Filip Nalaskowski*, Dagna Dejna**

Nicolaus Copernicus University

Online Candidate Registration database educational study potential – desk research analytical method

KEYWORDS

education, research, Online Candidate Registration, IRK, archeology, desk research

ABSTRACT

Filip Nalaskowski, Dagna Dejna, *Online Candidate Registration database educational study potential – desk research analytical method*. Culture – Society – Education no. 2(18) 2020, Poznań 2020, pp. 439–453, Adam Mickiewicz University Press. ISSN 2300-0422. DOI 10.14746/kse.2020.18.16.2.

The text presents and describes the desk research analytical method. As an example of the practical application of the described method and a little-known data source, a short analysis of the database of Online Registration of Candidates is presented. The components of the base are presented step by step, along with its information potential. A separate part is the presentation of the possible potential of IRK for educational research – in this case 16.2 years of candidates for archaeological studies at the Nicolaus Copernicus University. Analysis of the candidates' origin, study potential, determination to take a specific major, etc.

Introduction

The desk research method comprises of drawing from the existing data in order to include them within the research process. Existing data sources display numerous advantages, also for social sciences scholars. Primarily, the usage of the afore-

----

 * ORCID: 0000-0001-6973-9960.

** ORCID: 0000-0003-0722-1992.

mentioned data is related to unsubstantial financial costs and organisational effort. The method generates little cost, does not burden the environment, and does not increase the epidemic hazard. In the face of the current social challenges and hazards, the arguments mentioned deserve attention.

Secondary data analysis may be a complementary method within the process of data acquisition and scholarly issue solving; it may also be the primary and sole information source – as it is employed with the project described in the text. The great advantage of the desk research method is that it is possible to conduct analyses on enormous research samples. The disadvantage, in relation to our scholarly endeavour, is that, as scholars, we had no impact on the form of the data gathering instrument, on the way of gathering and storing them, as well as, on the selection of variables.

The data source used for the purpose of the presented text is is the Online Candidate Registration database (Internetowa Rejestracja Kandydatów – later referred to as IRK), as well as, information coming from the University Recruitment Commission (Uczelniana Komisja Rekrutacyjna (later referred to as UKR). We have been granted permission to use and publish data coming from these sources, issued by the Student and Staff Policy Prorector of the Nicolaus Copernicus University in Toruń.

It seems that still, few scholars interested in the higher education aspects, or even in broader terms: education and social issues, have little awareness of the existence of the IRK databases, and their information potential. IRK is a database including information on millions of individuals, that may be the subject of interest of social sciences' scholars.

Each higher education candidate must provide personal data, but interestingly, information regarding provenance, educational path and scholarly results (e.g. exam or first degree studies results). Each of the individuals participating in the registration declares their own choices regarding the planned educational paths.

Importantly, the source described was barely a point of interest of any scholars[1], therefore they deserve profound examination. The following text includes information on what data IRK contains, how to acquire access to it, and what is its scholarly potential.

---

[1] The recruitment department of the Nicolaus Copernicus University in Toruń informed us that since the beginning of the functioning of the system at the university, i.e. since 2007, the authors of this article were the only ones to treat the IRK database as an information source for research.

Desk research and exemplifications of potential data sources

During the time of universal access to the Internet, research conducted within the expanded programme of public statistics (e.g. the Statistics Poland Institute), and primarily, the tendency to publish and share research results, desk research projects are primarily dependent on the skills and talents of a scholar. The social sciences studies, successfully adopt descriptions, summaries, reports, databases (e.g. Social Diagnosis, Polish General Social Poll, European Social Survey, World Value Survey, and International Social Survey Programme).

Such projects assume sharing not only reports, but also databases with the ability the create independent analyses. One of the challenges may be to familiarise with the database, the variable coding method, and to learn all analytical limitations (e.g. sample being too small, sample was not accomplished within a given area, lack of data regarding a particular year). Moreover, a number of projects emerge, commissioned by public institutions, that are published and universally accessible (e.g. Office of Competition and Consumer Protection, Polish Agency of Entrepreneurship Development) (Bednarowska, 2015).

School Factors for Educational Efficiency (szkolne uwarunkowania efektywności kształcenia – later referred to as SUEK)

Secondary sources created with the contribution of institutions that conduct educational studies, are filled with data waiting to be analysed.

A good example of such a database waiting to be interpreted (or, more precisely, waiting for further interpretations) is a database, created for the purpose of studies conducted by the Educational Research Institute, regarding the school factors of educational efficiency. Works of the research team, especially commissioned to the task, were focused on identifying the key factors that determine educational efficiency in primary schools. The SUEK research was a longitudinal, qualitative study, conducted on a representative, nationwide sample of primary schools. The primary study questions were dedicated to the relations between the features of school environment and the efficiency of education in primary schools. The Educational Research Institute uploaded a gigantic database, the scholarly and analytic potential of which cannot be overestimated. It is an effect of the work of a research team, the effort of which was to reach 177 primary schools, and examine 5500 students. The databases presented there are organised in a chronological order, considering the research phase, and in

the order of particular respondents including: students, parents, their teachers and (school) principals. Using the provided databases is conducted under the condition of prior contact with the authors and providing footnotes that would contain an appropriate clause regarding data sources[2].

Other

A valuable source of information regarding the educational and social potential is the Technology and Science report by GUS. The report presents statistics compared within a 10–15 year span for Poland, as well as, for selected EU countries, regarding funding for research and development, personnel in research activities, human research for science and technology.

Additionally, the source of information regarding particular voivodeships is provided by regional technology transfer networks (Regionalne Sieci Transferu Technologii – RSTT) and the regional innovation support system (Regionalny System Wspierania Innowacji – RSWI), as well as, databases regarding technology and innovative solutions of academic centres and research institutes.

In the case of information regarding technology there is an assortment of operational databases, e.g. the database on modern technologies maintained by the Ministry of Economy, the Technologies, Enterprises and Innovative Products database maintained by the Information Processing Centre, Regional Technology Transfer Networks.

In the case of databases dedicated to the structure and the extension of funds contributed to the development of research, one should mention the Ministry of Science and Higher Education databases (Polish Science database, the POL-on system, the projects submitted for NCN and NCBiR), as well as, regional databases, the Science and Business Łódź Voivodeship Portal Database[3].

All the aforementioned databases are a potential information source, which may be used for the purposes of any scholarly idea.

POL-on

Before we describe the Online Candidate Registration database, one should also mention a database, or more precisely, an IT system called POL-on and indicate the possible scope and aim of using the data gathered there.

---

[2] http://eduentuzjasci.pl/bazy-danych.html [02.03.2020]
[3] https://www.gig.eu/sites/default/files/attachments/przetargi/2013_03_27_struktura_bazy_danych_zweryfikowana_fin.pdf [04.03.2020]

Subject/scope of data:
– non-public universities and unions of non-public university registry;
– comparison of departments;
– registry of higher education institutions;
– registry of scientific units;
– church institutions;
– rights of units to grant scientific degrees;
– comparison of individuals with scientific degrees and titles.

The potential scope and aim of using the aforementioned data is primarily to gather comprehensive information regarding the institution of higher education. Due to the data it is possible to define what institutions, what degrees, what scientists and areas of science accomplished in our country are accomplished, also in the regional aspect. The abovementioned allows to diagnose the social and human resources, as well as, infrastructural resources, which may be considered in determining the endogenous potential of particular regions and the country itself[4].

Basic information and the principles of the Online Candidate Registration

Nicolaus Copernicus University recruitment is a holistic process related to accepting the candidate for studies, whilst the registration is just but the first step of recruitment and is performed online via IRK.

Filling the forms while submitting is the first phase of the procedure. Within the next phase one must apply for a given faculty, pay the recruitment fee, and after receiving information, that one was qualified for studies, one must provide the complete assortment of documents for UKR[5].

Online Candidate Registration is a system serving the purpose of upgrading the higher education recruitment process in public universities in Poland. It is related to recruitment for uniform studies of first and second degree, and to PhD schools. The aforementioned includes full-time and part-time studies. In fact, all modes with the exclusion of post-diploma studies, or other types of open courses.

The system has existed since 2005 and currently encompasses 39 Polish universities. The originator of the system was the University of Warsaw, and the current administrator and the owner is the Intercollegiate Computerisation Centre

---

[4] https://www.gig.eu/sites/default/files/attachments/przetargi/2013_03_27_struktura_bazy_danych_zweryfikowana_fin.pdf [04.03.2020]

[5] https://www.umk.pl/kandydaci/niezbednik/faq/ [10.02.2019]

(Międzyuczelniane Centrum Informatyzacji – MUCI)[6]. In order to evaluate the scope of the base, one must consider the fact, that at the Nicolaus Copernicus University, the system has been operational since 2007, and approximately 20 000 individuals register to it yearly (provide data). On the nationwide scale, it yields millions of users – through many years of the system being operational.

Each candidate must create an account in the system[7]. The users log in using their personal identification number (PESEL), which provides their individual, unique login. They choose the password themselves. Secondly, the user provides a number of information about himself, using forms seen on the screen. These are:

– additional information: names, birth date, gender, nationality;
– contact information (address, email, phone number);
– education information (secondary, higher, educational olympic games, or certificates);
– other data, such as identification information or military service records;
– photo – used to create student identification document.

Following these activities, the candidate may choose the faculties which he intends to apply to. Regarding the mode and degree, the procedure may vary. For example, first degree studies may include a system of priorities. The candidate selects a random number of faculties, which he intends to attend to, prioritising them (the degree of interest/demand). Additionally, candidates are obligated to pay a fee for each selected faculty. In a result of the recruitment process, the candidate is accepted to the faculty with the highest priority, with the number of points that allow for acceptance (Figure 1). Recruitment points are estimated on the basis of grades from the high school exams, entry exams, or the result of first degree studies (in the case of applying for the second degree).

For the accomplishment of the aforementioned aim it is imperative to place data allowing to estimate points. In terms of first degree studies, the issue is significantly simplified by the KReM system (High School exam Nation Registry – Krajowy Rejestr Matur), which downloads all the grades automatically (Figure 2) – meaning that the data requires no further verification. The remaining candidates (with „old" high school exams, foreign high school exams, or grades from first degree studies) must input their grades and results manually – the data is verified during the document submission phase, by the University Recruitment Commission members.

---

[6] muci.edu.pl [10.01.2020]

[7] Description and illustration are referred strictly to the IRK system applied in 2020 – the version with an older interface. The layout was subject to change – an assortment of data gathered, as well as functioning principles had not changed.

Figure 1. Example „Priorities" tab

Figure 2. Example „Exam results"

Winners of school olympics and sports, can often experience exemptions in the recruitment process (free acceptance to selected faculties, or being free of recruitment fees). After declaring to have such a right, the candidates are obligated to provide documents that prove the fact, in order for them to be verified by commissions.

Additionally, candidates are obligated to pay a fee for each selected faculty. In the system, the candidates may review their payments and recruitment results.

Moreover, the system is a platform of contact between the candidates, and the staff that manages recruitment at universities.

Database characteristics – detailed information regarding the Online Candidate Registration system

Within the recruitment process, data required to accept candidates is acquired. The majority of data is provided by the applicants themselves, e.g. personal data, or diploma results from first degree faculties, while a portion is filled automatically – „downloaded" from the system from within (in the case of high school exam grades).

Data provided by candidates

The majority of data is provided on the „My personal data" tab (Figure 3). The information is divided into the following tabs and questions:

Figure 3. „My personal data" sub-tab

Basic data:
– PESEL – Personal Identification Number (filled by the system by default),
– name,
– second name,
– last name,
– birth date (filled by the system by default using PESEL),
– gender (filled by the system by default using PESEL).

Contact information:
– e-mail,
– telephone number,
– cellphone number.

Address:
– address of residence,
– address type (selection: city; town, village, foreign address),
– postal code,
– voivodeship (selected from the list),
– city,
– street,
– street number,
– apartment number.

Education data:
– high school graduation (selected from the list; input of a school that is not on the list);
– city (filled by default based on the selection of names from the list);
– high school graduation (earliest);
– student individual document number (if the candidate had already been a student of a university);
– document allowing to attend studies (list of documents: new high school exam (passed since 2005) or IB, or EB; old high school diploma (prior to 2005); foreign high school diploma; bachelor or similar diploma; master or similar diploma; foreign diploma allowing to attend second degree studies; school olympics of finalist title);
– number of selected document;
– issue date;
– institution that issues the document (name and city in terms of high school exams and diplomas, a drop-down list is given).

Documents/attachments (alternatively, in this part the candidates may sent an e-version of essential documents).

Other, additional personal information:
– identification type,
– identification number,
– country that issued the document,
– birthplace,
– country of origin,
– nationality,
– family name (in terms of women who are married),
– father's name,
– mother's name.

Additional recruitment information:
– Do you intend to stay at the dorm? (selected from the list)

Information regarding military service:
– military service status (selected from the list),
– military category (selected from the list),
– WKU (Wojskowa Komisja Uzupełnień – Army Recruiting Command) (selected from the list).

The assortment of the aforementioned data is supplemented with:
– high school exam results, university exams or results from first degree studies,
– selected faculties.

Database access and legal aspects

Access to the described database is given to, by definition, the employees of universities involved in the recruitment process, i.e.: the university management staff, administration departments responsible for recruitment, recruitment and (at times) exam commissions, as well as, application administrators. In terms of UMK, the access is given to 100 individuals yearly.

A question arises: what about individuals that are not part of universities, intend to work on the described database, e.g. using it for research purposes? In terms of UMK and other state universities, the JM rector is in charge of the data.

Additionally, one must note, that from the formal point of view one may not perform scholarly work on personal data, as the applicants had not given permission within the system. According to experts, RODO regulations would impede the IRK study related *stricte* to personal data[8]. However, one may perform collective ("statistical") research, which is allowed by regulations – there is no strict application of the RODO regulations.

Authorship example of a practical application of the Online Candidate Registration database

In 2017, we had both conducted studies on the basis of the Online Candidate Registration Nicolaus Copernicus in Toruń database. The aim of the study was an attempt to build the features of candidates applying for UMK archaeological studies. The scope of the studies would encompass years 2007–2017 – therefore 11 recruitment classes (recruitment on stationary first degree studies).

For the analyses introduced, the archaeology faculty is interesting, as it presents itself as an example of a faculty, where applications dwindled during the decade – from 120 to 20 candidates of the first year. Moreover, in the opinions of the lecturers themselves, the "quality" of candidates has lowered significantly. Research allowed to verify the aforementioned theses, and help in searching solutions to improve the situation. The aforementioned considerations had become a subject of scientific texts, that are currently under the review process.

In the result of the data examined, it was possible to achieve an assortment of information, and establish certain agreements. Here are the example data:
– number of recruitment individuals,
– gender percentage,
– types of graduated high schools,
– average high school exam results, as well as, results for particular (important for archaeology) subjects such as e.g. history, geography, or biology,
– city of examined residents (where they graduated), the distance between high schools and Toruń,
– place of residence (city, town),
– age of candidates,
– other faculties selected by the candidates.

---

[8] Information provided directly by the UMK personal data protection – Jakub Rutkowski.

The above mentioned information belong to the more simple, strictly statistical assessments, however, as previously mentioned, the database provides the ability to construct more complex categories, therefore – pursuing more difficult questions.

– What faculties provide competition for archaeology, and how they compare to archaeology?
– What are the forecasts regarding numbers of interested applicants and the "quality" of candidates?

Interesting, core categories, that were the subject of the study were: *strength of the candidate* and *strength of motivation*. *Strength of the candidate* is a value describing, based on the data available, the intellectual and cultural potential of the candidate. In other words, it is a material for an academically strong student. The second category – *strength of motivation* – responds to the question how the candidate is determined to undertake particular studies. Did he appear by chance or found no better alternative?

The initial study was dedicated to archaeology, but the developed mechanism and analysis measure may be applied to build characteristics of all remaining faculties within academic education.

The first category was, in fact, that which the academic teachers named the "quality" of students. It was estimated using components such as: the type and "quality" of graduated high school, the number of high school exams, the average result of all high school exams, the number of recruitment points achieved within IRK, and the fact of being a winner in school olympics. The second category, named "the strength of motivation", was built on the basis of the following data: distance from Toruń, means of assessing priority faculties during recruitment, class year.

The information acquired gave basis for formulating careful conclusions. Below, we mention some of the assessments, that were built entirely with the use of IRK data.

1. The number of applicants may be placed within three phases: prosperity, sudden fall, low stagnation/bounce. This is comparable to 2007–2009 recruitment (prosperity) of average 103 accepted, 2010–2014 (sudden fall) from 90 to 19, and 2015–2017 (stagnation and delicate bounce) 18–22 candidates.
2. The first 5 years were dominated, regarding candidates, by visitors (100 and more kilometres), and in the second period, a reverse tendency occurred – people from the region would dominate (under 100 km).
3. Additionally one may say, that on the scale of 11 years analysed, the archaeology faculty is slightly dominated by females, as 56% (392) candidates accepted were female.

4. 24% of examined individuals (168 individuals) were graduates of the best high schools within their voivodeship (first 15 schools according to the Perspektywy ranking).

5. Only more than a half of candidates would graduate basic exams on history – one would seem that it is a crucial subject for archaeology. However, 10% would graduate "exotic" art history. Identified as similar at times, and helpful for archaeology: biology and geography, respectively 9% and 30%.

6. The analysis of data presented, demonstrates that in the general overview, archaeology students not higher than above results in high school exams. They exceed at history, geography, social studies and English. They display the most issues with strict and natural sciences (mathematics, chemistry, biology) and German – they fall below the average.

7. Very interesting assessments are brought about with the analysis of priorities regarding faculties. In fact, among the young archaeologists, their discipline would always win within the priorities competition. A group of faculties that would break the trend is relatively low: law, conservation, internal security – they were most often displayed as alternatives to archaeology. The most common competition for archaeology was history – almost 15% of individuals would also apply for it, however: none of them would make it a high priority (sic!) – none of the young archaeologists would prefer to study history.

In reference to the *strengths* category, it was made possible to determine, that the 2007–2010 prosperity period (large number of applications) was accompanied by a subtle fall of the strength of motivation to study archaeology and intellectual/cultural strength of the candidates. The declining trend is correlated with the more dynamic fall of the described strengths among candidates. After the crisis year of 2014, a subtle increase in the number of candidates (by approximately 10–15%) is apparent, as well as, a visible increase of motivation and strength. In other words, during the 3 recent years, there have been less students, however, they have increased motivation and high value of intellectual and cultural strength, in comparison to previous years.

Weighing in variables – an attempt of establishing own methodological instrumentarium

Searching for weight values for described variables, we often had to refer to arbitrary resolutions, based on our pedagogical knowledge and the experience in conducting similar studies.

The first variable was the quality (strength) of the graduated high school. We assumed that the quality of the graduated high school is of importance, and may be a variable to estimate the *candidate strength*. High school is a certain *greenhouse*, in which the candidate grows. The offer it creates for the students, the potential of stimulating their development, its scientific atmosphere, its achievements, will translate to student development, the future higher education candidate. By analogy, weaker schools, creating worse conditions for development, will "spawn" weaker candidates. All this will be revealed not only in easily measurable indicators, such as high school exam results.

The variable discussed, was organised in three phase, based on the *Perspektywy* higher education rankings. Graduates of the first 5 schools in a given voivodeship received +25 points, by analogy the graduates of the last 5 schools in the ranking, received –25 points, while the remaining – points. Point values were established on the basis of evaluation and proportion. We did not want to promote (or punish) thoroughly with this research result. We have decided (arbitrarily), that the strength of the variable may be compared to c. 1/3 average high school exam points (75 pts. in our study).

Constructing a formula to estimate the values of subjects selected for the high school exam –a peculiar level of difficulty – we have used the formula known from IRK. Each faculty has the right to present own criteria of accepting candidates, computing their recruitment points. Naturally, these principles vary slightly within the UMK offer spectrum. The formula which we used, is a variation of the formula of computing points for one of the faculties of our university. The change consisted of expanding the application of the formula on all subject options within the high school exam, and not only the selected ones, as previously.

While defining the value crucial to the description of the candidate, the distance from UMK, we have set thresholds ranging from 20 km, 20–100 km, and more than 100 km. The first value is represented by locals – the 20 km distance was assumed as the maximum distance of the functioning of public communication in Toruń – i.e. the basic condition of the contact of a young person with the city. The second value, by analogy, demonstrated the range of local traffic (railbuses, PKS buses, suburban connections), as well as, denoted a default range of UMK influences. In this range, UMK is still the most often chosen university, however, the necessity of far-fetched drives or moving causes it to be a less obvious choice, in comparison to "local" candidates. The last value – more than 100 km – consists mostly of people, for whom Toruń is not the closest academic centre. It consists of people who, on the road to UMK, "pass" other universities, to which they have natural access.

Summary

The database presented creates large scholarly, analytical and interpretative capabilites. It is a unique assortment of educational and – in more broad terms – social data regarding young people. Without a doubt, its use provides an opportunity to expand numerous social, pedagogical studies, and for some it may be the basic data source.

The described databased has a much better potential to conclude on the condition of particular faculties and universities, than it is currently employed. It is not the only one. All other aforementioned bases constitute an enormous collection of data demanding to be processed and interpreted.

In the face of increasingly popular social unrest, increasing inequality, ecological crisis before our eyes, and the totally unpredictable epidemic hazard, the ability to apply the desk research method, and to utilise it for own research seems reasonable and safe.

As authors, we intend, so that the text and fragmentary research become inspiration for using IRK in a similar manner, as *terra incognita* in methodological inquiries in Poland.

Bibliography

Bednarowska Z. (2015), *Desk research – wykorzystanie potencjału danych zastanych w prowadzeniu badań marketingowych i badań społecznych*, „Marketing i Rynek", 7, pp. 23–24.