UNIVERSITY OF TARTU
Institute of Computer Science
Innovation and Technology Management Curriculum

Ayaz Karimov

Business process optimization in online educational platforms within the scope of localization and its impact on the learning experience of students: localization of Khan Academy content into Azerbaijani

Master's Thesis (20 ECTS)

Supervisor: Irene-Angelica Chounta
Supervisor: Alexander Nolte

Tartu 2021

Business process optimization in online educational platforms within the scope of localization and its impact on the learning experience of users: localization of Khan Academy content into Azerbaijani

Abstract: The localization of the world-class free online educational platforms brings many benefits to the students and teachers, nonetheless, it demands time and capital resources to localize any content. The localization of Khan Academy content into the Azerbaijani language is one of the examples of that case.

The aim of this research is to optimize the localization of Khan Academy content into the Azerbaijani language and evaluate the impact of the optimization on the learning experience of users. For this purpose, after designing and analyzing the current localization process, it has been redesigned, and the new process has been implemented to see the difference between the previous and updated localization process. Having finished redesigning, its impact on the learning experience of users has been measured based on the survey results.

Keywords: localization optimization, learning experience improvement, business process redesign, BPMN, Khan Academy
CERCS: P170 Computer science, numerical analysis, systems, control

Äriprotsesside optimeerimine veebipõhistes haridusplatvormides lokaliseerimise valdkonnas ja selle mõju kasutajate õpikogemusele: Khani Akadeemia sisu lokaliseerimine aserbaidžaani keelde

Kokkuvõte: Maailma tasemel tasuta veebipõhiste haridusplatvormide lokaliseerimine loob õpilastele ja õpetajatele mitmeid eeliseid, kuid kogu sisu kohandamine nõuab aega ja kapitali. Khani Akadeemia sisu lokaliseerimine aserbaidžaani keelde on selle heaks näiteks.

Käesoleva uurimistöö eesmärk on optimeerida Khani Akadeemia sisu lokaliseerimist aserbaidžaani keelde ja hinnata optimeerimise mõju kasutajate õpikogemusele. Seepärast on peale käesoleva lokaliseerimisprotsessi põhjalikku kavandamist ja analüüsimist see ümber kujundatud ning uue versiooni rakendamisel jälgitud erinevusi eelmise ja uuendatud lokaliseerimisprotsessi vahel. Peale protsessi ümberkujundamist mõõdeti ka selle mõju kasutajate õppekogemusele.

Võtmesõnad: lokaliseerimise optimeerimine, õpikogemuse parandamine, äriprotsesside ümberkujundamine, BPMN, Khani Akadeemia

CERCS: P170 Arvutiteadus, arvuline analüüs, süsteemid, juhtimine

1. Introduction

The recent alterations - such as changes in organizational structure or changes in organizational goals and/or process [25-26] - can lead to dramatic changes in the implementation of the business process management. Its application expanded from the industrial field to many sectors, including education [2]. When the companies experienced the benefits of business process optimization, the educational institutions have also commenced using it to improve their processes. Educational stakeholders, such as institutes and private organizations, apply business processes [11]. Its primary implementation is observed in online educational organizations. Here, we mean the organizations that offer web-based learning [9]. Since it provides online study materials, web-based learning is also known as online learning or e-learning [27].

1.1 Online education and Khan Academy

Online education is an instructional delivery system that encloses any pattern of learning that happens via the Internet. Practical and impactful platforms should be developed for achieving the goals in the online education sector. A plethora of online education platforms, as well as tools, have emerged. Online educational institutions are the ones that follow the principles of online education mostly or totally [17]. One of those institutions is Khan Academy. It is an online platform that offers free courses from kindergarten to university students. Khan Academy contains various subjects such as mathematics, computer science, art, history [7].

1.2 Motivation of research

The research aimed to design, analyze and improve the current localization process of the Khan Academy content into the Azerbaijani language. By doing so, we aimed to measure the effectiveness of this redesigned process. The optimization of the localization process is crucial from various aspects, such as lacking online platforms in the Azerbaijani language and the number of students who cannot access education with high quality. All of the contents of Khan Academy are in English, and it signifies that it is hard to use the platform if their English level is not sufficient. One of the goals of Khan Academy is to localize its content into other languages, and Azerbaijani is one of them. The localization of Khan Academy into the Azerbaijani language can bring many benefits because the number of online education platforms for the school students is not more than 10. Additionally, all of these platforms are paid, excluding the platform of the Ministry of Education in Azerbaijan and the Khan Academy website[1]. It means that the localization can help the students and teachers who live in the edge parts of Azerbaijan to learn and educate at a world-class level anywhere without financial expenditures.

[1] az.khanacademy.org

Moreover, if the Khan Academy Azerbaijani team can localize more content with good quality, the platform can help more students, parents, and teachers. Khan Academy has supported the learning experience of high school students in Turkey by 22% from the perspective of the learning duration. Therefore, it might bring high values for the Azerbaijani high school students as well. Hereby, this academic research aims to optimize the localization process in a way that will help the stakeholders to achieve their goals. These goals are chiefly allocating less time and human capital while improving the quality of the localized content.

Even though the localization of the Khan Academy content into Azerbaijani can bring a myriad of advantages, a major problem is the project`s cost. Here, the cost is time and human resources allocated for the project. The localization procedure started in 2017, and we encountered many bottlenecks, such as spending much time on the proofread of the content, saving the translation strings within the localization of the content. Localizing the contents fast and qualitatively is the main goal of the team. Nevertheless, the team cannot meet the key performance indicators, such as the number of localized videos, exercise currently. As a result of the research, we defined the following research questions to answer to address these problems:

- Research question 1 (RQ1): What is the "as is"[2] process of Khan Academy Azerbaijan within the localization framework?
- Research question 2 (RQ2): What are the dominant quantitative and qualitative aspects of work processes in Khan Academy Azerbaijan?
- Research question 3 (RQ3): How can we redesign current business processes in order to improve efficiency in terms of success?
- Research question 4 (RQ4): What is the difference between the current and redesigned business process based on the success measurements?
- Research question 5 (RQ5): What is the impact of the redesigned localization process on the learning experience of students on the platform?

The research on process optimization at educational organizations mainly focused on higher education process optimization. Moreover, the localization optimization topic has not been researched on a large scale. One reason for this is that localization started to play a significant role only in the early 2000s [9].

We investigated the research questions as a case study, and a mixed-methods approach (qualitative and quantitative) has been implemented. Within this framework, we hold 2 workshops and two interview phases. After holding two phases of interviews (Phase 1 interviews: to collect the feedback about the "as is" process; Phase 2 interviews: to collect the

[2] An "as is" business process defines the current state of the business process in an organization [44].

proposals about the proposals to redesign the localization process), we sent an online survey to the parents who are part of the Khan Academy Azerbaijan Community. Subsequently, based on the survey responses, we analyzed the impact of the redesigned process on the performance of students. For this, we took into account the time that students spend on each topic, the feedback on the quality of localized content.

The rest of the thesis is structured as follows. In Section 2, the concepts of business process management and optimization, localization of content, and the practices of utilizing process optimization in educational institutes and organizations are introduced. Section 2 also embraces the related work and depicts the research gap this paper aims to cover. In Section 3, the research settings and the methodologies used to answer the research questions have been explained. Section 4 describes the results of each research question. In Section 5, we provided information about the results and the limitations of this research. Finally, we conclude the paper in Chapter 6.

2. Background

In this section, we present background knowledge that helps to understand the content of this thesis. Particularly, Section 2.1 provides information about Business Process Management (BPM) and Optimization. The notion of the BPM lifecycle and its implementation usage within this research has also been explained. In Section 2.2, the utilization of the BPM within the education domain, and in Section 2.3, we described the notion of localization and the example of the Khan Academy platform. Lastly, in Section 2.4, an overview of the related work is described concerning process optimization in educational institutes and organizations, such as optimizing the process within the scope of selecting educational programs, smart curriculum design, localization of Khan Academy content into the Chinese language. We have also mentioned the research gaps and how this research will focus on the elimination of them.

2.1 Business process management and optimization

Business process management concentrates on using business processes as one of the crucial contributors to achieve the organizations` aims through analyzing lasting performance management, improvement, and governance of crucial business processes [3]. Hereby, business processes are considered as a series of activities that are performed by a stakeholder or a group of stakeholders. Each activity within the business process is a task. The main goal of the BPM is to analyze the current situation and define the parts that can be improved within the business processes [4].

For more than two decades, business process management has been a well-known concept. It began with an emphasis on organizational measurement and quality improvement. Still, BPM

has since evolved into an IT-driven proficiency, recognizing the importance of technology in process understanding and enhancement. As a result, it is expected that current workflow-centric representations of processes will give way to a more business-focused view of procedures, accountability, and evaluation. By using BPM, we can set the goals for the future and improve current processes. Such things are essential to become a leader in the given domain. Having the data, and more importantly, knowing how to read and convert it to solve given problems, is an essential feature. Business process optimization, which helps improve current processes, is a crucial road toward developing a suitable system to be used. The field of education becomes more and more popular over time, but it has many weak points that need to be solved or optimized. Thus, BPM can be beneficial to achieve such goals [29].

BPM is a critical tool for improving traditional learning methods by allowing for the simulation and implementation of complex processes. It is a solid framework that distinguishes process-specific organizations, which is one way to enhance e-business processes and make learning easier [28].

Business process optimization (BPO) is one of the critical elements of the BPM itself, and it is the discipline of adjusting processes to minimize input and maximize efficiency. By doing so, the key aim of the organization is to increase productivity throughout the company. However, in the vast majority of organizations, some bottlenecks create the problems [1]. They disable the emergence of the optimum business processes, and the BPO concentrates on eliminating those bottlenecks. Having mentioned that, it is quite significant to follow all the steps of the BPM lifecycle to optimize any processes (Figure 1).

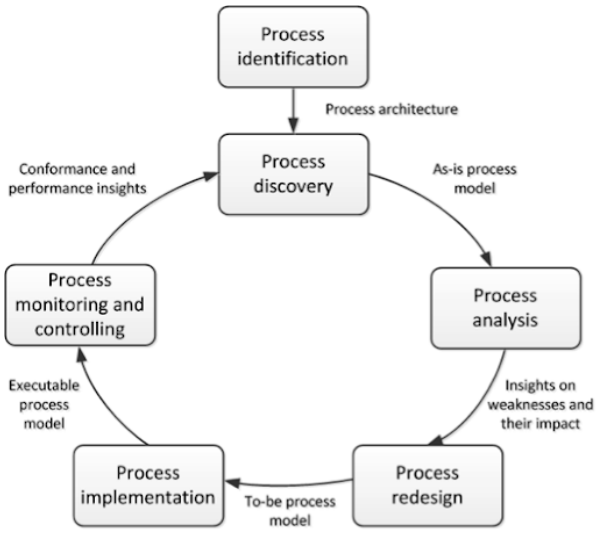


Figure 1. The BPM lifecycle [1]

BPM lifecycle consists of six major steps, and within this empirical research, the whole six steps were gone through. The lifecycle commences with the “Process Identification”. This phase plays a significant role because it enables us to define the context. In order to be successful at this phase of the BPM lifecycle, the identification of the organization's business processes, prioritizing those processes according to the particular criteria are the main activities. In this phase, the organization should also define those particular criteria and the business strategy that will be utilized for the process identification.

As a second step of the BPM lifecycle, “Process Discovery” is addressed. It is one of the patterns of process mining along with conformance checking and model extension/ enhancement [14]. Process discovery is also known as the “as-is” process model since it demonstrates the visualization of the current business process [15]. Furthermore, for the discovery, some techniques can be implemented, such as observation, interviews. Within this empirical research, we will implement individual interviews, group workshops. After the business process discovery activities, the tasks are noted. Then, specific tools are utilized to represent the business process visually. For this research, Signavio is used, which is a web-based business process modeling tool [16]. For the Business Process Model and Notation (BPMN), those four chief elements are used: activity, events (start and end), gateway, and sequence flow (Figure 2).

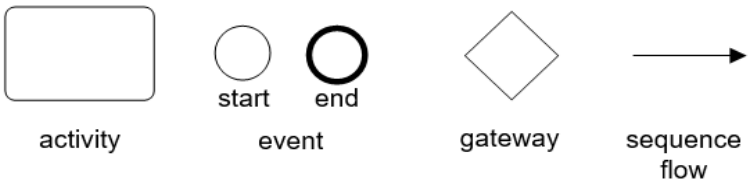


Figure 2. BPMN model elements

Activity is a task that stakeholders need to fulfill. Events are signals that something happens. Mostly used event types are “start” and “end” events that depict what triggers the business process and the organization's goal accordingly. Sequence flows are used to show how the activities, events, and gateways relate to each other. Gateways are about controlling as well as merging and splitting the process flow for the decisions. In addition to those four core elements, two more key visualizations play a crucial part in the business process design. They are pools and lanes, and they are utilized to define responsibilities and organizational boundaries within the process landscape. Hereby, a pool is a unit with well-defined organizational limits to its environment. Lanes are contained by another lane or a pool. Lanes within the same pool communicate without any restrictions [16].

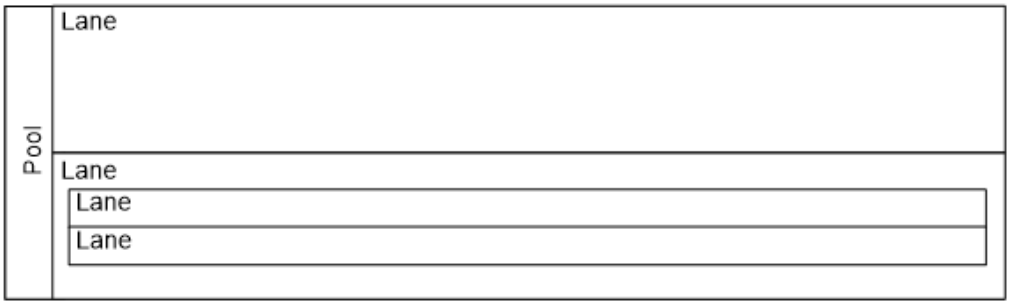


Figure 3. Pools and lanes in BPMN model

The third step of the BPM lifecycle is the "Process analysis". At this phase, bottlenecks are defined, and their impact on the business is calculated. For this, two various analyses are applied: 1) qualitative; 2) quantitative. Within the scope of the qualitative analysis, a few methods are applied, such as value-added & waste analysis, root-cause analysis, Pareto analysis, issue register. At this research issue register will be used as the technique to analyze the business process. In the issue register analysis, the problem`s name, explanation of the problem, data or hypotheses about the problem, and the problem`s qualitative and quantitative impact should be mentioned. Subsequently, from the perspective of quantitative analysis, flow analysis will be implemented. Flow analysis aims to calculate the cycle time to perform all the tasks within the business process [1].

The next step is the "Process redesign", and there, "as is" process changes to the "to be[3]" process. Within the process redesign, there are two major approaches which are exploitative redesign (transactional) and explorative redesign (transformational). While the transformational redesign tries to aim at radical innovation, in the case of the transactional redesign, it is more incremental for solving the problems. Transactional redesign method, particularly the heuristic redesign, is implemented within this research. There are nine heuristics that are used in this process, and they are task elimination, task composition/ decomposition, triage, re-sequencing, parallelism enhancement, process specialization/ standardization, resource optimization, communication optimization, automization [1].

The fifth step of the lifecycle is the "Process implementation". It is also known as the "to-be executed" model since it is the business process model that can be implemented to check the impact of the process redesign on the success measurements, business objectives. The key point here is to ensure the model is unambiguous and does not contain any doubts [1].

Last but not least, the "process monitoring and controlling" is the last phase of the BPM lifecycle. At this phase, organizations use different kinds of dashboards, such as operational,

[3] A "to be" business process defines the future state of a business process in an organization [44].

tactical, and strategic, in order to get performance insights [1]. For this, a few tools can be used, and one of the examples to this is Apromore. Apromore is a collaborative business process analytics platform supporting the full spectrum of process mining functionality [15]. All these phases, excluding the last one, are gone through within the research which is about the online educational institutions` process optimization.

2.2 BPM for online learning contexts

BPM has been used in the education domain more and more to optimize the workflows, improving the quality of learning. In online learning, students can learn anything from the internet. However, it is quite challenging to determine what they want to learn and how they want to learn (i.e., doing exercises, watching videos). BPM has been used to create a customized learning roadmap for the students to guide them within the learning process [36]. It has been utilized to design the learning environment based on the needs, requirements of the students, and adaptive assessments. This is a type of assessment that adapts to the test taker's ability level, and it is becoming more popular thanks to the integration of technology into education [44].

BPM has also been used indirectly to improve the quality of the learning process as well. One of the examples of this is improving any parts that mainly contribute to the learning process. The optimization of the localization process can bring benefits. Because it can also lead to the usage of less resources, such as time, human capital. Thus, implementing BPM for the optimization of the localization process may also benefit online learning indirectly.

2.3 The localization of content

The localization can be defined as the translation of any content or platform into a language in a way that crosses cultural boundaries. Hereby, the cultural boundaries can be anything from the names to the punctuations [37]. The main difference between the translation and localization is that while the content is directly translated in the translation process, in the localization process names, proverbs can be changed. For instance, if we translate "John had 5.2 apples." then we will not consider any points, and it will be "Conun 5.2 alması var." in the Azerbaijani language. However, if we localize the same sentence, then we will use local names and mathematical punctuation. Instead of using "John", the local name "Vasif" would feel more local. Furthermore, since in the Azerbaijani language, commas are used to show the decimals, then instead of keeping the dot, it will be replaced by a comma. As a result, the localization of the same sentence will be "Vasifin 5,2 alması var."

In the 1980s and 1990s, localization had become crucial for the software and hardware companies when they wanted to internationalize their businesses. Moreover, since the

technology has been integrated with other domains such as healthcare, finance, these sectors also implemented localization practices. One of the examples of this is the education domain. Once online learning started to grow rapidly in the 2000s, in parallel with the emergence of educational technology companies, their internationalization strategies were also developed [38]. Khan Academy is one of these companies. Even though the company was founded in 2008, its localization started in 2010 with the Turkish language. Furthermore, Azerbaijani advocates started to begin on the localization in 2017, and the team is actively localizing the content from the original page of the Khan Academy. In this research, we focused on the localization process optimization of the Khan Academy content into the Azerbaijani language.

2.4 Process optimization in educational institutes and organizations

In [19], the main focus of the research was to design a model for time optimization of educational processes to enhance students` creativity. As a result, they achieved the allocation of more budgets than they used to, thanks to the educational processes optimization. It also triggers the opportunity to anticipate a variety of educational signals modality for educational and methodical guidelines. For this, they determined an ensemble of information states, and this state concentrates on the consecutive alteration of creativity level. This academic research supports the optimization of educational processes from different perspectives since the main methodology will be the implementation of the business process optimization lifecycle through the whole phase. Moreover, in this paper, the chief aspect is online education within the framework of the localization process.

The chief concentration area in [20] was implementing the smart curriculum design and block-chain to optimize the process in the high educational institutions, particularly the universities. The challenge was the optimization of processes in the universities since those are being done in paper form. The final aim of the research was to suggest blockchain-enabled verification of education as well as other credentials, data analytics, and decision support for process optimization. QualiChain was the selected project to be focused on that paper, and the project has already completed its first step, which was to get feedback from the users and research requirements. This academic paper focuses mainly on online educational institutions, and the main focus area is the school students since 92% of the Khan Academy users are school students in Azerbaijan [21]. This research paper also focuses on comparing the impact of localization on learning outcomes.

In [22], Olga V. Stukalova investigated process optimization in higher education institutions, and that research paper defined the factors that directly impact productive optimization. Those factors are providing independence in selecting educational programs regarding the community needs; improvement and application of innovative learning technologies covering the chief

practice-oriented sectors of modern human life; continuity and staging of organizational and structural reform of the higher education system. This research paper focuses on both theoretical and practical aspects of the optimization in the non-profit online educational institutions.

Together with John Hilton and Sarah Harper, Allen Rao analyzed the main challenges, current situation, and quantitative impact on the students of localizing the Khan Academy content into Simplified Chinese language in [23]. They defined the current views, localized contents, and for this, they have analyzed the dashboard of the Simplified Chinese version of the website. After analyzing the current situation, this academic paper redesigns and implements the new processes to see the effectiveness according to the various success metrics, which are human capital allocation, time, and the quality of content.

In [24], localization of the digital content to use in the secondary schools was analyzed, and within this framework, various data visualizations were prepared based on the localized content. They analyzed the impact according to the region, content, localization pattern, type of the content. That research helps the reader to comprehend the success of the localized content. Moreover, this research focuses on the localization of the Khan Academy content into Azerbaijani, and it also applies the entire business process lifecycle from the current business process identification to the monitoring of it.

3 Methodology

This section describes how the research had been set and the methodologies that we have implemented to answer the research questions. First, we conducted workshops and interviews to collect the responses of the stakeholders. Then we redesigned the current localization process. Lastly, we compared the previous and the redesigned processes based on the duration spent on the localization process, the number of stakeholders involved, and the learning experience of the users.

To answer the research questions, a mixed-methods approach (qualitative and quantitative) has been implemented (Figure 4). As a quantitative method, we have held surveys, and based on this, we analyzed the collected data. As a qualitative analysis, we have held two workshops and two different types of interview phases. These different interview phases were as follows:

1. Phase 1 interviews: to collect the feedback about the “as is” process
2. Phase 2 interviews: to collect the proposals to redesign the localization process

We have selected workshops to collect the overall feedback of the stakeholders, and the reason we have selected the interviews as a part of the research methodology was the benefits of the semi-structured interviews, such as understanding stakeholders` opinions in detail. Since we interviewed small groups, it was also helpful for accurate screening. Furthermore, we selected the surveys as qualitative parts of this research. Because it is the fastest way to collect respondents` feedback and find the average thoughts on the process improvements were feasible.

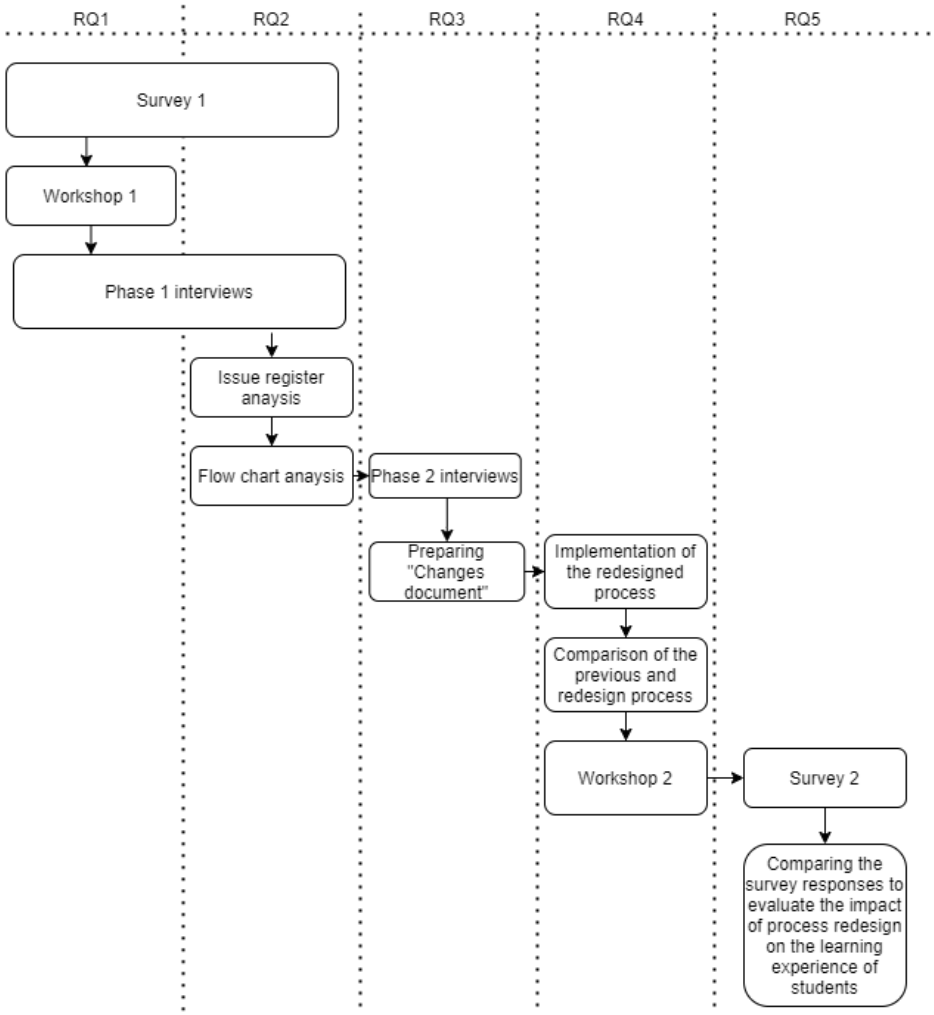


Figure 4: The research process

The research started with the survey to answer the RQ5. It was in the Azerbaijani language to collect the data about the learning experience and feedback on az.khanacademy.org. This survey has been shared with the users of the platform.

Then, to answer the RQ1, we moved to the first workshop in which the stakeholders gave information about their tasks and performance duration. Based on the information collected

from the workshop, we had designed the "as is" process. Then, we documented the process of content localization as it was carried out by the organization when this research started. After this, we started to conduct Phase 1 interviews with eleven stakeholders. In these interviews, we asked them to give their feedback about whether there should be any edits on the "as is" process design, and we have also asked them to mention the bottlenecks. In addition to the interviews, to answer the RQ2, we also analyzed the process both from qualitative and quantitative aspects through holding issue registers and flow chart analysis. In the issue register, we registered the issues and their potential impact on the number of users and the duration spent performing the tasks. In the flow chart analysis, we calculated the cycle time of each process. After process analysis, we stopped collecting the data from the survey, and we collected 127 responses. Then, to answer the RQ3, we started to conduct the Phase 2 interviews in order to collect the information about the potential changes in the "to be" process design. Within the Phase 2 interviews, we conducted eight interviews and summarized the collected responses in the "changes document". In this document, we have mentioned which changes we decided to implement by highlighting the reasons.

After preparing the "changes document" to answer the RQ4, we have invited stakeholders to localize new content (Algebraic expressions) based on the redesigned localization process. Firstly, we briefly explained the redesigned process to the stakeholders, and their questions have also been answered in this session. Subsequently, we gave them one week to test the process. After being sure that all stakeholders met with their tasks, we asked them to localize the content. Then, we compared the difference between the previous and redesigned localization process from two perspectives: the number of people involved and the cycle time. Then we conducted the second workshop in which they gave their feedback about the redesigned process.

Last but not least, after implementing the redesigned process, to answer the RQ5, we sent the same survey to the website users to compare their responses with the first survey. In the second survey, the number of responses was 107. We have compared the learning experience of users by asking about the quality of visuals and texts on the content, the quality of exercises, and articles from the locality perspective.

In the localization of Khan Academy content into the Azerbaijani language, two localization tools are used, Crowdin[4] and Amara[5]. Crowdin is used to localize the textual content such as exercises, articles, and Amara is used to localize the video contents. In this research, we analyzed both of these two tools because both of them play an essential role in the complete

[4] https://crowdin.com/
[5] https://amara.org/

localization of the content, and we could not ignore any of them. When the user opens the website (az.khanacademy.org), then they see videos, exercises, and articles together as one content. Thus, if we considered only one of the tools, we would not be successful in analyzing the optimization of the entire process.

3.1 Research setting

This section describes the activities taken to set up the research. This setup includes the information about Khan Academy, its localization, the selection of the subject, and contents to use in the implementation of the redesigned process.

Khan Academy

Khan Academy is an online educational institution that provides free world-class education anywhere for anyone who can understand English or any other available languages on the website (www.khanacademy.org). The company was founded by Sal Khan in 2008, and there are more than 60 topics, subjects under the major spheres such as art history, science, economics, computing, maths, history, and more, including K-14 and test preparation (SAT, Praxis, LSAT). Khan Academy has language advocates in more than 50 countries who localize the content into different languages [7]. The localization procedure is the process in which the content, product, or service is adopted. The localization brings benefits to the organizations, such as the increase in the number of users, the rise of revenue, and thus, it is one of the important processes within the organization. Within this framework, it is quite significant to design the business process optimization in online educational institutions since it could lead the organizations to prepare the content faster without allocating financial and human resources [18]. Khan Academy Azerbaijani is not an official part of Khan Academy company. It is operated by two people (Ayaz Karimov and Leyla Gasimova) who lead the localization process in Azerbaijan. This empirical research aims to optimize the Khan Academy Azerbaijani business process within the scope of the entire content localization into the Azerbaijani language.

Why have we selected Mathematics?

To measure the impact of the redesigned localization process of Khan Academy content into the Azerbaijani language, we have selected “Algebraic expressions” to compare with the “Negative Numbers.” Before discussing why we have chosen “Algebraic expressions” as content to compare, we would like to mention the primary motivation for selecting the Mathematics content. The primary motivation points to select this content were:

1. Mathematics is the only content that the Khan Academy Azerbaijani team localized into the Azerbaijani language, one of its subtopics fully and that fully translated topic is “Negative numbers.”

2. It is the most required [42] and used topic for the students [7].

How did we decide on choosing the contents to compare?

In each Khan Academy content, there are exercises, articles, descriptions, and videos to be localized. Within this framework, the data related to Arithmetic and Algebra Basics can be found in this table.

Firstly we have decided to choose the "Negative numbers" because that is the only localized content by the Khan Academy Azerbaijani team. Secondly, for the comparison, we have chosen "Algebraic expressions" from "Algebra Basics" as a chapter to localize to compare the result of localization process optimization. The reason why we have chosen those subjects ("Negative numbers" and "Algebraic expressions") to compare is that they include roughly the same number of exercises, articles, descriptions, videos and possess a similar number of words. It will help us to measure the difference based on the success measurements. Figure 7 depicts the difference between the number of words in "Negative numbers" and "Algebraic expressions" contents.

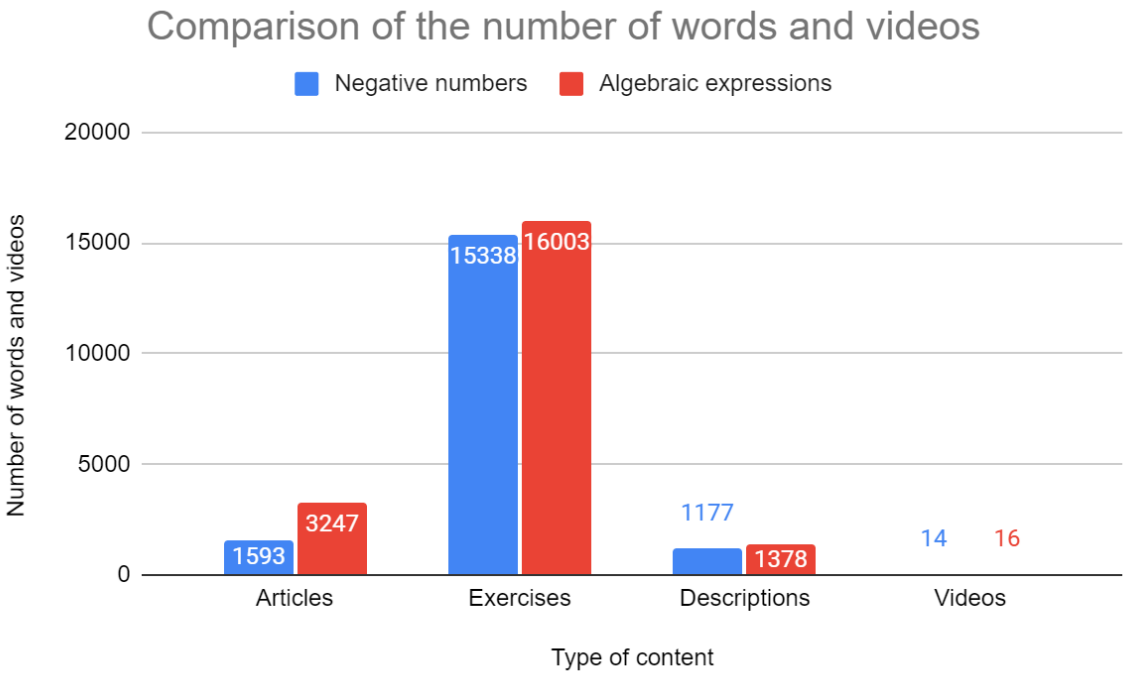


Figure 7. The comparison of the number of words, videos in "Negative numbers" and "Algebraic expressions"

Success metrics

Within the scope of the localization of the Khan Academy content into the Azerbaijani language, there are three primary success metrics that the team focuses on. One of these success metrics is the number of stakeholders involved. It is important for the team to use minimum input and achieve maximum outcome. Because it helps them to localize more contents by allocating less human capital. Within this framework, the duration spent on the localization is the second success metric. Localizing the content faster lets the team create more contents. Based on the experience of the team, more content on the platform drives the attention of more users [8]. As it also serves the mission of the Khan Academy to provide a free, world-class education for anyone, the team tries to onboard more users every day. The last success metric is the learning experience of the users. This is about the feedback of the users on the platform. Moreover, it is one of the team's key aims to produce high-quality content so that the users are satisfied with their experience on az.khanacademy.org.

3.2 Research question 1

In order to answer the first research question, we conducted a workshop and Phase 1 interviews to collect the feedback about the “as is” process with 11 stakeholders, such as the translator, the team leader, who actively participated in the localization process. We selected the stakeholders based on their experiences. It takes around six months to understand the localization process, and thus, we have selected the stakeholders who have at least six months of experience in the team. Moreover, we have also selected the participants based on their position in the team. Currently, there are 11 positions in the team, and we invited one person per position. In order to invite the stakeholders, we have contacted the language advocate. She has directed the stakeholders who meet the requirements mentioned previously and can dedicate their time to the research. Firstly, we have held the workshop with all (11) stakeholders simultaneously, and the workshop commenced after giving them brief information about the purpose of the research. Hereby, we would like to mention that one of the co-authors of this research is the project manager at Khan Academy Azerbaijani. Thanks to his 3-year-experience, he was already familiar with the localization process, and a year ago, he designed the “as is” process. That signifies that the team had experienced the business process design a year ago. However, the main purpose of this workshop was to define the main tasks and detailed information about each task, such as the owner of the task, with whom the stakeholder collaborates before or after fulfilling this task to observe the current localization process.

We have utilized Miro [33] within this workshop, which is an online visual collaboration tool. The workshop continued for 80 minutes, and the stakeholders participated voluntarily in the workshop. We have announced the workshop in the Khan Academy Azerbaijani community.

Two weeks in advance, we have recorded and sent basic tutorials about the platform. The stakeholders who have participated are Azerbaijani language advocate, Crowdin translator, project manager, Amara translator, content specialist, Crowdin proofreader, Amara proofreader, team leader, video editor, Crowdin coordinator, dubber, Amara coordinator. There was one person per role, and in total, there were 11 stakeholders involved in this research (Table 1).

Interviews/ Workshops	**Event**	**Participants**	**ID**
Phase 1 Interviews	Interview 1	1	I01 (Dubber)
	Interview 2	1	I02 (Language advocate)
	Interview 3	4	I03.1 (Crowdin proofreader), I03.2 (Crowdin translator) I03.3 (Amara proofreader), I03.4 (Crowdin coordinator)
	Interview 4	1	I04 (Team leader)
	Interview 5	1	I05 (Amara translator)
	Interview 6	1	I06 (Content specialist)
	Interview 7	1	I07 (Amara coordinator)
	Interview 8	1	I08 (Video editor)
Workshop 1	Workshop 1	11	W01 (Dubber), W02 (Language advocate), W03 (Crowdin proofreader), W04 (Crowdin translator), W05 (Amara proofreader), W06 (Crowdin coordinator), W07 (Team leader), W08 (Amara translator), W09 (Content specialist), W010 (Amara coordinator), W011 (Video editor)
Phase 2 Interviews	Interview 1	1	I11 (Dubber)
	Interview 2	1	I12 (Language advocate)
	Interview 3	4	I13.1 (Crowdin proofreader), I13.2 (Crowdin translator) I13.3 (Amara proofreader), I13.4 (Crowdin coordinator)
	Interview 4	1	I14 (Team leader)
	Interview 5	1	I15 (Amara translator)
	Interview 6	1	I16 (Content specialist)
	Interview 7	1	I17 (Amara coordinator)
	Interview 8	1	I18 (Video editor)
Workshop 2	Workshop 2	11	W11 (Dubber), W12 (Language advocate), W13 (Crowdin proofreader), W14 (Crowdin translator), W15 (Amara proofreader), W16 (Crowdin coordinator), W17 (Team leader), W18 (Amara translator),

			W09 (Content specialist), W110 (Amara coordinator), W111 (Video editor)

Table 1. Workshop and interview participants

Having introduced the research and its main purpose shortly, we have started to ask the previously-prepared questions. There were four basic questions that we wanted to get answers to prepare the draft version of the “as is” process, and those questions were as follows:

1. What is your position at Khan Academy Azerbaijani?
2. What is your main task?
3. What are the other tasks that you do?
4. With whom do you collaborate before and after fulfilling this task?

These questions were important for us to know and understand the tasks of each stakeholder to design the “as is” localization process. It is significant to design the “as is” localization process punctually. Because it helps to define the bottlenecks. Within this framework, we asked these four questions to understand who the stakeholders are, their tasks, and how they interact to fulfill them.

Having held the workshop, we have designed the new “as is” process based on the responses. As mentioned above, the team had already experienced the business process design a year ago, and we have also referred and used the help of that previously-designed “as is” process. Furthermore, it is worth mentioning that when the team designed the localization process a year ago, they did not redesign it. Because the main purpose was to design the current localization process, and this design was used in the grant applications. Additionally, a year ago, the stakeholders focused on preparing the demo website, and redesigning the process was not their goal.

For the “as is” design, we have utilized the academic version of Signavio [16]. Signavio is a business process software, and from which we have used the standard Business Process Model and Notation (BPMN 2.0) diagrams.

In order to design the “as is” process of Khan Academy Azerbaijan within the framework of the localization process, we have implemented a flow chart technique [1] for process modeling. As a result, we have designed the “as is” localization process and described how the process happens.

Data collection

Having designed the “as is” localization process of the Khan Academy Azerbaijani, we conducted Phase 1 interviews with all stakeholders in order to

1. get their feedback about the “as is” process,
2. learn the average time that they spend on doing their tasks individually,
3. the main bottlenecks both in each subprocess and platform.

The interview was semi-structured. In the Phase 1 interviews, there were eight interviews in general. We organized the interviews based on the roles so that some of the stakeholders, such as team leader, language advocate, dubber, video editor, content specialist, participated in the interview individually. At the same time, we have invited Amara translator and Crowdin translator to the same interview because their tasks were quite similar.

Each interview went on within 35-45 minutes on average, and the interviews were conducted on Zoom [31]. After getting the consent of the stakeholders in the form of signed documents, we have recorded the entire interviews. Then all interviews have been transcribed to be utilized for the coding scheme methodology [34]. This methodology is one of the ways to analyze the qualitative content.

Data analysis

All the interviews were conducted in Azerbaijani and were also transcribed into Azerbaijani to not lose the meaning of words. Moreover, while scheme coding[6], we have translated the quotes from the interviewees. We have categorized the responses into three categories:

1. *Platforms*: we defined this category in order to categorize the bottlenecks based on which platforms they occur
2. *Subprocess*: we defined this category in order to categorize the bottlenecks based on which subprocesses that occur
3. *Subprocess edit*: we defined this category because after designing the “as is” process, we have also asked the stakeholders whether the designed localization process is the same one they currently implement, and stakeholders proposed edits in certain parts of subprocesses. We categorized their edit proposals based on the subprocess within the coding scheme.

To analyze the interviews in this phase, we have used Google Sheets [35], an online spreadsheet app that lets users create and format spreadsheets and simultaneously work with other people.

[6] The coding scheme can be accessed here: https://zenodo.org/record/4753755#.YJwvRqgzZPY

In this scheme coding, we have utilized the template below (Table 2). In this scheme coding, the coding rule was decided in a way that if the stakeholder has mentioned the same or similar comments about the given an example, then we gave “1” for this stakeholder, and the stakeholder has never mentioned the same or similar comment, then we gave “0” for this stakeholder for that example.

ID	Category	Examples	Stakeholder 1	Stakeholder 2	Stakeholder 3	Total
The ID of examples	Category of the example	The comment that was mentioned in the interview (Stakeholder, Interview number)				

Table 2. The frame that was used within scheme coding

For example, in Table 3, we gave the example in which the Crowdin translator mentioned the bottleneck which occurs in the Crowdin platform. Thus, we put his example in the “Crowdin” category by adding his example in the “Examples” column. We have mentioned his position and the ID of interview. Then, while analyzing the project manager`s interview transcription, then we see a similar comment about the same example. Thus, we gave “1” to the project manager and “1” to the Crowdin translator since he has also mentioned the example in his interview. However, the dubber has never mentioned the same or similar comment about that example in this case. Therefore, we gave “0” for the dubber. Finally, the last column depicts the total number of codes which is “2” in this example.

ID	Category	Examples	Project manager	Crowdin translator	Dubber	Total
1	Crowdin	"The title of the exercise does not match with the one in Khan Academy translation platform" (Crowdin translator, Interview 4)	1	1	0	2

Table 3. Example for the frame that was used within scheme coding

Subsequently, based on this qualitative content analysis, we have altered the proposed parts of the localization process and collected the data about the main bottlenecks for each platform as well as the subprocess. Moreover, the last version of the “as is” process was shared with all stakeholders via email, and all of them confirmed it.

3.3 Research question 2

In order to answer RQ2, we have implemented both the qualitative and quantitative methods for analyzing the process.

3.3.1 Qualitative (issue register) analysis

From the perspective of qualitative process analysis, we have selected the "issue register" analysis. Issue register is a type of qualitative process analysis to categorize identified issues as part of "as is" process modeling [1]. The main reason we have selected the issue register analysis for the qualitative process analysis was its ability to provide detailed information about each issue and its quantitative and qualitative impact. We have used the coding scheme results about bottlenecks from Phase 1 interviews and our observations to hold the issue register analyses.

Here, in the issue register analysis, after naming, describing, and assigning the ID for each issue, we have mentioned the potential impact both from qualitative and quantitative aspects. We used the data collected previously in the Phase 1 interviews, Workshop 1, and our assumptions to calculate the quantitative aspects.

3.3.2 Quantitative (flow) analysis

Moreover, from the perspective of quantitative analysis, we have conducted the flow analysis. Flow analysis has different techniques such as flow analysis for cost, critical path method, cycle time efficiency, calculating cycle time [1]. Within this research, we have implemented the cycle time calculation because time is the main success factor of the stakeholders. One of the main objectives of this research is to redesign the localization process that helps stakeholders finish the localization within a shorter timeframe. In order to calculate the cycle time for each subprocess, we first made the table mentioning which stakeholder performs which task and how much time they spend on this task.

While implementing the cycle time calculation flow analysis, we utilized the following rules [1]:

1) If the tasks are performed one after the other, we summed the time they spent doing their tasks (Figure 5). For this, we used the formula below:

$$CT = \sum_{i=1}^{n} T_i \qquad (1)$$

in which:

- CT (cycle time): the total duration that spent in the process
- T_i: the set of tasks

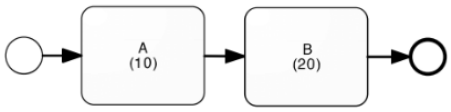


Figure 5. Fully sequential process model (durations of tasks in hours are shown between brackets) [1]

2) In addition to the sequential processes, there were also processes with the exclusive (XOR)[7] gateways. One of the main points was to know how frequently each branch of the XOR gateway is taken (Figure 6). The stakeholders also answered this frequency within the Phase 1 interviews. Furthermore, in order to calculate that kind of processes, we implemented the formula below (2):

$$CT = \sum_{i=1}^{n} p_i \times T_i \qquad (2)$$

Where CT is the cycle time of the process, p_i is a branching probability, and T_i is the cycle of these nested processes.

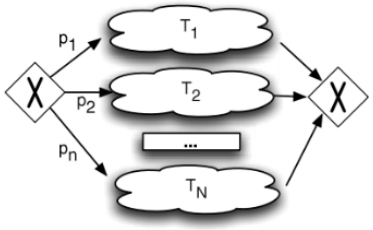


Figure 6. XOR-block pattern [1]

3.4 Research question 3

Section 3.4 describes the methodologies that used to answer the research question 3.

3.4.1 Phase 2 interviews

In order to answer that research question 3, we have conducted Phase 2 interviews with 11 stakeholders to get their feedback about solving the current bottlenecks. The interviews were in a semi-structured format, and our primary question was about asking the question whether they have any proposals to solve the bottlenecks or not. We notably held the Phase 2 interviews because, within the Phase 1 interviews, the stakeholders mentioned several times that they also know the solutions to most of the problems, but they could not implement those solutions due to other workloads. In Phase 2, we have conducted eight interviews with all (eleven) stakeholders in the Azerbaijani language, and the interviews continued around 35-45 minutes

[7] The exclusive (XOR) gateways are used when a process splits into several paths and only one of them can be active [5].

on average. The stakeholders were the same people who participated in the Phase 1 interviews. After getting consent to record the interview, we took the recordings, and they have been used for transcription. Even though the transcriptions were in the Azerbaijani language, we have translated the quotes from the interviews and used them in English in the coding scheme table. To analyze the interview results, we have implemented the same methodology, which was implemented in answering the research question 1 called “coding scheme.”

While implementing the coding scheme methodology, we utilized the same table (Table 2). Coding was done as it was in research question 1. It signifies that when there is any statement by a stakeholder, and if another stakeholder mentions exactly or similarly the same statement, then we coded it as “1” on their column. If there was no similar statement from that stakeholder, we coded it as “0”.

3.4.2 Changes document

Moreover, we have prepared the changes document based on the Phase 2 interview results, issue register, and flow chart analysis to define the alterations in the localization process.

While proposing any changes, we have utilized the template from [1], which was as follows:

[The ID of the change]. [The title of the change]
Description: the brief description of the potential change and the problem.
Justification: explanation of how the change can be implemented to solve the problem and which heuristic we refer to.
Result: the potential quantitative and qualitative result of the change.

At the end of the changes document, we also mentioned the statements from the stakeholders that were not implemented. Hereby, we have highlighted the reason why it was not applicable within this research. After preparing the changes documents, we took the approval from all stakeholders that the changes are feasible, and they agree with all the changes.

3.4.3 Redesigning the process

Subsequently, we redesigned the localization subprocesses based on the changes document. In order to redesign, we have used the same tool called “Signavio” as it was in research question 1, and the technique was a flow chart since it lets us make any changes if needed.

After defining the changes, the stakeholders have been invited to participate in the localization of the new content. To compare the impact of the redesigned process, “Algebraic expressions” topics from the Math content on the Khan Academy platform have been selected.

3.5 Research question 4

In the RQ4, we analyzed the changes from three perspectives: human capital, cycle time, and the quality of content to compare the difference between the previous and redesigned process. Furthermore, we conducted the workshop to collect the feedback of the stakeholders about the redesigned localization process.

3.5.1 Comparison of the previous and redesigned process

To measure the difference between the previous and redesigned process, we have compared the success metrics of the Khan Academy Azerbaijani. These success factors are as follows:

1. the duration spent on the localization of the content,
2. the number of people involved in the localization of the content,
3. the quality of the content.

In order to calculate the duration of the localization process, we have implemented Formula 1 and Formula 2 based on whether the process is sequential or containing any XOR gateways. Having collected all the data about the localization process, we have visualized them to depict how much time it used to take and what it currently takes to localize the content. From the perspective of human capital allocation, we demonstrated how many people got involved in the previous and redesigned process. Last but not least, to understand the quality change, we have asked the research question separately, which is addressed in research question 5.

3.5.2 Workshop 2

Moreover, we conducted the second workshop with the same eleven stakeholders to get their feedback about the redesigned process. In the workshop, we asked the stakeholders` opinions about localizing content based on the redesigned system. The workshop continued for 50 minutes, and it was entirely in Azerbaijani. In the workshop, we used the same tool, which is Miro. The responses of the stakeholders have been added to the sticker board.

3.6 Research question 5. The impact of the redesigned process on the learning experience of users

Before redesigning the localization process and after doing so, we have shared the survey with Khan Academy Azerbaijani users to measure the impact of the redesigned method on students' learning experience. Our purpose was to understand how much the redesigned localization

process has impacted users' learning experience. Hereby, as a learning experience, we considered any interaction that contributes to students' learning on the Khan Academy platform (az.khanacademy.org).

As mentioned above, we have conducted two surveys: the first one was shared with the users before redesigning the process, and the second one after redesigning the process. The survey was shared through different channels such as the Khan Academy Azerbaijani Facebook community, the current email list of Khan Academy Azerbaijani users, social media channels of the Khan Academy Azerbaijani. To collect the data, we have used Google Forms, and the data was stored in Google Sheets. 127 users have filled the first survey, and 107 users filled the second one. Once collected, the data has been checked, and the unrelated responses were removed. There were two and one responses in the first and second surveys, respectively, in which the respondents mentioned different feedback about their volunteering experience in the comments section, and we have found it as an unrelated response. Since the number of unrelated responses was only three, we have decided to remove them rather than adjust.

Moreover, we have also cleaned 17 responses in total by editing the spelling mistakes. All of these spelling mistakes were in the question about the residency of respondents, and these mistakes were minor ones such as instead of writing the “Moscow'' (or “Moskva'' in Azerbaijani), they wrote it like “ Moscwo”. After editing these spelling mistakes, we have compared the results of the first and second surveys.

4 Results

In this section, the results of each research question are presented. The results description follows the order of the research questions defined in Section 1.2 of this thesis. Section 4.1 describes the results of the first Workshop 1, which was organized to design the "as is" design, the explanation of the "as is" design, and the Phase 1 interviews which were about editing the "as is" localization process design and finding the bottlenecks based on the feedback of the stakeholders. In Section 4.2, we presented the results process analysis from a qualitative (flow) and quantitative (issue register) perspective. Section 4.3 explains the results of Phase 2 interviews in which we have collected the proposals of stakeholders about the potential improvements in the "as is" process, the changes document which explains the accepted and rejected proposals for the redesigned process, redesigned subprocesses. In Section 4.4, we compared the previous and redesigned process from the perspective of the duration spent on the localization process and the number of stakeholders involved. Lastly, in Section 4.5, we measured the impact of the redesigned process on students' learning.

4.1 Research question 1

In this section, we describe the results of the Workshop 1, Phase 1 interview, and the “as is” localization process of the Khan Academy content.

4.1.1 Workshop 1

As a result of the first workshop, we collected the responses from all stakeholders about their tasks and contact persons to fulfill these tasks. The final board (Figure 8) was the primary source to design the “as is” process.

Figure 8. Workshop 1 sticker board

4.1.2 “As is” localization process[8]

The localization process of Khan Academy into Azerbaijani commences with selecting the content, and it finishes when the content is localized totally. The entire localization process consists of six major subprocesses, which are as follows:

1. Assign content tasks
2. Translate Crowdin content
3. Translate Amara content
4. Recreate videos
5. Control quality
6. Deliver video

[8] https://zenodo.org/record/4761968#.YJ5fXKgzZPY

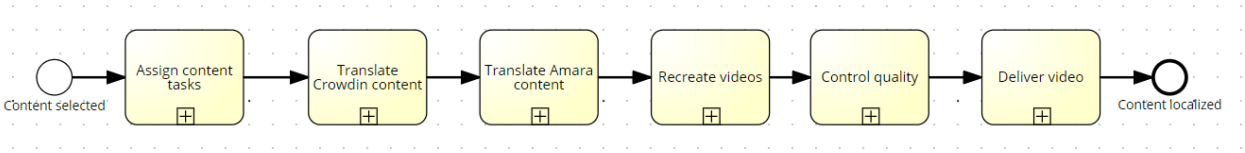


Figure 9. "As is" localization process of Khan Academy into Azerbaijani language

All the information about the task's duration, the possibility of any cases happening was collected within the Workshop 1 and Phase 1 interviews.

Subprocess 1. Assign content tasks (Figure A)

The project manager is responsible for selecting the content, and after choosing the content, he assigns the team that will be responsible for the localization of that content. The project manager works closely with all teams. Thus, he directly assigns the content based on the need of the team. For instance, if the Math team is finishing the current localization of any content, then the project manager assigns the content a few days in advance for this team. Having chosen the content and assigning it to the team leader, the project leader notifies the team leader, and then the team leader notifies the Crowdin coordinator. The coordinator works on the Google Sheets to assign each volunteer to specific topics. On the Khan Academy platform, each topic has some exercises, descriptions, articles, and the coordinator assigns each person to the content they are responsible for translating. The Crowdin coordinator also defines the deadline to each volunteer based on the number of words in the content.

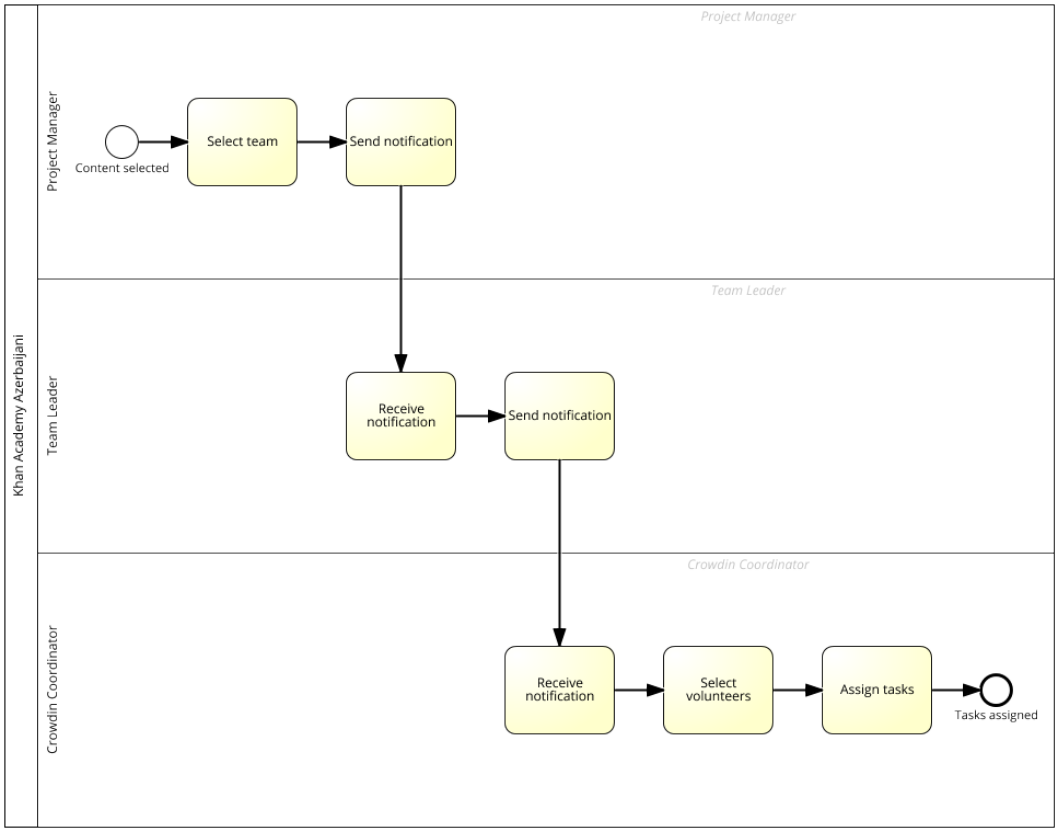


Figure A. The "as is" diagram of the "Assign content tasks" subprocess

Subprocess 2. Translate Crowdin content (Figure B)

When the Crowdin coordinator assigns the task to the volunteer, they also notify the translator by sending an email. Sometimes, after getting the notification from the Crowdin coordinator, the Crowdin translator encounters the challenge to find the exercise on the translation platform. The main reason for this is that the name of the content assigned to the volunteer does not match the one that the Crowdin translator searches. In these cases, the Crowdin coordinator takes the link of the task and sends it to the Crowdin translator directly. This problem occurs 0.3% of the time. In general, Crowdin translators can easily find the task, and then they start translating the task. 99.2% of tasks contain both texts and photos. The Crowdin translator is responsible for translating for both of them, and only after translating both of them can they save their translations. The translation duration of the task varies based on the number of words that they need to translate. However, in total, it takes 30 minutes minimum and eight days maximum.

There is a tool called “Smart translations,” which helps the Crowdin translator translate the same upcoming words at once. The Crowdin translators and Crowdin coordinator mentioned the weak aspects of Smart translations because if the translator does not translate the string correctly, then the entire upcoming string translations will not be well-translated. Moreover, in some cases, mistranslation can happen because of the difference in grammatical structures of the Azerbaijani and English languages. For example, in the Azerbaijani language, the suffixes are added after the numbers. Nevertheless, in English, those suffixes are generally prepositions. Therefore, it creates the problem because even if the Crowdin translator translates the string correctly, the “Smart translations” do not consider those suffixes, and in Azerbaijani, the suffixes are changed in each number. Thus, Crowdin translators and Crowdin coordinators find this a bit challenging to use.

When the Crowdin translators save the translations, the strings may not be saved in 0.7% of all cases, and for this, they need to wait one hour to resave the translations. This situation also occurs due to the browser, and changing it helps to solve the problem. Having saved the translation, the Crowdin translators notify the coordinator, and some team leaders check the content whether it is fully translated or not. It is not a mandatory step, but some team leaders do this not to have any problems within the proofreading process. Then, the team leader notifies the Crowdin proofreader to start the proofreading process. The proofreader is also responsible for checking the local perspectives of the content as well. For example, if the Crowdin translator has utilized any foreigner names, then it is the proofreader`s responsibility to edit it. Thus, it takes a bit more time to proofread rather than translate the content. The proofreader checks both textual and non-textual materials, and if there are any mistakes, they edit them. If not, then they approve the translations. The entire proofread process of the task differs according to the number of words in each task accordingly. It takes two hours minimum and a month maximum to proofread the content. When the Crowdin proofreader approves the translation, they notify the team leader about this, and after this, the localization on Crowdin finishes.

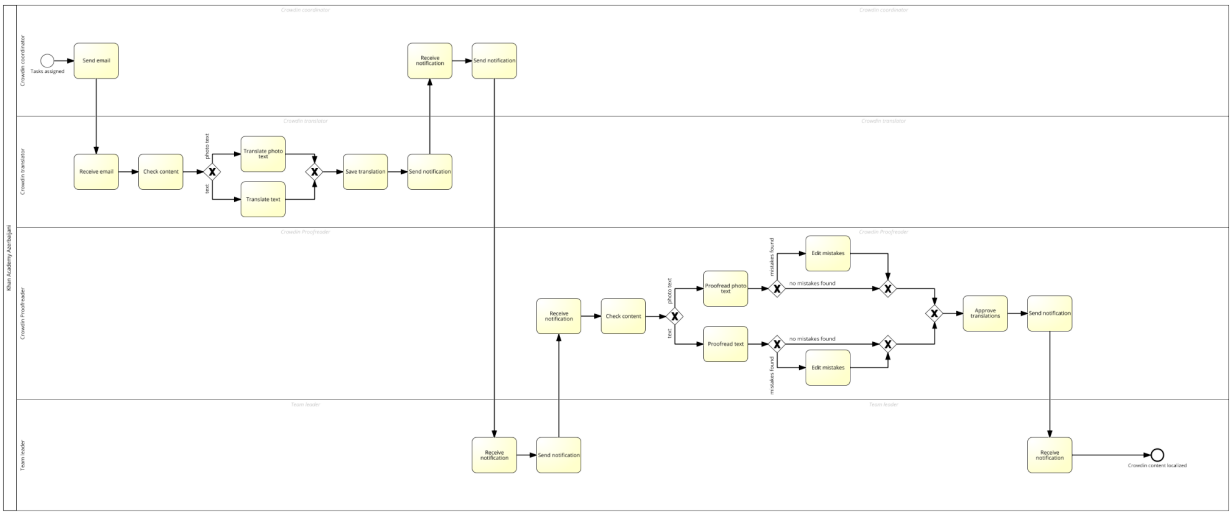


Figure B. The "as is" diagram of the "Translate Crowdin content" subprocess

<u>Subprocess 3. Translate Amara content (Figure C)</u>

When the content is localized on Crowdin, the team leader notifies the Amara coordinator to start the localization of the content on Amara. Then, the Amara coordinator notifies the Amara translator to start the translations of videos. When the Amara translator gets this notification, they check whether the video is on Amara or not because all of Khan Academy videos do not exist on Amara or have been uploaded to the platform with another name. In these cases, the translator notifies the problem to the Amara coordinator, and the Amara coordinator reports it to the team leader. Next, the team leader sends a notification to the project manager about the problem. The project manager reports the problem to the HQ Employee, and this person edits the video on the system and sends the notification to the project manager. The project manager notifies the team leader, and they notify the Amara coordinator. After this, the Amara coordinator notifies the Amara translator about the update. After this, the Amara translator starts the translations of the assigned videos. Hereby, in 4% of all cases, two challenges occur. The first one is about videos that do not have English subtitles. When the video does not have the subtitles, then the translation errors and the duration for translating that video increase. The team generally does not solve this problem, and the Amara translator is expected to translate the video even though the subtitles do not exist. The second problem is technical. When the sentence is translated into Azerbaijani, then the length of the sentence might be a bit longer. In this case, Amara does not allow the Amara translator to save the translations. For this, the Amara translator needs to break subtitles into sections. It requires four times more time than the casual translation, and it is again the Amara translator`s responsibility to translate even if this problem emerges.

Furthermore, when the Amara translator saves the translation, they notify the Amara coordinator about it, and the Amara coordinator sends the notification to the team leader. The team leader notifies the Amara proofreader about the update and the Amara proofreaders, and if there are any mistakes, they edit those mistakes, and they also need to check whether the content is translated by considering the local issues. Then, having proofread the content, the Amara proofreader approves the translations and notifies the team leader.

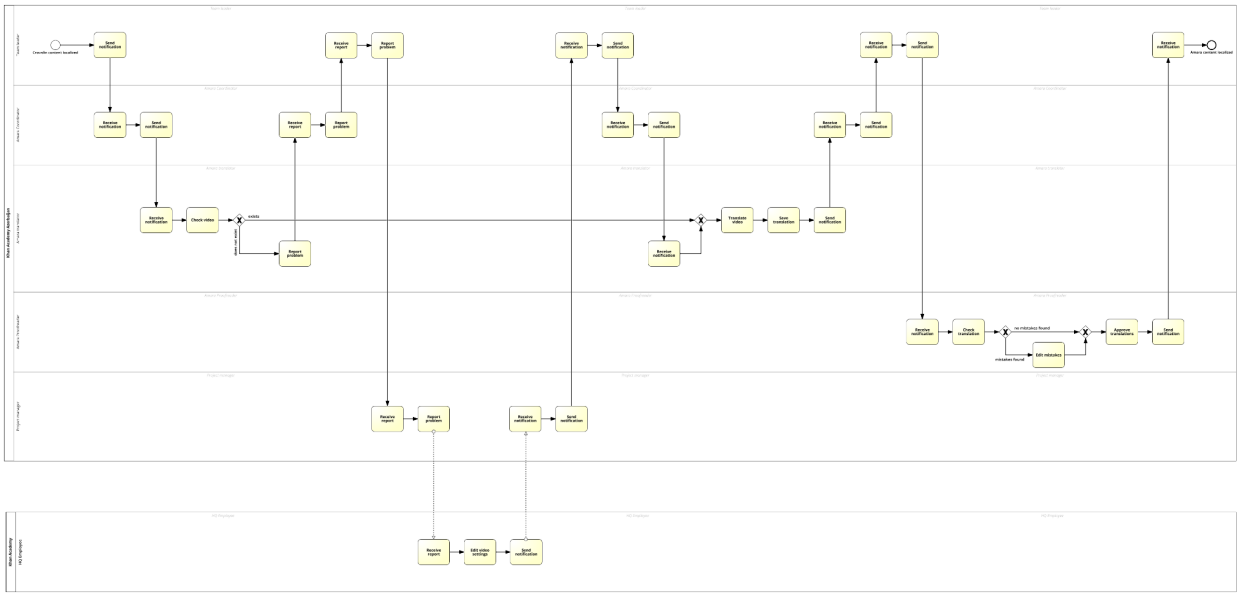


Figure C. The "as is" diagram of the "Translate Amara content" subprocess

Subprocess 4. Recreate videos (Figure D)

By recreating the videos, we mean dubbing and editing the original videos of Khan Academy. The recreation of the videos starts when the team leader notifies the project manager about the localization of Amara content, and the project manager notifies the dubber about it. The dubber, firstly, checks whether the script of the videos in Azerbaijani is on Amara or not. If not, the dubber notifies the problem to the project manager, and then the project manager notifies the problem to the team leader. After this, the team leader sends the notification to the Amara coordinator, and the Amara coordinator reports the problem to the Amara translator. When the Amara translator gets the notification about the script, he edits the mistakes and sends the notification about it to the Amara coordinator. Then the Amara coordinator notifies the team leader about the script update. The team leader sends the notification to the project manager, and the project manager notifies the dubber.

In 0.002% of all cases, the errors still exist even if the Amara translator edits the script. In this case, the problem is solved as mentioned above. After being sure that the script is in Azerbaijani, the dubber commences dubbing the videos. According to the length and difficulty

of the video, it takes one hour - eight days to dub the videos. Once the dubbing is ready, the dubber sends audio to the video editor, and the video editor checks the software every time, whether it works or not. In 8.2% of all cases, the software does not work, and he needs to get in touch with the project manager to get the keyword to use the software. Only with this keyword it is possible to start editing the videos. While editing, in 22% of all cases, the audio does not match with the video. In these cases, the dubber is asked to redub the video. Within the localization process, Azerbaijani translation might be a bit longer than how the same sentence is expressed in English. Therefore, the dubber should be careful that what she is dubbing should match with what is on the video screen. After redubbing the video, the dubber sends the audio to the video editor, and the video editor adds the audio to the video. Having done this, the video editor sends the notification about the recreated video.

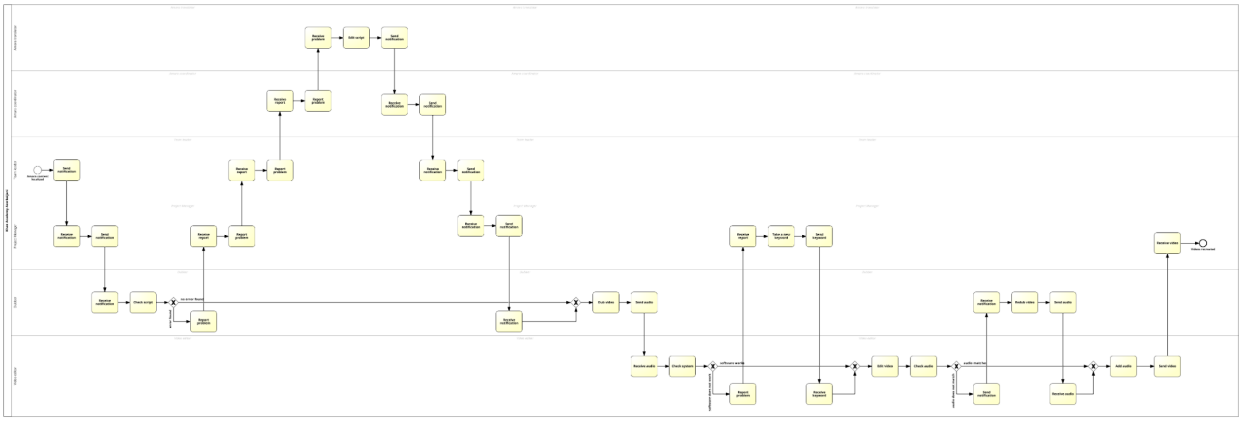


Figure D. The "as is" diagram of the "Recreate videos" subprocess

<u>Subprocess 5. Control quality (Figure E)</u>

There are mainly three stakeholders in quality control activities: the project manager, content specialist, and Azerbaijani language advocate. All of them check the content separately. Firstly, the project manager checks the content, and if he finds any mistakes, he reports the problems to the video editor, and the video editor fixes those problems. Having fixed those errors, the video editor sends the video to the project manager. Secondly, the content specialist does the same quality control checking, and any edits again are sent to video editors, and the video sends back the edited video to the content specialist. Finally, the Azerbaijani language advocate is the last person who does the checking process, and the same quality control takes place. After making all these quality checking processes, the Azerbaijani language advocate marks the video as checked.

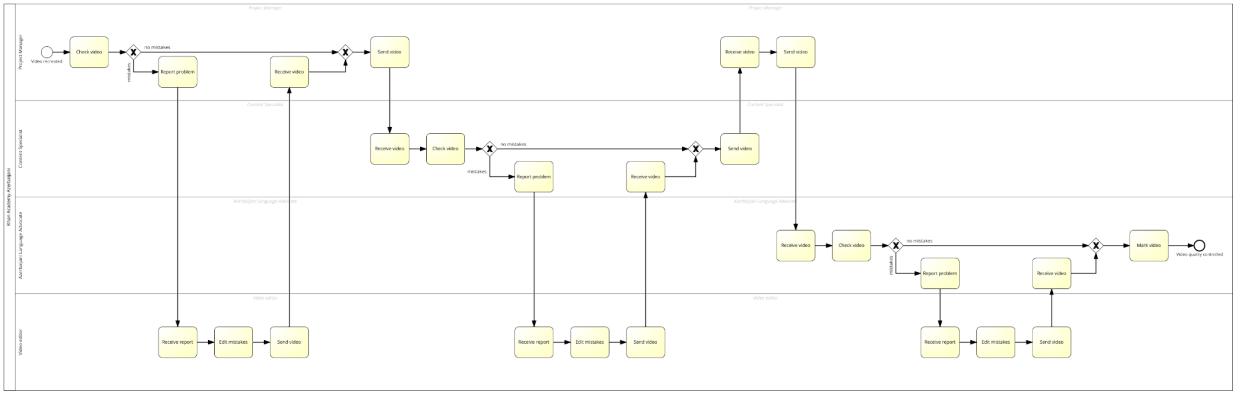


Figure E. The "as is" diagram of the "Control quality" subprocess

6. Deliver video (Figure F)

Having controlled the quality of the video, the Azerbaijani language advocate sends the video to the project manager, and the project manager uploads the video on YouTube. Khan Academy uses YouTube as the primary source of uploading its videos. After downloading the video, the project manager sends the link of that uploaded video to the Azerbaijani language advocate to map the videos on the Khan Academy database. *Mapping* is the process in which the localized videos are uploaded to the Khan Academy database. When the advocate is mapping, she needs to check whether all the videos are mapped or not because, in 26.2% of all cases, she encounters problems while mapping due to different reasons such as uploading the previously uploaded videos, entering the wrong video ID, internal technical error. When the error occurs, the advocate tries to analyze why it happened and upload the missing videos. In general, the mapping of 10 videos takes around 30 minutes to do so without any errors, but when the errors occur, this duration can go up to three hours. Actually, due to that problem, more than 150 videos were not mapped correctly. After mapping the videos, the entire content is marked as a localized content.

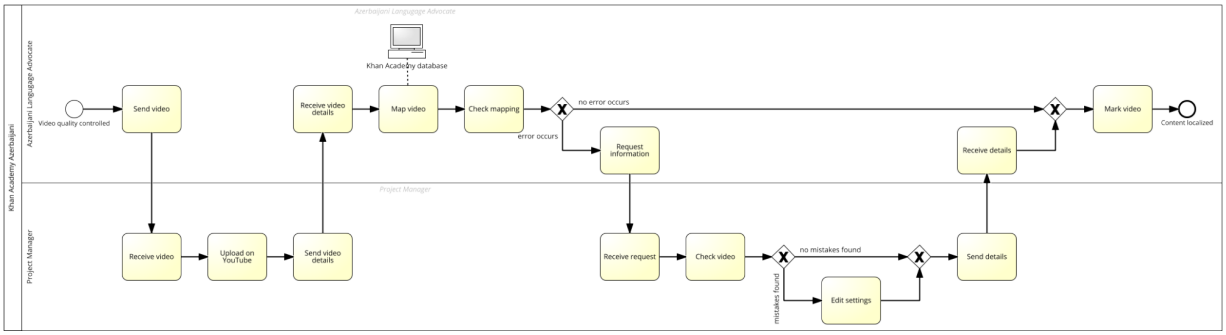


Figure F. The "as is" diagram of the "Deliver video" subprocess

Based on the Phase 2 interviews conducted, we have detected the main bottlenecks. We have summarized all of them in the Miro, and the interview coding scheme results were depicted

based on the template below (Figure 10):

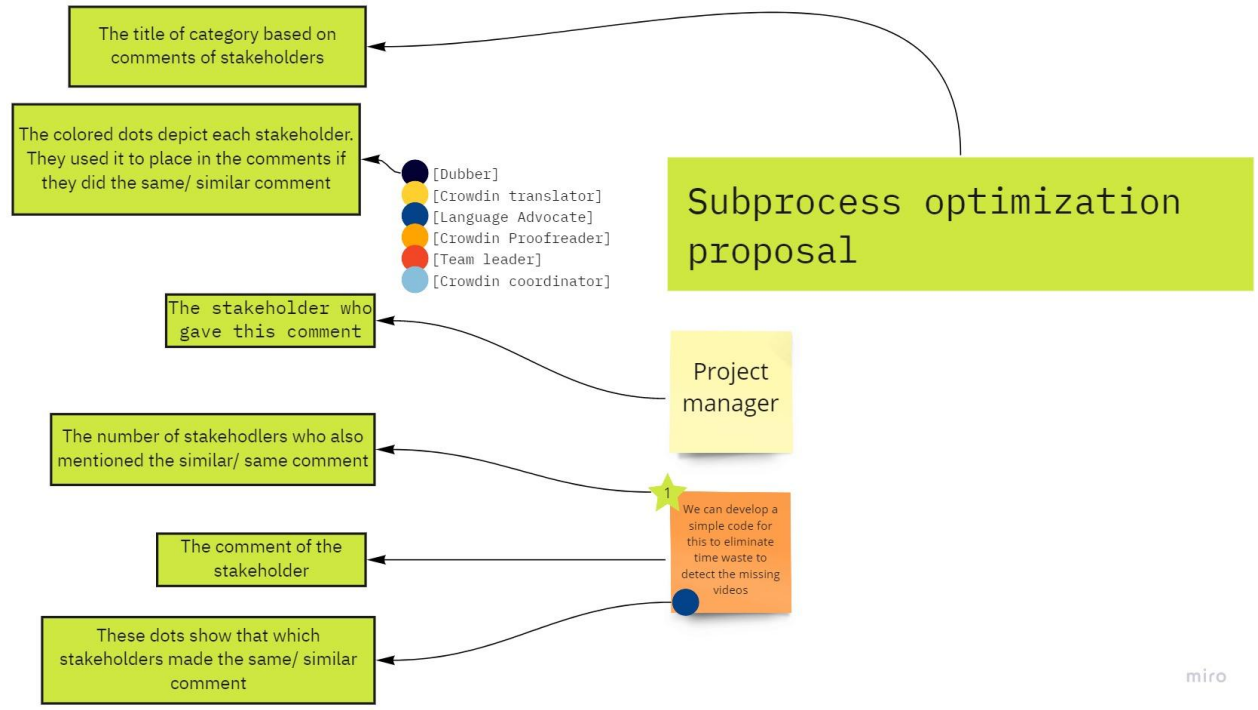


Figure 10. The framework to depict interview results

The first category in the Phase 2 interviews was whether the stakeholders did not agree with any parts of the "as is" design (Figure 11). Having implemented the coding scheme, we have found out that both Amara and Crowdin translators have to contact the Amara and Crowdin coordinators to solve their problems. The Crowdin translator proposed this change, and the other five stakeholders mentioned the change: *"I always get in touch with the Amara coordinator in case I would like to mention any problems" (I05).* The second change proposal was by the dubber, and she proposed changing the subprocess so that the video editor sends back the audio if there are any problems with this. The language advocate also supported her proposal. The third change proposal belongs to the team leader and is supported by the project manager. She proposed that she never contacts the translators. Instead, she first informs the team leaders, then the team leader notifies the translator; *"I also deliver the information to the team leader, and I do not contact the translators" (I04).* The subsequent change was about the software key problem that was not mentioned in the first version of the "as is" process. Last but not least, Amara translator has also mentioned a change which was about solving the problems. All of the five proposals were correct, and we have updated the "as is" localization process based on these proposed edits.

Figure 11: Subprocess edit proposals

The second category within the scheme coding was the platforms in which the bottlenecks occur (Figure 12). Within the interviews, stakeholders mentioned the bottlenecks, and we have categorized them based on the platform. Firstly, the language advocate mentioned two bottlenecks: 1) mapping is time-consuming, and the project manager also supported it: *"While doing the mapping we lose quite a lot of time" (I03.1)*; 2) texts on the video screen cannot be translated on Amara. Then there was also a statement by the Crowdin translator about the title of the tasks. She mentioned that the Crowdin coordinator sometimes makes a mistake while writing the task`s title on Google Sheets. Both Crowdin coordinator and team leader have approved the problem: *"The title on Translation portal and the exercise itself are not the same always, and because of this, volunteers might encounter difficulties to find the task" (I03.4); "You need to spend 5-6 minutes just to find the assignment sometimes" (I04).* Thirdly, there were also statements from the Crowdin coordinator and Crowdin proofreader about misusing the strings and inefficiency of "Smart Translations" on Crowdin, respectively. The Crowdin coordinator`s statement was supported by the language advocate as well, four stakeholders (Amara translator, Amara proofreader, Crowdin coordinator, Crowdin translator) supported the Crowdin proofreader`s statement: *"Smart translations might lead to having much more mistakes if not used correctly" (I03.3); "Smart translations make the mistakes when it comes to the suffixes in Azerbaijani language" (I03.2).* Finally, the Amara translator made the final statements, and they were about finding the videos on Amara and not possessing the subtitles in English on Amara. Four other stakeholders supported the first statement, and the main bottleneck was that the stakeholder spent quite a lot of time finding the video on Amara.

Nevertheless, sometimes it cannot even be found due to various problems such as the title or ID of the video. The second statement was about not being able to provide the subtitles in English.

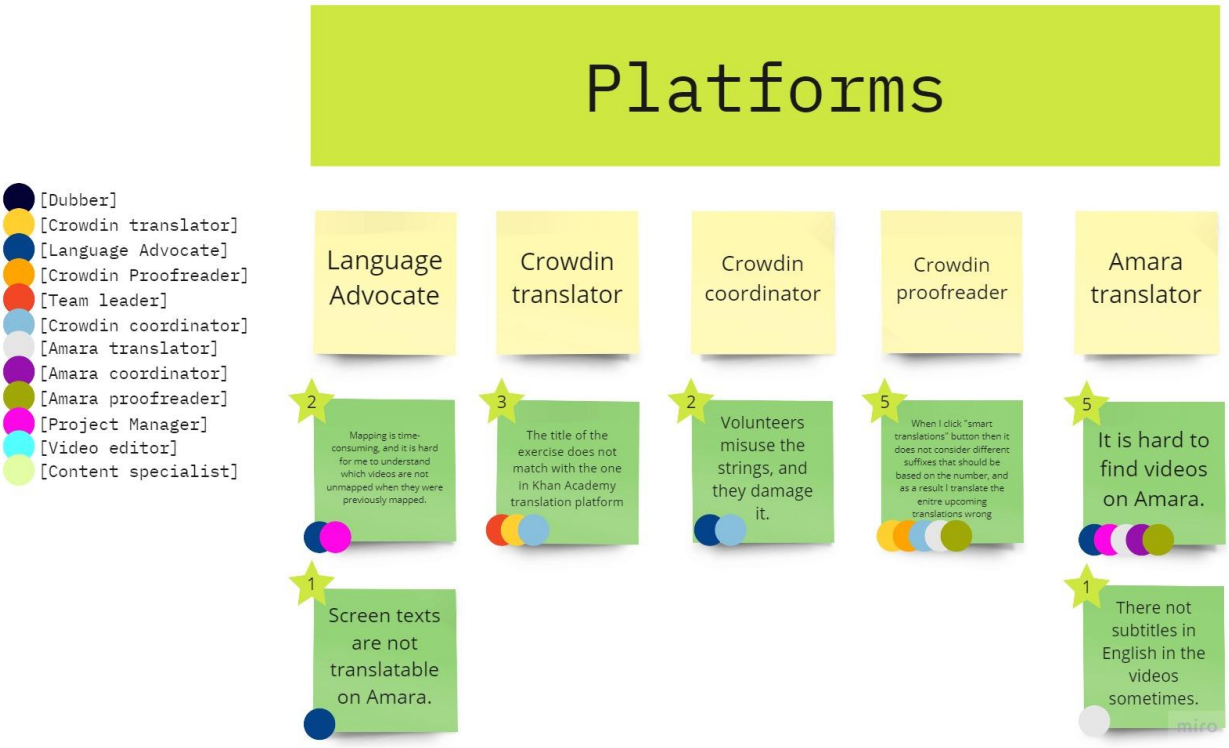


Figure 12. Interview results about the main bottlenecks that occur in platforms

The third category in the coding scheme was dedicated to showing in which subprocesses the bottlenecks occur and by whom the statement was made (Figure 13). Initially, the team leader has mentioned the lack of proofreaders in the team, and the project manager and Amara proofreader supported her statement: *"We wait a lot for the proofreader to fulfill their tasks" (I03.3).* Then, the language advocate made two different bottlenecks: 1) while mapping, she misses some videos, and finding the videos which were unmapped is time consuming, and that problem was also mentioned by the project manager and the video editor: *"While mapping, if the video was previously downloaded, then we cannot define which video was mapped, and which one was not mapped." (I08).* 2) the translation on Amara is harder than the one in Crowdin, and four other stakeholders also highlighted that bottleneck: *"Translating a video is much more difficult than the translating the text" (I05); "The translator makes more mistakes if they translate the videos even though they were quite good at Crowdin" (I03.3).* Subsequently, the Amara translator has mentioned that she wastes quite a lot of time while waiting for the response from other stakeholders. Four other stakeholders have mentioned similar comments about this bottleneck: *"I am just telling the same problem to the team leader, and we waste time to wait for the answer" (I07); "The translator needs to wait a lot to get the response from the coordinator, team leader, and project manager" (I03.4).* Within the interviews, Amara

proofreader has also mentioned that compared to the Crowdin, the same translators spend much more time fulfilling their tasks, and the main cause for this was explained that Amara gives 14 days to complete the translation. However, this duration is relatively long compared to our team requirements. Both content Specialist and Crowdin proofreader have mentioned similar statements: *"There was one volunteer who used to finish the tasks on Crowdin within two days generally, but on Amara, she waits the 14th day to come, then she submits the translation" (I03.1).* The video editor brought the last bottleneck, which was about editing the same video three times.

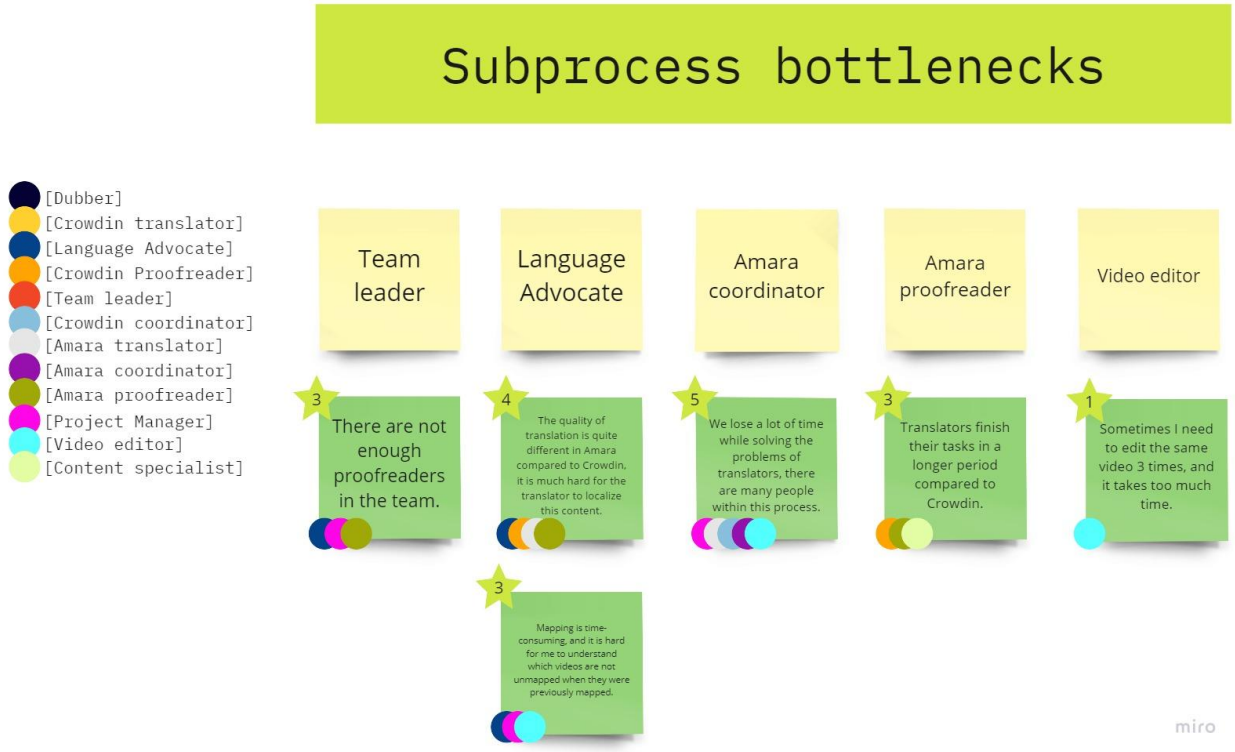


Figure 13. Interview results about the main bottlenecks that occur in subprocesses

Based on the bottlenecks mentioned above, in Research Question 2 and Research Question 3, we have analyzed more of the problems and redesigned the system.

4.2 Research question 2

In this section, we described the results of the issue register and flow chart analysis.

4.2.1 Issue register analysis

The results of the issue register analysis and the detailed version of the analysis can be found here. Based on the issue register analysis, we have defined twelve issues that were 1) translating Crowdin first, then Amara content; 2) mixing the name of tasks; 3) misusing the localization platforms; 4) low quality of the localization; 5) finding videos on Amara; 6) the damage of the Crowdin`s proposed translations on the quality of translation; 7) getting the software key; 8) being not able to track the unmapped videos; 9) spending too much time on proofreading; 10) spending additional time for checking the same document three times; 11) notifying problems; 12) not translating the texts on video screens. After defining these issues, we have also measured their potential quantitative and qualitative impact. From the perspective of qualitative impact, we found that these issues result in demotivating translators, unsatisfied users, extra work for the quality control team, the bad learning experience of learners, and churned users. Moreover, from the perspective of quantitative impact, we mainly focused on time and the number of lost potential students because, for the Khan Academy Azerbaijani team, the primary success metric is to localize the content as fast as possible so that they can drive the attention of more users. While calculating the time lost for each issue, we also noted the assumptions and the data that we used. With this, we found out that the team wastes 14137.65 minutes translating Crowdin first, then Amara content, and one another finding was that the damage of Crowdin`s proposed translations on the quality of translation caused the loss of 40 students each month and waste of 30 minutes.

4.2.2 Flow analysis (cycle time calculation)

Having done the cycle time calculation technique of the flow analysis, we have found the general time spent on the localization process and each sub-processes cycle time duration. The general localization process took 48516.45 minutes which equals 33.692 days. Moreover, the cycle time for the “Assign content tasks” was 438 minutes, and for the “Translate the content” subprocess, it took 14137.65 minutes. Subsequently, it follows with 20199.8 minutes for the “Translate Amara content”, 4380.6 minutes for the “Recreate videos”, 5783.8 minutes for the “Control quality”, and lastly, 3576.6 minutes for the “Deliver video” (Figure 14).

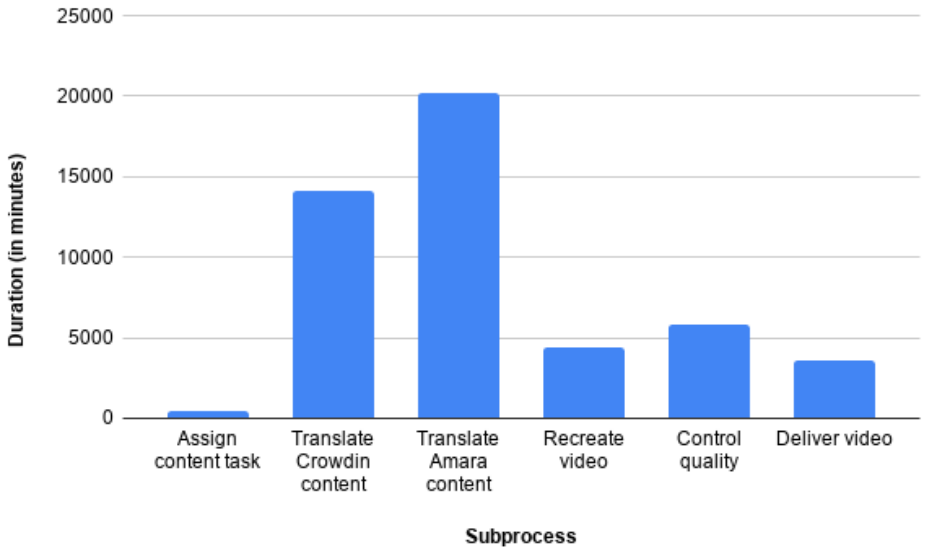


Figure 14. Cycle time duration of each sub-processes

4.3 Research question 3

Section 4.3 describes the results of Phase 2 interviews, “changes document,” and redesigning the localization process.

4.3.1 Phase 2 interviews

Phase 2 interview results are depicted in Figure 15 based on the template Figure 10. In the second phase of interviews, we collected the stakeholders` proposals about the potential changes in the localization process. One of the most mentioned points at this phase of interviews was that it should be the same proofreader on Amara and Crowdin. Six other stakeholders supported this statement by the Crowdin coordinator: *"If there is the same proofreader on both platforms, then he/ she can also use the same names, terms so that there are no different usages" (I16).* Moreover, there was another statement by the development of the simple code to eliminate the waste within the mapping process, and this proposal was supported by the language advocate as well: *"Writing a simple code can be done by any of the team members, and it would save much time" (I12).* Subsequently, there were also statements about creating a dictionary and submitting it into Crowdin, so that the proposed translations by Crowdin can be better: *"Once we tried adding the basic dictionary of the terminologies in Algebra, and it helped us to save time, maybe we can create a better dictionary, and train the Crowdin system" (I14).* Last but not least, there were also statements such as eliminating project managers, solving the problems, localizing Amara and Crowdin contents parallelly.

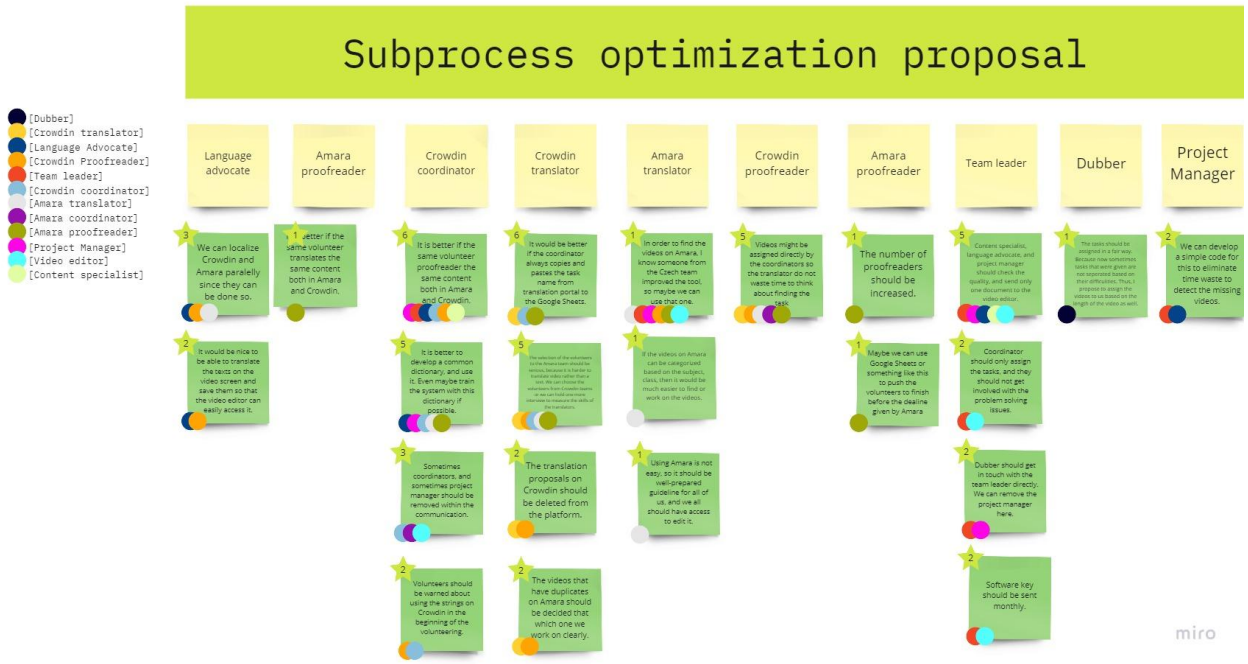


Figure 15. Interview results about the proposals for the redesigning

4.3.2 Changes document[9]

We collected the proposals from the interview, and we also selected the bottlenecks based on the issue register and flow chart analysis. Then, all these bottlenecks have been summarized in the changes document. In general, there were 16 change proposals from the stakeholders, and due to different reasons, we could not implement all of them. Figure 16 depicts the proposed changes that have been implemented in the redesign process. The accepted proposal was as follows:

- *Translating Amara and Crowdin content parallelly:* Firstly, the Crowdin team localizes the content, and then the Amara team starts the localization. However, both teams can localize the content parallelly. In this change, the language advocate proposed to perform these sub-processes parallelly.
- *Text on the video screen to be translated on Amara:* While translating the script of videos, translators ignore the part about the text on the video screen. This situation leads to a decrease in the quality of videos. In order to solve this problem, we can add the new column on the Google Sheets of the team, and the translator can add the text on the screen there. These texts are generally short (4-6 sentences), and it would not create complications on Google Sheets.

[9] https://zenodo.org/record/4737473

- *The same proofreader is responsible for proofreading the same content on Amara and Crowdin:* On the Khan Academy platform, the content on Amara and Crowdin is the same. However, the responsible proofreaders are different. Thus, sometimes proofreaders might utilize the terminology differently. It increases the workload of the quality control team. Nevertheless, if the same person is responsible for Amara and Crowdin content, then there will be the same terminology or names used in videos and exercises.
- *Training Crowdin proposals before the localization:* It is possible to train the algorithm of the Crowdin platform in a way that proposes the correct terminology. In order to do this, it will be enough to make the list of terminologies in English and their translation into the Azerbaijani language.
- *Eliminating both coordinators and project manager within all problem-solving issues:* When there is a problem, the translators get in touch with the coordinators. Nevertheless, the coordinators are only delivering the same issue to the team leader, and thus the coordinators can be removed from the process (happens in the "Translate Crowdin Content" and "Translate Amara Content" subprocesses). Similar to the previous issues, the dubber needs to communicate with the project manager to inform the translator, and hereby, the project manager can be removed from the process since he only informs the exact problem ("Recreate videos").
- *Mapping the videos through a code:* While mapping the videos on the Khan Academy Database, the language advocate can use the basic coding to map the videos, and if there are any, to detect the missing videos.
- *The entire quality control team gives one feedback document*: The language advocate, content specialist, and project manager give the three different feedback and edit proposals to the video editor, and he needs to change it three times. However, the language advocate, content specialist, and project manager can work on the same document, and send only one editing feedback to the video editor.
- *Using the video finding tool on Amara:* In order to find videos on Amara, the tool by the Khanova skola team will be used.
- *Increasing the number of proofreaders*: There are only two proofreaders for the Math content, and some of the translators have enough experience and knowledge to be a proofreader. Therefore, upgrading the positions of four well-experienced translators from translator to proofreader will help speed up the localization process.
- *Software key access is given to the video editor*: The video editor will have access to the internal communication board where the software key is updated. Furthermore, instead of the video editor getting in touch with the project manager each time to get the software key to use the video editing software, he will access it through this board if needed.

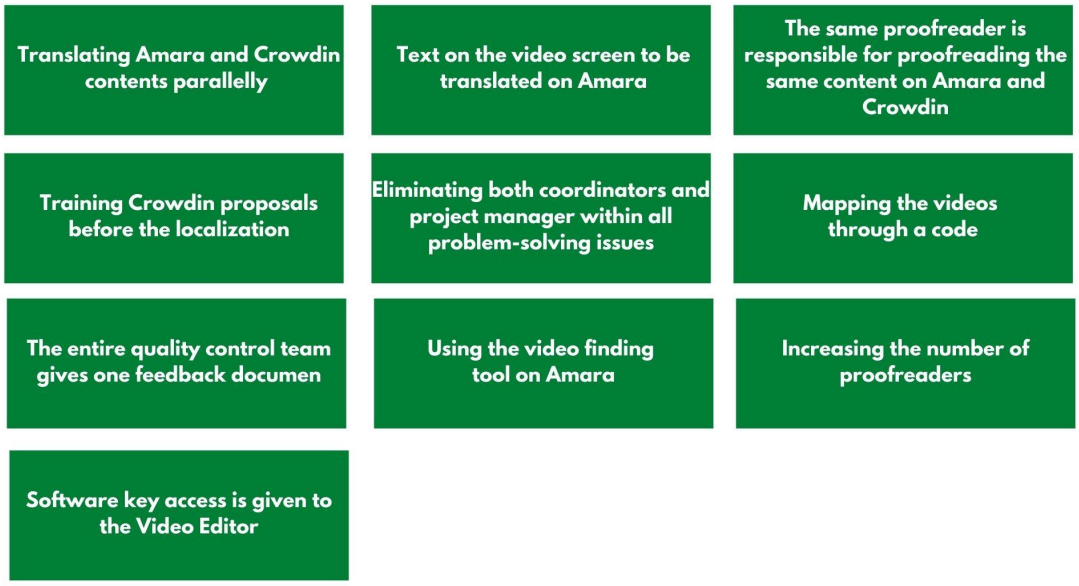


Figure 16. Accepted proposals for the redesigned process

Furthermore, there were also six proposals that we did not implement in the redesign due to different reasons such as time constraints, not being connected to the localization process, legal and administrative limitations (Figure 17). The detailed information of why we have rejected these proposals can be found in the "changes document."

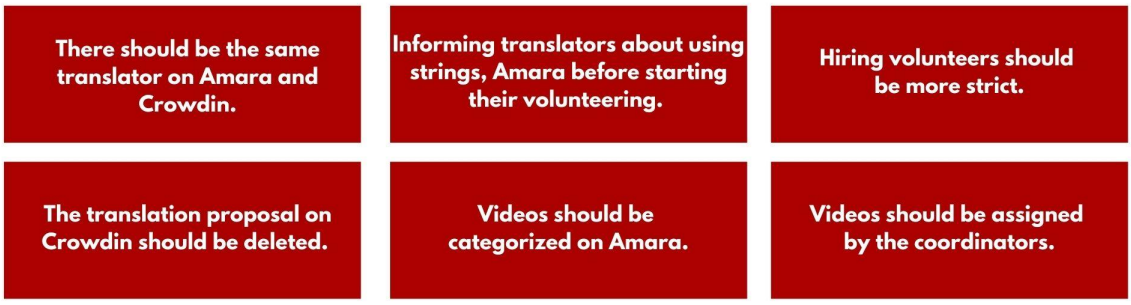


Figure 17. Rejected proposals for the redesigned process

4.3.3 Redesigning process[10]

Based on the change proposals that were accepted (Figure 16), we redesigned some parts of the "as is" process, and the changes on the localization process were as follows:

1. In the "as is" localization process, the Crowdin localization is performed previously. However, in the "to be" process, both Crowdin and Amara localization is performed

[10] Redesigned localization process: https://zenodo.org/record/4762179#.YJ56nqgzZPY

simultaneously (Figure 17).

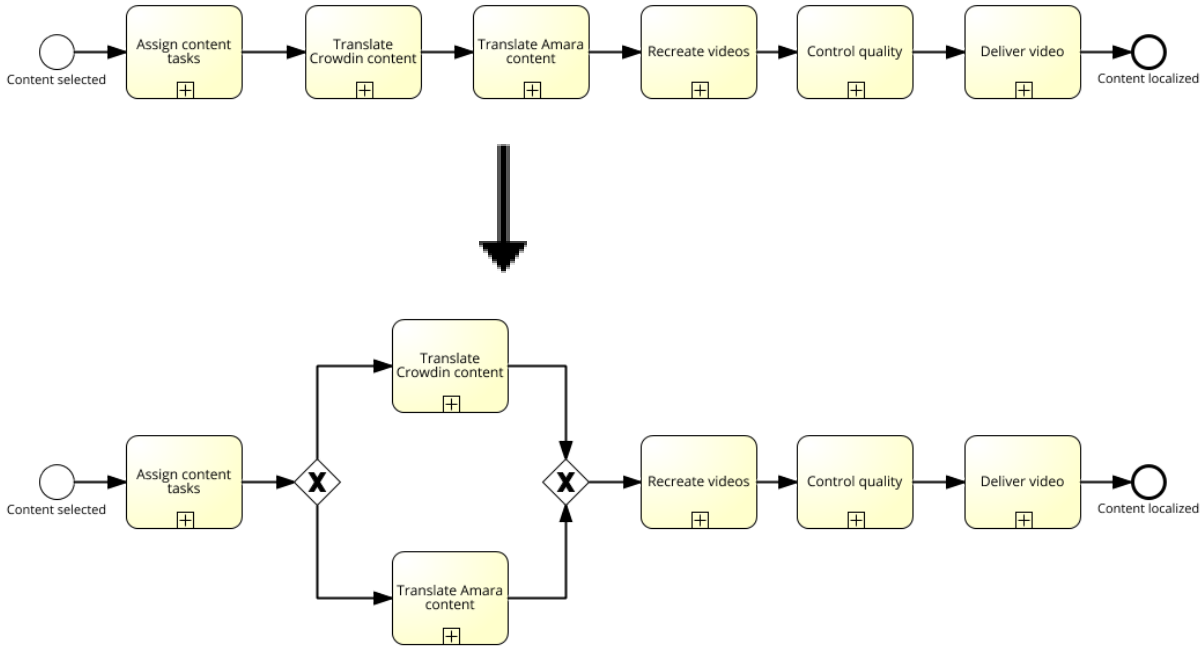


Figure 17. The difference between "as is" and redesigned processes in the localization process

2. The coordinator and project manager are removed from the tasks that were not crucial stakeholders both on Amara and Crowdin localization subprocesses (Figure 18 and Figure 19).

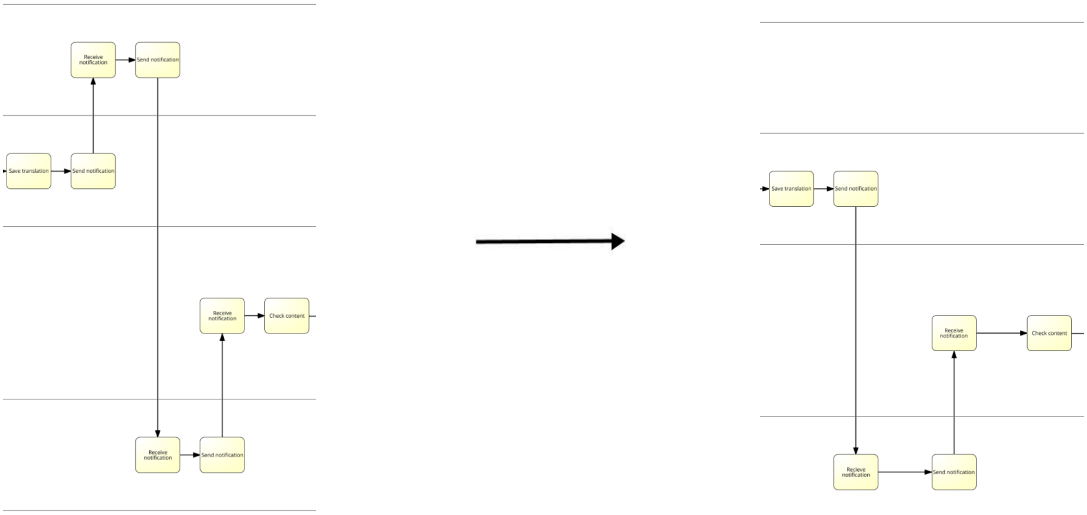


Figure 18. The difference between "as is" and redesigned processes in the "Translate Crowdin Content" subprocess.

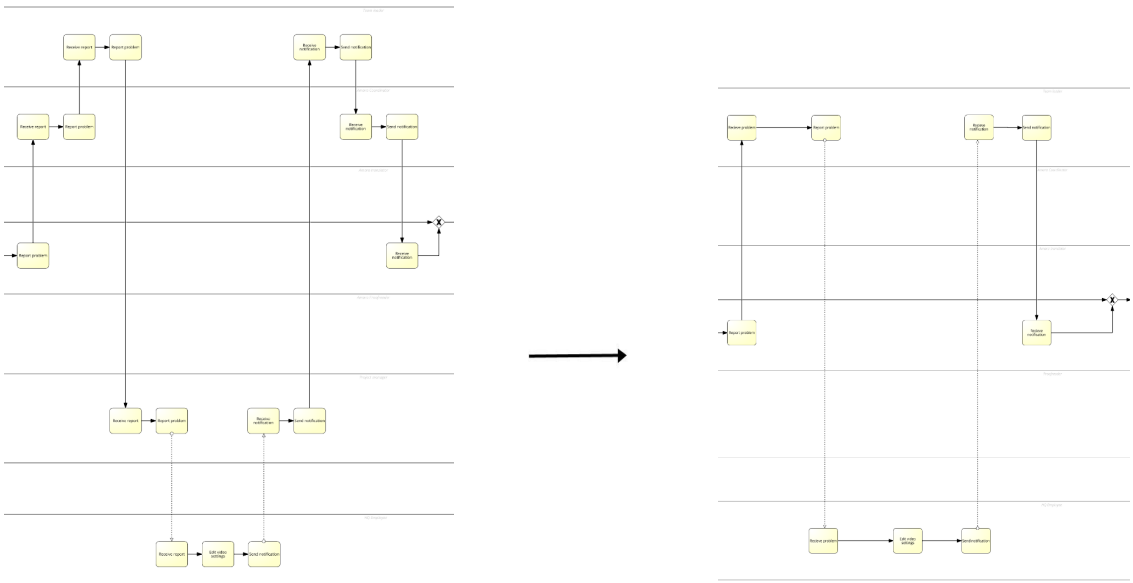
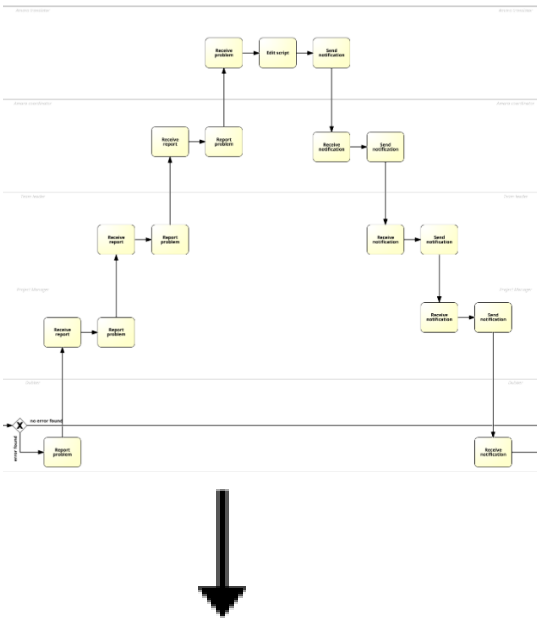


Figure 19. The difference between "as is" and redesigned processes in the "Translate Amara Content" subprocess

3. In the "Recreate videos" subprocess, there were two changes. The first one was eliminating both the project manager and Amara coordinator within the problem reporting. The second change in the process was about deleting the tasks of the video editor to get the software key because after giving him access to the platform, he will be able to get the software key without doing extra tasks (Figure 20).



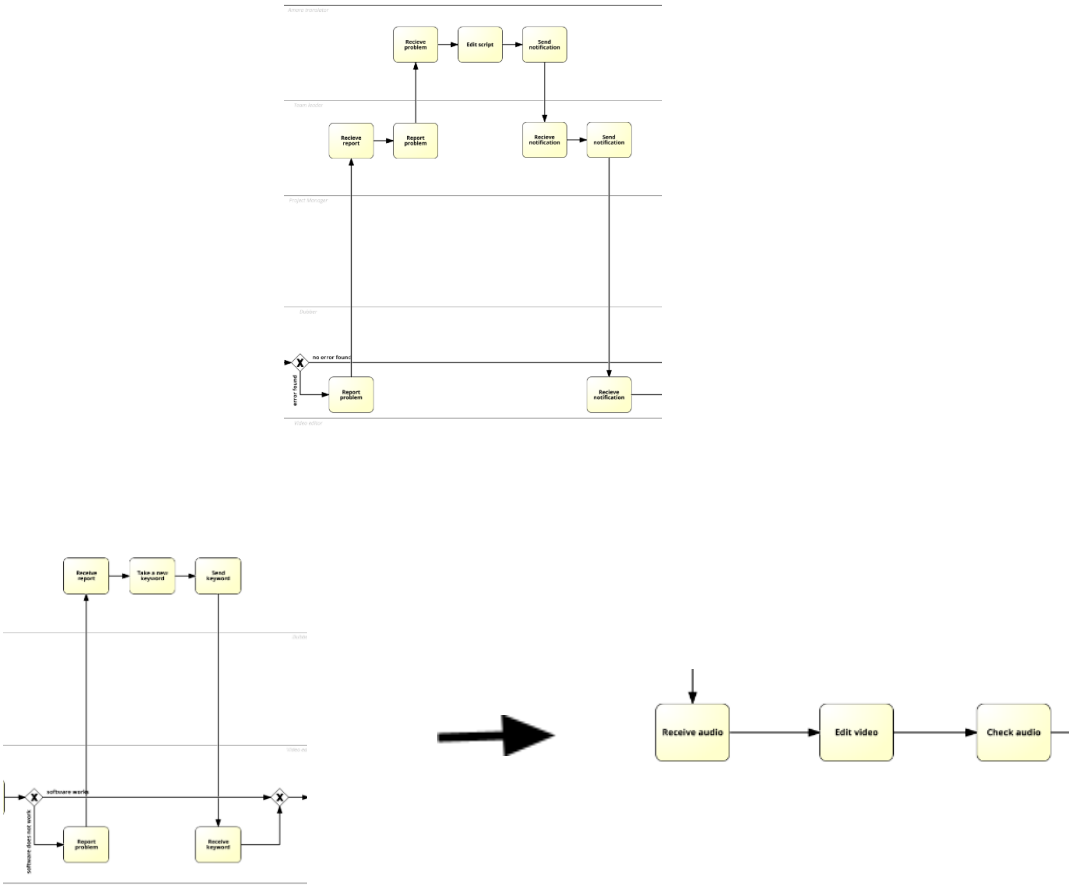
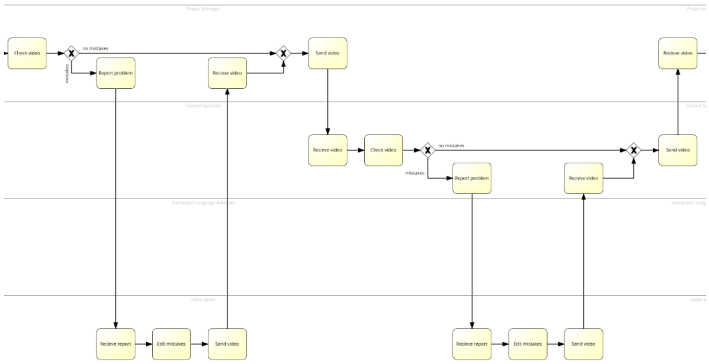


Figure 20. The difference between "as is" and redesigned processes in the "Recreate videos" subprocess

4. In the "Control quality" subprocess", instead of all three stakeholders controlling the quality of videos and video editors to edit them separately, the process has been altered in a way that all of these three stakeholders provide the feedback on the same document and the video editor implements changes only once (Figure 21).



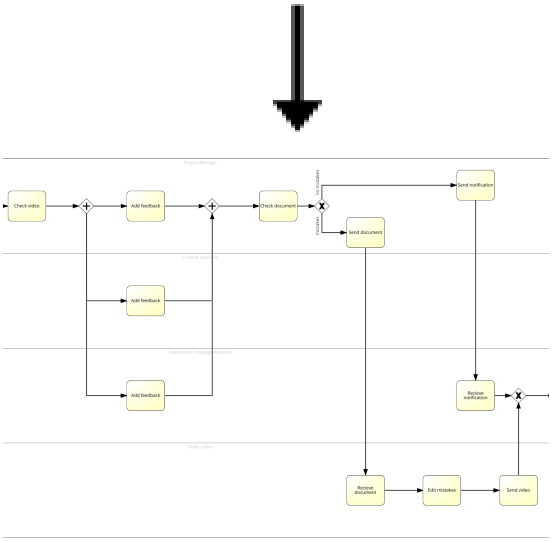


Figure 21. The difference between "as is" and redesigned processes in the "Control quality" subprocess

4.4 Research question 4

Section 4.4 describes the comparison of the "as is" and redesigned process.

4.4.1 Comparison of the previous and redesigned process

In research question 4, the previous and redesigned localization processes were compared based on the two success metrics: the number of people involved in the localization of the content and the duration spent on the localization of the content. From the human capital perspective, there was a reduction, because previously there were two proofreaders: Amara proofreader and Crowdin proofreader. Nonetheless, in the redesigned localization process, those two proofreader titles were removed, and there is only one title regarding the proofreading, which is "proofreader" (Figure 22).

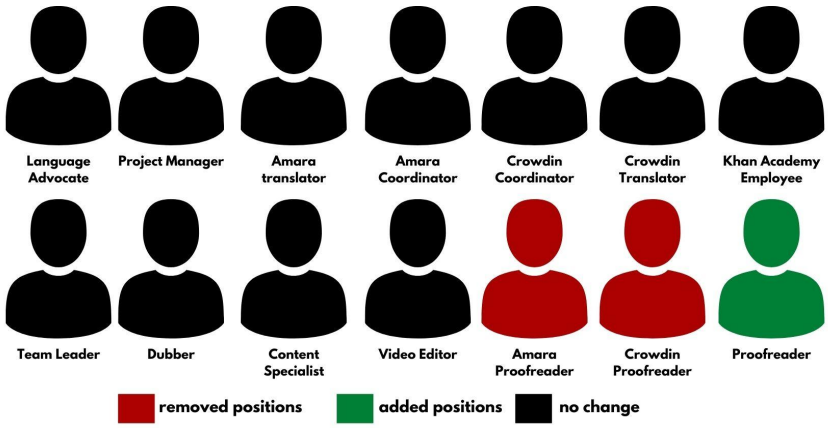


Figure 22. The stakeholders change after the redesigned process

Subsequently, we compared the total duration spent on the localization of the contents. Figure 23 depicts the comparison of the total time spent on the localization of each subprocess. Since there were no changes or improvements in the “Assign content tasks” subprocess, the duration remained stable. Nevertheless, removing some stakeholders (project manager and Amara coordinator), implementing the developments to automate the process triggered reduced cycle time in the other five subprocesses. This differences between the “as is” and the redesigned process were 11597.64, 1465.8, 749.6, 798.8, 167.4 minutes for the “Assign content tasks,” “Translate Crowdin content,” “Translate Amara content,” “Recreate videos,” “Control quality,” “Deliver video” subprocesses respectively.

Furthermore, the total cycle time for the redesigned process was 31197.2 minutes. Because in the redesigned process, “Translate Crowdin content” and “Translate Amara content” subprocesses were performed simultaneously. In this case, since the “Translate Amara content” subprocess takes longer, its cycle time was considered while calculating the general cycle time of the localization process. The cycle time of the previous localization process was 48516.45 minutes, and it signifies that there is a 17319.25 minutes difference between the previous and the redesigned localization process in general, which is roughly equal to 12 days.

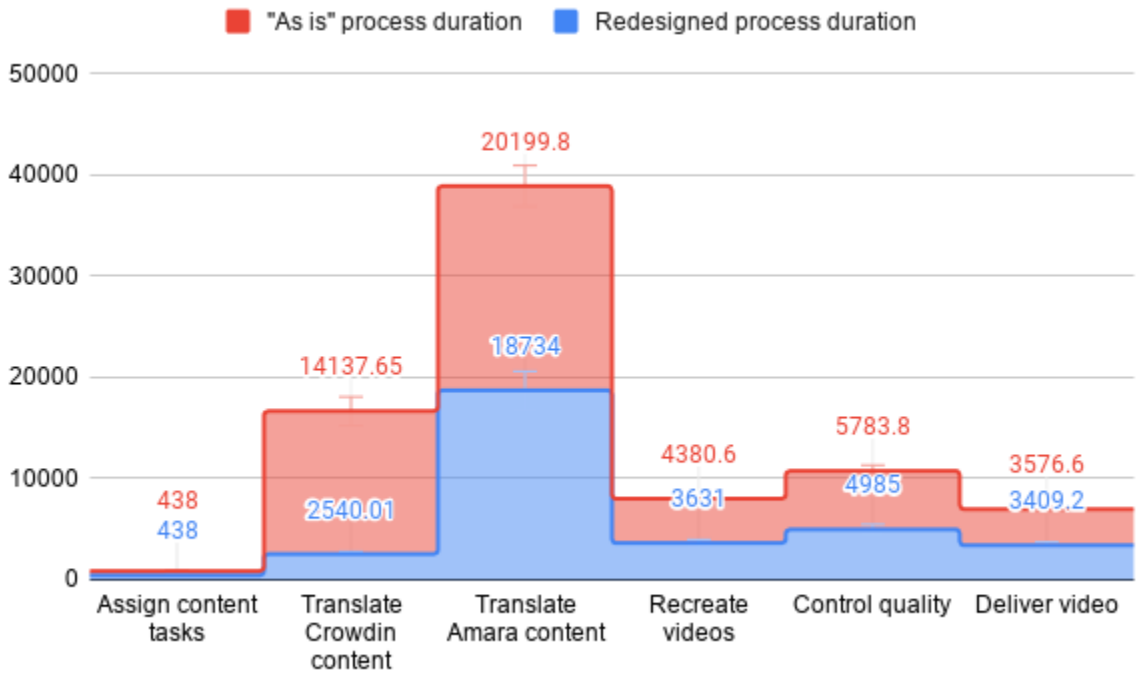


Figure 23. The cycle time of the “as is” and the redesigned process

4.4.2 Workshop 2

In Workshop 2, we have conducted a workshop with all (eleven) stakeholders to get their feedback. As a result, the stakeholders mentioned positive comments about all changes in the redesigned localization process, such as *"It helped a lot when the same proofreader checked the content both on Amara and Crowdin" (W110); "Removing the coordinators and project managers saved at least two days each time whenever we want to solve the problem, and it also impacted on the motivation of translators. Now, they do not need to wait a lot" (W13).* Moreover, the stakeholders have also mentioned potential changes for future improvements in changing roles, assigning different tasks for different stakeholders, removing some positions.

4.5 The impact of the redesigned process on the learning experience of users

Firstly, we analyzed the general overview of the responses, and we compared the first and second survey responses in Table 4.

The first survey	The second survey
The average age was 19.	The average age was 21.
73.8% of respondents identified themselves as women.	62.2% of respondents identified themselves as women.
77.4% of respondents were students.	68.5% of respondents were students.
The main responses were from Baku (71 responses) and Gandja (seven responses).	The main responses were from Baku (66 responses) and Gandja (eight responses).
Around 92.8% of respondents spend 60 minutes or less learning any topic on the platform on average.	Around 95.4% of respondents spend 60 minutes or less learning any topic on the platform on average.

Table 4: General overview of the respondents` background

Next, we asked which topic of content the users most wanted to see. It turned out that most users voted for physics, algebra, and computer science. At the moment, the main focus of all teams is precisely aimed at such topics for content so that the goals of Khan Academy Azerbaijani and the users` desires coincide (Figure 24).

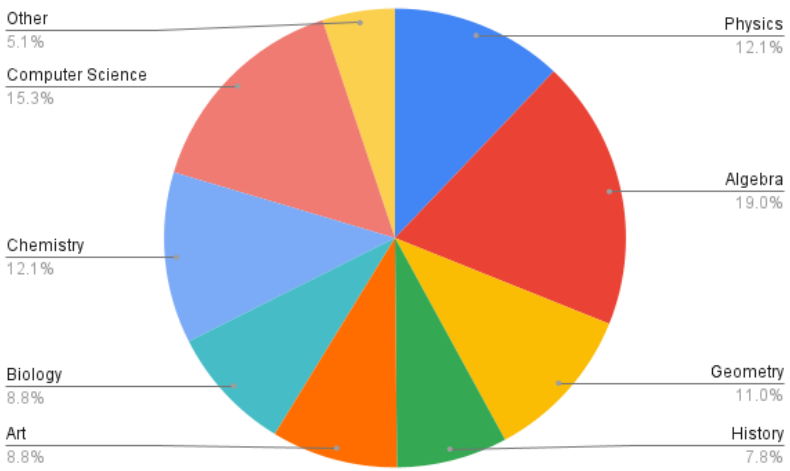


Figure 24. The contents that users most wanted to see

Next, we asked the same questions in the first and second surveys to observe the changes in the learning experience of users. We found the average of all responses to each question. In both surveys, the users selected "1" if they gave "Definitely not" feedback to the question, and opposite to this was to provide "5", which means "Absolutely yes" to the questions. Firstly, we asked to learn how often the users encounter translation errors on the platform, and we have found that the score decreased from 2.5 to 1.9 (Figure 25). It signifies that the content localized according to the redesigned process positively impacted the number of translation errors. Subsequently, there was a positive change in the adaptation of the learning materials to local culture from 3.6 to 4.1 (around 87.8%), the translations of non-textual (i.e., pictures) materials from 3.7 to 4.5 (82.2%).

Furthermore, there was no change from the perspective of the irrelevance of learning materials for Azerbaijan. It remained 2.1. This situation means that 42% of content might be irrelevant. The respondents mentioned that these are about the lack of exercises on the localized content, the need to increase the number of localized content for the higher or secondary high school students.

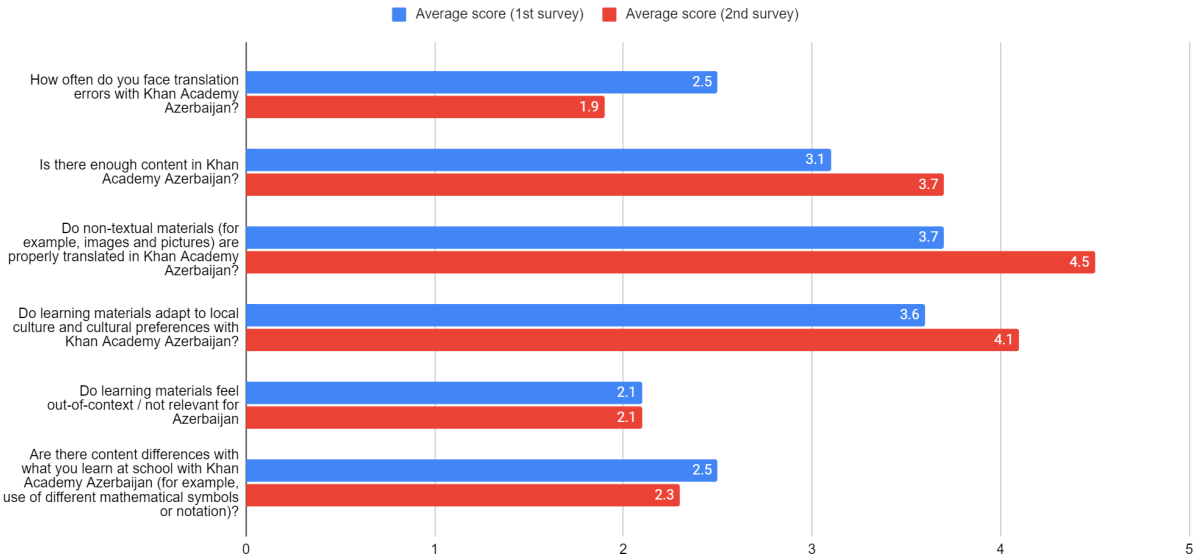


Figure 25: The comparison of the first and second survey

5 Discussion

In this section of the research, we discussed the results, limitations of the research questions defined in Section 1.2 of this thesis.

5.1 Research question 1

In the first research question, we focused on the design of the current "as is" localization process of the Khan Academy Azerbaijani team. The team already has previous experience with the BPM design. Thus, it was easy to lead them in the workshop and the interviews. At some points, the stakeholders mentioned opposite comments about the performance of the tasks. However, at last, all of them agreed with the last process design.

Moreover, the "as process" design was quite crucial because we analyzed and redesigned the new process based on the "as is" process. Within this framework, all the information collected from the interviews helped us to detect the bottlenecks. The designed "as is" process contained the exact subprocesses of [39]. In [39], there are old and new models, and the oldest one follows the same processes as it was in the "as is" process of the Azerbaijani localization.

5.2 Research question 2

In RQ2, we analyzed the process both from the qualitative and quantitative perspectives. As a result of issue register analysis (qualitative analysis) (Figure 13), we noted 12 issues, and from the perspective of quantitative analysis, we calculated the cycle duration of each subprocess,

then defined the general cycle time. Holding these two analyses helped us focus on which processes we need to focus on redesigning the process. We have also considered the value-added, waste, and root cause analysis. However, since these analyses do not provide detailed information about the impact of the issue in detail, we did not select them.

Before commencing to hold the issue register analysis, it is beneficial to conduct the stakeholder analysis [1]. Nonetheless, since we have already held the workshop with all stakeholders to understand their positions, tasks, and contact people within the process and previous coworking experience with the stakeholders within two years helped us to have a strong stakeholders` information.

Both simulation and queueing analysis were also considered. We did not choose the simulation though it is one of the most popular ones [1] because of its design characteristics for the more extensive processes and organizations. The localization of Khan Academy content into the Azerbaijani language consists of shorter processes compared to other corporations. Secondly, queueing analysis was not selected either because within the localization process the stakeholders do not encounter certain queueing, and basically, they start working on the content whenever they finish the previous tasks. Last but not least, there was also the reason that we needed to hold the analysis during a shorter period so that we can have enough time to redesign and analyze the redesigned process. Compared to queueing and simulation analysis, the flow analysis is more straightforward. Therefore, it can be conducted within a short period of time. It also gives information about cycle time which is the main factor we are looking for from this analysis.

While implementing the flow analysis, there are a few limitations addressed [1], and below we have given the answers to why the flow analysis was applicable in this research.

- The first limitation is the structure of the process. However, since the utilized process design in this research is block-structured and only used XOR and AND gateways, it was acceptable to use the flow analysis.
- The second limitation could be the estimation of the average cycle time for each task. As mentioned in [1], one way to tackle this problem is to conduct interviews or workshops. In the second interview, we have asked each stakeholder the average duration of the particular task.
- The third factor could be the behavior of the process based on load. Nevertheless, since the Khan Academy Azerbaijani team works on a volunteer basis and accepts the responsibility from the beginning, the workload does not change. Thus, alterations of the load are not a subject to consider in this research.

Hereby, we have also analyzed the interviews held, particularly for getting the stakeholders` proposals about process optimization. We could not implement the six proposals, and the main reasons for this were time, finances, human resources. However, it would be quite possible to update some subprocess if there were no such constraints. As mentioned in [40], while holding more than three process analyses may be time-consuming. However, being able to carry both quantitative and qualitative process analyses could explain different bottlenecks.

5.3 Research question 3

In RQ3, we concentrated on the redesign of the process and the direct implementation of this. We designed the process based on the process analysis and stakeholders` proposals, and all stakeholders agreed with the updated localization process. After this, stakeholders localized another content, and before starting the localization process, we already expected to decrease the duration by two days. With the integration of the content specialists and quality controllers, the redesigned process followed a similar update pathway as it was in [39]. Furthermore, there were different points between this research and [39]. For instance, in [39], it was expected that content creators could move back and ask translators to retranslate the content. Nevertheless, in the localization process that we focus on in this research, all the sub-processes happen to be like sequences. It signifies that if the translator makes a mistake, it will be corrected by either proofreader or quality controllers.

5.4 Research question 4

In RQ4, our focus was to compare the previous and redesigned localization processes. Before the implementation of the redesigned process, we already knew that the number of people involved would be decreased, and we anticipated that there would be a reduction in the general cycle duration. Nonetheless, the stakeholders expected this reduction around 5-6 days, and the actual reduction was 12 days roughly. This improvement is crucial for the team because it will allow them to produce one more content each month, and it means that there will be 34% growth in the number of localized content each year. We have also held interviews with the stakeholders, and they all agreed and gave positive feedback about the redesigned process.

5.5 Research question 5

Last but not least, in RQ5, we measured the impact of the redesigned localization process from the learners` perspectives. In these surveys, we observed the positive changes in the responses of people. The survey also showed some points that we were not searching for in the beginning. For example, it just showed how the source of problems changed after this redesigned process.

Moreover, it also gave insights for the team`s future roadmap, such as the importance of focusing on the Mathematics, Computer Science team. The research [41] emphasized the potential positive impact of the localized international content in the local language. Therefore, it is essential to ensure that the localization optimization process also brings advancements in the quality of localized content. Hereby, in this research, we have reached this goal based on the responses of the platform users.

5.6 Limitations

Within this academic research, we have defined three main points that threaten the research's validity: selecting the stakeholders, selecting subjects (Mathematics), the data collection both in the interviews and surveys, and the data analysis.

Within the interviews and workshops, we invited the same eleven people to participate. However, since each stakeholder brings various thoughts to the research, it would be more efficient to invite more stakeholders rather than limiting it to eleven. Furthermore, more participants for the same positions could help us to understand the bottlenecks better. Therefore, to mitigate this problem, we have invited the stakeholders who have at least six months of experience in the team, and each stakeholder's position was different from the other.

Furthermore, in this research, we have concentrated on the localization process of the Khan Academy Mathematics subject into the Azerbaijani language, and the implementation of this optimized localization process for another platform or the subject can be subject to bias. Hereby, it is worth mentioning that the research findings are not generalizable beyond the context of this study.

In the interview and workshops, the collected responses can be subject to bias because, in this kind of research, it is anticipated that the interviewees may answer the questions to damage the quality. Because stakeholders also responded based on their perspectives which is inevitably subjective and, as such, subject to bias. Additionally, workshops also induce bias because people react differently when they are in a group. To mitigate the bias in the workshops, we conducted the interviews along with the workshops. We also organized workshops and interviews in a semi-structured format to adjust the flow based on the responses. Moreover, we also held one interview per one (maximum three) stakeholder. It helped us to understand the answers of each stakeholder in detail. Last but not least, we would like to highlight that we did multiple rounds with the same people, which provided the chance for them to respond to our interpretation.

6 Conclusion

The localization of the educational contents eliminates the limitations for the students who want to access world-class education for free. However, the localization process generally requires almost the same amount of time spent on creating the content. Because, while localizing the educational contents, it is better to create the same video by adding locality such as names, company names. Within this framework, the importance of business process optimization (BPO) should be highlighted. Nowadays, educational organizations have also commenced utilizing BPO to accelerate their activities through keeping or improving the quality. Furthermore, localization is one of the activities that can be accelerated by implementing the BPO to increase learners' learning experience.

In this academic research, we have focused on localization process optimization and its impact on the learners` learning experience. For this, we selected the localization process of Khan Academy into the Azerbaijani language. Having conducted Workshop 1 and Phase 1 interviews, we designed the "as is" process of the localization, then we focused on analyzing the bottlenecks. In addition to the data collected in the Phase 1 interviews, we have also analyzed the process both from qualitative (issue register analysis) and quantitative (flow analysis). As a result, we have found that the main bottlenecks were solving the daily problems of translators, checking the quality of the video, and misusing the localization tools. In the next step, we collected the potential changes in the "to be" process. All the proposals are documented in the changes document in which we highlighted the approved and rejected changes. The number of approved changes was 10, and some of them were about using the video finding tool on Amara, mapping the videos through a code.

Meanwhile, we rejected six change proposals, and the reasons for this were different such as time and financial limitations. After creating the changes documents, we redesigned the localization process, and eleven stakeholders have started localizing the selected content based on the redesigned localization process. When we compared the previous and redesigned processes, we have found that by using the redesigned localization process, we have reduced the number of stakeholders (from 13 to 12), and the time duration for the localization decreased by 12 days approximately. Last but not least, by organizing two surveys, we checked the difference in the learning experience of users. Here, we observed the positive change in the users` responses about the quality of the localized non-textual (i.e., pictures) materials and the adaptation of the learning materials to local culture.

From the perspective of future work, we defined that this research can be extended in three aspects:

1. The localization process optimization research can be conducted in other subjects such as Chemistry, Computer Science. In this research, we optimized the localization process for Mathematics. However, different subjects require various processes. For instance, in the localization of the Computer Science content, it is essential to consider the localization of codes, but this is not even a topic to be considered in Mathematics. Thus, one of the future works can be done only focusing on Computer Science content localization.
2. Additionally, this research can be elaborated by focusing on the localization in other languages such as Chinese, Arabic. There are different grammar and lexical rules in each language, and it may change the localization process. For example, in some languages, such as Chinese, localization, and implementation in this language is more complicated than in other languages. Thus, as future work, the localization process of the same subject in other languages can be conducted.
3. The localization process of Khan Academy in other languages for the same subject (Mathematics) can be done. Since the team sizes and the percentages of the localized content are quite different in each national localization team of Khan Academy, the process can be different. As future work, the localization process optimization can be done for another national localization team of Khan Academy.

References

[1] Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2013). Fundamentals of Business Process Management. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-33143-5

[2] Brocke, J. vom, & Mendling, J. (2017). Business Process Management Cases – Learning from Real-World Experience. BPTrends.Com.

[3] Business Process Management: Practical Guidelines to Successful Implementations. (n.d.). Routledge & CRC Press. Retrieved 12 May 2021, from https://www.routledge.com/Business-Process-Management-Practical-Guidelines-to-Successful-Implementations/Jeston/p/book/9781138738409

[4] Business Process Change Online by Paul Harmon | Books. (n.d.). Retrieved 12 May 2021, from https://www.scribd.com/book/282652130/Business-Process-Change

[5] Brocke, J. vom, & Rosemann, M. (2015). Business Process Management. https://doi.org/10.1002/9781118785317.weom070213

[6] Niedermann, F., Radeschütz, S., & Mitschang, B. (2011). Business Process Optimization Using Formalized Optimization Patterns. BIS. https://doi.org/10.1007/978-3-642-21863-7_11

[7] Khan Academy | Free Online Courses, Lessons & Practice. (n.d.). Khan Academy. Retrieved 12 May 2021, from https://www.khanacademy.org/

[8] Official Khan Academy translations—Translation Project on Crowdin. (n.d.). Crowdin. Retrieved 12 May 2021, from https://crowdin.com/project/khanacademy

[9] Ivana, D., & Arba, R. (2014). Business Process Modeling in Higher Education Institutions. Developing a Framework for Total Quality Management at Institutional Level. Procedia Economics and Finance, 16. https://doi.org/10.1016/S2212-5671(14)00779-5

[10] Ranganathan, C., & Dhaliwal, J. (2001). A survey of business process reengineering practices in Singapore. Information & Management, 39, 125–134. https://doi.org/10.1016/S0378-7206(01)00087-8

[11] Rhee, S.-H., Cho, N., & Bae, H. (2010). Increasing the efficiency of business processes using a theory of constraints. Information Systems Frontiers, 12, 443–455. https://doi.org/10.1007/s10796-008-9145-9

[12] Stoitsev, T., Scheidl, S., Flentge, F., & Mühlhäuser, M. (2008). Architecture for End User-Driven Composition of Underspecified, Human-Centric Business Processes. 165–172.

[13] George, M. (2003). Lean Six Sigma for service: How to use lean speed and Six Sigma quality to improve services and transactions.

[14] Aalst, W. M. P., & Der, V. (2011). Process mining: Discovery, conformance and enhancement of business processes. Springer. https://doi.org/10.1007/978-3-642-19345-3

[15] Home. (n.d.). Apromore. Retrieved 12 May 2021, from https://apromore.org/

[16] The Only All-in-One Business Process Software | Signavio. (n.d.). Retrieved 12 May 2021, from https://www.signavio.com/

[17] Li, C.-S., & Irby, B. (2008). An Overview of Online Education: Attractiveness, Benefits, Challenges, Concerns and Recommendations. College Student Journal, 42(2), 449–458.

[18] Brandon, D. (2001). Localization of web content. Journal of Computing Sciences in Colleges, 17(2), 345–358.

[19] Gafarova, E. A., Belevitin, V. A., & Smyrnov, Y. N. (2017). Modeling educational process optimization in the development of individual creativity by extending modality skills of the learners. Mechanics and Advanced Technologies, 0(80). https://doi.org/10.20535/2521-1943.2017.80.113250
[20] Kontzinos, C., Markaki, O., Kokkinakos, P., Karakolis, V., Skalidakis, S., & Psarras, J. (2019). University process optimisation through smart curriculum design and blockchain-based student accreditation. 93–100. https://doi.org/10.33965/icwi2019_201913L012
[21] Khan Academy is now available in Azerbaijani! (n.d.). Retrieved 12 May 2021, from https://blog.khanacademy.org/khan-academy-is-now-available-in-azerbaijani/
[22] Stukalova, O. V. (2018). Optimization of educational process in higher education institutions: Modern trends and prospects. Turkish Online Journal of Design Art and Communication. 8(SEPT), 1016–1023. https://doi.org/10.7456/1080SSE/140
[23] 'Khan Academy Videos in Chinese: A Case Study in OER Revision' by Allen Rao, John Hilton III et al. (n.d.). Retrieved 12 May 2021, from https://scholarsarchive.byu.edu/facpub/3381/
[24] Chowdhury, M. D., Al-Mahmood, A., Bashar, M. A., & Ahmed, J. U. (2011). Localization of Digital Content for Use in Secondary Schools of Bangladesh. In Online Submission. https://eric.ed.gov/?id=ED523765
[25] Business process management—Lessons from European business | Semantic Scholar. (n.d.). Retrieved 12 May 2021, from https://www.semanticscholar.org/paper/Business-process-management-lessons-from-European-Pritchard-Armistead/a6a0384d7bf8ddad303034fe691f324734409568
[26] Koliadis, G., & Ghose, A. (2006). Relating Business Process Models to Goal-Oriented Requirements Models in KAOS. In A. Hoffmann, B. Kang, D. Richards, & S. Tsumoto (Eds.), Advances in Knowledge Acquisition and Management (Vol. 4303, pp. 25–39). Springer Berlin Heidelberg. https://doi.org/10.1007/11961239_3
[27] McKimm, J., Jollie, C., & Cantillon, P. (2003). ABC of learning and teaching: Web based learning. BMJ (Clinical Research Ed.), 326(7394), 870–873. https://doi.org/10.1136/bmj.326.7394.870
[28] Rădescu, R., & Ardelean, T. (2019, April 11). E-business learning tool for online banking based on BPM (Business Process Management) https://doi.org/10.12753/2066-026X-19-047
[29] Reijers, H. A., van Wijk, S., Mutschler, B., & Leurs, M. (2010). BPM in practice: Who is doing what? Proceedings of the 8th International Conference on Business Process Management, 45–60.
[30] Helic, D., Hrastnik, J., & Maurer, H. (n.d.). An Analysis of Application of Business Process Management Technology in E-Learning Systems. 6.
[31] Video Conferencing, Web Conferencing, Webinars, Screen Sharing—Zoom. (n.d.). Retrieved 12 May 2021, from https://zoom.us/
[32] Weston, C., Gandell, T., Beauchamp, J., McAlpine, L., Wiseman, C., & Beauchamp, C. (2001). Analyzing Interview Data: The Development and Evolution of a Coding System. Qualitative Sociology, 24, 381–400. https://doi.org/10.1023/A:1010690908200
[33] Miro | Online Whiteboard for Visual Collaboration. (n.d.). Retrieved 12 May 2021, from https://miro.com/app/dashboard/
[34] Mayring, P. (2000). Qualitative Content Analysis. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [On-Line Journal], Http://Qualitative-Research.Net/Fqs/Fqs-e/2-00inhalt-e.Htm, 1.

[35] Google Sheets: Free Online Spreadsheets for Personal Use. (n.d.). Retrieved 12 May 2021, from https://www.google.com/intl/en/sheets/about/
[36] Adesina, A., & Molloy, D. (2011). A Business Process Management based Virtual Learning Environment—Customised Learning Paths. 1, 365–368.
[37] Parrish, D., Luther, J., Shadbolt, D., Gray, K., Bennett, B., Jones, D., & Compton, S. (n.d.). MultiLingual Computing & Technology. 25.
[38] Dawson, G. S. and D. G. and S. (2015). Preparing for the Digital University: A review of the history and current state of distance, blended, and online learning. https://ictlogy.net/bibliography/reports/projects.php?idp=3138&lang=ca
[39] Vergidis, K., Tiwari, A., & Majeed, B. (2008). Business Process Analysis and Optimization: Beyond Reengineering. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 38(1), 69–82. https://doi.org/10.1109/TSMCC.2007.905812
[40] Melão, N., & Pidd, M. (2000). A conceptual framework for understanding business processes and business process modelling. Information Systems Journal, 10(2), 105–129. https://doi.org/10.1046/j.1365-2575.2000.00075.x
[41] Cheng, Y. C. (2006). New Paradigm for Re-engineering Education: Globalization, Localization and Individualization. Springer Science & Business Media.
[42] Ministry of Education of Azerbaijan. (n.d.). Retrieved 12 May 2021, from https://edu.gov.az/en/
[43] Cyr, D., & Trevor-Smith, H. (2004). Localization of Web design: An empirical comparison of German, Japanese, and United States Web site characteristics. Journal of the American Society for Information Science and Technology, 55(13), 1199–1208. https://doi.org/10.1002/asi.20075
[44] Aguilar-Savén, R. S. (2004). Business process modelling: Review and framework. International Journal of Production Economics, 90(2), 129–149. https://doi.org/10.1016/S0925-5273(03)00102-6

Appendix

Appendix I: The calculation of cycle time

In order to calculate the cycle time, we have collected the data about each task`s owner, duration, and the related document can be found here. Then, we calculated each subprocess, and the calculations for each of them were as follows;

Assign content tasks.

The duration of tasks in the "Assign content tasks" depicted in Figure 8. Since the "Assign the content tasks" subprocess is a sequential process, we implemented the formula (1), and the calculation was as follows:

$$CT_1 = 5 + 2 + 2 + 180 + 2 + 240 + 7 = 438\,(minutes)$$

Translate Crowdin content.

The duration of tasks in the "Translate Crowdin content" depicted in Figure 9. "Translate Crowdin content" process possesses XOR gateways, and thus we will implement the formula (2) for XOR-type processes, and for the sequential ones, we will use the formula (1). Hereby, it is important to mention that in case of XOR gateways calculations if there is no task as following part of the upper or lower branch, then we will multiply the probability with 0. Based on that information, the calculation will be as following:

$$\begin{aligned} CT_2 = {} & 120 + 2 + 10 + 0.1 \times (20) + 0.9 \times (2880) + 5 + 2 + 2 + 120 + 2 + \\ & + 360 + 2 + 480 + 0.1 \times (1440 + 0.05 \times (430) + 0.95(0)) + \\ & + 0.9(10080 + 0.3(650) + 0.7(0)) + 25 + 10 + 2 = 14137.65\,(minutes) \end{aligned}$$

Translate Amara content.

The duration of tasks in the "Translate Amara content" depicted in Figure 10. In order to calculate the cycle time of "Translate Amara content" subprocess, we implemented both formula (1) and formula (2), the calculation was follows:

$$\begin{aligned} CT_3 = {} & 130 + 2 + 60 + 2 + 180 + 0.35 \times (10 + 2 + 120 + 2 + 400 + 2 + \\ & + 2880 + 2 + 1440 + 2 + 2 + 2380 + 2 += 360 + 2 + 300 + 2) + \\ & + 0.65 \times (0) + 14400 + 10 + 2 + 2 + 300 + 2 + 300 + 2 + 1440 + \\ & + 0.6 \times (0) + 0.4 \times (1440) + 20 + 2 + 2 = 20199.8\,(minutes) \end{aligned}$$

Recreate videos.

The duration of tasks in the "Recreate videos" depicted in Figure 11. Calculation of the "Recreate videos" was as below, and since this subprocess possesses both XOR gateways, and sequential processes, both (1) and (2) are implemented:

$$CT_4 = 180 + 2 + 1440 + 2 + 600 + 0.9 \times (0) + 0.1 \times (2 + 2 + 1440 + 2 + 360 + \\ + 2 + 300 + 2 + 480 + 2 + 2 + 120 + 2 + 240 + 2 + 1440 + 2) + 360 + \\ + 10 + 2 + 120 + 0.2 \times (5 + 2 + 1440 + 2) + 0.8 \times (0) + 480 + 30 + \\ + 0.3 \times (2 + 2 + 360 + 10 + 2) + 0.7 \times (0) + 300 + 10 + 2 = 4380.6\,(minutes)$$

Control quality.

The duration of tasks in the "Control quality" depicted in Figure 12. Furthermore, the calculation of the "Control quality" subprocess was done by utilizing formula (1) and formula (2), because it holds both sequential processes and XOR gateways:

$$CT_5 = 1440 + 0.85 \times (0) + 0.15 \times (2 + 2 + 600 + 10 + 2) + 360 + 2 + 480 + 0.9 \times (0) + \\ + 0.1 \times (2 + 2 + 480 + 10 + 2) + 2800 + 2 + 360 + 2 + 120 + 0.95 \times (0) + \\ +0.05 \times (2 + 2 + 300 + 10 + 2) + 60 = 5638.8\,(minutes)$$

Deliver video.

The duration of tasks in the "Deliver video" depicted in Figure 13. Hereby, the cycle time calculation for the "Deliver video" subprocess was done by into consideration both formula (1) and formula (2) owing to possessing both sequential and XOR gateways, and the calculation was as follows:

$$CT_6 = 180 + 2 + 2880 + 2 + 2 + 120 + 60 + 0.55 \times (0) + 0.45 \times (10 + 2 + 300 + \\ + 0.6 \times (0) + 0.4 \times (360) + 10 + 2) + 120 = 3576.6\,(minutes)$$

Appendix II: License

Non-exclusive licence to reproduce thesis and make thesis public

I, Ayaz Karimov,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to

reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

Business process optimization in online educational platforms within the scope of localization and its impact on the learning experience of students: localization of Khan Academy content into Azerbaijani,

supervised by Irene-Angelica Chounta, Alexander Nolte.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Ayaz Karimov

13/05/2021