

DATA MINING – VIVET PROJECT

Žarko Bogićević¹; Milorad Murić²;

¹ Technical school, Užice, Serbia, zarko1993@hotmail.com

² High school of vocational studies, Užice, Serbia, muricmilorad@gmail.com

Abstract: *The main goal of this paper is to analyze the data collected from piloting the VIVET Erasmus+ project in Italy, Bulgaria and Serbia. The Virtual Internships for Vocational Education and Training project (VIVET) aims to develop and promote the model for virtual internships in VET. The VIVET project envisions a virtual platform that provides a framework for carrying out tasks, learning and communication. Students, their teachers, and mentors from companies during the internship can perform and monitor activities envisaged by the plan and the program of the internship. Apart from the segment dedicated to the tasks and activities, students communicate with their teachers and mentors through various communication channels within the platform, but also use the lessons and materials from the platform. The platform has been acknowledged by the Ministry of Education in Serbia and will be included in further training of teachers and mentors in schools across Serbia. This platform is entirely new and in order to provide the best possible experience for its users and to optimise it for large scale use, an analysis of the data collected during piloting is critical.*

Keywords: VET, education, VIVET, virtual internship

1. INTRODUCTION

International internships are gaining more importance in the context of internationalization of higher education and globalization of our professional world. Traditional international internships or even internships in different cities within the same country, where the learner has to travel are not always feasible because of financial, geographical, social or other reasons (e.g. disability, social exclusion, legal constraints, etc.). There are plenty of difficulties to overcome, starting from those related to the lack of communication between the learner and the company and the institution of higher education. In order to solve this, virtual internships (mobility) is defined as “a set of ICT supported activities that realize or facilitate international, collaborative experiences in a context of teaching and/or learning”[1]. This can be used to support physical internships (blended) or to realize the entire internship online (virtual). One of the main channels that can be useful for conducting an internship experience through ICT is an on-line platform that can provide the necessary information, material and documents to put in place a virtual internship. This platform would be used to monitor the virtual internship step by step from its plan to evaluation.

Many educators, employers and students are not familiar with this form of internship or they do not recognize the value or learning benefits of such internships. These sorts of virtual internships are often considered as less valuable than actual internships that require presence on site of an organization. This has built a negative perception of virtual internships in general. In some fields, virtual internships are the best and most affordable solutions for people who want to gain experience (e.g., marketing, branding, software design, programming, website design, graphics design etc.). However, due to the lack of familiarity and experience of many educators these sorts of projects get poorly executed or they do not get executed at all. This creates opportunities for small to medium size companies to include virtual interns to develop their customer databases, parts of their website (or a whole website), create marketing materials and so on.

However, it is necessary to point out that fully virtual internships are not feasible in all study fields. For example, most manual professions (carpentry, health services, metallurgy etc.). Due to the way these internships work it is not possible to achieve the required learning goals of the internship through virtual activities. There is also the challenge of accreditation of the virtual experience in the country that it has taken place together with the possible extra work load caused by the required detailed preparation of a full virtual internship.

Teaching staff may be resistant to the use of ICT, which is why planning a virtual internship needs to take into account the setting up of workshops, or similar training activities, to familiarize everyone with the tools and the platform that will be used for the virtual internships.

Recognition of virtual internships in a specific country could be problematic due to the diversity of common standards of recognition in the educational department. There is also the issue of certain documents/agreements that are put in place for minors etc.

In order for an international virtual internship to be successful it will be necessary to have a intermediate knowledge of, at least the English language, if not the local one by all the actors involved in the virtual internship.

2. VIRTUAL INTERNSHIPS AND THE STATE OF VET

According to the EU-VIP[3] project partnership, virtual internships can be described as “a set of ICT (Information and Communication Technology) supported activities that realize or facilitate international, collaborative experiences in a context of teaching and/or learning”.

It comes out clearly that ICT related skills can be identified as the central core of a virtual internship. They are required for any type of virtual internship and students interested should at least possess an intermediate level of knowledge of ICT related skills such as basic knowledge of computers, productivity software, electronic communication skills, internet skills, etc.

From this, two types of virtual internships can be identified:

- a) Fully virtual internships that are completely done in an on-line environment, starting from the preparation phase all the way until the final evaluation and assessment of the virtual internship.
- b) Blended virtual internships where on-line virtual experiences are implemented with occasional or temporary visits to the company which provides the internship.

The European Credit System for Vocational Education and Training - ECVET[4] structures the learning outcomes into three separate levels and defines them as follows:

- Knowledge – means the outcome of the assimilation of information through learning. Knowledge is the body of fact, principles, theories and practices that is related to a field of study or work.
- Skills – means the ability to apply knowledge and know-how to complete tasks and solve problems.
- Competence – means the proven ability to use knowledge, skills and personal, social and methodological abilities in study or work situations and in professional and personal development.

In education, a learning outcome is a specification of what a student should learn as the result of a period of specified and supported study. The learning activity could be a course, a module or an entire programme. Learning outcomes can help guide students in their learning by explaining what is expected of them. They also help teaching staff focus on exactly what they want students to achieve in terms of knowledge, understanding and skills which helps defining the assessment criteria more effectively. This also provides a useful guide to potential employers about the general knowledge, skills and understanding that a graduate will possess.

The virtual internship envisions three roles- the VET school, the Company and the VET student. When a School decides they wish to set up a virtual internship, especially in the case regarding Mentors from the Company and VET Students, it is necessary to understand why they should choose this framework. Some of the main advantages are:

- Confident atmosphere;
- Flexible conduction of supportive sessions;
- No geographical or time constraints;
- Cost efficient – no transport costs;
- No environmental barriers;
- No third parties can disrupt the internship;
- People with disabilities may work from the comfort of their homes and do not need personal assistance in most cases;
- New ICT skills acquired by all parties involved.

The introduction of virtually supported coaching, mentoring and peer feedback or group work that the virtual exchanges provide can contribute to an overall enhancement of the learning experience being offered. So, before starting a virtual internship there are some steps that are suggested. The staff should get plenty of time to practice with the chosen ICT infrastructure in order to build up competence before beginning the virtual internship. There should be a few samples and examples so that the staff can know what to expect. Providing a well serviced help desk well before the virtual internship is critical, this means that all the initial problems will be ironed out immediately.

In the school year of 2015/2016, 189 084 students were enrolled in upper secondary VET programmes. Approximately 80% of the VET students follow a four-year programme. VET students represented nearly 75% of the total enrolment in upper secondary education. VET programmes are offered in 327 upper-secondary schools, 74 programmes with a duration of three years and 112 programmes with a duration of four years [2].

3. THE PROJECT PLATFORM

The platform for Virtual internship should be an online platform that will support all of the virtual internship needs in terms of communication and activity support. The platform is playing a supportive role as it needs to facilitate the kind of activity in which companies, schools and students are involved. It should be an on-line infrastructure that replaces traditional face to face activities by those supported by Information and Communication Technologies. This means that face to face meetings are replaced by synchronous video and audio conferencing and asynchronous, one on one or one to many, electronic messaging.

There are some essential points that need to be taken into account:

- **Accessibility** – Choosing a technology that is accessible to all parties involved. Some companies use their own systems and platforms, such as Office 365, and they might be willing to open those up to users from outside of the company, in this case, students and teachers.
- **Usability** – Choose a technology that is user friendly and reliable. Also, it is recommended that all stakeholders have a stable internet connection.
- **E-literacy** – Preferably work with tools already known to the actors involved. If this is not the case, making sure all the actors learn to use these tools, provide training if possible.
- **Test and support** – Testing the tools beforehand and making sure that everyone involved who to turn to in case of technical difficulties.

Synchronous tools allow communication between users at the same time while asynchronous tools allow for communication independent of time which tends to be more effective, in some cases, as it gives time to formulate opinions and other information in a structural way. These tools create a feeling of decreased distance between the users.

Here are some steps on how a virtual internship can be put in place. The virtual internship can be split into three phases – before, during and after phase.

1. **Planning the Virtual Internship** – The VET school and the Company create a project plan that includes activities, duration, aims, goals, deadlines, required knowledge etc. Both parties choose their representatives (a teacher and an employee) that will register together on the platform. They check the materials and fill out the content on the platform.
2. **Recruitment and matching** – The VET school and Company advertise the virtual internship and gather candidates. Then they choose the students that respond to the requirements. The Company checks if the student’s profile matches their needs. Once everything is settled, the company gets in touch with the student.
3. **Introducing the VET students to the platform** – The first meeting should be a face to face meeting between the VET School Teacher, Mentor from the Company and the student(s). This meeting could be held using the platform itself or any other form of video or audio communication. The student(s) are introduced to all of the activities expected of them.
4. **The Agreements** – This step should cover all of the documents required for the virtual internship such as start and end date, legal agreements, insurance, data protection (GDPR, especially if minors are involved).
5. **The virtual internship** – All of the activities are performed during this step, the tasks and deadlines are assigned, monitoring of the development is done in this step as well. The mentor and teacher assist the student(s) if they face specific issues and difficulties.
6. **Conclusion** – The student(s) provide personal reports about their virtual internships experience and the teacher with the student(s) checks the tasks carried out and the knowledge acquired.

E- learning platforms such as Moodle and similar, typically facilitate communication between the students and the education center.

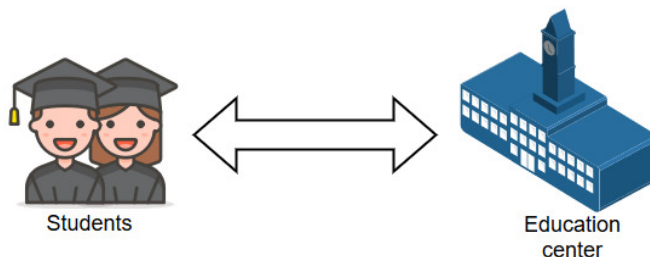


Figure 1: Typical e-learning experience

Virtual internships envision a third role in this already familiar scenario – The Company. With this third role included, the platform must facilitate all of the activities envisioned within the virtual internship.

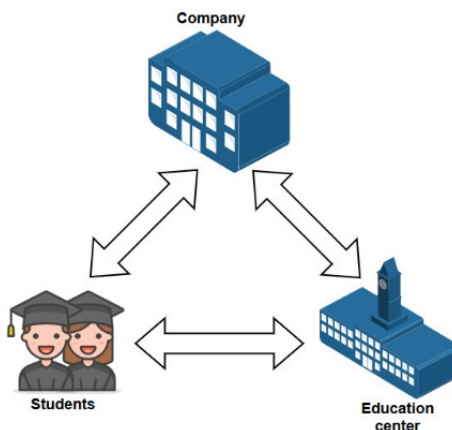


Figure 2: Virtual internship roles

4. PLATFORM REQUIREMENTS

Considering that the platform is based off of an open source platform – Chamilo [4], the requirements for running the core of the platform can be checked on their official requirements documentation located on their website.

When working with a web based platform, two factors come into play [5]:

- 1) The server side requirements
- 2) The client side requirements

Both of these requirements need to be met in order for the platform to function properly, so it is highly recommended to have, at the very least, the recommended level of requirements, as the minimum requirements cannot facilitate a virtual internship due to the inability to load all of the platform modules, however, the core functionality is still operational even on the minimum requirements.

Here are the recommended requirements to use the platform on the client side:

- Pentium II CPU
- 1GB RAM
- 50MB of HDD space for caching
- An OS capable of supporting modern Web standards
- A modern Web browser such as Chrome, Firefox etc.

The server side requirements are variable, depending on how many users the platform should facilitate, as with the increased number so do the requirements increase. The recommended requirements are:

- A single dedicated CPU core
- 256MB of RAM (apart from the server OS requirements)
- 250MB of disk space (just for the core of the platform)
- A network connection of at least 10mbps
- Windows Server OS or Linux server OS

These requirements are rather basic and are sourced from the official Chamilo[4] website. While the requirements seem basic for a Windows based computer, they are reasonable for a Linux based machine. However, for fluent use, a modern processor is recommended, as modern browsers have a lot of features and require more memory and processing power. For example, a low end Intel Celeron processor, found in the cheapest laptops would be more than enough for all the platform features. As features are added to the platform, so do the requirements on both the server and client side increase.

As with every web based service, the amount of dedicated resources depends entirely on the amount of users and active modules. Due to the nature of the virtual internship, a single school or institution should be fine with roughly 10GB of HDD space, with monitoring and occasional checking on the memory usage. For example, a school that creates virtual internships for graphics companies should be prepared to extend the space available, as the file types, typically associated with graphics software, tend to take up more memory space than files used commonly, such as word, PowerPoint, excel etc.

The recommended RAM memory per active user on the platform should be around 20MB of RAM. This means that 10 simultaneous users would consume 200MB of RAM, however, keep in mind that as a user issues a page load and it completes, the connection is closed and the amount of memory can be given to another user and so on.

The platform for the VIVET project is hosted on a Linux based server. The server is required to have PHP 7 or newer, Apache, MySQL and additional software depending on the modules installed. There is also support for Apache OpenMeetings [6]. This option allows for simultaneous video and audio conferencing, including file sharing, whiteboard, instant messaging, file explorer, calendar, screen sharing and so on. In order to integrate this option a server with 4 to 8 cores and 8GB of RAM, this increases the requirements and server costs by a huge margin. In order to have

these features, but still remain cost efficient, options such as Skype, Viber and Whatsapp have been added to the platform as well.

5. METRICS AND ANALYSIS

Based of of the collected metrics during the piloting phase of the project[7] some improvements can be made on the server side of things in order to create a better virtual internship experience. There are several important parameters to look out for.

The location of the server for the virtual internships will be key to how fast and how responsive the platform feels. For example, if the virtual internship platform is hosted for a single school/university, the server could be located at that school/university which would have the best response times and best ping possible. If it is nation wide, the server should be hosted on the main network node, for example, in Serbia the main network node is located in Belgrade. If the server is EU based, then the optimal response time from each country would be if the server is located in either Germany or Luxembourg.

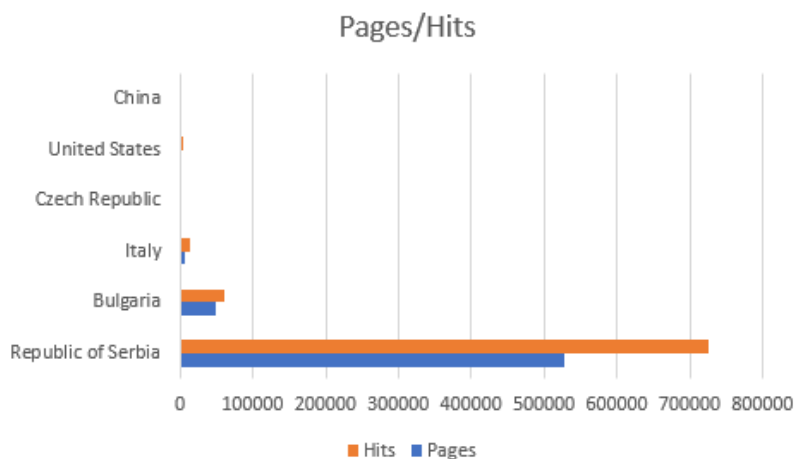


Figure 3: Pages/Hits per country

As seen from the previous figure, most of the traffic is coming from the Balkan area so the server can be located in that region in order to accommodate minimum response times.

Since the project started the platform has been set up for internal use first, with the version for hosting purposes has been set up around the end of July and has been internally tested and modified until internal tests were complete. After that, mentors from the companies were brought in to review the platform for further modifications, this was done in early September. After these reviews final modifications were made before the piloting process started. There was another meeting held with training for the mentors and part of the school staff so that they would familiarize themselves with the platform before the piloting starts and the students register.

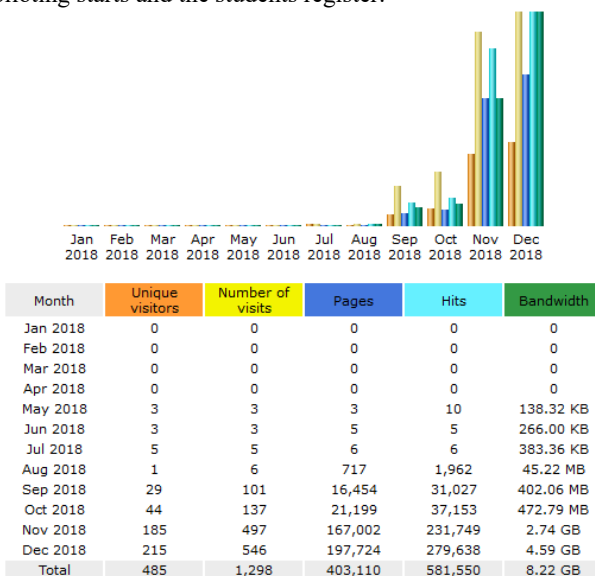
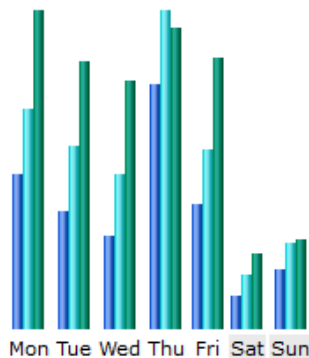


Figure 4: Overview of active months, 2018

Analyzing which days of the week are the ones with the most traffic would be the next step. This analysis allows for a deeper understanding of when the platform is most used and when the ideal day for maintenance or upgrades would be.



Day	Pages	Hits	Bandwidth
Mon	1,441	2,060	32.65 MB
Tue	1,091	1,701	27.54 MB
Wed	868	1,440	25.52 MB
Thu	2,289	2,971	30.92 MB
Fri	1,168	1,672	27.77 MB
Sat	307	496	7.68 MB
Sun	557	801	9.18 MB

Figure 5: Platform usage days of the week, 2018

From the previous figure shown it is obvious that the days with the least amount of traffic are Saturday and Sunday. However, small scale maintenance could be done in least active hours, if it is something that requires very little or no down time at all, it would be ideal to implement those changes without waiting for the least active days.

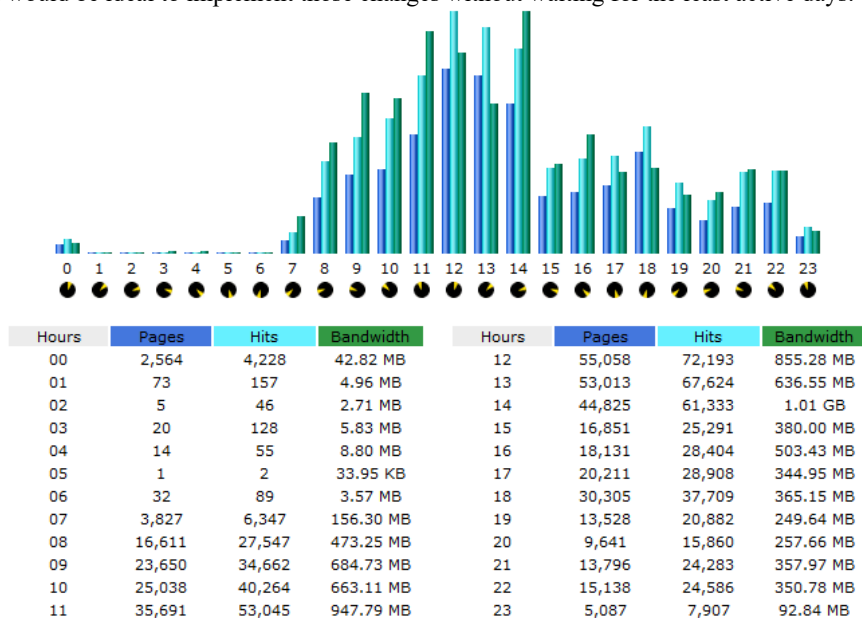


Figure 6: Platform usage during hours of the day, average for 2018

This indicates that regular maintenance, with very little to no down time at all, could be done in late night hours (after 11PM) or early morning hours (before 8AM).

Analyzing the visit duration allows for monitoring of the website content. Generally, short visits are not favorable for any website, especially not for a virtual internship website.

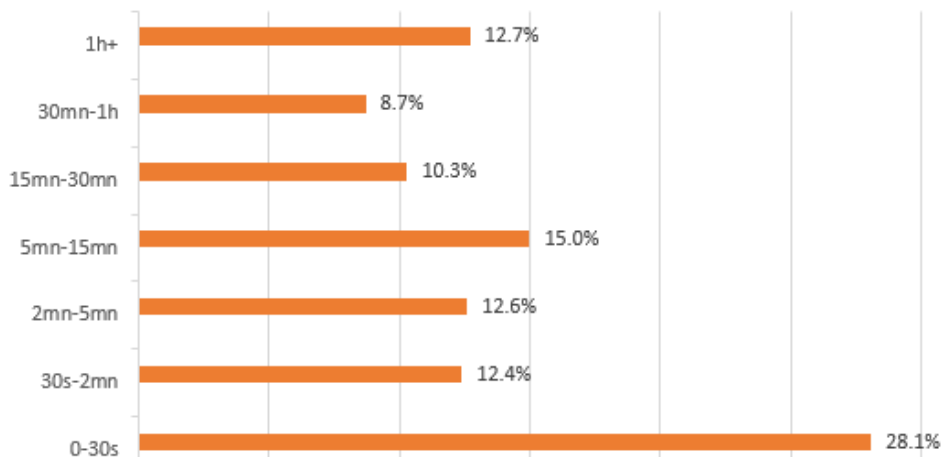


Figure 7: Visit duration

Looking at the visit duration chart, 28.1% of the visits were not productive visits, this accounts for visits by users that spent very little time on the platform and left quickly and visits by bots. Visits that range from 30 seconds to 15 minutes account for 40% of platform visits. These visits are considered productive as this is more than enough time for a student to check the latest notifications, download the required materials, post light files to the platform, chat with users or respond to personal messages or forum discussions. Visits ranging from 15 minutes to more than one hour account for 46.7% of total visits. These visits are considered very productive as they allow for much larger file uploads, detailed work and much more detailed interaction with the platform itself.

The project platform contains all different kinds of file types. Almost all types can be hosted on the platform, however, the file types that dominate are the ones that the platform uses for its purpose. Reviewing these types, it is obvious that the most common ones are web related. The most common file type is php with a 68.2% of bandwidth usage and a definitive highest number of hits. The next file type is png and jpeg, image files, these files consume 18% of the bandwidth. And coming in third it's javascript with a 8.4% bandwidth usage. In order to improve the server response, efficient coding in php and further php.ini file modifications should be made, as it is the dominant file type. Allocating a larger amount of memory and longer load times ensures the scripts execute completely before closing the connection. Modifying image files and properly saving them for web use will shorten load times, even though they are 18% of the bandwidth, they account for most of the slowdowns during load times.

Analyzing the operating systems used will allow for further optimizations depending on which OS is used the most. This also shows how and from which devices virtual internships are used the most. Statistics show that roughly 50% of all traffic on the internet comes from mobile devices [8]. When virtual internships are involved, that is not the case. Windows accounts for 93.2% of traffic, which makes sense, as all school and university computers run on Windows. Linux accounts for 5.5%, out of which 5.2% come from Android devices. iOS holds 0.9% and Macintosh 0.1% of total visits. This shows that virtual internships are almost completely run on full desktop machines that are using Windows, while only a small percentage of people access virtual internships through their mobile devices, in most cases check only their notifications and writing messages. The platform is fully responsive and adapts to any device screen size, so the data was not influenced by the platform being unoptimized for mobile devices. It is also noteworthy that under the Windows operating system, Windows 10 was the most dominant with a 75.2% share, followed by Windows 7 at 14.6%, Windows Xp at 1.7% and Windows 8 and 8.1 at 1.4%.

Optimizing the platform for browsers is crucial in order to maintain the lowest load times and maximum compatibility across the browsers that are being used.

Table 1: Browser share in user access

Browser	Google Chrome	Microsoft Edge	Mozilla Firefox	Opera	Apple Safari
Share	66.1%	19.3%	10.8%	2.5%	0.9%

6. CONCLUSION

The potential of virtual internships available for students enrolled in a VET environment has not been fully explored. This is a great opportunity for people with disabilities and those with geographical, financial, social or other difficulties, due to which they cannot undertake a complete national or international internship or simply cannot attend for a long period of time. A virtual environment is a place where the distance can be shortened and if done correctly it could provide a natural feeling of cooperation. The students themselves will get an enhanced experience which could improve their CV's.

The blended model could prove to be the most used one as it allows for a partial virtual experience and will save costs and time when it comes to the school and the companies. For example, a metallurgy student could complete an entire “safety at work” internship before coming to the company, this ensures that all students are aware of the dangers in said company, it saves time for company employees teaching them face to face about something crucial before coming to the company itself. Another example would be graphics design students who could complete their projects on-line using the virtual internship platform and at the end of their projects they would go to the printing company and see their projects come to life.

The Virtual Internship can be successful if the school and the company mentors guarantee an appropriate level of tutoring for the students, the ICT structure is familiar to everyone, the content is relevant and peer communication and exchange is guaranteed.

Based on the previously collected data We may see that platform use varies from targeted groups of users. The institutions in Serbia are piloting with VET students in VET schools which shows that this is the largest user group, followed by people with disabilities piloted in Bulgaria and university students piloted in Italy.

The main goal of this paper is to analyze and further optimize the platform, however, only limited optimization may be done from a technical point of view based on the data collected. Real world usage still depends on the feedback which will be provided using questionnaires given to teachers, company mentors and VET students at the end of the school year. Based on their feedback, further improvements and features will be added for the next school year.

Until the school year ends, the user activity will be monitored in great detail, and based on the user’s behavior and feedback provided, the platform will be developed into its next version.

REFERENCES

- [1] <https://provipproject.files.wordpress.com/2013/10/edulearn.pdf> (25.03.2019)
- [2] <https://ec.europa.eu/programmes/erasmus-plus/projects/eplu-project-details/#project/504502-LLP-1-2009-1-BE-ERASMUS-ECUE> (25.03.2019)
- [3] http://www.cedefop.europa.eu/files/5556_en.pdf (25.03.2019)
- [4] <https://chamilo.org/en/> (25.03.2019)
- [5] https://11.chamilo.org/documentation/installation_guide.html#1.1.Pre-requisites (25.03.2019)
- [6] <https://openmeetings.apache.org/> (25.03.2019)
- [7] <http://www.platform.vivet-project.eu/> (25.03.2019)
- [8] <https://www.statista.com/statistics/277125/share-of-website-traffic-coming-from-mobile-devices/> (25.03.2019)