important? • Will the requirements to enable the sharing of questions across institutions to be important? Will there be any security challenges if we do so? • Are there any security threats of paper exams that have been removed or reduced by digital exams? Common questions to vendors, process managers and system managers: • What are the most important cheating/security risks to e-exam technologies during on-campus digital exams, especially using BYOD devices?

References

Achimugu, P., Selamat, A., Ibrahim, R., & Mahrin, M. N.r. (2014). A systematic literature review of software requirements prioritization research. *Information and Software Technology, 56*(6), 568–585.

Adebayo, O., & Abdulhamid, S. M. (2014). E-exams system for Nigerian universities with emphasis on security and result integrity. *International Journal of Computer Integrated Manufacturing, 18*(2), 1–12.

Alier, M. F., Guerrero, M. J. C., Gonzalez, M. A. C., Penalvo, F. J. G., & Severance, C. (2010). Interoperability for LMS: The missing piece to become the common place for e-learning innovation. *International Journal of Knowledge and Learning, 6*(2–3), 130–141.

Anney, V. N. (2014). Ensuring the quality of the findings of qualitative research. *Looking at trustworthiness criteria, 5*(2), 272–281.

Baró-Solé, X., Guerrero-Roldan, A. E., Prieto-Blázquez, J., Rozeva, A., Marinov, O., Kiennert, C., et al. (2018). Integration of an adaptive trust-based e-assessment system into virtual learning environments—the TeSLA project experience. *Internet Technology Letters, 1*(4), e56.

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *Qualitative Report, 13*(4), 544–559.

Bilen, E., & Matros, A. (2021). Online cheating amid COVID-19. *Journal of Economic Behavior & Organization, 182*, 196–211.

Boeije, H. (2002). A purposeful approach to the constant comparative method in the analysis of qualitative interviews. *Quality and Quantity, 36*(4), 391–409.

Boezerooy, P., Cordewener, B., & Liebrand, W. (2007). Collaboration on ICT in Dutch higher education: The SURF approach. *Educause Review, 42*(3).

Boussakuk, M., Ghazi, M. E., Bouchboua, A., & Ouremchi, R. (2019). Online assessment system based on IMS-QTI specification. In *2019 7th mediterranean congress of telecommunications (CMT)* (pp. 1–4).

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Brink, R., & Lautenbach, G. (2011). Electronic assessment in higher education. *Educational Studies, 37*(5), 503–512.

Butler-Henderson, K., & Crawford, J. (2020). A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity. *Computers & Education, 159*, 104024.

Buzzetto-More, N. A., & Alade, A. J. (2006). Best practices in e-assessment. *Journal of Information Technology Education: Research, 5*(1), 251–269.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. sage.

Chirumamilla, A., & Sindre, G. (2019). E-assessment in programming courses: Towards a digital ecosystem supporting diverse needs? In I. O. Pappas, P. Mikalef, Y. K. Dwivedi, L. Jaccheri, J. Krogstie, & M. Mäntymäki (Eds.), *Digital transformation for a sustainable society in the 21st century* (pp. 585–596). Cham: Springer International Publishing.

Chituc, C.-M., Herrmann, M., Schiffner, D., & Rittberger, M. (2019). *Towards the design and deployment of an item bank: An analysis of the requirements elicited*. Cham: Springer International Publishing.

Chituc, C.-M., & Rittberger, M. (2019). Understanding the importance of interoperability standards in the classroom of the future. In *Iecon 2019 - 45th annual conference of the IEEE industrial electronics society* (Vol. 1, pp. 6801–6806).

Cizek, G. J. (1999). *Cheating on tests: How to do it, detect it, and prevent it*. Routledge.

Cluskey, G., Jr., Ehlen, C. R., & Raiborn, M. H. (2011). Thwarting online exam cheating without proctor supervision. *Journal of Academy of Business and Economics, 4* (1).

D'Souza, K. A., & Siegfeldt, D. V. (2017). A conceptual framework for detecting cheating in online and take-home exams. *Decision Sciences Journal of Innovative Education, 15*(4), 370–391.

Dagger, D., Connor, A. O., Lawless, S., Walsh, E., & Wade, V. P. (2007). Service-oriented E-learning platforms: From monolithic systems to flexible services. *IEEE Internet Computing, 11*(3), 28–35.

Dawson, P. (2016). Five ways to hack and cheat with bring-your-own-device electronic examinations. *British Journal of Educational Technology, 47*(4), 592–600.

Dick, J., Hull, E., & Jackson, K. (2017). *Requirements engineering*. Springer.

Dick, M., Sheard, J., Bareiss, C., Carter, J., Joyce, D., Harding, T., et al. (2002). Addressing student cheating: Definitions and solutions. *ACM SIGCSE Bulletin, 35*(2), 172–184.

Dyer, W. G., & Wilkins, A. L. (1991). Better stories, not better constructs, to generate better theory: A rejoinder to eisenhardt. *Academy of Management Review, 16*(3), 613–619.

Ferdiana, R., & Hoseanto, O. (2018). The implementation of computer based test on BYOD and cloud computing environment. *International Journal of Advanced Computer Science and Applications, 9*(8), 121–124.

Firesmith, D. (2003). Specifying good requirements. *Journal of Object Technology, 2*(4), 77–87.

Fitzharris, R., & Kent, S. (2020). Adoption of bring-your-own-device examinations and data analytics. In D. Ifenthaler, & D. Gibson (Eds.), *Adoption of data analytics in higher education learning and teaching* (pp. 327–348). Cham: Springer International Publishing.

Fluck, A. E. (2019). An international review of eExam technologies and impact. *Computers & Education, 132*, 1–15.

Fluck, A., Adebayo, O. S., & Abdulhamid, S.i. M. (2017). Secure e-examination systems compared: Case studies from two countries. *Journal of Information Technology Education: Innovations in Practice, 16*(1), 107–125.

Fluck, A., & Hillier, M. (2017). eExams: Strength in diversity. In A. Tatnall, & M. Webb (Eds.), *IFIP advances in information and communication technology* (pp. 409–417). Cham: Springer International Publishing.

Fluck, A., Pálsson, H., Coleman, M., Hillier, M., Schneider, D., Frankl, G., et al. (2017). eExam symposium: design decisions and implementation experience. *11th IFIP TC 3 world conference on computers in education*.

Fluck, A., Pullen, D., & Harper, C. (2009). Case study of a computer based examination system. *Australasian Journal of Educational Technology, 25*(4).

Foss-Pedersen, R. J., & Begnum, M. E. N. N. (2017). Universell utforming og digital eksamen i UH-sektoren: 5 anbefalte tiltakspunkter. In *Norsk konferanse for organisasjoners bruk at IT* (Vol. 25).

Frankl, G., Schartner, P., & Zebedin, G. (2012). Secure online exams using students' devices. In *Proceedings of the 2012 IEEE global engineering education conference (EDUCON)* (pp. 1–7).

Frosini, G., Lazzerini, B., & Marcelloni, F. (1998). Performing automatic exams. *Computers & Education, 31*(3), 281–300.

Glaser, B. G., Strauss, A. L., & Strutzel, E. (1968). The discovery of grounded theory; strategies for qualitative research. *Nursing Research, 17*(4), 364.

Gustafsson, J. (2017). *Single case studies vs. multiple case studies: A comparative study*.

Hillier, M., & Fluck, A. (2013). Arguing again for e-exams in high stakes examinations. *Electric dreams. Proceedings Ascilite*, 385–396.

Huszti, A., & Petho, A. (2010). A secure electronic exam system. *Publicationes Mathematicae Debrecen, 77*(3–4), 299–312.

Immonen, A., Ovaska, E., Kalaoja, J., & Pakkala, D. (2016). A service requirements engineering method for a digital service ecosystem. *Service Oriented Computing and Applications, 10*(2), 151–172.

Isaias, P., Miranda, P., & Pífano, S. (2019). Framework for the analysis and comparison of e-assessment systems. In *ASCILITE 2017-conference proceedings-34th international conference of innovation, practice and research in the use of educational technologies in tertiary education* (pp. 276–283). Australasian Society for Computers in Learning in Tertiary Education (ASCILITE).

Jakimoski, K. (2016). Challenges of interoperability and integration in education information systems. *International Journal of Database and Theory and Application, 9* (2), 33–46.

JISC. (2007). Effective Practice with e-Assessment: An overview of technologies, policies and practice in further and higher education. In *Joint information systems committee*.

Kaiiali, M., Ozkaya, A., Altun, H., Haddad, H., & Alier, M. (2016). Designing a secure exam management system (SEMS) for M-learning environments. *IEEE Transactions on Learning Technologies, 9*(3), 258–271.

Kallio, H., Pietila, A. M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing, 72*(12), 2954–2965.

Kerssens, N., & Dijck, J.v. (2021). *The platformization of primary education in The Netherlands*. Learning, Media and Technology.

Korstjens, I., & Moser, A. (2018). Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. *The European Journal of General Practice, 24* (1), 120–124.

Kuikka, M., Kitola, M., & Laakso, M.-J. (2014). *Challenges when introducing electronic exam* (Vol. 22), 2014,.

Ladonlahti, T., Laamanen, M., & Uotinen, S. (2020). Ensuring diverse user experiences and accessibility while developing the TeSLA e-assessment system. *Engineering data-driven adaptive trust-based e-assessment systems*. Springer.

Laine, K., Sipilä, E., Anderson, M., & Sydänheimo, L. (2016). Electronic exam in electronics studies. In *SEFI annual conference 2016*.

Lauesen, S. (2006). COTS tenders and integration requirements. *Requirements Engineering, 11*(2), 111–122.

Llamas-Nistal, M., Fernández-Iglesias, M. J., González-Tato, J., & Mikic-Fonte, F. A. (2013). Blended e-assessment: Migrating classical exams to the digital world. *Computers & Education, 62*, 72–87.

Llorens, F., Molina, R., Compañ, P., & Satorre, R. (2014). *Technological ecosystem for open education* (pp. 706–715). IDT/IIMSS/STET.

Luo, M.-Y., & Lin, S.-W. (2013). From monolithic systems to a federated e-learning cloud system. In *2013 IEEE international conference on cloud engineering (IC2E)* (pp. 156–165). IEEE.

Mellar, H., Peytcheva-Forsyth, R., Kocdar, S., Karadeniz, A., & Yovkova, B. (2018). Addressing cheating in e-assessment using student authentication and authorship checking systems: Teachers' perspectives. *International Journal for Educational Integrity, 14*(1), 1–21.

Melve, I., & Smilden, B. (2015). *ICT architecture for digital assessment*. UNINETT.

Menge, F. (2007). Enterprise service bus. In *Free and open source software conference* (Vol. 2, pp. 1–6).

Naous, D., Giessmann, A., & Legner, C. (2020). Incorporating the voice of the customer into mass-market software product management. In *Proceedings of the 35th annual ACM symposium on applied computing* (pp. 1397–1404).

Okada, A., Noguera, I., Alexieva, L., Rozeva, A., Kocdar, S., Brouns, F., et al. (2019). Pedagogical approaches for e-assessment with authentication and authorship verification in Higher Education. *British Journal of Educational Technology, 50*(6), 3264–3282.

Okada, A., Whitelock, D., Holmes, W., & Edwards, C. (2019). e-Authentication for online assessment: A mixed-method study. *British Journal of Educational Technology, 50*(2), 861–875.

Piotrowski, M. (2011). Qti: A failed E-learning standard? In L. Fotis, G. Steve, & P. Elaine (Eds.), *Handbook of research on E-learning standards and interoperability: Frameworks and issues* (pp. 59–82). Hershey, PA, USA: IGI Global.

Queirós, R., Leal, J. P., & Paiva, J. C. (2016). *Integrating rich learning applications in LMS*. Singapore: Springer Singapore.

Robson, C. (2002). *Real world research: A resource for social scientists and practitioner-researchers*. Wiley-Blackwell.

Sandkuhl, K., & Lehmann, H. (2017). *Digital transformation in higher education–The role of enterprise architectures and portals*. Digital Enterprise Computing. DEC 2017.

Sclater, N. (2007). The demise of eAssessment interoperability? *WWWrong, 317*.

Sclater, N., Low, B., & Barr, N. (2002). *Interoperability with CAA: Does it work in practice?*.

Severance, C., Hanss, T., & Hardin, J. (2010). Ims learning tools interoperability: Enabling a mash-up approach to teaching and learning tools. *Technology, Instruction, Cognition and Learning, 7*(3–4), 245–262.

Striewe, M. (2019). *Components and design alternatives in E-assessment systems*. Cham: Springer International Publishing.

Tomberg, V., Savitski, P., Djundik, P., & Berzinsh, V. (2017). *Design and development of IMS QTI compliant lightweight Assessment delivery system*. Cham: Springer International Publishing.

Twining, P., Heller, R. S., Nussbaum, M., & Tsai, C.-C. (2017). Some guidance on conducting and reporting qualitative studies. *Computers & Education, 106*, A1–A9.

Uden, L., Wangsa, I. T., & Damiani, E. (2007). The future of E-learning: E-Learning ecosystem. In *Digital EcoSystems and technologies conference, 2007. DEST '07* (pp. 113–117). Inaugural IEEE-IES.

Urquhart, C., Lehmann, H., & Myers, M. D. (2010). Putting the 'theory' back into grounded theory: Guidelines for grounded theory studies in information systems. *Information Systems Journal, 20*(4), 357–381.

Veiga, W., Campos, F., Braga, R., & David, J. M. (2016). A Software Ecosystem approach to e-Learning domain. In *Proceedings of the XII Brazilian symposium on information systems on Brazilian symposium on information systems: Information systems in the cloud computing era -, ume 1* pp. 574–581). Santa Catarina, Brazil: Brazilian Computer Society. Florianopolis.

Weishaupl, E., Yasasin, E., & Schryen, G. (2018). Information security investments: An exploratory multiple case study on decision-making, evaluation and learning. *Computers & Security, 77*, 807–823.

Wibowo, S., Grandhi, S., Chugh, R., & Sawir, E. (2016). A pilot study of an electronic exam system at an Australian University. *Journal of Educational Technology Systems, 45*(1), 5–33.

Wills, G. B., Davis, H. C., Gilbert, L., Hare, J., Howard, Y., Jeyes, S., et al. (2009). Delivery of QTIv2 question types. *Assessment & Evaluation in Higher Education, 34*(3), 353–366.

Yin, R. (2002). *Case study research: Design and methods* (Vol. 5). SAGE Publications, Inc.

Yin, R. K. (2017). *Case study research and applications: Design and methods* (6th ed.). Sage publications.