

## Research objective

The research objective is an analysis of **institutionalization of political parties since the democratic transition in five Central European countries (Poland, Czech Republic, Lithuania, Slovakia and Hungary) as reflected in parliamentary discourse and voting**. The goal will be accomplished using a set of data mining methods of analysis of textual (parliamentary debates) and numerical (roll-call voting) data on an innovative database which will be created in the project. We will analyse in particular:

- how and by which parties new political issues were introduced to public discourse;
- how the semantic content of individual issues, their context and salience changed in the parliamentary debates;
- what is the relationship between the level of cohesiveness in parliamentary speeches regarding key political issues and party capability of influencing the behaviours of other political parties;
- what are relationships between: party cohesion in debates, unity in legislative voting and party splits;

## Research project methodology

The methodological aim of the project is the **evaluation of existing methods of defining the ideological positions of individual members of the parliament (MPs) and parliamentary clubs and - on that basis - modification and creation of new methods for simultaneous analysis of parliamentary debates and roll-call votes, while taking into account metadata related to MPs and debates and macro-level variables**. This will increase the accuracy of analysis and prediction of the behaviours of MPs and the activities of parliamentary clubs thanks to the use of various statistical and machine learning techniques for classifying, analysing links and reducing the dimensionality of textual (parliamentary speeches) and numerical (voting) data.

The practical objective - necessary for both theoretical and methodological objectives, but also a goal itself - is the **creation of the up-to-date cross-referencing database covering: 1) the complete, structured collection of speeches** (transcripts of proceedings and speech metadata) in the parliaments of Poland, Czech Republic, Lithuania, Slovakia and Hungary - approximately 1.5 million individual speeches, **2) the complete results of roll-call votes in the parliaments** - approximately 1 million individual MPs' votes, **3) full available information on MPs** including their socio-demographic data and political career, **4) relevant information from the ParlGov database**, including data on elections, composition of parliaments and governments, and several variables related to ideological positioning of the parties.

## Expected impact of the research project

**Innovation of the proposed project lies in flowing from the joint analysis of parliamentary debates, roll-call votes, and political metadata on micro and macro levels**. Several groups of techniques of quantitative analysis of textual and numerical data (classification, dimension reduction and network analysis) from the fields of statistics, machine learning and natural language processing will be used in the study. Some of them were already used to analyse parliamentary data, and application of some of them, to the author's best knowledge, will be an innovation in political science.

Data will be analysed in a comparative perspective, taking into account macro variables in order to isolate universal effects. Due to the lack of such comparative analyses, the project is innovative in theoretical and empirical terms, and its results will expand the current state of knowledge and, therefore, **enrich Polish and global science in the field of research on parliamentary elites and party systems**. The impact of the project's results on the development of the research field and scientific discipline will also **stem from the creation of a unique database**.