



Research & Methods
ISSN 1234-9224 Vol. 23 (1, 2014): 115–134
The Ohio State University
Columbus, Ohio, USA
Institute of Philosophy and Sociology
Polish Academy of Sciences, Warsaw, Poland
www.askresearchandmethods.org

Construction and Use of a Relational Database to Analyze the Careers of Professional Politicians in Poland, 1985–2007*

Justyna Nyćkowiak

Institute of Sociology, University of Zielona Góra

This article describes a unique version of a longitudinal database featuring the Eastern Europe Parliamentarian and Candidate Database (EAST PaC). This version contains contextual and biographical information on career politicians from Poland, 1985–2007. I define career politicians as those who, three times during the period covered in the data, were granted a parliamentary seat in either the upper (Senat) or lower house (Sejm). As part of this project, I reconstructed the political biographies of career politicians. Based on practical experience in using these data, I propose the use of relational databases as a best practices approach to manage data of this type. I describe the need for relational databases in general, and describe in detail how to apply them to longitudinal data consisting of parliamentarians and their contexts.

Key words: professional politicians; parliamentarians' careers; relational database; longitudinal data.

Acknowledgements

This article was funded by Poland's National Science Centre (decision number 2012/05/E/HS6/03556) for the project, „Who Wins and Who Loses in the Parliamentary Elections? From Formal Theory to Empirical Analysis.”

This article proposes a particular approach to the process of gathering information and data management for research that involves longitudinal data analysis. I use the example of a scientific project that aimed to analyze the institutional contexts of political careers' paths in Poland, 1985 to 2007. As part of this project, I also reconstructed the political biographies of professional politicians.

Career paths are complex and multidimensional (Arthur, Hall, Lawrence 1989). An efficient use of political institutions allows certain people to create a uniquely political career, and in the extreme of this, such people can be called ‘professional politicians’. Professional politicians can be characterized by several sets of variables related to their demographic characteristics, parties to which they belong, and the history of their participation in elections, etc. (Norris, Lovenduski 1995). These sets of variables are related in a complicated way, and depicting these relations needs to be based on certain rules (cf. Ragin 2006). To prepare and compile empirical material that accounts for political careers, one needs to find, collect and process a lot of interconnected information and, thus, build a database.

Considering that data regarding career paths are supposed to document events across time and changes of conditions, the best solution is to design a relational database. The main reason to choose this form of a database is its inherently flexible connections between information coming from different levels of the analysis. When taking up projects – both quantitative and qualitative – social researchers have to decide on a number of issues that will enable them to collect high quality data. This is facilitated by the involvement in the process of research stages specifically focused on data collection methods, its processing and storage. Thanks to such a construction, a range of collected data within a database can form a solid research foundation.

The acquisition, collection and processing of data is one of the key issues faced by the social sciences. The social sciences are challenged by the need to harmonize data that has been collected for years under a number of research projects. Every new research project requires gathering large amounts of data from multiple sources and describing its structure and type, as well as gaining access to it. An important obstacle is the time-consuming nature and high cost of their transformation and harmonization.

The tools that allow one to design a scheme for collecting and processing data to their optimal use are database systems. Broadly speaking, a database is a collection of information about the facts: (1) about specific phenomena, logically connected, meaningful in a strictly defined domain; (2) collected on purpose and in an organized way, according to a predefined scheme; (3) coming from a clearly described source, that guarantees the quality of data. A database is a structure composed of different elements that are used to store information about particular categories. Thanks to their design, databases allow for relatively quick sorting of records according to specific classifications. Moreover, they enable one to search for information in the indicated fields. To facilitate the use of databases, one can apply appropriate filters that allow submitting information in a schematic and clear way.

For the needs of the scientific project regarding professional politicians’ careers in Poland since 1985, there was a unique database prepared – ‘Professional

politicians 1985-2007'. The database collects data scattered throughout varied sources, and thus does not include 'new' information. Rather, it is information newly processed and organized in a pattern that allows for dynamic analyses. It accounts for a time aspect and changes that biographies of specific parliamentarians, especially professional politicians, might have undergone.

PROFESSIONAL POLITICIANS

During radical political transformations, when political institutions undergo a reshaping process, knowledge about how political institutions function is a special and valued type of knowledge that, to certain political entrepreneurs, might bring them varied types of benefits (Przeworski 1991; Pereira, Maravall, Przeworski 1993; Mach 2000; Shabad, Słomczyński 2000; Wnuk-Lipiński 2001). Making use of that knowledge may foster professionalism of actions of those people who pursue a political career (Post 1995). Activities of some of these politicians may be self-conscious in that they deliberately forge their careers based on a system of political institutions (cf. Borchert, Zeiss 2003; Shabad, Słomczyński 2004; Shabad, Słomczyński, Zieliński 2008). A political professional may take on a variety of political activities¹, and at the same time develop a set of potential rewards for the most persistent political players, or for those who are most knowledgeable of the rules of the political game (cf. March, Olsen 1984, 2006; Wesolowski 2001).

I defined professional politicians as those who, three times during the period 1985-2007, were granted a parliamentary seat in either the upper or lower house. This choice is related to the evaluation of politicians who, after their first successful election run, repeated their electoral success no less than twice, or were focused on repeating the success in the future²; these events could not be merely accidental. Considering the high degree of variability of political preferences of voters, using the "three terms" threshold in the context of parliamentary elections in Poland strongly suggests that a politician had to undertake very effective steps to strengthen his or her electoral position in the eyes of voters and party structures alike. Not all politicians could be classified in the category of professional politicians. The concept of a "political elite" would be too general and not adequately capture the deliberate nature of politician's seeking a long career out of politics (see also Wesolowski 1998: 8-9).

Collecting data depicting the course of politicians' careers in Poland in the years 1985-2007 can provide answers to a number of important questions, such as: (1) To what extent elements of the electoral system determine the course of political careers (including modifying the strategies used by the candidates)? (2) To what extent elements of the party system determine the course of political careers (including modifying the strategies used by the candidates)? And (3) To

what extent individual attributes of candidates determine their electoral success chances and prolonging of political careers?

To explain the trajectory of political careers, one needs to account for three levels of analysis: (1) institutions, specifically the party and electoral systems; (2) the organization of political parties, and (3) individual characteristics of the politicians themselves (Anderson 2007; Heath 2007; von Beyme 2005). In contrast to classical studies carried out in accordance with the logic of social mobility, the approach presented here focuses on examining the relationship between an event and its historical context through time. Models of the course of career paths must take into consideration micro, mezzo and macro variables. It is essential to carefully determine which mechanisms and on what terms within this process occurs. This requires us to create a specific kind of database, since only a combination of these assumptions with a statistical procedure allows the formulation of legitimate conclusions. This calls for event history analysis (cf. Raciborski 2006; Bolero 2007; Chmielewski 2010). Its role is mainly to model the duration of processes, and therefore it seems to be a good tool for explaining and predicting events within the course of political careers.

Event history analysis is derived from the sociology of mathematics and its applications include the assessment of demographic behavior and their impact on events within careers. It refers to the dynamics of the process, which is examined from the viewpoint of an individual in terms of: (a) time of one's stay in certain states (e.g., periods of holding a political office, breaks from participation in elections); (b) transitions between states (e.g., types of transitions – change for higher/lower position, a change in the level of education); and (c) a sequence and timing of certain events (e.g., periods of work in parliamentary committees and holding specific functions in them) (Balicki 2006).

The study of events needs to be divided into several research tasks:

- clear recognition and interpretation of events and conditions,
- examination of a specific event type (for example, a win/loss in parliamentary elections),
- study of time duration of a given state (e.g., holding a seat within a certain period of time) or the waiting time for an event to occur (for example, the waiting time for taking part in another election),
- inspection of the number of events that occurred within a specified period of time (e.g., the number of held posts in the period of time),
- investigation of the sequence and intensity of events (e.g., the sequence of held posts),
- study of the events' relationship (e.g., the relationship between the number of electoral successes and occupying a specific place on the electoral list).

PROGRESS OF THE PROJECT AND THE EMPIRICAL MATERIAL

According to the scheme presented above I began preparations for the project with a preliminary analysis of statistical data (coming from the POLCAN database, a unique data set of electoral activity of all candidates for the parliament since 1985, constructed under the direction of Kazimierz M. Słomczyński and Goldie Shabad within the CONSIRT³ program that was carried out by the Departments of Sociology and Political Science of The Ohio State University and the Institute of Philosophy and Sociology, and the Graduate School for Social Research at the Polish Academy of Sciences). The POLCAN dataset allows to trace exactly who and when appears in the parliamentary elections in the years 1985-2007.

After 2007, POLCAN has been folded into a new database EAST-PaC, or the East European Parliamentarian and Candidate database, funded by Poland's National Science Centre and includes similar data from the Ukraine and Hungary (see electoralcontrol.org). As I am discussing a project already completed prior to the creation of EAST PaC, I will refer to the Polish candidate data as POLCAN. However, I am currently applying the relational database concept developed prior to the new EAST PaC data.

The following information from the POLCAN database has been used in the project:

- Candidate's ID (the lower house of the parliament (Sejm) 1985-2007, the upper house of parliament (Senat) 1989-2007)³,
- Candidate's sex,
- Candidate's age (birth date),
- Candidate's election result (Sejm 1985-2007, Senat 1989-2007),
- Candidate's electoral ward (Sejm 1985-2007, Senat 1989-2007),
- Candidate's place on the voting list (Sejm 1985-2007, Senat 1989-2007),
- Candidate's number of votes (Sejm 1985-2007, Senat 1989-2007),
- Candidate's political party (Sejm 1985-2007, Senat 1989-2007).

By conducting basic statistical analyses one was able to distinguish categories of people who were then included in the main study sample. Given the nature of the POLCAN database, the information available within and its purpose, it was deemed necessary to broaden the scope of the data collected (modifying the POLCAN database with additional variables would not yield the expected results due to its design). This database contains information about all the candidates in the elections in the years 1985–2007 (42,385 candidates), thus adding to it the variables describing the course of careers of a specific category of politicians would expand the base with multiple repetitions of the 'not applicable' or 'not available' values. This would pose a significant problem for the analyses of the

full set. As mentioned in the introduction, such situations in which the traditional hierarchical model database turned out to be an obstacle were the impulse to create new database schemas, including relational and post-relational ones.

Figure 1 POLCAN database – attributes' view

	Name	Type	Width	Dec.	Label	Values	Missing	Colu...	Align	Measure
1	ID	Numeric	8	0	CANDIDATE UNIQUE ID CODE	None	None	8	Right	Scale
2	YEARR	Numeric	4	0	2007 ELECTION-PARTICIPATED IN 2007 SEJIM ELECTIONS	None	None	8	Right	Scale
3	V3R	Numeric	2	0	DISTRICT-ELECTORAL DISTRICT INDIVIDUAL IS FROM 2007	None	None	8	Right	Scale
4	siedzibaR	String	20	0	siedziba	None	None	20	Left	Nominal
5	nrlistyR	Numeric	2	0	Nr listy	None	None	8	Right	Nominal
6	V4R	String	10	0	KW-LIST-PARTY ELECTORAL LIST AFFILIATION 2007	None	None	24	Left	Nominal
7	LISPOS07	Numeric	2	0	2007 POSITION ON SEJIM ELECTORAL LIST	None	None	8	Right	Scale
8	V9R	Numeric	1	0	SEX 2007	{1, MALE}...	None	1	Right	Nominal
9	R11R	Numeric	2	0	OCCUPATION-GR75 JOB TITLE 2007	{1, PRAC NAUKOWY}...	None	8	Right	Scale
10	V10R	Numeric	2	0	AGE -REPORTED IN 2007	None	None	8	Right	Scale
11	R10R	Numeric	4	0	YEAR OF BIRTHDAY	None	None	8	Right	Scale
12	V15R	Numeric	6	0	NUMBER OF VOTES-2007 ELECTION	None	None	8	Right	Scale
13	V14R	Numeric	1	0	ELECTED OR NON ELECTED--WAS ELECTORAL SEAT WON 2007	{0, NON ELECTED}...	None	1	Right	Nominal
14	R4R	Numeric	2	0	KW-LIST-PARTY ELECTORAL LIST AFFILIATION 2007	{1, KPN}...	None	5	Right	Nominal
15	CHAMR	Numeric	1	0	SEJIM ELECTIONS--CHAMBER BEING CONTESTED 2007	{1, POSEL}...	None	10	Right	Scale
16	YEARS	Numeric	4	0	2007 ELECTION-PARTICIPATED IN 2007 SENATE ELECTIONS	None	None	8	Right	Scale
17	V3S	Numeric	2	0	DISTRICT-ELECTORAL DISTRICT INDIVIDUAL IS FROM 2007	None	None	8	Right	Scale
18	siedzibaS	String	20	0	siedziba	None	None	20	Left	Nominal
19	V4S	String	10	0	KW-LIST-PARTY ELECTORAL LIST AFFILIATION 2007	None	None	24	Left	Nominal
20	LISPOS07	Numeric	2	0	2007 POSITION ON SENATE ELECTORAL LIST	None	None	8	Right	Scale
21	V9S	Numeric	1	0	SEX 2007	{1, MALE}...	None	1	Right	Nominal
22	R11S	Numeric	2	0	OCCUPATION-GR75 JOB TITLE 2007	{1, PRAC NAUKOWY}...	None	8	Right	Scale
23	V10S	Numeric	2	0	AGE -REPORTED IN 2007	None	None	8	Right	Scale
24	R10S	Numeric	4	0	YEAR OF BIRTHDAY	None	None	8	Right	Scale
25	V15S	Numeric	6	0	NUMBER OF VOTES-2007 ELECTION	None	None	8	Right	Scale
26	V14S	Numeric	1	0	ELECTED OR NON ELECTED--WAS ELECTORAL SEAT WON 2007	{0, NON ELECTED}...	None	1	Right	Nominal
27	R4S	Numeric	2	0	KW-LIST-PARTY ELECTORAL LIST 2007	{1, KPN}...	None	5	Right	Nominal
28	CHAMS	Numeric	1	0	SENATE ELECTIONS--CHAMBER BEING CONTESTED 2007	{2, SENATOR}...	None	10	Right	Scale

Source: Own compilation based on POLCAN

In the second phase of the project a database was constructed for which the information regarding three hundred and eighteen people was collected, all of whom met the condition of getting into the parliament at least three times in the years 1985-2007. A political biography of each of these individuals has been described using a number of events illustrating the course of their political careers. All events were analyzed at intervals no longer than a parliament term (simultaneously, they constitute measurement periods). Due to the number limitation of examined units' categories, each person with relevant features could be included in the study to the extent to which the course of his or her political career had been documented in the sources used.

Table 1 Content of the Sejm’s members archive data (terms IX–VII)

		term IX	term X	term I	term II	term III	term IV	term V	term VI	term VII
Personal data	full name	✓	✓	✓	✓	✓	✓	✓	✓	✓
	date of birth	✓	✓	✓	✓	✓	✓	✓	✓	✓
	place of birth	✓	✓	✓	✓	✓	✓	✓	✓	✓
	marital status		✓	✓	✓	✓	✓	✓	✓	
Education and occupation	education		✓	✓	✓	✓	✓	✓	✓	✓
	school graduated		✓	✓	✓			✓	✓	✓
	university degree		✓	✓						
	occupation				✓	✓	✓		✓	✓
	languages known		✓	✓	✓	✓				
Participation in elections	date of affirmation	✓		✓	✓	✓	✓	✓	✓	
	parliamentary experience		✓	✓	✓		✓			✓
	list			✓	✓	✓	✓	✓	✓	✓
	party (elections)		✓	✓	✓	✓				
	electoral ward	✓	✓	✓	✓	✓	✓	✓	✓	✓
	number of votes						✓	✓	✓	✓
Parliamentary activity	parliamentary club		✓	✓	✓	✓	✓	✓	✓	✓
	speeches during parliamentary sessions			✓	✓	✓	✓	✓	✓	✓
	interpellations		✓	✓	✓	✓	✓	✓	✓	✓
	queries		✓	✓	✓	✓	✓	✓	✓	✓
	queries regarding current events					✓	✓	✓	✓	✓
	statements				✓	✓	✓	✓	✓	✓
	votings					✓	✓	✓	✓	✓
	parliamentary offices								✓	✓
	list of associates							✓	✓	✓
	Membership in committees and parliamentary groups	committees	✓	✓	✓	✓	✓	✓	✓	✓
sub-committees				✓	✓	✓	✓	✓	✓	✓
permanent parliamentary delegations and bilateral groups									✓	✓
parliamentary groups									✓	✓
Financial declarations	financial declarations						✓	✓	✓	✓
	register of benefits						✓	✓	✓	✓
Additional information	website									✓
	email									✓

Source: Own compilation based on <http://orka.sejm.gov.pl/ArchAll2.nsf>

The three fundamental sources of data that constituted the basis for supplementing the information previously collected were the archives of the Sejm (www.sejm.gov.pl), the Senat (www.senat.gov.pl) and the National Electoral Commission (www.pkw.gov.pl). Therefore, a relational database scheme was developed, that contained information collected partially from the POLCAN database and then supplemented by data from the analysis of the documents produced by the institutions and regarding the progress of professional politicians' careers.

Construction of databases on deputies and senators generates a number of problems, concerning both the range and the process of information collection, as well as the degree of its processing and utilizing⁴. Table 1 presents the contents of the information archive about the deputies along with alterations of the scope within the information collected. They point to significant deficiencies and inconsistencies in the way of presenting profiles of deputies. There are several fundamental problems to be solved while starting the work with the parliament⁵ databases: the data contained therein apply only to persons who have received a mandate, at the same time they do not provide information about what has happened to a given person in the periods before, after or between terms. Databases of the Sejm and the Senat differ in the respect of both the content and structure of the information stored in them (which is also true for successive terms). However, data provided by the National Electoral Commission apply to all candidates, but it lacks a number of basic socio-demographic data. The structure and format of the data collected by the NEC, the Archives of the Sejm and the Senat requires numerous efforts to identify relevant data and transform them in such a way as to form a coherent nature of the examined units.

'PROFESSIONAL POLITICIANS 1989-2007' RELATIONAL DATABASE'S DESIGN

The answer to the problems associated with the creation and management of databases, due to increasing amounts of data, the diversification of the types of data and attempts to maintain consistency in the data structure, was a simpler and easier to use relational data model (Codd 1970). To describe the rules of creation and action of relational databases one needs the following notions:

- A table (a collection of records describing objects),
- A record (a line in a table),
- A field (the smallest part of a record storing a piece of information),
- Data (the smallest units of information about an object),
- Type of data (a form of information recording),
- Data format (a way of submitting and displaying of data),
- A relation (connections between data, described after establishing main and

- foreign keys providing a logical structure to collected data),
- A query (a form allowing searching and combining information stored in the database),
- A filter (serving to limiting a range of search for information to specifically defined information).

In case of the database ‘Professional politicians 1985-2007’ we deal with a relational database that presents data in a form of records or lines, going by the name of “tuples.” This type of a database is based on a relational model, where every relation is described through a header and content. Therefore, we receive a collection of attributes and a group of their values. It means that data are presented in a form of a table with columns dedicated only to a specific type of data. The content of every tuple describes a relation between data in specific cells (e.g., a politician with a given ID has a specific level and type of education). In the database there were used the following types of data:

- Int – positive integers (e.g., a number of received votes);
- Varchar(n) – a string of ‘n’ characters (e.g., name and surname);
- Datetime – date/time (e.g., a date of foundation or disbanding a political party);
- Bit – 0 or 1 (yes/no);
- Default – after adding a line, by default a tuple in a table receives a value without the need to be explicitly submitted;
- NULL – submitting a value for a field is not obligatory.

Construction of the database can be described in the following way:

$$R = \{X, Y, Z\}$$

$$X = \{x_{1n}, x_{2n}, \dots, x_{318n}\}$$

$$n = \{1, \dots, 24\}$$

X – a collection of information concerning a person

$X_{(1-318)n}$ – people

n – person’s attributes

$$Y = \{y_{1m}, y_{2m}, \dots, y_{25m}\}$$

$$m = \{1, \dots, 12\}$$

Y – a collection of information regarding a party system

$y_{(1-25)m}$ – elements of a party system

m – attributes of party system’s elements

$$Z = \{z_{1o}, z_{2o}, \dots, z_{8o}\}$$

$$o = \{1, \dots, 12\}$$

Z – a collection of information regarding an electoral system

$z_{(1-8)o}$ – election

o – attributes of a parliamentary election

Taking under consideration a number of sets (groups of possible values of data) X, Y, Z, the R Relation is a 'n-member' set: $\langle x_{1n}, y_{1m}, z_{1o} \rangle$ where x_{1n} is an element of the X set, y_{1m} is an element of the Y set, z_{1o} is an element of the Z set. Sets X, Y, Z are the domains of the R Relation. The level of the R Relation equals 'n' – that is a number of areas of the Relation.

Person's attributes: 1) sex, 2) date of birth (additionally date of death if applicable), 3) place of birth – name of town, 4) place of birth – size of town, 5) marital status (changes within time of measurements), 6) level of education (changes within time of measurements), 7) type of education (changes within time of the measurements), 8) type of university (regarding a person with the higher education), 9) name of university town (regarding a person with the higher education), 10) practiced profession before the first election that granted one a parliamentary status, 11) practiced profession before the latest election that granted one a parliamentary status, 12) political party membership (changes within time of the measurements), 13) participation in parliamentary elections (during measurements), 14) number on an electoral list (during measurements), 15) electoral ward (during measurements), 16) number of obtained votes (during measurements), 17) obtaining parliamentary seat (during measurements), 18) parliamentary committees membership (during measurements), 19) types of committees membership (during measurements), 20) leading parliamentary committees (during measurements), 21) occupying a position (president, prime/vice minister, minister, speaker/deputy speaker (both lower and upper house) – during measurements), 22) taking into account trust to politicians in polls (during measurements), 23) results from trust to politicians' polls (during measurements), 24) own website (till the end of 2010).

Attributes of party system's elements: 1) political party – name (abbreviation), 2) political party – year of foundation, 3) political party – year of disbanding, 4) political party – ideological core, 5) political party – orientation (socioeconomic, governmental, axiological), 6) political party – participation in elections (during measurements), 7) political party – list number (during measurements), 8) political party – election result – number of votes (during measurements), 9) political party – election result – percent of votes (during measurements), 10) political party – election result – number of seats (during measurements), 11) political party – participation in government forming, 12) act regarding political parties (changes within time of measurements).

Attributes of a parliamentary election: 1) electoral committee – name, abbreviation, 2) electoral committee – type, 3) electoral law – types of electoral

lists (during measurements), 4) electoral law – electoral thresholds (during measurements), 5) electoral law – method of converting votes into seats (during measurements), 6) electoral law – electoral law act (during measurements), 7) election process – number of votes (during measurements), 8) election process – number of candidates taking part in an election (during measurements), 9) election process – number of electoral committees taking part in an election (during measurements), 10) election process – number of electoral committees whose candidates won the seats (during measurements), 11) election process – electoral attendance (during measurements), 12) election process – preterm election (during measurements).

In the database ‘Professional politicians 1989-2007’ the following types of relation can be found:

- one to one (1 – 1) – this kind of relation means that for each record in the B table there is a corresponding record in the A table, e.g., Political Party – Ideology;
- one to many (1 – n) – this kind of relation means that for each record in the B table there are corresponding records in the A table, e.g., Person – Occupation;
- many to many (n – n) – this kind of relation means that for many records there are many other corresponding records (cf. Whitehorn, Marklyn 2003: 99-111).

Creating a coherent structure using as much information required forming collective tables with particular characteristics. The table is described with the use of a so-called main key, which designates a specific set of attributes providing particular and clear features for every record of a relation. “Such a key consists of at least one field (column). None of the values of the main key can be empty. Every line in a table must have a unique value of the main key. The main key is defined during a phase of defining the table.” (Whitehorn, Marklyn 2003: 93-95). In the database used in this project, a constant value is “ID,” in which there is a code stored that clearly defines a given record. In such a case, thanks to using a foreign key, it was possible to reference quickly to data contained within a specific record via another tables. The foreign key is a combination of one or more attributes of a table, which is defined in two or more relations. “Foreign keys are not a necessary element of tables. In other words, every table must possess a defined main key, but all of them do not have to define foreign keys. If there exists a relation between two tables, one of them must have a main key, on basis which there are drew specific data from the second table. In practice, it happens that the majority of tables have foreign keys and it is fully acceptable to have more than one foreign key in a table. If so, a table must have connections with several other tables” (Whitehorn, Marklyn 2003: 97-98).

The content of the tables, which were converted into the ‘Professional politicians 1989-2007’ database appears as follows:

- **Committee** (Committee ID, Committee abbreviation, Committee name, Committee type),
- **Election** (Election ID, Time of measurement ID, Electoral ward ID, Method of votes counting ID, Senat, Attendance, Number of electoral committees taking part in an election, Number of electoral committees taking the seats, Number of candidates taking part in an election, Number of votes, List type ID, Electoral threshold for a political party, Electoral threshold for a coalition, Comment),
- **Electoral committee** (Electoral committee ID, Electoral committee Type ID, Electoral committee name),
- **Electoral Committee type** (Electoral committee type ID, Electoral committee type name),
- **Electoral Law Act** (Electoral law act ID, Electoral law act name),
- **Electoral poll** (Electoral poll ID, Electoral poll date),
- **Electoral threshold** (Electoral threshold ID, Electoral threshold name),
- **Electoral ward** (Electoral ward ID, Electoral ward name),
- **Government** (Government ID, Prime Minister, Government From, Government To, Comment),
- **Ideology** (Ideology ID, Ideology name),
- **Level of education** (Level of education ID, Level of education name),
- **List type** (List type ID, List type name),
- **Marital status** (Marital status ID, Marital status name),
- **Method of counting votes** (Method of counting votes ID, Method of counting votes name),
- **Occupation** (Occupation ID, Occupation name),
- **Occupation type** (Occupation type ID, Occupation type name),
- **Occupation type** (Occupation type ID, Occupation type name, Comment),
- **Orientation** (Orientation ID, Orientation name),
- **Parliament** (Parliament ID, Time of measurement ID, Change of political parties’ act, Comment, Preterm election),
- **Person** (Person ID, Place of birth ID, Person name, Person surname, Person date of birth, Person date of death, Person sex, Person website, Comment),
- **Place of birth** (Place of birth ID, Place of birth name),
- **Political party** (Political party ID, Ideology ID, Political party name, Political party abbreviation, Political party date of foundation, Political party date of disbanding, Political party articulation, Political party nomination, Political party – socioeconomic dimension, Political party – political dimension, Political party – axiological dimension, Comment),

- **Time range of measurement** (Time range of measurement ID, Time range of measurement From, Time range of measurement To, Comment),
- **Type of education** (Type of education ID, Type of education name).
- The attributes of people, elements of the party system and elections were linked to each other creating sets of information that make it possible to freely use the query syntax as per one's intended test hypotheses. In order to achieve it, the following relations were formed:
- **Committee – Parliament – Person** (Parliament ID, Person ID, Committee ID),
- **Education – Level** (Person ID, Level of education ID, Time of measurement ID),
- **Education – Type** (Person ID, Type of education ID, Time of measurement ID),
- **Electoral Committee – Measurement** (Electoral committee ID, Time of measurement ID),
- **Person – Election** (Person ID, Election ID, Electoral ward ID, Electoral committee ID, List type ID, Person – Election – List number, Person – Election – Number on the list, Person – Election – Number of votes, Person – Election – Percentage of votes),
- **Person – Government** (Government ID, Person ID),
- **Person – Measurement** (Time of measurement ID, Person ID, Marital status ID),
- **Person – Occupation** (Person ID, Occupation ID, Occupation type ID),
- **Person – Occupation** (Time of measurement ID, Occupation type ID, Person ID),
- **Person – Parliament** (Parliament ID, Person ID),
- **Person – Political party** (Political party ID, Person ID, Time of measurement ID),
- **Political Party – Election** (Political party ID, Election ID, Electoral committee ID, Political party – Election – Result, Political party – Election – List number, Political party – Election – Percentage of votes, Political party – Election – Number of votes for a list, Political party – Election – Government, Comment),
- **Poll – Person** (Electoral poll ID, Person ID, Electoral poll result),
- **Poll – Political party** (Electoral poll ID, Political party ID, Electoral poll result).

Thanks to such a construction of a database, it becomes possible to look for information, for example, according to the following principle: There is a person $X = \{x_{1n}, x_{2n}, \dots, x_{318n}\}$, with a feature $n = \{1, \dots, 24\}$, who in an election $Z = \{z_{1o}, z_{2o}, \dots, z_{8o}\}$, described by $o = \{1, \dots, 12\}$, represented a political party $Y = \{y_{1m}, y_{2m}, \dots, y_{8m}\}$, which has an attribute of $m = \{1, \dots, 12\}$.

Due to the fact that ‘Professional politicians 1989-2007’ database describes data undergoing changes throughout time, in the majority of cases we deal with a situation n-n, e.g., person-occupation: one occupation could be held by many people, and a person could occupy many positions throughout many years. In case of relation n-n (many to many), there are additional tables required, in which one joins together main keys from A and B tables into pairs, so that one is able to know which data can be combined. In such tables we can put extra information, e.g., a date from – to. Thanks to that, data in the database should not repeat themselves.

During the database design process, as well as the data collection phase, there occurred situations when it seemed logical to supplement the scheme, even if the collected data were not subsequently fully used. This is yet another argument in favor of the collection of information in the form of a relational database that can be expanded or modified later on, but primarily it can be used for additional in-depth analyses.

SUMMARY

The project presented in this article has provided material applicable to conduct the analysis of the course of political careers in the years 1985-2007 in Poland. Analyses concern, in particular: the history of running for parliament seats by professional politicians, the history of their successes and failures along with electoral effectiveness, socio-demographic factors impacting the professional politicians’ careers’ course, attributes of the party system that affect the chances of pursuing a successful political career, the characteristics of the electoral system and their impact on prolonging political careers. Against this background one can present the course of politicians’ careers, taking into account the stage of preparation to practice politics. Confrontation with competitors during the electoral game and the very political career’s course are viewed from their position in the parliamentary and governmental structures.

By considering the variables specific for shifts concerning both electoral and party systems, and also by putting them in the broader context within the scope of considerations about political system’s studies, elements of electoral law and the functioning of the election market, it becomes possible to identify the reasons for which the Polish electoral system adopted specific solutions determining the course of political careers. Those explanations assume intentional actions of political actors, as well as institutional actors organizing the new constitutional rules together with mechanisms for selecting the election winners. Describing characteristics of shifts within electoral and party systems in Poland in the years 1985-2007 is a prerequisite for the proper selection of variables and indicators of the course of careers. This creates the need to describe the tools available

for a politician constructing one's parliamentary career, and thus it also gives opportunity to answer the question about what macro-factors build a frame for the potential use of these tools. This allows for a later analysis of the effectiveness of their use and the scope of introduced modifications. It adds a possibility of a relatively free-form interpretation of the rules that govern them. Analyses of this type let one conclude in a legitimate way not only what attributes make people proficient within the political arena, but also what characteristics of professional politicians can be regarded as favoring the continuation of political careers and modeling their progress in the future. Dependencies are designed to emphasize an active role of parliamentary candidates in the process of making use of solutions as far as the electoral law's field and the political parties' market situation are concerned (Nyćkowiak 2013).

The completion of the basic goal of the project, which was to analyze the institutional determinants of political careers' courses in the years 1985-2007 in Poland and a reconstruction of professional politicians' political biographies, was made possible by the use of a particular look at the data sources, their structure and capacity for harmonization. This article presents the basic principles of the research project, which uses the opportunities offered by relational databases. It was described how the database was designed to allow a flexible use of the collected data for the analyses. The concept, scheme and functionality of the aforementioned database prove how beneficial it can be when a research project successfully uses advanced application tools⁶.

Therefore, one has to point out advantages of collecting and managing data in a form of systems. Among others, these are:

- eliminating multiplied registrations of duplicated data,
- limiting space required to store collected data,
- increasing a logical control of data collections, an effective access control,
- a possibility of presenting complicated relations within data and an efficient obtaining of data from the database via an optimization of queries,
- providing an option of recovering data thanks to backup copies that take into account consecutive stages of the process of creating the database,
- adapting the database interface for needs of the users with different requirements and varied levels of a preparation to use the database (Elmasri, Navathe 2005: 34-40).

These advantages justify the inclusion of the preparation of databases' phase into the majority of large research projects. Therefore, it is recommended to use relational databases designed for the needs of researches within the social sciences, especially in the case studies taking under consideration the longitudinal type of data.

NOTES

- 1 Electoral successes did not have to be consecutive.
- 2 What, in an indirect way, although in accordance with the logic of institutions' functioning, limits equal access and competition. As rightly pointed out by Wesołowski, when a politician reaches one's goal of becoming a parliamentarian, then one begins one's "political socialization" that would allow becoming a professional politician in the future (increasing one's chances in comparison to people within politics that do not get into the parliament despite numerous run in elections) (See Wesołowski 1998: 9). The key question is, of course, how to determine boundary conditions that would classify someone as a politician, and at the same time would let to form a category of professional politicians. Following the lead proposed by Włodzimierz Wesołowski, and referring to the concepts of Max Weber, one of possible solutions would be to turn to the selected institutions of political life that have the greatest potential for decision-making and without which it is impossible to conduct any systematic policy of the state (Wesołowski 1998). Therefore, it is assumed that the key institution being both a cause and a goal of the people who pursue political careers in Poland after 1989, is the parliament (cf. Borchert, Zeiss 2003).
- 3 The CONSIRT Program has been established by the Departments of Sociology and Political Science of The Ohio State University and the Institute of Philosophy and Sociology, and the Graduate School for Social Research at the Polish Academy of Sciences (IFiS PAN and GSSR PAN, respectively). Source: consirt.osu.edu
- 4 Naturally, there are certain doubts about the quality of the data contained in official documents. Most often, one's attention is drawn to the facts that: (1) the range of data registered by institutions does not meet the needs of social researchers, who do not find interesting information within such data, (2) many institutions do not modify the range of collected data, although they are notified that a given system is insufficient for many reasons, (3) the access to official documents is either limited or completely denied. Often, it happens due to the fact that by their very nature, and in the mind of their creators, official documents are not meant to be used for researches, thus they do not have to be made available for such activities, (4) data registered by institutions are selective – they 'see' only those facts that are likely to happen due to the institution's activity profile or an organizational structure.
- 5 The website of the Sejm in The Deputies Archive tab (<http://orka.sejm.gov.pl/ArchAll2.nsf>) gives access to a search engine that allows to

find parliamentarians through several interesting criteria: term of the parliament, full name, sex, education, electoral ward, party club/deputies' club/group, function within the parliament and other information included in the description of a parliamentarian. With such a search engine one can, for example: 1) select all parliament members, who only had secondary education during a term (a problem: What about the people whose education has changed between terms? A solution: one should check the level of education of each person during each term, to find out if and when the level of education has changed), 2) select all members, who won their seats running from the electoral lists of The Democratic Left Alliance (a problem: What about people whose party affiliation changed between terms and those that left and came back to their party? A solution: one should check party affiliation of each member in each term, in which one won a mandate, to find out whether there were any changes in their party affiliation)

- 6 Presented below are examples of queries which yielded results that in turn, allowed processing data that later provided obtaining answers to selected research problems:

Example 1 Query selects parties participating in elections to Sejm/Senat during the period of measurement.

```
SELECT      PartiaPolityczna.pp_Id, OkresPomiaru.op_Id
FROM        PartiaPolityczna INNER JOIN
           Partia_Wybory ON PartiaPolityczna.pp_Id = Partia_Wybory.pp_Id
           INNER JOIN
           Wybory ON Partia_Wybory.wy_Id = Wybory.wy_Id INNER JOIN
           OkresPomiaru ON Wybory.op_Id = OkresPomiaru.op_Id
WHERE       Wybory.wy_Senat = 0 · przy wyborach do Senatu '1'
ORDER BY   PartiaPolityczna.pp_Id, OkresPomiaru.op_Id
```

Example 2: Query selects professional politicians who began their political careers in 1989.

```
SELECT      Osoba.os_Id, Osoba.os_Imie, Osoba.os_Nazwisko
FROM        Osoba INNER JOIN
           Osoba_Wybory ON Osoba.os_Id = Osoba_Wybory.os_Id INNER JOIN
           Wybory ON Osoba_Wybory.wy_Id = Wybory.wy_Id INNER JOIN
           OkresPomiaru ON Wybory.op_Id = OkresPomiaru.op_Id
WHERE       (OkresPomiaru.op_Id = 2)
AND         Osoba.os_Id NOT IN
           (
SELECT      Osoba.os_Id
```

```

FROM      Osoba INNER JOIN
          Osoba_Wybory ON Osoba.os_Id = Osoba_Wybory.os_Id INNER JOIN
          Wybory ON Osoba_Wybory.wy_Id = Wybory.wy_Id INNER JOIN
          OkresPomiaru ON Wybory.op_Id = OkresPomiaru.op_Id
WHERE     (OkresPomiaru.op_Id = 1)
          )

```

Example 3: Query selects people presenting their level and type of education, participation in elections to Sejm/Senat, place on the list and the number of votes received.

```

SELECT    OkresPomiaru.op_Id, Osoba.os_Id, Osoba.mu_Id, Edukacja_Poziom.
          pw_Id, Edukacja_Rodzaj.rw_Id, Osoba_Wybory.ow_NrNaLiscie,
          Osoba_Wybory.ow_LiczbaGlosow
FROM      Osoba INNER JOIN
          Edukacja_Poziom ON Osoba.os_Id = Edukacja_Poziom.os_Id INNER
          JOIN
          Edukacja_Rodzaj ON Osoba.os_Id = Edukacja_Rodzaj.os_Id INNER
          JOIN
          OkresPomiaru ON Edukacja_Poziom.op_Id = OkresPomiaru.op_Id AND
          Edukacja_Rodzaj.op_Id = OkresPomiaru.op_Id INNER JOIN
          Wybory ON OkresPomiaru.op_Id = Wybory.op_Id INNER JOIN
          Osoba_Wybory ON Osoba.os_Id = Osoba_Wybory.os_Id AND Wybory.
          wy_Id = Osoba_Wybory.wy_Id
ORDER BY  Osoba.os_Id

```

Example 4: Query selects political parties according to the level and type of education of professional politicians representing a given organization.

```

SELECT DISTINCT  PartiaPolityczna.pp_Id, Osoba.os_Id,
                 PoziomWyksztalzenia.pw_Id,
                 RodzajWyksztalzenia.rw_Id, OkresPomiaru.op_Id
FROM      Osoba INNER JOIN
          osoba_partia ON Osoba.os_Id = osoba_partia.os_Id INNER JOIN
          PartiaPolityczna ON osoba_partia.pp_Id = PartiaPolityczna.pp_Id
          INNER JOIN
          OkresPomiaru ON osoba_partia.op_Id = OkresPomiaru.op_Id INNER JOIN
          osoba_pomiar ON Osoba.os_Id = osoba_pomiar.os_Id AND
          OkresPomiaru.op_Id = osoba_pomiar.op_Id INNER JOIN
          edukacja_poziom ON Osoba.os_Id = edukacja_poziom.os_Id AND
          OkresPomiaru.op_Id = Edukacja_Poziom.op_Id INNER JOIN
          PoziomWyksztalzenia ON edukacja_poziom.pw_Id =
          PoziomWyksztalzenia.pw_Id INNER JOIN
          edukacja_rodzaj ON Osoba.os_Id = edukacja_rodzaj.os_Id AND
          OkresPomiaru.op_Id = Edukacja_Rodzaj.op_Id INNER JOIN
          RodzajWyksztalzenia ON edukacja_rodzaj.rw_Id =
          RodzajWyksztalzenia.rw_Id
WHERE     (OkresPomiaru.op_Id=5) AND (PartiaPolityczna.pp_Id=1)
GROUP BY  PartiaPolityczna.pp_Id, Osoba.os_Id, PoziomWyksztalzenia.pw_Id,
          RodzajWyksztalzenia.rw_Id, OkresPomiaru.op_Id
ORDER BY  Osoba.os_Id, OkresPomiaru.op_Id

```

REFERENCES

- Anderson, C. J. 2007. "The Interaction of Structures and Voter Behavior." In: Russell J. Dalton, Hans-Dieter Klingemann (eds.). *Oxford Handbook of Political Behavior*, Oxford University Press.
- Arthur, M.B., Hall, D.T., Lawrence, B.S. 1989. "Generating new directions in career theory: the case for the interdisciplinary approach." In: Michael B. Arthur, Douglas T. Hall, Barbara S. Lawrence (eds.). *Handbook of Career Theory*, Cambridge University Press.
- Balicki, A. 2006. *Analiza przeżycia i tablice wymieralności*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Bolero, F. 2007. *Should I Stay or Should I Go? An Event History Analysis of Congressional Carter Decisions In Multi-Member Districts*, Prepared for delivery at the 65th Annual Conference of the Midwest Political Science Association, Chicago, April 12-15.
- Borchert, J., Zeiss, J., eds. 2003. *The Political Class in Advanced Democracies. A Comparative Handbook*, Oxford University Press.
- Chmielewski, P. 2010. *New Institutionalism: a Platform for Productive Integration in Social Sciences*, Warsaw Forum of Economic Sociology, 1(1): 9-88.
- Codd E. F. 1970. "A Relational Model of Data for Large Shared Data Banks." *Communications of the ACM*, 13(6): 377-387.
- Elsuari, R., Shamkant B. Navathe, 2006. *Wprowadzenie do systemów baz danych*, Wydawnictwo Helion.
- Heath, A. 2007. "Perspectives on Electoral Behavior." In: Russell J. Dalton, Hans-Dieter Klingemann (eds.) *Oxford Handbook of Political Behavior*, Oxford University Press.
- Mach, B.W. 2000. "Nomenklatura, Communist Party Membership, and Advantage on the Labor Market." In: Kazimierz M. Słomczyński (ed.). *Social Patterns of Being Political. The Initial Phase of the Post-Communist Transition in Poland*, IFiS Publishers, Warsaw.
- March, J. G., Olsen J.P. 1984. "The New Institutionalism: Organizational Factors in Political Life." *The American Political Science Review*, 78(3): 734-749.
- . 2006. "Elaborating the "New Institutionalism." In: R.A.W. Rhodes, Sarah A. Binder, Bert A. Rockman (eds.). *The Oxford handbook of Political Institutions*. Oxford: Oxford University Press.
- Norris, P., Lovenduski, J. 1995. *Political Recruitment. Gender, Race and Class in British Parliament*, Cambridge University Press.
- Nyćkowiak, J. 2013. *Perspektywa neoinstytucjonalna w analizach przebiegu karier politycznych*, Acta Universitatis Lodziensis, Folia Sociologica, 46: 75-87.
- Pereira, L.C., Maravall, J.M., Przeworski, A. 1993. *Economic Reforms in New Democracies. A Social-Democratic Approach*, Cambridge University Press.
- Post, B. 1995. „Aktorzy polskiej sceny politycznej – amatorzy czy profesjonalści?" In: Włodzimierz Wesołowski i Irena Pańków (eds.). *Świat elity politycznej*, IFiS PAN Warszawa.
- Przeworski A. 1991. *Democracy and The Market. Political and Economic Reforms in Eastern Europe and Latin America*, Cambridge University Press.

- Raciborski, J. 2006. *Elity rządowe III RP 1997-2004: portret socjologiczny*, Wydawnictwo TRIO.
- Ragin, C. C. 2006. "Set Relations in Social Research: Evaluating Their Consistency and Coverage." *Political Analysis* 14: 291-310.
- Shabad G., Słomczyński K.M. 2000. "Political biographies and Support for Systemic Change." In: Kazimierz M. Słomczyński (ed.). *Social Patterns of Being Political. The Initial Phase of the Post-Communist Transition in Poland*, IFiS Publishers, Warsaw.
- Shabad G., Słomczyński K.M. 2004. "Inter-party Mobility Among Parliamentary Candidates In Post-Communist East Central Europe." *Party Politics*, 10(2): 151-176.
- Shabad G., Słomczyński K.M., Zieliński J. 1999. "Fluid Party System, Electoral Rules and Accountability of Legislators In Emerging Democracies. The case of Ukraine." *Party Politics*, 14(1): 91-112.
- von Beyme, K. 2005. *Współczesne teorie polityczne*, Wydawnictwo Naukowe Scholar, Warszawa.
- Wesołowski, W. 2001. "Change and Continuity: Four Polish Parliaments 1989-2001." *Central European Political Science Review*, 2(4): 25-50.
- Wesołowski, W., Post B., eds. 1998. *Polityka i Sejm. Formowanie się elity politycznej*, IfiS, Warsaw.
- Whitehorn, M., Marklyn B. 2003. *Relacyjne bazy danych*, Wydawnictwo Helion.

Justyna Nyćkowiak received her Ph.D from the Institute of Philosophy and Sociology, Polish Academy of Sciences in 2012. She is an assistant professor in the Institute of Sociology of the University of Zielona Góra.

E-mail: jjustan@gmail.com