Provision of Academic Data for Research: A Step for Academic Success

ABSTRACT

One of the most widely researched issue on higher education relates to exposed paths that lead to academic success. Nowadays information systems represent an essential part of the education sector in many universities. In particular, the increasing of the number of students in higher education in Portugal leads to the progressive increase of student's data contained in institutional information systems. This article aims to include the use of student data contained in the Fénix information system of a Portuguese higher educational institution, turning their information in an internal tool, to assist analysis and studies of internal academic success. Providing relevant data for queries can offer a clearer explanation of the paths to success in higher education. It can also help the identification of problems and failures aiming to support a more effective intervention of institutions and researchers to improve the academic success.

Keywords: Higher education, Academic Success, Information Systems, Data Presentation, Web Application

1. INTRODUCTION

Studies concerned with finding solutions that can increase students' success represent an increasingly important subject on education, especially for countries committed to improving the efficiency of their own higher education institutions (York, Gibson & Rankin, 2015; Kuh, Kinzie, Buckley, Bridges & Hayek, 2006; Kolster & Kaiser, 2015). It is the case of Portugal, where the participation of students in higher education institutions has been increasing progressively and therefore, the graduate opportunities have been multiplied (Martins, Carvalho, Ávila & Costa, 2017).

Improving student success is a subject that remains a concern of governments, education policymakers, institutional leaders (rectors and principals), practitioners and researchers (Kahu & Nelson, 2018). A vast amount of studies is concerned with the global valuation movement and increasing pressures for institutions to highlight student's academic success (York et al., 2015). Particularly in the field of learning and assessment, information technology (IT) could be a key tool to help researchers find ways to establish strategies and policies that promote the academic success through academic data analyses, in order to deal with the pressures of responsiveness in higher education institutions (Campbell, DeBlois & Oblinger, 2007).

Although there are international and national statistical platforms that provide data for study, currently in Portugal these tools only provide access to aggregations of cases provided from the combination of datasets, known as macrodata (Hand, 1992), that cannot allow a user to analyse internal academic registries. The existence of an academic research domain that is interested in carrying out studies that reveal responsible factors for student success and failure, supports and validates the creation of an internal

Portuguese platform that provides academic information from microdata, the individual information of users such as score or age (Hand, 1992).

2. BACKGROUND

A few years ago, data did not occupy a fundamental role in the analysis and decision-making for high educational institutions (Earl, 2005). Not only were there little data related to educational institutions, the ones that existed were difficult to manage, change, and enable meaningful cohorts between datasets (Matters, 2006). Currently, due to the technological advances related to collecting, sharing and representing massive amounts of information, the extensive use of data to inform improvement plan policies and for decision making is now conceivable in several areas such as healthcare (Frati, 2019), education (Gauld & Whatley, 2017; Ling, 2019), security (Krishnamurthy & Desouza, 2014), among others (Earl, 2005; Matters, 2006). According to Earl (2005), "data provide tools for the investigation necessary to plan appropriate and focused improvement strategies. Synthesizing and organizing data in different ways stimulates reflection and conjecture about the nature of the problem under consideration. Over time, this process gives rise to defensible plans for changes" (p. 7). Directing the focus to the Education sector, a diversity of data can be captured linked to the learning context (Matters, 2006).

Information systems have now become an essential part of the education sector in several universities (Campbell et al., 2007; Delavari, 2008; Gauld & Whatley, 2017). This factor leads to a rise in the capability to assemble, offer, and represent a massive quantity of data about the whole constitution of the institutional body and its interactions with educational systems. Hence, the responsibility to manage, maintain and integrate large volumes of data has significantly increased for academic institutions (Lotsari, Verykios & Panagiotakopoulos, 2014; Daud, Aljohani, Abbasi, Lytras, Abbas & Alowibdi, 2017). It also happens that in the scope of educational research, this large quantity of data translates into a priceless asset of highly relevant information about students, which can be used as study tools to reveal patterns (Lotsari et al., 2014).

Although owning data is a great start, it is not enough. According to Earl (2005), under these circumstances, higher education institutions need to go beyond "data-rich to being information-rich and knowledge-rich as well" (p. 8). The process of knowledge formation results from the organization and contextualization of information. In turn, relating and giving a meaning to information generates knowledge to act properly (Earl, 2005). Earl (2005) suggests that "the challenges come in deciding what data are appropriate and useful for their purposes, ensuring the quality of the data and doing the kinds of analyses and interpretations that will help them make sense of the data" (p. 9). Implicitly associated with the value that the data can possess, is a judgment on the selection of data according to its quality, the organization of the information, as well as, reflection on the results and application of measures with knowledge (Earl, 2005).

However, it is essential to recognize that for universities, it is not a common practice to use their own data in such ways. Neither all applicable data are collected by institutions, and the existing ones are often not collected for specific study purposes (González, 2009). A research about 100 community colleges across the USA revealed that "in the colleges' efforts to increase student success, it became clear that it was necessary for everyone to understand the types of data needed and exactly how such data was to be used" (González, 2009, p. 2).

The Fénix information system typifies a good example of how web technologies increasingly represent an integral part of educational field in Portugal and how their use

has a great positive impact. This system grants, through institutional certified users (students, teachers, staff and candidates) a variety of services and features that are present in the daily life of all those who study or work in a large number of higher educational institutions that use the Fénix. Beyond all the features available, Fénix has an extensive database where all the information is stored about each institution. It is the case of Iscte – Instituto Universitário de Lisboa (ISCTE-IUL), which holds the record of all students who enrolled at this university, not only personal data but a variety of data regarding the entire course of their students. However, we intended to focus all interest on data within the framework of academic success. Higher Education Institutions must understand the types of data needed from what is available, and exactly how they should be used in order to allow the establishment of good analysis, and in turn, reliable conclusions (González, 2009). From this point, we conducted interviews to gather the needs of potential users regarding data that will be of greatest interest to consult and what data should be used for the purpose of the analysis of success in Higher Education Institutions.

3. REQUIREMENTS GATHERING

The requirements gathering focused on two core issues: what do we need vs. what do we have (Figure 1). The first question resulted in a process of several interviews. We wanted to understand from different perspectives what would be the definition of academic success, as well as the main points for measuring academic success/failure from the perspective of teachers and staff belonging to a Higher Education Institution. For the second question, we focused on the relationship between the overall interview outcome and the supply of data from ISCTE-IUL available at Fénix system database, as it is the most used academic information system in Portuguese higher education institutions to store information about students and their academic path. Finally, we developed a platform that added the information resulting from this process.
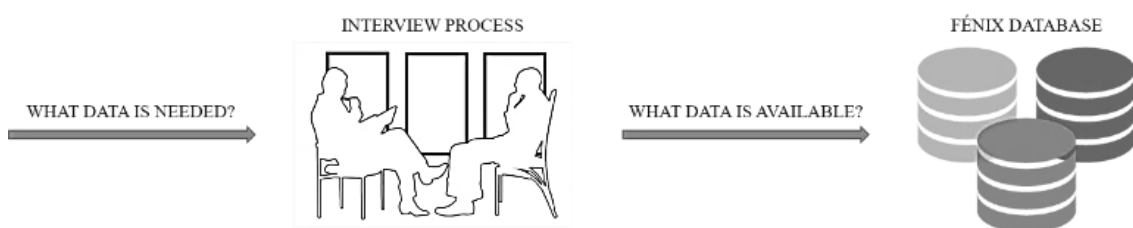

Figure 1- Requirements gathering process sequence

3.1. Interview Process

This new platform was created as a tool to support research and institutional monitoring of academic success. For the creation and design of the platform, a set of interviews was carried out to complete and consolidate the guidelines written in the literature on the subject. The interviewees have responsibilities and experience in coordinating curricular units and courses. Consequently, they knew the information in the Fénix administrative database. This procedure served as an exploratory basis, to define

3

different perspectives regarding the theme of academic success, having a secondary use in this work.

The interviewees were six researchers or teachers (mainly those who have course and curriculum unit coordination responsibilities), familiar with the type of information contained in Fénix system and identified as potential users of the proposed platform.

The interviews were guided by a script composed of five questions:

1. What is your definition of academic success?
2. What are the factors that are most involved or condition academic success?
3. Which types of students may be more vulnerable to failure?
4. Given the universe of the Fénix system, what data seem to be most relevant to establishing indicators of academic success?
5. Comments on the possible usefulness of having a platform that allows people to consult and download academic data related to academic success.

3.1.1. Interviews' Outcomes

To establish a more wide-ranging view, we interviewed six potential users who agreed to participate in the interview and will use the platform in the coming years. The interviewees are teachers of the institution in different academic areas and oriented towards the analysis of success (either through research or positions where this is relevant), namely Sociology, Information Systems, Computer Science and Electronic Systems Engineering, which ended up reflecting certain discrepancies in the definitions of success.

According to the interviews' outcomes about question 1, some interviewees see the definition of academic success as achieving the degree for the stipulated time, depending on the statute of the student (part time, full time, student worker or international student). Others tended to distinguish multiple success levels based on the number of completed courses per year versus the stipulated ones, from the moment the student is approved for the next curricular year.

Others viewed academic success as the predominance of the set of skills and knowledge (theoretical and practical) acquired during a course. For others, academic succes was a combination of student performance, acquisition of knowledge, continuity of studies or pursuit of a professional career in the technical area of training. Finally, some argued that success is the result of a very comprehensive enriching experience in the academic context that includes acquiring a set of academic goals (completing courses on time, acquiring a degree) as well as the acquisition of features related to personal (personal development), social and intellectual skills.

Concerning the question about the factors that may influence academic success (question 2), the responses we got evoked:
- Satisfaction of the student with the frequency of the course;
- Demographics factors;
- Enrollment choices;
- Integration in the university environment (student involvement in the institution and social relations with teachers and students);
- Family background (parents' education and occupation);
- Student work methodology (planning and time management, study habits, autonomous and group work capability);
- Student motivation;
- Aptitude for the training area;

- Institutional factors (evaluation methods, teaching approaches, methods and pedagogical practices, provision of suitable spaces for study and digital content repository);
- Characteristics of the course (difficulty and requirement of the course);
- Characteristics of the teacher (ability to transmit knowledge, agility and availability to clarify any doubts)
- Socioeconomic factors.

In addition, the interviewees were asked about the types of students that may be more vulnerable to failure (question 3). The answers were grouped into: International Students (Portuguese language barrier, different assessment and teaching methods), displaced students, students with a vulnerable social background, students with low schooling family background, students with economic shortages, students with low motivation, students who do not attend classes, students with low previous education preparation, students with difficulties in adapting to assessment, teaching methods and practices, students with integration difficulties and students who are unsure about the choice of the course.

After the collection of definitions and factors related to academic success, interviewees were asked which data would be most relevant to establish indicators of academic success and which information could be associated to this concept (question 4). The different data sets proposed in the responses were as follows:

Table 1- indicators identified by interviewees considering their knowledge on the Fénix information system

Data Set	Indicators identified by interviewees
Demographic factors	Age
	Sex
	Nationality and second nationality
	Place of residence
Family background	Parents' education and occupation
School experience (prior to attendance in university)	Specifics related to previous education (Public / Private, specificity area)
	Higher education ingression grade
Academic satisfaction and choices	Number corresponding to the ingression degree preferences
Academic experience and path	Enrolment
	Registration in degrees and curricular units
	Attendance
	Access to higher education (normal, post-work)
Academic performance and achievement	Grades in curricular units
	Degree grade point average
	Approval rates of courses
	Number of courses completed per year
Special condition or statute	Student statute (working student, international student, displaced student, scholarship student)
	Erasmus students
Educative tools	Course's content repository

When asked about the usefulness of the platform (question 5), most of the interviewees considered quite interesting the idea of having an interface that allows to consult and export institutional data related to academic success. Considering that the

Fénix system contains a large volume of information, we can take advantage of it to provide easy access to consult institutional data, looking for solutions that are not currently available for their regular use, which subsequently could introduce more effective measures to combat school failure and dropout in higher education.

3.1.2. What data is needed?

By grouping the proposed data sets into categories, it ends up reflecting an easier and structured analysis (Table 2).

Table 2- Variables proposed through the interviews by analysis dimensions (identified from literature)

	Dimensions	Indicators	Additional Information
Pre-Academic Experiences	Socio-demographic characterization	Sex	---
		Age	---
		Place of Residency	---
		Nationality and second nationality	---
	Family background	Parents' education and profession	---
	Previous education	Precedent degree designation and conclusion year	---
		Previous education sector	Public, private or both
		Specificity of study' area or field of study	(Ex: Economy, arts, etc.)
		Candidacy preferences	
		Access Type	Normal regime, post-work regime
		Higher education ingression grade	---
Academic Experiences	Students' profile	Working student	---
		Partial time student	---
		Dislocated student	---
		International student	---
		Scholarship	---
		Outgoing mobility	Erasmus students
	Persistence	Enrollments	Number of enrollments per degree
		Retention in curricular Units	Uncompleted curricular units
		Attendance	Attendance in classes
	Academic Achievement	Curricular units' grades	---
		Curricular units completed per year	---
		Degree grade point average	---
	Degree's proprieties	Approval rates to curricular units	---
		Curricular unit's content repository	---
		Degree Completion Rate	---

3.2. Data selection process

After gathering definitions, indicators and potential sources of information that could be correlated with academic success, it was necessary to understand which data from Fénix system database were associated to the collected factors. We answered this question by analyzing the related variables available in Fénix (Table 3). The information presented in Table 3 allowed us to advance to the next steps in the development and design of the data consultation platform.

3.2.1. What data is available?

Table 3 – Needed variables available in Fénix system database

Variables	Additional information
Sex	Variable related to students' gender
Place of Residency	Variable related to students' place of residency
Year of Birth	Variable related to students' year of birth
Nationality	Variable related to students' nationality
Second Nationality	Variable related to students' second nationality
Parents' Profession	Variables related to the occupation of the students' parents
Parents' Education	Variables related to the education of the students' parents
Marital Status	Variable related to students' marital status
Precedent degree designation	Variable related to the designation of the students' previous degree
Precedent degree conclusion year	Variable related to the conclusion year of the students' previous degree
Previous education sector	Variable related to the type of school attended in high school (Ex: Public, private or both)
High School Completion Course Type	Variable related to the high school course or area, which appears in DGES's [Direção-Geral do Ensino Superior] candidacies (Ex: Scientific-Humanistic course, equivalence, etc.)
Candidacy preferences' order	Variable related to the candidacy preferences' order
Higher Education Ingression Grade	Variable related to the overall access grade
Access Type to Higher Education	Variable related to the students' type of access to higher education (Ex: General access regime, part-time regime, degree holder, etc.)
Special Education Needs	True, if the student has special needs status.
Displaced Student	True, if on enrollment via DGES the student indicated that he/she was displaced
Transport Supplement	True, if the student has supplement for transport
Accommodation Supplement	True, if the student has supplement for accommodation
Scholarship	True, if the student has any sort of scholarship (FCT or SAS)
Iscte Financial Support	True, if the student has ISCTE-IUL financial support
Student Statue	Variable related to the students' statute in higher education system (Ex: SAS or FCT grant owner, Working student, Partial time student, International Student, etc.)
Outgoing Mobility Period	Indicate if the student at ISCTE-IUL will be attending a semester or a year at a non-Portuguese higher education institution
Degrees' Designation	Variable related to degrees' designation
Degrees' Type	Variable related to degrees' type (Ex: Degree, master's degree, PHD degree, etc.)
Degrees' School	Variable related to ISCTE-IUL degrees' school (Ex: ISTA, ECSH, EG, ESPP, IPPS, etc.)
Registration State	Variable related to the students' registration state (Ex: Registered, concluded, interrupted, quitted, etc.)
Curricular units' Designation	Variable related to the curricular units' designation
Curricular units' ECTS Credits	Variable related to the curricular units' ECTS credits
Curricular units' Grade Scale	Variable related to the curricular units' grades scale (Ex: 0-20, etc.)
Curricular units' Grade Value	Variable related to the students' final grades to curricular units

Considering the variables mentioned in the interviews and correlating them with some of the variables in the Fénix database (Tables 2 and 3), it was possible to verify in advance that even though Fénix is not a platform for the purpose of an academic success-focused tool, it collects a lot of information from the students that it can have a lot of potential to help monitor academic success.

4. DATABASE DESIGNING

In order to create a platform for the purpose of making data available for consultation, it was necessary to develop a database in which the system can perform the requested queries (Figure 2). Initially, Fénix was designed for importing data contained in own database but always as an independent service. This required the designing and structuring of a smaller database than the Fénix database, containing only the relevant and desired information for management purposes.

Figure 2- Main platform database design process sequence

It is an essential condition that the information is properly structured for a good performance of any information system. A good structure makes the management of the information system easier and more intuitive to guarantee its maintenance and consultation (Ramos, 2008). Over the years, several modeling languages have been developed and used to improve the structure and organize information. The relational model is one of the most favored, as it is supported by most information systems development tools (Ramos, 2008). Also, this preference is essentially due to the simplicity and ease of use embodied in this model. In the relational model, the information is structured in relationships, also referred to as tables (Ramos, 2008). Given all these features, we propose an approach to managing the data of this web application through the relational model. We designed a model that not only supports the importation of the Fénix data, but also makes it available for consultation on our platform. We connected both systems by establishing relationships between tables and allowing the crossing of relevant student information to enhance studies related to academic success.

After designing the relational database, a system software that allows to create and manage databases is needed and therefore we used the RDBMS model (Relational Database Management Systems). In order to interact and communicate with the data stored in these systems, we used the Structured Query Language (SQL), which is the default programming language for dealing with relational databases. The system chosen to create and manage the relational database was MariaDB (2020) because it is a fast, scalable and robust open source relational database management system and it also provides a SQL interface for accessing data.

5. WEB APPLICATION

After we have implemented and connected both databases, the next steps were dedicated to the design and development of the platform.

5.1. System Design

At this stage, we planned the system design and architecture to meet the needs of end users: institutional leaders, practitioners and researchers. As such, we considered the various types of requirements outlined as well as some conditions of usability and functionality to potential users.

To design a system architecture, we had taken into account the intended functionality in the system and also factors such as integration of services. We concluded that the best solution to meet all project requirements would be to deploy a web application. Therefore, the entire structure and architecture of the system was focused on the integration of systems, a required feature to implement this type of application software.

A web application is a software application that runs on a remote server. In most cases, the web browsers are used to access web applications on a network, such as the Internet. In this particular scenario, the web server worked on a local server for security and testing purposes and it communicated with the web browser through the user interface, as well as the web application. To integrate the main functionality of the system (make data available for consultation), the web application communicated with the databases.
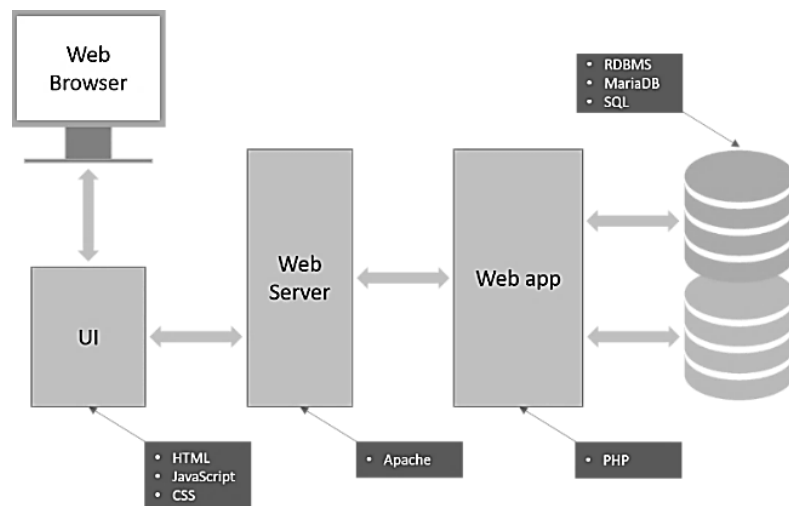
Figure 3- System architecture

5.2. System Implementation

The implementation of the web application was essentially divided into two basic processes: front-end and back-end development. The front-end (UI – User Interface) refers to the presentation layer of a web application. This process is also called the client side because it includes everything related to the user views and experiences with direct contact, like images, navigation menus, text, colors, buttons, among others. In front-end process, we used three languages, namely: HTTP, CSS and JavaScript.

The back-end (server side) allows us to create components and features that are accessed indirectly by a user through a front-end application. Succinctly, it is the part of the web application that the user does not see, and which is accountable for ensuring that everything on the user interface really works, including the management and interaction of information systems (databases). For the back-end development process, we used two languages, namely: PHP and SQL. To integrate all the intended architecture in accordance with the presented languages and systems, we used XAMPP, which is an open source cross-platform web server solution package consisting primarily of the Apache HTTP server, MariaDB database, and script interpreters written in several programming languages, including PHP, in a complete package.

The implementation of the system went through the development of the user interface and back-end simultaneously. Initially, the Fénix system was used as a graphic aid in the development of the user interface, to resemble the front-end in both services. We wanted to convey the feeling to the user that the service was an extension of Fénix. These similarities were most visible in structure, navigation menus, colors, and logos.

5.3.   Data anonymization procedures

Developing a web application whose goal is to create new perspectives for visualizing, extracting and analyzing data also means that ethics/privacy issues will arise. Guaranteeing and safeguarding the privacy of all those whose information is collected and displayed is vital (Piety, Hickey & Bishop, 2014). Therefore, one of the core parts of the design was to normalize and anonymize the data according to the current policies and rules on the processing of personal data relating to people in European Union (GDPR - General Data Protection Regulation, 2016).

In our project, since one of the most important requirements of the system was to ensure the preservation of student anonymity, at least three restrictions have been incorporated. The first constraint starts at the database level. Student identification data such as name, student number, email, mobile phone, among others, were never part of any of the information systems used during the project. The other restriction makes it impossible to cross-check data that directly identifies the student, such as candidacy identification numbers or registration identification numbers. And as the previous restrictions do not fully guarantee the non-access to the identification of specific student cases, the other restriction works in the return of each query, making them available only if the consultation returns a minimum number of cases. This restriction means that the system does not allow to show minority cases in which some students could easily be identified. In addition to above listed these restrictions, the system only grants access to ISCTE-IUL certified users with the internal domain email and who are authorized to query the data.

6. RESULTS

The result of this project was a functional web application with a feature/indicator set. It is a finished product, fully functional, with a basic structure composed of components that are easy to learn and interact.

The web application starts with a login as only authorized users can have access to the application's features since the information contained is sensitive and critical. This system should only be accessible to users who will make the correct use of the data, as HEI administrators, teachers and researchers, in the sense that they will not attempt to

take advantage of the system for negative practices such as the purpose of identifying and exposing particular student cases.

The implemented system compares many variables that were considered relevant for academic success studies, and it can establish and perform an extremely important role in helping this type of studies and investigations. This feature is called as "Free query" and it is the main functionality of this system since it aims to enable cross-reference of a student dataset. All available data were considered relevant during the requirements gathering and database design, for finding new standards and paths of academic success. Its structure consisted of a selection panel (side navigation menu), in which each title (pre-academic and academic experiences) includes a set of dimensions and its respective dataset (Figure 4).

Figure 4- Free query navigation menu

Each dataset of each dimension is hidden by definition, but it can be expandable. All these sets can be flagged or related through the checkbox. Some of them have, in addition to the checkbox, a selection filter to provide filtering on time of the search and performing a more dynamic search execution.

Since some of the variables may not be so semantically clear to the user, each variable has a brief description inserted in the help hint button. These features are visible in the Figure 5.

Figure 5- Free query features

In addition, we can change the range of years during which the selected variables should be consulted. These queries are returned in the form of a table, with the possibility to export to an excel function, beyond the additional option to save the query. Figure 6 illustrates an example.

Figure 6- Free query searches example

This web application provides a semantic contextualization of all variables and results of all information that can be crossed. This feature is called "Metadata" and it allows the user to understand the information about all possible outcomes, as well as the theoretical framework of the variables (Figure 7).

Figure 7- Metadata example

The web application also offers a feature named aggregated data (macrodata) which provides a set of pre-defined indicators within the context of higher education as well as academic success. To include these indicators, we have developed a side navigation menu consisting of subgroups to allow an easy interaction between the user and the service. In this feature, we used the DGEEC (2020), PORDATA (2020) and Eurostat (2020) platforms as a reference. Table 4 and Figure 8 show the result of the indicators established that were based on those found in these platforms.

Table 2- Macrodata established indicators

Enrollments	Number of enrollments	• per year; • per year and sex; • per year and nationality; • per year and degree school; • per year and degree type;
	Number of enrollments for the first time	
	Number of enrollments in the 1st year	
	Number of enrolments in ISCTE as first option	
Graduates	Number of Graduates	
Dropouts	Number of Dropouts	
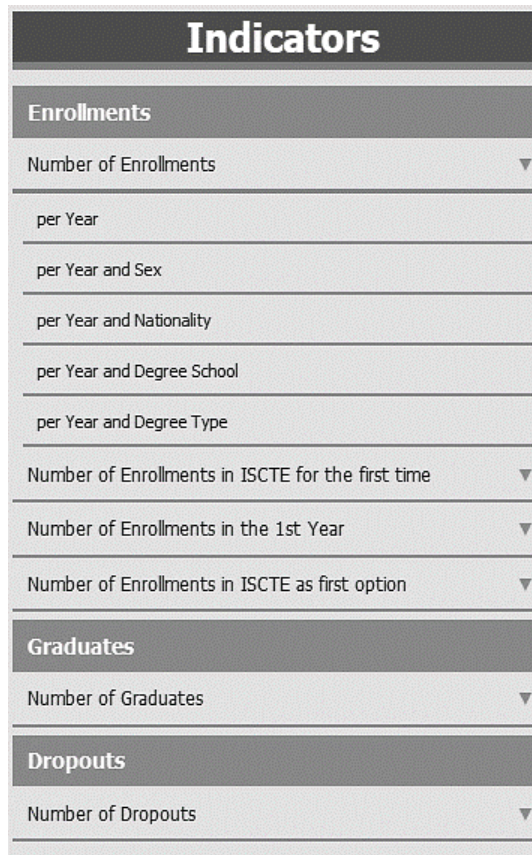
Figure 8- Macrodata navigation menu

Similar to the "Free query" feature, these queries are returned in the form of a table, with the possibility of exporting to the excel function and saving the query. The following example illustrates this feature (Figure 9).

Graduates in ISCTE
(per Year and Degree Type)

Bachelor (BSc)	Master Degree (MSc)	Integrated Master Degree	PHD Program	Post Graduation Degree	Specialization
7	0	0	0	1	0
9	0	3	2	13	0
373	19	5	3	7	0
604	70	4	4	30	0
704	408	9	2	90	0
733	801	34	5	601	0
837	731	55	26	437	7
809	601	39	56	394	30
881	659	33	87	470	22
873	679	60	82	455	77
885	668	51	92	528	39
889	640	42	60	430	63

This query returned: **12 cases**

Figure 9- Macrodata searches example

In addition to all the features presented, the web application also includes the "latest searches" and "account settings" that allow the user to achieve uniqueness and identity in their session (Figure 10).

Latest searches

Name	Date	
table-exportation	2019-09-25 20:12:23	View search
table-exportation	2019-09-19 14:41:23	View search
table-exportation	2019-09-19 14:15:16	View search
table-exportation	2019-09-19 14:14:27	View search
table-exportation	2019-09-19 14:14:00	View search
table-exportation	2019-09-19 11:00:50	View search
table-exportation	2019-09-10 19:26:53	View search
table-exportation	2019-09-10 15:23:18	View search
table-exportation	2019-08-19 14:58:17	View search

Figure 10- Latest searches feature

After defining the architecture, we ensured that the databases were implemented correctly and worked with the system, testing all features. We also verified that the databases have been configured and adjusted properly, from the login process to the correct communication to any tool that the user has contact within his session.

7. CONCLUSIONS

This article presents a tool that allows the availability of academic data for interactive consultation, aiming, above all, to frame the scope of consultations in the context of academic success. This information concerns indicators to measure success such as the completion of course units, conclusion of graduation (etc) - but also some of the characteristics of students with different pathways of success.

In Portugal, higher educational institutions collect a lot of student information. By leveraging this aggregate data to assess for specific problems within the institution, it may be possible to establish more effective measures to combat failures in specific cases. In order to allow stakeholders to analyze and identify some types of academic performance standards within the higher educational institutions, we have developed this web application which represents a set of research aid features. Offering a user interface through the browser allows direct access to authorized users to a set of information imported from the Fénix system, one of the most academic information systems in Portugal. In addition to the visualization and allowing a much deeper and advanced use, the system grants the exportation of queries to ".xls" compatible software, such as Excel.

Through the completion of this project, it was possible to conclude that academic success is a matter of enormous relevance and concern not only for educational policy makers and governments, but also for institutional leaders, practitioners and researchers. The whole objective of the project proposal has been achieved and we hope that this web application can play a very important role as a research aid tool related to academic success. However, new challenges were identified, as well as some limitations, which will be discussed in the following section.

8. LIMITATIONS

We have identified some limitations throughout the project. The first is related to the design and implementation of the Fénix. This academic information system was not developed to be fully dedicated to studies and research on academic success and consequently much of the stored information cannot be repurposed to be used in the context of such systems. This leads to the second and third limitations. Many of the fields and variables in the Fénix database were not designed for further analysis, which means that some of them are not required or mandatory. Consequently, during the data comparison, there is a great loss of information, which greatly hampers those who wish to analyse the complete data sample of the institution and not just a few portion.

Some fields in the Fénix database are poorly filled and not standardized. There are anomalies and inconsistencies that could limit not so much the availability of information, but above all, further analysis and system implementations that use Fénix such as our proposal in which the selected and anonymized information relies on Fénix data.

9. FUTURE WORK

As mentioned earlier, studies concerned with finding hypotheses to produce contexts that foster student success are increasingly important topics, especially for countries interested in and committed to improving the efficiency of higher education institutions (York et al., 2015; Kuh et al., 2006; Kolster & Kaiser, 2015). Within this framework, it is necessary to continue our project, in order to improve and update it.

The system should continue to receive Fénix imports each year to further expand the sample of cases and to reveal even more details of successful developments. It should also continue to be used and tested so that errors and failures can be found and improved.

There must be further indepth investigation. We propose to research more indicators to be established, since our study probably did not focus on all indicators of academic success that could still be obtained from the Fénix.

10. REFERENCES

Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic Analytics: A New Tool for a New Era. *EDUCAUSE Review, Vol. 42 No. 4*.

Daud, A., Aljohani, N. R., Abbasi, R. A., Lytras, M. D., Abbas, F., & Alowibdi, J. S. (2017). Predicting Student Performance using Advanced Learning Analytics. 415-421.

Delavari, N., Phon-Amnuaisuk, S., & Beikzadeh, M. R. (2008). Data Mining Application in Higher Learning Institutions. *Informatics in Education - International Journal, Vol. 7, No. 1*, 31-54.

DGEEC [Direção-Geral de Estatísticas da Educação e Ciência]. (2020). *Direção-Geral de Estatísticas da Educação e Ciência*. Retrieved November 2018, from Dgeec.mec.pt: Available at: http://www.dgeec.mec.pt/np4/dgeec

Earl, L. (2005). From Accounting to Accountability: Harnessing Data for School Improvement. *2005 - Using data to support learning*. Retrieved from https://research.acer.edu.au/research_conference_2005/8

Frati, F. Y. (2019). Using an inquiry-based learning approach to support engagement with information and scholarship in health care education. *Education for Information*, (Preprint), 1-8.

Eurostat (2020). *Your key to European Statistics*. Retrieved from https://ec.europa.eu/eurostat/web/main/home

Gauld, C., & Whatley, P. (2017). Distance learning for information professionals: A practical, reality-driven model for postgraduate education. *Education for Information, 33*(2), 121-133.

González, K. P. (2009). *Using data to increase student success: A focus on diagnosis.* Lumina Foundation for Education.

Hand D.J. (1992) Microdata, Macrodata and Metadata. In: Dodge Y., Whittaker J. (eds) Computational Statistics. Physica-Verlag HD

Kahu, E. R., & Nelson, K. (2018). Student engagement in the educational interface: understanding the mechanisms of student success. *Higher Education Research & Development, Vol. 37, No. 1*. doi:10.1080/07294360.2017.1344197

Kolster, R., & Kaiser, F. (2015). *Study success in higer education: Male versus female students.*

Krishnamurthy, R., & Desouza, K. C. (2014). Big data analytics: The case of the social security administration. *Information Polity, 19*(3, 4), 165-178.

Kuh, G. D., Kinzie, J., Buckley, J. A., Bridges, K. B., & Hayek, C. J. (2006). *What Matters to Student Success: A Review of the Literature.*

Ling, X. (2019). Model for evaluating the art education teaching quality with uncertain information. *Journal of Intelligent & Fuzzy Systems, 37*(2), 1967-1972.

Lotsari, E., Verykios, V. S., Panagiotakopoulos, C., & Kalles, D. (2014). A Learning Analytics Methodology for Student Profiling. In A. Likas, K. Blekas, & D. Kalles, *Artificial Intelligence: Methods and Applications* (pp. 300-312).

MariaDB (2020). *Maria DB Enterprise Open Source Database & SkySQL MariaDB Cloud*. Retrieved from https://mariadb.com/

Martins, S. d., Carvalho, H., Ávila, P., & da Costa, A. F. (2017). Policies for Widening Participation and Success Factors in Portuguese Higher Education. *Creative Education, Vol.8 No.02*. doi:10.4236/ce.2017.82018

Matters, G. (2006). *Using data to support learning in schools : students, teachers, systems.* Camberwell, Victoria: Australian Council for Educational Research .

Piety, P. J., Hickey, D. T., & Bishop, M. J. (2014). Educational data sciences: Framing emergent practices for analytics of learning, organizations, and systems. *4th International Conference on Learning Analytics and Knowledge*. doi:10.1145/2567574.2567582

PORDATA. (2020). *Fundação Francisco Manuel dos Santos*. Retrieved from https://www.pordata.pt/

Ramos, P. N. (2008). *Desenhar Bases de Dados com UML.* Edições Silabo.

Site Map - Fénix - ISCTE-IUL. (n.d.). Retrieved from Fénix ISCTE-University Institute of Lisbon: https://fenix.iscte-iul.pt/siteMap.do?locale=en_EN_ISCTE

York, T. T., Gibson, C., & Rankin, S. (2015). Defining and Measuring Academic Success. *Practical Assessment, Research & Evaluation, Vol.20 No.5*. Retrieved from http://pareonline.net/getvn.asp?v=20&n=5