

Which requirements need to be matched to allow for sound capability testing, in form of decentralized e-exams?

Introduction

The examination process is often a **tedious task** for those who are in charge. Great amounts of time go into organizational problems. Digitizing exams would resolve many of these problems. A step towards electronic examination would make the process more flexible, scalable and resource-efficient. Meanwhile, leading to a more accurate depiction of a students' competence.

Taking the current pandemic under consideration, it may not be an option to just move from paper to e-exams. Exams thus must be conducted decentralized, i.d. students take their exams at home.

It is important to notice, that *Decentralized E-exams (DE-exams)* differ from *paper based exams (PB-exams)* and even from *centralized e-exams (CE-exams)* in some key points. Foremost, the examiner has less control over the environment the exam is taken under. This raises questions about exam integrity and fairness. These questions must be addressed through careful conceptualization of questions and intelligent software design.

Digitizing exams is no new idea. Although, most concepts and implementations focus on conducting e-exams in the same environment as *PB-exams*. With regard to COVID-19 this is not an option. Of course some *DE-exams* are **already conducted** today. These exams are, for the most part, making use of a **proctoring system**. In such a system a supervisor can access the examinees' device, can monitor all their activity and will watch them through their webcam. This proctoring process is costly. It hardly scales and still easily can be fooled. Further, test-taking applications are found in many *LMS' (Learn Management Systems)* such as Ilias, Moodle or Blackboard. Unfortunately, most often these applications focus on student self-assessment. They also majorly **vary in quality and utility**. As they are integrated in a complete LMS, changing to the *best* implementation is in many cases not an option. Last, as exam data is **highly confidential**, there is a strong argument to be made against closed source solutions. It is crucial to know exactly how the used application works and how data is handled. Adding, open source projects are less prone to major security issues as the development can leverage the crowdsourcing capabilities that an open source system provides.

An overview of criteria and derived questions

The main criteria emerge as:

- The electronic examination process must be advantageous compared to paper based examination.

- The exam has to have a high level of integrity. Students cannot fake their identity, cheat, cooperate, use disallowed aid or access exam questions beforehand.
- The exam questions must aim to fairly extract a students' skill level.
- Students must be treated equally in the answer checking process.
- The e-exam taking process must be at least be as stress-free as the paper base equivalent.

Having identified these main criteria, I will ask *sub-research questions* on how to meet these criteria. In doing this I will give insight into how other systems perform.

The case for e-exams

E-education is a much discussed topic. Most of the educational material has become digital. Still, paper based exams are the way to go, when it comes to assessments in German higher education. Although some universities among these are the FU Berlin and the Johannes Gutenberg University Mainz have implemented some way of e-assessment, it still lacks wide application.

In its research question this thesis ask about how to improve upon centralized e-exams. This implies, that e-exams in general are superior to paper based exams. There is good reason for this implication. It is a good idea to take a look at these reasons before moving on.

Why move from paper based to electronic exams in the first place? inherent vs. earned advantages

Put simply central e-exams replace the pen and paper of paper based exams with a computer or tablet. To give intuition of the advantage of e-exams i will focus on *bring-your-own-device* (BYOD) e-exams. In BYOD exams the device the student is taking the exam on is not provided by the university authority, students bring their own tablet or laptop to take the exam on. The BYOD concept is not new. For example Robert Peregudof talked about the BYOD implementation of the University of British Columbia in the conference.

It is important to note that moving from paper based exams to e-exams should be a strict upgrade. Still, paper based exams and e-exams are different in some key points. Some things that were considered best practice must be reevaluated and rethought in order to adequately fit the e-exam context. As I will show these changes do not limit the examination process but rather improve it.

When compared to paper based exams e-exams provide many advantages as discusses in their book. I will divide these into two different types. The first type are inherent advantages. These are advantages that arise from solely digitizing a paper based exam. These advantages mainly effect the efficiency of the assessment process.

The other type of advantages are created through additional considerations and by rethinking how paper based exams approach the examination process. These advantages arise mainly from improvements in testing accuracy and equality.

In order to later assess the quality of existing software solutions, after each section I will derive requirements. These requirements are needed to make use of the theoretical advantage.

Exams as a logistics problem

Thinking about the inherent advantages of e-exams, the logistical implications of exams pop into mind. To illustrate this we use the examination process at the KIT as an example. Although some steps may differ from university to university, the gist remains.

Exams must be printed and stapled. As, exam taking students are numerous, it is common for exams to take place at many different sites. Therefore, on the test day exams must be carried out to the test site. Here, the Ciw of the KIT recommends at least two supervisors per test site. On test site, exams must be distributed to students. After the exam is written, exams are collected and counted. They are then carried back to a central location, where they remain until correction. For answer checking, correctors come together, again at a central location, where they then are able to correct the exam. After correction a grade for students is published via internet. Succeeding, an exam revision for students is planned and executed. Lastly exams are archived in their paper form.

It is not surprising that handling large amounts of paper results in logistics overhead. Removing the paper, subsequently removes much of the logistics overhead. In an e-exam all of the exam data is digital. There is no printing, since exams are directly transferred to the students device. Further students answers can automatically be retrieved after the exam is over. Exam answers are digitally available. The correction of answers is no longer bound to a certain site but can be done remotely. Exam results and feedback can be directly issued to the students. Thus, revision can also be realized decentrally. The digital exam data can then be easily archived.

This comparison illustrates the advantage of e-exams with regard to logistics. Not only is there less movement of employees but more importantly there is no movement of physical paper. Further we find advantages in the digital archiving. Data can be stored space efficiently with no need for large physical archives. Adding, archives are much safer as redundant backups are feasible and cheap.

Requirements:

- Finished exams can be collected remotely.
- Exam data is stored at a central instance and can be accessed remotely.
- Exam data can be archived.

The hassel of handwriting

Checking an exam for correctness is one of the most time-consuming process in conducting an exam. Moving away from paper can reduce this time drastically making use of two things: First, some question types, such as multiple choice can be checked automatically. This is an immediate improvement over correcting these questions by hand. Second, as exam answers are available in digital text reading and checking answers is easier. Answers must not be deciphered, correction of exams can be done faster.

Further, the replacement of handwriting with digital text has implication for the equality of answer checking. As <James 1927> shows students with bad handwriting get categorically worse grades than students with better handwriting. E-exams can help to make the correction process fairer and less time consuming.

Requirements:

- Question types that are able to be checked automatically are checked in such a way.

Statistics on the fly

Thinking of the digital nature of exam data another advantage emerges. The digital nature of e-exams allow for a fast creation of statistics. Whereas in paper exams every piece of information must be manually digitized, e-exams are digital out of the box. Thus, analysis of exams becomes more feasible. Thinkable are statistics about general performance, but also analysis of specific questions, or student groups. As the exam is the sole indicator of a students understanding of the matter at hand, it is of utmost importance to understand where students struggle and what they are capable of. Having easy access to exam data yields the possibility of both better exams and better courses.

Requirements:

- Possibility to create a statistics automatically.

Drinks at the pool

Part of the complexity and time intensiveness of exam creation lays in creating appropriate questions. Although many courses are not unique to one university, sharing of test question is not common. In paperbased exams there are no standards and there is no suitable collaboration platform. In paper based examination there is no real foundation for sharing and reusing exam questions.

E-exams have to make use of an software artefact in order to leverage their theoretical benefits. Such a artefact allows for an enforcement of a shareable general format. If users create exams in a specific software, implementing a standard is fairly straight forward. Such a standard already exists under the name of <QTI 2.2>. Having a standard allows for educators to collaborate to create exams.

Further this collaboration must not be limited by a single exam with only a few questions. Exams can come from question pools — a small subset of questions is selected from a way larger superset. The development and maintenance of a large question pool is very time intensive for a single person, that is why collaboration is so important. Projects in open source software show how collaboration of many can function with great success.

Large question pools do not only help in exam creation, as they allow to produce exams automatically, it also prevents students from knowing the exam beforehand. This will be further elaborated in the section about cheating.

Partial Open Book Exams

Until now I discussed advantages of e-exams that require no rethinking of exam content. E-exams enable assessments to more closely depict a student's actual skill level. As the states in their paper, examinations should support the purpose of university to produce highly capable individuals. As much of a student's academic success is based on their performance in exams, students are highly incentivised to focus their studies on the given exam format. This interdependency between knowledge acquisition and examination shows the importance of exam design and poses the question of what and how to test.

Questions can be categorized into main groups :

- **(Semi) Closed questions**, which mainly revolve around the demonstration of factual knowledge. Solutions are given by using a format like multiple choice [closed] or simple text input semi-closed
- **Competence questions**, which are suited to test for a certain practical skill. Solutions are given in form of an implementation of the specific task at hand. (e.g. "Using the software, implement an e-exam about e-learning.")
- **Essay-type questions**, which are suited for assessing transfer knowledge and understanding. Solutions are given by free text input. (e.g. "Give reasons why subjects in computer engineering are especially well suited for e-exams.")

The last two types can also be set in the context of an open book exam. In which students have access to any resource they feel they need in order to solve the question at hand. These open book tasks most closely resemble real world problems and thus are highly suitable for assessing a student's capabilities.

Looking at classical paper based exams we find that they often focus either on (semi) closed questions or on open book questions. As students have access to all relevant facts during the entirety of open book exams, testing for fact knowledge in these cases is rendered meaningless.

Still fact knowledge cannot be neglected. For example, consider a doctor, who needs his cheat sheet, in order to conduct open heart surgery. Or an accountant who has to look up the meaning of working capital, every time it shows up. Fact knowledge in the above cases is necessary, because it is needed immediately. A

doctor who does not know his facts may lose lives, an accountant who does not know his vocabulary is nearly useless in any kind of negotiation. The constraining factor in both cases being time. These examples show that in order to assess for capabilities a marriage of (semi) closed questions and open questions is necessary.

With E-exams it is fairly easy to implement *partial* open book exams that allow for both open and closed task formats. E-exams can achieve this by enacting per question time constraints. Closed questions are imposed with a short answer frame that leaves no time for looking up the actual answer. Further, access to question must be restricted after the time frame has expired.

Cheating and randomization

When thinking about any assessment the consideration and handling of academic dishonesty (e.g. cheating in an exam) is one of the most important parts. Moving from paper to e-exams poses the questions what parts – if any – must be adjusted to accommodate for changed circumstances and environments.

In his paper poses seven fields of possible cheating in exams which he then evaluates by accuracy and perceived severity. Six of which are relevant for this thesis purpose. The fields can be described as follows:

Student cooperation:

- **Knowing the questions** Learning what is on an exam from someone who has already taken it.
- **Cooperation with outsiders** Helping someone else cheat on an exam.
- **Cooperation with fellow examinees** Copying from another student on an exam with their knowledge.

Use of disallowed aid:

- **Exploit environmental circumstances** Copying from another student on an exam without their knowledge.
- **Use of unauthorised crib/cheat notes** Bringing prepared cheat notes to use in the exam.
- **Use of electronic, unauthorized aid** Using e.g. a smartphone to google or review lecture material.

Before thinking about how to obviate these cheating scenarios an important statement must be made: Cheating cannot completely be eliminated. There are always ways for students to engage in cheating. E-exams cannot change that. But compared to paper based exams some measures against cheating may be more effective.

Knowing a question. As the question finding process in a time consuming process, a strategy can be to keep questions as secret as possible and reuse them throughout many exams. This is a rather ineffective strategy as platforms such as studydrive often provide comprehensive protocols from memory of the given exam. E-exams can choose a different approach. Instead of having few questions

and keeping them secret, e-exams can leverage collaborative question pools. At a certain point it becomes unfeasable for students to prepare for every available question as question pools grow larger.

Cooperation with outsiders and fellow examinees. For closed questions this cooperation can be prevented by using strict time restrictions. As already stated above these questions fall in the category *Either you know the answer or you don't* there is no need for a lengthy reflection period. As time constraints are tight, there really is no way of communicating with others and solving the question. For more open question types the time limitation is not as tight. At the same time answers require more in depth considerations. To ensure that students write down their own ideas and cannot share their thoughts, the copy and paste functionality can be disabled. Further, e-exams can easily be randomized, thus preventing students from peeking or signaling solutions to each other.

Exploit environmental circumstances. Again randomization can solve this problem. As questions appear in a different order for each student even multiple choice questions can not simply be copied.

Use of unauthorized cheat notes or electronic aid. Following the argument made about partial open book exams we find that besides time constraints no additional measures must be enforced. Cheat notes really are not of much help if there is no time to use them.

We find these cheating scenarios to be to a large degree managed by e-exams. Still, as specific software is in use the degree of cheating must constantly be assessed and measures against bugs or security flaws must be identified.

From central to decentral e-exams

In the previous section I talked about advantages of central e-exams compared to paper based exam. One key characteristic of e-exams is that they are issued via the internet. We can use this to remove any local component of e-exams, thus making them decentralized. In the following, I will talk about advantages that decentral e-exams have and the implications of decentralizing exams on cheating and infrastructural considerations.

E-Exams as an logistics problem

As seen above, e-exams can be immensely more efficient in terms of logistics, compared to a paper based exam. There are however still some shortcomings. First, consider the allocation of students to their respective test taking location. As different students take part in different exams this allocation grows ever more complex and requires lots of planning of a central authority. Making e-exams decentral allows for this process to be streamlined enormously.

Second, for supervision of exams human resources are needed. The for example recommends two supervisors per location. E-exams can free up some of these resources. To do this we need to take earlier considerations about partial open

book exams into account. As I laid out, Partial open book exams allow for tests that do not ask for supervision, making existing supervision obsolete.

When talking about decentralizing e-exams I think it is important to say that the goal is not to force students to write their exams at home. There still can be space reserved for exams by the university. The management and supervision of these spaces just becomes magnitudes more easy.

Cheating

The main points that were made about cheating in section , still hold true when talking about decentralized exams. A big difference however can be found in the authentication of students. In a central exam each students identity is checked by an examinee. The students signature and the student identification card provide identification. In decentralized exams this direct checking is no longer possible. There is no satisfactory solution to remotely checking a students identity. Still, e-exams have ways to deter this kind of dishonesty.

A first measure might be to make use of integrated webcams and microphones of the devices at hand. This low key supervision creates a barrier. Students get a feeling of doing really committing academic fraud. This is important, as students who really put energy into faking their identity can fool the decentral system just as easily as the central system. The sole existence of any measures makes the students behave more honest. Just as video surveillance makes crime less common at public places . This video and sound data can be reviewed if needed.

Infrastructure

I talked about decentralized exams in form of BYOD exams. These devices are heterogeneous. They run different operating systems, are produced by different companies and have different specifications. It makes little sense to develop proprietary software for each device. E-exam can although make use of browser technology. As modern web-applications do not lack speed or functionality and can be adopted cross platform. The software is hosted at a central instance where it can be maintained and improved.

Of course e-exams need an internet connection. Still, connection issues are not a big issue. Through advances in web technologies it is feasible to make web applications available in offline use. Even if the device loses power, exams can be resumed after a reboot.

Software

In the previous section I laid out areas in which e-exams can improve upon the examination process. As I already discussed the actual accuracy of these advantages heavily depends upon the software that is used in order to implement an e-exam. In the following I will take a look at common solutions:

- Ilias
- Moodle
- Blackboard
- LPlus
- OpenOlat

I will measure their performance