MEDem Concept and Methodology

(a) Concept

MEDem's overall concept is the *provision of harmonised interoperable data accessible to all user communities* interested in the study of electoral democracy. This is oriented foremost towards political scientists, but it will also allow for inter-disciplinary research by sociologists, economists, legal experts, and researchers in international relations, media studies and political psychology, among other disciplines.

Building on the broad existing knowledge-base from past research and existing projects, this action will address the efficiency gap produced by the scattered community and data collection landscapes, and will enable better study of the interface between democratic procedures and policymaking.

Europe is home to many research programs/infrastructures for collecting data directed at better understanding electoral democracies. Historically these developed around national election studies. To summarize what was already said, over time these national research projects have widened their focus from studies of citizens to studies that include the candidates and parties these citizens vote for and the media that report on the behaviour and utterances of parties and candidates. In addition, many additional projects focus on other components of electoral democracies, such as parliaments, governments, and political parties. Projects also became more comparative and started to coordinate across countries on parts of data collections. While less established in many cases, various comparative research projects also exist that monitor parliaments, government and the media. These projects are important for studying the link between what happens at elections and the inter-election period, and facilitate the comparative study of important democratic processes beyond elections.

The *ambition of MEDem is to bring together all this existing wealth of data into a future infrastructure* that will focus on future data collections that are, independently from MEDem, already funded through various mechanisms. Practically this means that the existing and future data need to be harmonized and linked to satisfy four underlying research perspectives: the long-term perspective, the comparative perspective, the multi-actor perspective and the multi-level perspective. Specifically:

i.   *Long term perspective*

Many phenomena can only be studied through the perspective of time, allowing researchers to understand the importance of path dependency, the potency of tipping dynamics, and the time variation of incubation periods for ideologies, social movements and social change. Fifty years of studying voting behavior using election studies has cumulated in a significant wealth of data that can be the basis of

time series. A specific Project called True European Voter (TEV) Cost-Action has already given the first results by harmonising 30 years of electoral data that will be a core component of MEDem.

ii.   *Multi-actor perspective*

For scientist to answer electoral democracy research questions they need *information on the character of parties and candidates, the functioning of parliaments and governments, the attitudes and behaviours of voters, the election results and media coverage of political activities*. Preferably, all these types of data need to be connected to each other. Practically these connections happen on the basis of key variables. A typical example is connecting a voter with the party he or she voted for on the basis of their policy preferences. Party facts is a project that allows connection between various ways a party has been coded and enables easy linking between party data and other types of data. Party Facts will be incorporated in MEDem.

iii.   *Comparative cross-country perspective*

The study of elections cannot be confined in individual countries, where many important features are fixed, but need to *be studied in a comparative framework*. To provide adequate variance in the character of parties and candidates, the functioning of parliaments and governments, and in media coverage of their activities and pronouncements, data manifestly needs to be made available on multiple democracies covering as many countries as possible over as long a time-span as possible. Inter-relationships and inter-dependence of European democracies today can only be researched with the use of high quality comparable data. Here there are significant research works and ongoing projects that MEDem will build on – for example, the CSES, MARPOR, CCS, CAP and NES projects. These projects are associated to MEDem and will provide their already harmonised data as additional core components of MEDem.

iv.   *Multi-level perspective*

Governance within the European Union is of a multilevel nature. Voters are clustered in municipalities, regions, countries, but are also represented by transnational party groups in the European Parliament. The relevant contexts, being geographical or political, vary depending on the research question at hand. However, the ideal is to be able to take all relevant contexts into account. For this, it is necessary to have a complex data structure in place that can be fed with high quality interlinked data. MEDem will provide the means for researchers to link data from various actors as they need to.

MEDem is an integration exercise of various national and comparative data collection projects on different political actors. In addition, it will connect a variety of data collection programs focused on contextual data regarding the economic and political circumstances in which elections are conducted or

that result from those elections (most importantly the details of election outcomes per district and of the governments that form on the basis of such election outcomes).

The following comparative projects have been involved in the conceptualization of MEDem:

Monitoring Voters	Monitoring parties and Candidates
Comparative Study of Electoral Systems (CSES)	Manifesto project (MARPOR)
VAA Research Network – Data on parties and candidate	Comparative Candidate Survey (CCS)
European Election study (EES)	VAA Research Network – Data on party positioning
National Election Studies in many European countries	Chapel Hill Expert Survey (CHES)
	Party Facts
	Constituency-Level Election Archive (CLEA)
Monitoring institutions	**Monitoring media and policy agendas**
Parliamentary Research Network	Media Research Network
Parliaments and Governments Database (ParlGov)	Comparative Agendas Project (CAP)
Executive Approval Project (EAP)	

Central to MEDem operations is the leveraging of stakeholder knowledge, motivation and activity. In the long run we expect MEDem to be a *user-powered infrastructure* whose purpose is primarily to coordinate existing data-collection activities directed at monitoring the good-functioning of democratic processes of governance, resulting in the re-direction of those data-collection activities into more productive channels. A critical feature of MEDem, designed to maintain stakeholder enthusiasm and commitment, is that its operations are in no way top-down. MEDem does not direct the coordination of data collection by stakeholders; that is done by stakeholders themselves. What MEDem provides are the institutional structures that make coordination possible.

It is important to stress that *MEDem itself will not be responsible for data collection*. Rather, it *co-ordinates existing pan-European distributed data-collection centres and projects and encourages new initiatives as necessary to provide nodes in all European countries*. Thus it builds on existing projects as far as possible, also taking active steps to encourage the establishment of new cross-country research projects that would fill gaps in the MEDem data structure (for example in media studies and parliamentary behaviour studies that have up until now no established cross-national data providers).

MEDem will also *use the existing and well-established research archive infrastructure CESSDA ERIC* and its national service providers to make data FAIR (Findable, Accessible, Interoperable, and

Reusable) when it establishes a new and innovative online access facility as a Portal to the linked data for the electoral democracy research community. It will coordinate with ESS ERIC and SHARE ERIC to create synergies in methodological standards and in tools to support data collection.

(b) Methodology

Each of these data collection enterprises are currently conducted separately. They follow their own internal logic using data collection instruments specific to each project and vary in the range of countries they cover. This makes it difficult or impossible to use the data collected from different projects in conjunction with each other. Let us imagine one simple example. Three different projects collect data about the activity of political parties. Project 1 focuses on data regarding candidate attitudes and behaviours using a survey where candidates are the respondents. Project 2 collects data about policy positions of political parties using the coding of their published party manifestos. Project 3 codes media reports on the activities of political parties and candidates. Using the entity of a political party, it should be simple to combine the data from the candidates of the party, with the media reports on the party and the party policy positions. That would allow a full report on the said party. This idea, simple as it sounds, is often not straightforward because of different historical traditions in the collection of each type of data. Much research potential is locked away beyond use because of these data incompatibilities.

MEDem seeks to *reduce and eventually eliminate these debilitating incompatibilities by introducing procedures for harmonizing the data* (*ex-post* in a first step and *ex-ante* in a second step) that flow from different data collection infrastructure/programs in different countries and Europe-wide. Just as importantly it seeks to provide a "viewing window" (more properly a "data portal") through which the contents of the resulting complex data structure can be viewed and/or engaged by citizens and scholars. MEDem aims to enable cutting-edge European research regarding the functioning of European democracy in its broadest sense.

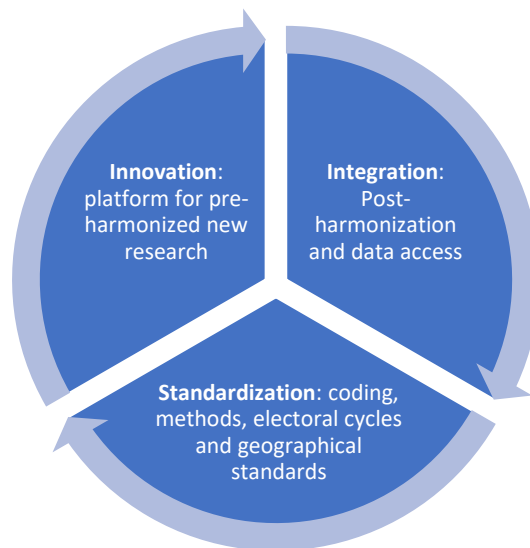Comparative social science projects mainly have three tasks:

1. *They integrate and document collected data*, which is then made available to researchers
2. They *define standards* for data collection that are used in different countries and /or by different teams
3. They *organize communities*, projects, institutions and researchers to make sure that data is collected, archived, and used and to promote innovation in research.

While the above three key functionalities are found in all existing comparative projects, *MEDem is now raising this functionality to a higher level so that research-potential that exists in data already being*

produced by the different projects is fully realized through the establishment of cross-project linkages. This means that the data already being produced in the different projects become usable in a straightforward fashion across existing projects. In some cases MEDem also facilitates the creation of collaborative projects, as it does for the National election studies (NESs). Beyond the core NES data already harmonized by CSES – which is already a comparative project – remaining NES data still stands in need of harmonization if its full value is to be realized. And in some fields, such as media studies or parliamentary studies, projects are not yet at the necessary level of maturity to be integrated into MEDem. In such fields, MEDem will support the creation of strong comparative projects.

MEDem seeks to provide these three functionalities across projects: to provide data integration and harmonization and make data usable across projects, to define standards and help integrate and, to some extent, to organize the different sub-communities (where such organization is currently lacking). MEDem will establish the provision of the following key services to the scientific community:

MEDem core services to the research community

Integration of existing data

i. Post-harmonization: *Integration of different existing data sets by making them comparable and linkable*. In some cases (for example CSES, CCS, CAP, MARPOR etc.) country-comparable (comparative) data already exists, in other cases (such as NES data beyond that already harmonized by the CSES, media data, parliamentary data) comparative data do not as yet exist beyond individual scholarly projects, each largely project-specific in concepts and coding. For both types of harmonization

a multi-method approach will be employed that will build on a "conceptual map". Based on this map, MEDem will start a deliberation process seeking agreement between different projects on common grounds for concept and scale harmonisation. At a second step, the project will seek to "translate" original coding categories into common measurement scales tuned to particular outcome (dependent) variables. Public good surveys designed for wider communities have the problem that it is impossible to anticipate which outcomes scholars might wish to study. A similar problem will be faced by MEDem. To solve this the community will have to pre-decide which outcomes and processes will be studied and harmonised, based on long established research questions regarding electoral democracy, but also innovative questions that can only be answered because MEDem standards have been adopted.

ii. Access: Data Access in in general is not an issue for the electoral studies community. Most of the data are available for download either through data archives or project websites. The problem is mainly that researchers have to have the knowledge of where the data are located before they can find them. There is *no overview or single point from whch to search and browse* them. The second problem is that the data are not always correctly documented or not documented in a standard way. Thus, even though one can access them, usability is not always ensured. MEDem will provide a single entry point for data users – even though data may be stored at different places, mainly existing national data archives. It will assure that existing data follows a common documentation standard so that data that originates from different projects can be analysed in conjunction. Also MEDem will provide documentation standards, training and how-to's to support the community in making the most of the existing data by ensuring maximum transparency and replicability of harmonisation processes. This will help all current and future researchers to evaluate the quality of the datasets, of the linkage mechanisms and of the assumptions and code used to harmonise and merge them.

Standardization for future data collections

The further development of the conceptual map mentioned above will specify how specific key concepts that can work as linking variables should be coded; also how key concepts that appear in more than one type of data should be coded for eventual integration, using instruments suited to different data types. *The main goal here is to eventually make post-harmonization redundant* (to the extent of course that this is possible and plausible), as most of the new data produced under the auspices of MEDem would be pre-harmonized. These instruments will focus on four main domains:

i. Coding schemes on parties, constituencies, candidates etc. that allow for comparability and matching of existing data within the MEDem framework. Although some coordination activities have already

been conducted (e.g. the TEV Cost Action) coordination has so far only been sought within specific projects (e.g. National Election Studies);

ii. Methodological standards on data collections for different parts of MEDem. MEDem, in collaboration with the existing projects, will research the different ways in which data are collected and coded and develop guidelines for "best practice" and quality control for data collection.

iii. The cycle of elections in European countries will lie at the heart of pre-harmonization efforts, with some change in data collection instruments (Modules) from cycle to cycle in order to address evolving research questions. Since different countries hold elections at different frequencies (and some countries occasionally experience early elections), for some countries more than one election will fall within the period of currency of a particular Module.

iv. The changing shapes of geography: The way countries divide their administrative districts has fundamental impact on the way data are collected. A typical example is the shapes and sizes of electoral districts. We might collect over-time data on one district, but this district never remains the same, as governments freely re-draw the physical boundaries including more or fewer individuals. This imposes significant limitations to over time comparisons of the same geographical units – limitations that must be mitigated as far as reasonably possible by bridging methodologies that will be investigated by MEDem.