Outline "Monitoring Electoral Democracy (MEDem)"

February 2020

Table of Contents

1   Summary

Monitoring Electoral Democracy (MEDem) is an emerging European research infrastructure that links many existing comparative as well as national projects to study the functioning of electoral democracy in Europe and brings comparative research on European electoral democracies to a new level.

Understanding voters, parties, and elites (and how they influence policy-making) is crucial to monitoring the functioning of electoral democracy in modern societies. Even though research on electoral democracy is a well-established research area in Europe and elsewhere, further advancements in this field is still restricted due to incompatibility between data sources, projects and countries. The proposed research infrastructure for Monitoring Electoral Democracy (MEDem) is designed to overcome these obstacles and release the underlying potentials in already collected data. The central aim of MEDem is to increase the harmonization, integration and accessibility of data collections regarding various aspects of representative democracy in Europe. With increased coordination and integration, and with improved accessibility of already existing and future data collections, MEDem will provide excellent opportunities for a wider community both within academia and for the broader public to monitor the quality of electoral democracy in Europe.

MEDem is aiming for a status as a distributed European Research infrastructure, linking together existing national infrastructures and international networks. While this is an ongoing initiative, the formal goal of MEDem is to gain a position on the ESFRI roadmap in recognition of its central and strategic scientific importance for European social and political research.

2   The scientific case for MEDem

Well-functioning democracies lie at the foundation of modern Western societies as they provide for citizen representation in the executive and legislative branches of power. To understand in depth how citizens, elites, parliaments, governments and media interact and relate to each other in the democratic process is therefore central for the understanding of modern societies.

2.1   Challenged democracies and the rationale for MEDem

Democracies face a number of very evident challenges, to name a few:

➢ A wave of populism and nationalism is visible in Europe that challenge established institutions and parties;

➢ This development is accompanied by growing mistrust of public institutions and politicians in different countries and also, in some instances, by decreasing voter turnout and satisfaction;

➢ The digitalization of society imposes changes in interpersonal and political communication and can bring about changes to the way elections are being conducted – indeed to the very way electoral systems function;

➢ Established media systems are under pressure as they experience the erosion of their previous business models. Many people obtain information of political relevance through social media platforms, and this information is often selective and sometimes biased;

➢ Party systems, having been stable for a long time have recently been shaken in many countries with established parties struggling or disappearing and new parties and politicians appearing on the scene and even achieving government office;

➢ European societies are transforming due to processes such as migration, urbanization and demographic changes. European societies are also aging. The changing of the societal structure poses challenges in all functions of democracy;

➢ The future of the EU itself is put into question by popular sentiment as manifested in referendum and election outcomes;

> ➤ Fifty years of studying voting behaviour has taught us much about voters' party choices, but recent electoral surprises seem to have been associated especially with new departures in media usage and with unconventional candidacies that affect the context within which elections take place. Furthermore, new political conflict lines and changes in party systems have led to new coalition patterns and even government instabilities. In Europe, the multi-level democracy of the European Union presents a challenge in itself.

These matters and challenges cannot be studied from within the confines of individual countries — confines that limit the extent to which context can vary. Such matters can also not be studied from a single perspective, be it voters, or media, or candidates. To respond to these challenges policy-makers need to understand the causes and consequences of new political developments. For scientist to answer these questions requires solid data capturing change in democracies over time in many countries. 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. A single country perspective cannot meet this need. Also, European democracies today are interrelated and inter-dependent and must be researched accordingly.

2.2   The MEDem scope

The research infrastructure Monitoring Electoral Democracy (MEDem) releases the potential locked away in data that are already being collected. This potential is currently inaccessible due to the idiosyncrasies of country- and topic-specific data collection and coding conventions.

MEDem brings together scattered European infrastructures, research initiatives, centres, and projects under one umbrella to facilitate data harmonization, to enable linking and integration between various projects and data sources and to stimulate cooperation and collaboration. MEDem will thus contribute to innovation and excellence in the study of electoral democracy. It situates existing and future voting data within broader social and dynamic contexts of elite, party, parliament and government activities as well as those of the media in its broadest sense, over the passage of time. In broad terms, the infrastructure hosts the following main components.

Figure 1 Monitoring Electoral Democracy, the components
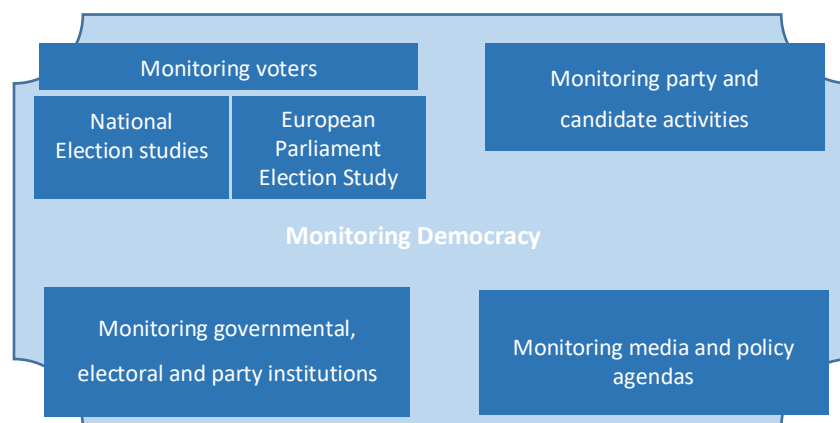
In this way MEDem enhances research capacity for monitoring the legitimacy and well-functioning of dynamic electoral processes – topics of interest to scholars well beyond political science, including those in economics, sociology and media studies.

Research on electoral democracy is a well-established research field throughout much of Europe. It is represented in journals, standing groups and sections in international organizations, summer schools and training centres and so on. These scholars exist in most EU member countries and beyond, and constitute a

well-integrated international community that encompasses political scientists, economists, sociologists, media studies, and more. Beyond academia, the research of this community is widely employed by political party organizations, public opinion research institutes, government departments and the media. It has promoted innovation in electoral campaign techniques and especially in providing policy and procedural advice to political parties, governments and Think Tanks.

Monitoring electoral democracy has recently come to the fore because of the strong performance of populist parties promoting illiberal political ideas. Because not all countries have seen similar developments, MEDem data might even permit scholars to devise political and institutional correctives for poor democratic performance. The importance of these questions is widely understood, as is the need for appropriate data in order to address them.

In addition to facilitating high quality academic research on European democracy, we expect MEDem to also be widely used beyond of the academic community by those who write reports for media, politicians and the public on election topics, summarizing the empirical data and research findings. Furthermore, a vital component of MEDem will build capacity and provide training for the next generation of researchers and experts on European democracy. Given this context, the impact of MEDem will be profound.

The research infrastructure Monitoring Electoral Democracy (MEDem),

➢ Brings together well functioning national and comparative data collection projects and links and coordinates existing endeavours as well supports innovation in research on electoral democracies;

➢ Allows for more, better and new comparative research linking a variety of information on electoral democracies in a comprehensive way;

➢ Assembles existing (and promotes new) long-term time series to study major developments in electoral democracies over time;

➢ Through a position on the ESFRI roadmap, gains recognition as being of strategic importance for the political science, political sociology, and communication studies research communities in Europe;

➢ Sets standards and develops instruments for data collection to allow for comparative research;

➢ Increases accessibility by providing a single data linkage and access point in collaboration with existing data archives;

➢ Strengthens existing national election studies and allied projects by connecting them to a stable European network of projects and scholars working collaboratively in this field as well as to providing training and building capacity;

➢ Providing a foundation for new academic research as well as reports and information on the functioning of electoral democracy in Europe to the wider community and public.

It is important to note that MEDem itself is not responsible for data collection. MEDem co-ordinates existing pan-European distributed data-collection centres and projects, encourages new initiatives as necessary to provide nodes in all European countries. Thus it builds on existing projects as far as possible.

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 to the linked data for the research community on electoral democracy. It will coordinate with ESS ERIC and SHARE ERIC to create synergies in methodological standards and in tools to support data collection, and support the establishment of the European Open Science Cloud (EOSC/SSHOC).

2.3   The current research landscape on electoral democracy and the gaps to fill

Europe is home to many research programs/infrastructures for collecting data directed at better understanding electoral democracies. Historically these developed on a national basis around the collection of data regarding citizen behaviour at the time of elections to national parliaments. These "National election studies" (NESs) are

located in almost all European countries. Some of these election studies have a history back to the 1950s and 1960s (such as Sweden, UK, Germany, Norway, Netherlands).

Nowadays election studies exist in almost all countries, but there is large variation in their institutionalization. Some have stable long term funding and are included on their national infrastructure roadmaps. But historically there are debilitating gaps in the data. Not all these projects have collected data for every one of their national elections since their founding. Often funding for particular election studies has depended on separate funding applications for each election and not all of these applications have been successful. However, where successful, the funding has virtually always been from national governments or their funding agencies. Thus, national governments in Europe have a long history of providing funds for national election studies in their countries.

Over time these national research infrastructures 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, though these data collection programs are sometimes also external and supplementary to the NESs of each country. In addition, many additional projects focus on other components of electoral democracies, such as parliaments, governments, political parties, and the media.

As a parallel and complementary process, initiatives to collect and analyse data cross-nationally have intensified over the last 30 years, from projects to harmonize existing election data, for instance through the True European Voter (TEV) Cost-Action, to comparative international projects such as the European Election Studies (EES), the Comparative Study of Electoral Systems (CSES), the Comparative Candidate Survey (CCS), the Comparative Agendas Project (CAP), the Manifesto Project (MARPOR), the Chapel Hill Expert Survey (CHES) and the network of Voting Advice Applications (VAAs). In addition, a variety of data collection programs also collect institutional 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). This list is far from complete, we do not even count the numerous individual research projects that also collected data at different points in time.

Comparative research on democracies is not limited to elections. While less established in many cases, various comparative research projects also exist that monitor parliaments, government and the media. These projects are important to create the link between what happens at elections and the inter-election period, and facilitates the study of important democratic processes beyond elections in a comparative way.

Because each of these (sets of) data collection enterprises is conducted separately, their data collection instruments are specific to each project and vary in the range of countries they cover. This makes it difficult or impossible to use the data collected within one country or project in conjunction with data collected within other countries and projects. Even within specific countries, using data about candidates in conjunction with data about the parties that field those candidates, along with the media that report on their activities and pronouncements, 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 aim 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 is used in different countries and /or by different teams
3. They organize communities, projects, institutions and researchers to make sure that data is collected and used and often, beyond that, they define future research agendas.

While the above three key functionalities exist 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 becomes usable in an easy way across existing projects.

In some cases MEDem also facilitates the creation of collaborative projects, as for the National election studies. Beyond CSES – which is already a comparative project – national data needs to be harmonized first in order to be able to use them in a comparative way. And in some fields, such as media studies or parliamentary studies, projects are not yet at the same level of maturity to be integrated into MEDem. In those fields, MEDem will support the creation of strong comparative projects.

MEDem is an umbrella organization that provides 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 proposes to divide these three main tasks into five different centres (see 3.3.2 below). These centres do not themselves take key scientific decisions: such decisions are taken by MEDem's Scientific Board (see 3.3.4 below), representing its member community. The centres, however, are involved in the preparation and documentation of the decisions and they support the Scientific Board with their specialized knowledge on the various aspects of MEDem's activities. Above all they provide organizational continuity and institutional memory essential to an enterprise that depends on the participation of shifting groups of scholars with expertise on and interest in different aspects of the overall enterprise.

3 The future MEDem organisation

MEDem will bring together existing national and comparative projects. While existing projects will continue to retain their independence, they will coordinate through MEDem in efforts to make data inter-linkable by agreeing on common standards of data collection with the ultimate goal of making redundant the current need for (expensive and inadequate) *ex-post* data integration of different data sources. MEDem will support the establishment and integration of comparative projects within the scope of MEDem, providing a single point of entry for data access and provide a platform for future comparative European research on electoral democracy.

The following section proposes a model of how MEDem may be organized in bringing together existing projects that all share ownership of MEDem.

Figure 2: Proposal for MEDem's structure
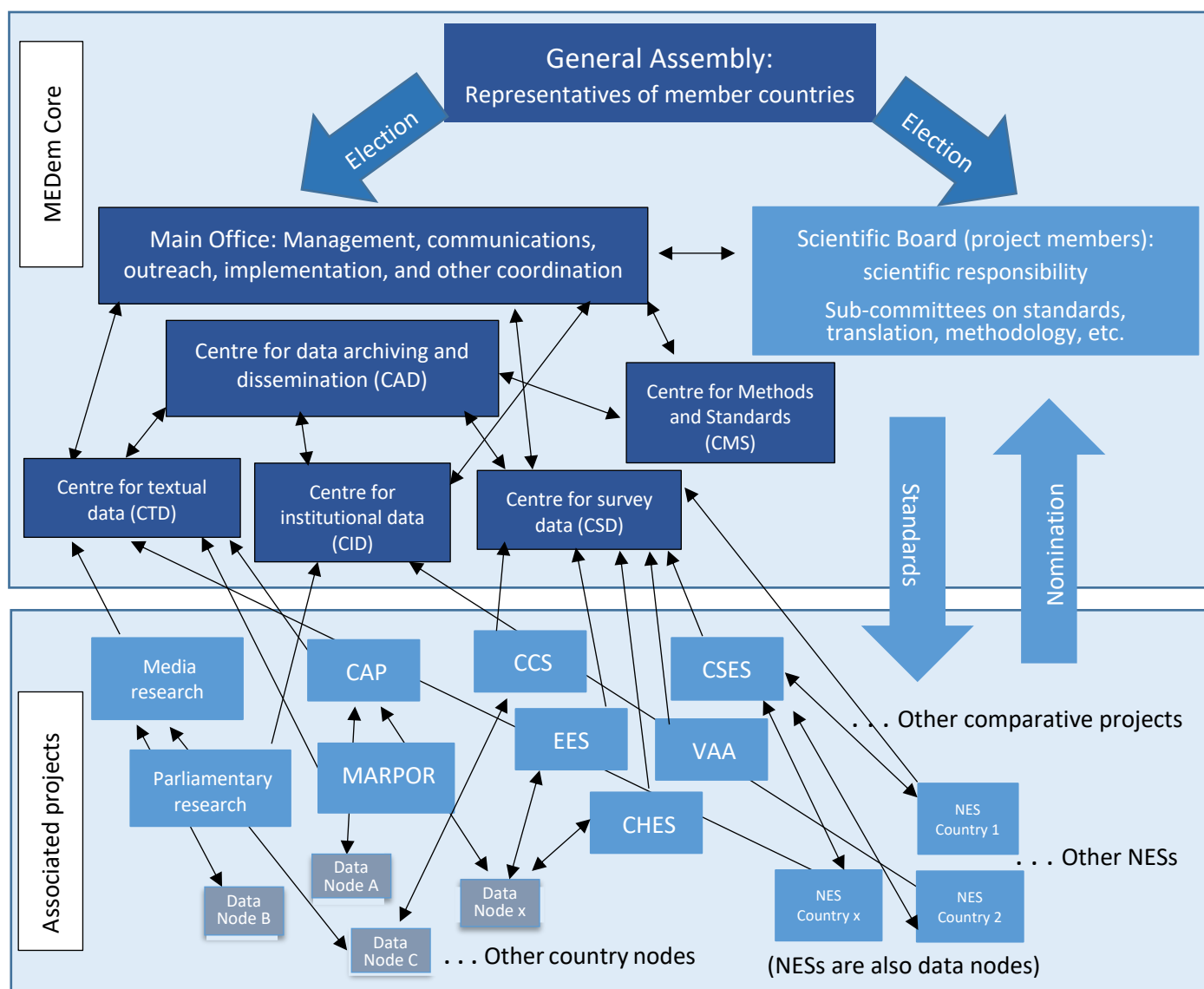
Abbreviations: European Election Studies (EES); Comparative Study of Electoral Systems (CSES); Comparative Candidate Survey (CCS); Comparative Agenda Project (CAP); Manifesto Research on Political Representation (MARPOR); Chapel Hill Expert Survey (CHES); network of Voting Advice Applications (VAAs); National Election Studies (NES).

MEDem members shall eventually be countries.[1] Members;

- ➢ Fund the national data collections for those parts of MEDem that exist within their countries;
- ➢ Commit to developing the data collection parts that do not yet exist in their country;
- ➢ Contribute to the overall MEDem operation costs of the main office and centres;
- ➢ Appoint a representative to the General Assembly.

---

[1] Should MEDem become an ERIC, MEDem will grant observer status to countries that cannot be full members, as is the case for other European infrastructures.

3.1 MEDem Services

MEDem will provide the following key services to the scientific community:

1. **Integration of existing data**

    a. *Post-harmonization*: Integration of different existing data sets by making them comparable and linkable. In some cases (for example CSES, CCS, CAP, MARPOR etc.) comparative data already exists, in other cases (such as NES data, media data, parliamentary data etc.) comparative data files still needs to be produced and established;

    b. *Access*: Providing a single entry point for data users. The data may be stored at different places, mainly existing national data archives, but MEDem will provide a single access point and assure that existing data follows a common documentation standard so that data can be used across projects.

2. **Standardization of future data**

    a. *Pre-harmonization:* Development of a "conceptual map" that specifies how data for increasing numbers of concepts should be coded for eventual integration, using instruments suited to different data types, thus eventually making post-harmonization redundant for new data produced under the auspices of MEDem. These instruments should facilitate comparative research in two main domains:

    b. *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);

    c. *Methodological standards* on data collection for different parts of MEDem. MEDem, in collaboration with the existing projects will develop rules and quality standards of data collection.

3. **Innovation:** MEDem will provide a platform and procedures to develop and implement joint collaborative research topics in one or several sub-fields associated within MEDem across different data collections.

In addition, MEDem actively supports:

- ➢ **Community integration and development**: the MEDem framework will bring together scholars from the different subfields in a comprehensive and permanent fashion that ensures inclusivity.

- ➢ **Integration for new comparative projects**: MEDem will support new research teams in their attempt to organize a sub-community related to MEDem and to organize the future data collections.

- ➢ **Competence building**: providing training and guidance for using MEDem data for both the scientific and non-scientific community.

- ➢ **Knowledge Dissemination**: communication activities directed to a scientific and non-scientific audience.

MEDem will be organized around the cycle of elections in European countries 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.

3.2 Existing nodes and projects

3.2.1 National nodes

National nodes are the lowest organizational component of MEDem. They are research teams located at national research institutions who are in charge of parts of the actual data collection within the MEDem

framework according to the standards set by the comparative projects and MEDem. Often national nodes are responsible for more than one part of the MEDem data collection. National nodes:

➢ Commit to conduct data collections for one or several different MEDem components within the MEDem framework, following jointly agreed overall MEDem scientific and methodological standards, building on existing relevant standards of the comparative projects (where these already exist);

➢ Secure funding for their national data collection component;

➢ Commit to working through the relevant comparative project in coordinating data collected at the national level, so as to be able to contribute the data to the comparative data collection;

➢ Commit to deliver data free of charge and without any embargo.

3.2.2 Existing projects linked to MEDem

Comparative projects

Existing comparative projects continue to be independent projects and they remain the main building blocks of MEDem. They will have rights and obligations. They shall,

➢ Ensure the appropriate functioning and funding for their own endeavour;

➢ Contribute to the development of joint standards and jointly supervised coordination of their existing data collections through MEDem, with their existing national data collection nodes, following jointly decided coding and methodological standards to allow for comparative (cross-national and cross-project) research;

➢ Commit to making resulting data available free of charge and without any embargo,

➢ Nominate members to the Scientific Board;

➢ Commit to open and transparent procedures for the definition of future data collections and to ensure that the ownership of the data collections rests with the scientific community,

➢ Sign an agreement between MEDem and the project that defines rights and duties of both partners;

➢ May receive funding for parts of their operations from MEDem.

The following comparative projects have been involved in the discussion to establish MEDem and may become part of MEDem in the future (non-exhaustive list):

➢ Comparative Study of Electoral Systems (CSES)

➢ Comparative Agendas Project (CAP)

➢ Manifesto project (MARPOR)

➢ Comparative Candidate Survey (CCS)

➢ VAA Research Network - Data on Party Positioning

➢ Chapel Hill Expert Survey (CHES)

➢ European Election study (EES)

➢ Media Research Network

➢ Parliamentary Research Network

➢ Party Facts

➢ Parliaments and Governments Database (ParlGov)

➢ Constituency-Level Election Archive (CLEA)

➢ Executive Approval Project (EAP)

National Election Studies

National election studies exist in many countries, in some for a very long time by now, however, with varying degrees of institutionalization. Because of their importance in the national academic environment and their

long tradition, national election studies are a central pillar of the MEDem data collection. National election studies have an independent existence as MEDem members apart from their membership of the CSES (and/or any complementary/successor future comparative project). National election studies,

➢ Contribute to the development of joint standards and jointly supervised coordination of their existing data collections through MEDem in coordination with the CSES and/or any complementary/successor future comparative project.

➢ Commit to conducting a national election study in their country for the national elections according to the defined joint standards,

➢ Secure funding for their national data collection,

➢ Nominate members to the Scientific Board;

➢ Provide their data for post-harmonization to the Centre on Survey Data as well as implement standards of data collection for future election studies in order to facilitate comparative (cross-national and cross-project) research.

➢ May receive funding for parts of their operations from MEDem (but not for the data collection itself).

3.3   MEDem Core Structure

3.3.1   The MEDem Main office

The main office is located in the leading country under the leadership of the MEDem Director. The main office:

➢ Is in charge for the overall coordination and operation of MEDem;

➢ Prepares the agenda and documentation of the General Assembly and implements its decisions;

➢ Prepares the agenda and relevant documentation of the Scientific Board;

➢ Coordinates with the units and centres and the management all work related to MEDem;

➢ Ensures, together with the management, the functioning of the centres;

➢ Is responsible for promoting MEDem and for reaching out to the academic community as well to the wider public;

➢ Is responsible for documenting and supporting the MEDem membership of projects and nodes;

➢ Organizes periodic external reviews of the project.

The Main Office is the operational hub of MEDem. It coordinates activities with its Competence Centres and supports meetings of the General Assembly and the Scientific Board. It prepares and implement decisions taken of the General Assembly and the Scientific Board. It runs joint EU and internal projects and maintains the MEDem website.

The main office is also responsible for external and internal communication, for contact and support with comparative projects and potential fundraisers. It also keeps contact with comparative projects related to electoral democracy for potential future partnerships, as well external organizations and networks such as ESFRI and the Social Science & Humanities Open Cloud – SSHOC.

The operations of the main office will to high degree be organized around annual organizational plans for the competence centres and work packages with annual monitoring and revisions.

Another vital function is on communication and outreach; within the MEDem network in order to strengthen and support comparative projects and nodes, with potential future partners, with other relevant research infrastructures and organizations as well as with policy-makers and the wider public. The MEDem main office will organize the annual meetings for the General Assembly and the Scientific Board, as well as regular meetings (web and irl) of the management board of MEDem. The main office is also responsible for organizing events for scholars interested in MEDem such as summer/winter schools, scientific workshops and conferences.

The outreach activities described above may be carried out in a separate Unit of the Main Office. The Unit may or may not be in the same physical location.

Outreach unit (OU)

The Outreach unit (OU) is in charge of supporting existing projects in:

➢ Reaching out to new countries and building new nodes for their data collection;

➢ Developing and strengthening the governance and organizational and operational capacity of projects;

➢ Finding sustainable funding for the comparative projects and their nodes where this does not yet exist.

➢ Implementing communication and knowledge-transfer activities together with the centres and projects involved in MEDem.

The management consists of the MEDem Director as well as the directors of the centres. The role of the management is to plan, execute and coordinate the MEDem workplan at the main office and the centres.

3.3.2 Competence Centres

Centre for data archiving and dissemination (CAD)

The CAD will be hosted by one or several established national social science data archive(s). The proposed host for the CAD is GESIS – Leibniz Institute for Social Sciences.

While all of the MEDem data may not be physically stored at a single location, the MEDem CAD will provide a single data access point for MEDem-related data. The main tasks of the CAD are to:

➢ Make data usable across projects, and datasets collected in the different projects, in collaboration with the other competence centres;

➢ Provide tools and instruments to link the data sources from different projects;

➢ Map the existing projects and data sources, and make data findable;

➢ Ensure that all the data is well documented and updated following a joint metadata and documentation standard and archived in established data archives, if not archived at the competence centre itself;

➢ Archive a substantial part of the data collection and provide links to data archived elsewhere;

➢ Contribute to developing post- and pre-harmonization procedures and standards for previous and future data collections jointly with the Centre for Methods and Standards;

➢ Advise the other centres, in close collaboration with the Centre for Methods and Standards, in their data harmonization and in the development of standards for data collection, data integration and data documentation;

➢ Support and train researchers in the use of MEDem data.

The proposed specific services for the CAD, on the basis of FAIR Data principles and the concept of a distributed infrastructure, are the following:

➢ Data Access and Dissemination

The CAD will develop an integrated database, which offers data user a central point of access for MEDem-related data. This involves the development and maintenance of a dynamic online interface that allows data users to interactively search for data and select subsets of the data by topic, time points, and geographical coverage. In addition, users can search for comparable variables across different studies. Entire studies, documentations or variables can be down-loaded either directly or by redirecting users to external repositories. This will include a service for making dynamically generated data subsets reproducible and citable. Long-term archiving of the data will be provided through the established services of GESIS DAS.

The data need to be archived in places where persistent identifiers are offered, but they do not need to be in one specific repository. The challenge is to identify all data that need to be included, and coordinate with the repositories where they are located in order to add the correct metadata that make sense for the community. Once this is done, metadata will be automatically harvested. Electoral data often contain sensitive information that has to be removed from the use files or made anonymous for data protection reasons (e.g. regional data). In cases where this is not possible or where there is a legitimate interest to take the variables into account in the analysis, we can distribute the corresponding datasets via the Secure Data Center (SDC) of GESIS.

➢ Data Harmonization and Visualization

An important task of the CAD will be the post-harmonization of MEDem-related data across sources/measures, space, and time. GESIS will develop and maintain a harmonization interface which allows user to easily harmonizes variables from various studies and generate the associated code. Many studies have already been harmonized to be linked together in order to produce workable datasets for publications. These harmonization codes could be used again for new research projects. Thus the harmonization interface will be accompanied by an online library of documented harmonization's efforts of existing data, where researchers can follow the logic of already used harmonization processes and download the code.

When using the aforementioned across-source harmonization approach, it becomes relevant to obtain easier access to the data regardless of its physical location. Ideally, data can be linked without having to manually download the data from their original repository. This becomes possible by the development of open APIs to offer standardized machine access to the participating repositories. Having such access options and online-accessible harmonization routines in place will for example allow the easy production of high-level visualizations from the harmonized data.

➢ Data Linkage

In addition to data harmonization and integration, a major task will be the linkage of elite and mass surveys to contextual characteristics. The CAD intends to integrate the corresponding contextual data in the integrated database, where user can either download ready-linked data sets via the online interface or generate customized datasets with specific contextual variables. It should be noted that these linked data sets may also be sensitive if it is possible to identify the survey respondents. However, this data can then also be offered via the SDC.

➢ Metadata and Documentation Standards

All the above technical developments are only effective if the data in question are very well described with machine-readable, structured metadata that make sense to the community. Therefore significant efforts will be dedicated to the development of a joint metadata standard and controlled vocabularies that enable the interoperability of all electoral democracy data and the correct functioning of the above tools. This also involves standardization of the content and layout of the data documentation.

Centre for Methods and Standards (CMS)

The *Centre for Method and Standards (CMS)* suggests the standards to be agreed by the Scientific Board, in close collaboration with the other centres. It implements the agreed standards to be used in future data collection in order to facilitate the later data integration and usage across projects. The proposed host for the CMS is the Department of Political Science of the University of Gothenburg.

The CMS's role is to reflect upon and prepare standards, so that the *decisions can be made by the Scientific Board*. CMS also provides expertise on methodology of relevance for MEDem operations, both in developing "best standards" as to collect and disseminate methodological expertise within the MEDem community. The main tasks of the Competence centre for Methods and Standards (CMS) include:

➢ To coordinate and promote developments and agreements of standards between the centres;

- To promote agreement about harmonization standards between the different areas covered in the Competence Centres and oversee the implementation of those standards;
- To organize negotiations between centres and between those and the national nodes in order to involve as many as possible future users in the deliberations on post-harmonization and pre-harmonization Standards of measurement: the development and harmonization of measurement instruments for making the different studies comparable (e.g. in elite and voter surveys, or CHES with voter surveys etc.), creation of scales and indices;
- Expertise on various modes of data collection, tests on new modes of data collection;
- Analysis of complex and heterogeneous datasets; defining data imputation and weighting of heterogeneous data sets;
- Develop techniques to cope with the new type of data MEDem aims at creating;
- Organize events for scholars interested in MEDem methodology, such as summer/winter schools, scientific workshops and conferences, jointly with the outreach unit of the main office.

At a later stage, MEDem shall also be used as an innovation platform for new research. This can include new types of data or modules of data collection to be integrated in the different sub-parts of MEDem. MEDem shall in the future, and in close collaboration with the existing projects, design and establish procedures (such as calls for modules, as well as review- and decision-making procedures) that allow for innovation and make sure that new scientific topics can be studied by using different MEDem components. CMS will also be an innovation platform for new research, this can be new type of data or for modules of data collection to be integrated in the different sub-parts of MEDem.

The CMS may organize its work in creating different units along the type of data, or the type of tasks. This will be defined at a later stage.

Centre on Survey Data (CSD)

The Centre on Survey Data (CSD) will coordinate and integrate all survey related data, mainly on voters and citizens, but also on candidates and MPs. The proposed hosts for the CSD is based on a collaboration between three institutions: FORS–Swiss Centre of Expertise in the Social Sciences; the Institute of Political Studies at the University of Lausanne; and the Department of Political Science and International Relations at the University of Geneva.

The main objectives of the CSD are the following:
- Coordinate and link national election studies and harmonize national election study data. For this part the centre will closely collaborate with CSES. There have been previous attempts to integrate national election study data from various countries and over several decades, most recently through the True European Voter Cost-Action project. The CSD will build on this, and on national election study data, to establish a set of time series that will be extended into the future as new editions of the data are collected;
- Promote the usability of survey data, also in cooperation with other competence centres, across datasets collected within the different projects;
- Contribute to the development of post- and pre-harmonization procedures and standards for previous and future survey data collections, in close cooperation with the CMS and CAD;
- Advise projects in the development of standards for data collection, data integration and data documentation that will facilitate data harmonization in the future;

The CSD may also be involved in supporting the preparation of the European Election Study (EES), but it will not be responsible for securing funding for the EES.

In broad terms, the tasks of the CSD include:

- Coordinating the different national and comparative projects based on survey data. This includes ensuring an information flow between the different projects on methodological, thematic and governance issues, together with other standards.
- **Integration of national election study data** which will happen in a two steps process: a) harmonization of national data; b) establishment of a comparative data file;
- **Integration of comparative data sets** which will include; the provision of keys (various codes that allow to link parties, countries, constituencies etc.); recommendation and tools how data should be aggregated when you want to use the data set across projects (for example if you want to match candidate information to a voter data set); Provision of guidelines and tools how different concepts used in the different surveys can be analyzed comparatively.
- **Documentation**: Establishment of a list of concepts, variables and scales used in different data collections (conceptual map), documentation standards for data collection.
- **Recommendations** for future **standardizations**. Based on the experience from the data integration work, the centre will also recommend standards on coding, concepts, questions, answer categories, methods of data collection etc. facilitate the data integration at a later stage.
- Produce **guides and training material** how to use the data sets in a comparative way across different data sources.

Centre on textual Data (CTD)

The proposed centre will rely on a decentralized structure between the University of Vienna, Department of Communication and Vienna Center of Electoral Research (VieCER), and the University of Amsterdam, Amsterdam School of Communication Research (ASCoR). The use of textual data for empirical research is rapidly growing. Because of this, MEDem plans to establish a Centre for textual data (CTD) that will:

- Provide coordination between different comparative projects involved in the collection of textual data, such as the MARPOR and CAP, as well as comparative media and parliamentary research projects that are under development;

- Promote the usability of textual data across datasets collected in the different projects;

- Document methodological and technical approaches for collecting, scraping and analysing digital textual data from a variety of sources, with a special focus on the multilingual nature of such raw data;

- Ensure that all the textual data is well documented and updated following a joint metadata and documentation standard, and archived in established data archives if not archived at the competence centre itself;

- Contribute to the development of post- and pre-harmonization procedures and standards for previous and future textual data collections;

- Advise projects in the development of standards for data collection, data integration and data documentation that will facilitate data harmonization in the future.

A number of challenges characterize the field of political communication and media analysis. First, the types of information researchers deal with are fundamentally different in shape, format and style, ranging from parties' election manifestos, speeches and parliamentary debates, to traditional news media coverage to political posts and debates in social media, from pure text-based data to multimodal information. Second, capturing such information environments across Europe inevitably deals with a multitude of political contexts and languages, and text analysis tools and data gathering methods are often very sensitive to such variation. Third, and importantly for the MEDem project, the field is very little institutionalized, at least across different types of political texts, and partly, such as in the case of media or social media, also within such types, with very little attention to common and comparable data collection tools and data sharing.

It is before such backdrops that the CTD seeks to address the following tasks:

➢ The CDT will act as a coordination point between existing projects in the area of political text analysis and thereby facilitate a common understanding of the challenges of comparative political text analysis and the opportunities of cooperation.

First and foremost, it will focus on bringing together existing comparative projects, such as the CAP or the MARPOR. Second, it will serve to support the establishment of networks of scholars working in the fields of parliamentary debates, traditional news media or social media, aiming at an institutionalization of such international networks. Networking activities rely on regular physical meetings (biyearly) of the main players in the different fields, of seeking the opportunities of international conferences to bring people together for short events, through workshops focusing on state-of-the-art methodological aspects of text analysis, and on communicative activities such as regular newsletters, a website and social media channels. The support of network structures aiming at institutionalizing comparative research projects and the coordination between existing projects should create critical support among data producer and user communities in the area of political text analysis for the MEDem

➢ The CTD will promote the usability of textual data among different data producer and data user communities, focusing on cross-data usability.

While some established comparative or national projects and data collections are well-established, with visibility in their subfields and beyond, and with a clear focus on data sharing structures, for other projects, in particular in less institutionalized areas, such openness towards data sharing is less evident. The CDT will push the idea of open science and measurement instrument and data sharing for the area of political text analysis, and will promote the value of data linkage and integration of political text data among the different communities, highlighting the added value of data linkage through workshops, conference presentations and exemplary studies/publications. Promotion of usability will also greatly hinge on activities described below.

➢ A major task of the CDT in the first years of operation will be to collect and document existing approaches to political text analysis in the various fields, ranging from sampling procedures via manual coding instructions to text-as-data tools being applied to political texts.

Building on the INCA structure the CDT seeks to build a measurement data base that integrates a comprehensive collection of existing operationalizations and measurements, including, if possible, assessments of the quality of such instruments. First, it systematizes sampling strategies being applied in different projects and works towards a better understanding of sampling procedures in the different areas of political text analysis. Second, it distinguishes coding instructions applied in projects relying on human coders, systematically documenting the measurement tools, their comparability across projects and languages, and their reliability. Third, it collects tools being developed and/or applied in computer assisted text-as-data approaches in the field of electoral communication, ranging from data collection tools (scraping and pre-processing) to analysis tools, again focusing on the comparability of such tools and their quality/ validity. A major challenge lies in the multilingual nature of comparative political text analysis, and the documentation will maintain original language tools as well as provide advanced translations of such tools into English in order to enable comparisons and work towards understanding the comparability of approaches. The documentation will be openly provided to the MEDem community. Such systematic documentation, continuously updated, will serve as major input towards further data harmonization and integration as described below. It will furthermore serve as point of departure for the identification of possible linkage points between different projects and data collections, and with other MEDem data collections.

➢ The CDT should become a major player in pushing the open data approach for the political text analysis community and providing standardization of data archiving.

To that end it will provide templates for data documentation practices for the different approaches of text analysis and push for standardization in the reporting of meta data and full transparency regarding data collection tools and their reliability/ validity. It will also address data archives in different countries that are part of CESSDA and promote openness and the development of routines for the archiving of text data files. The CDT

will coordinate communication between data producers and archivers, if necessary. All data produced by MEDem member projects coordinated through the CDT should be open.

- ➢ Based on the documentation described in task 3, the CDT seeks to develop post-harmonization procedures for existing data and, more importantly, promote pre-harmonization across different MEDem projects.

For post-harmonization, the CDT will use the inventory and documentation to identify all instances of data for which cross-project post-harmonization appears feasible. The CDT will construct and continuously update such data linkage database and share the task of building data integration tools with the Centre for data archiving and dissemination (CAD). For pre-harmonization, in close collaboration with MEDem Project PIs and the CAD, the CDT identifies, where possible, best practices regarding measurements and approaches, or develops and validate such measures and tools. These are shared and promoted across the community and thereby measurement harmonization of political text data collections is implemented in the multilingual European media environment. To facilitate the use of harmonized core measures of political texts in different contexts, while allowing for flexibility to integrate context specific measures, CDT will develop a tool in which core measures and instructions are available in all languages, and which can be used for either manual online coding of textual and audio-visual materials or for automated text and eventually visual analysis procedures. The tool will serve as international exemplary for a standardization and harmonization of text analysis approaches. This applies both to traditional, manual content analysis approaches, for which the CDT provides also training to coder trainers (national PIs) and advises in terms of data management, retrieval, etc. and for computer assisted approaches, for which the CDT facilitates training for the usage and application in regular workshops or tailored trainings.

- ➢ Given the fluidity of some political text data (e.g. online news, public social media posts) and the rather dispersed archiving of some political texts (e.g. parliamentary debates, news coverage, press releases), but at the same time the fact that the majority of the political texts are nowadays available in digital format, in the long run the CDT seeks to build a comprehensive, large scale system of data recording and linking to archived data across European member states.

Based on such system the CDT, in cooperation with the projects involved, will act at the forefront of integrating innovative, automated, computer-assisted approaches of text analysis to the multi-lingual reality of European democracies. Increasing application of computer assisted content/ text analysis approaches will facilitate cheaper, more efficient, and quicker content data collection, and in particular will ensure ready comparability of such data.

- ➢ Provided that multimodal communication has become central in electoral communication, in the mid-to-long term the CDT will develop/ apply approaches for automatic visual recognition and speech recognition allowing for more encompassing analysis of multimedia and multimodal data.

Centre on Institutional Data (CID)

The CID is responsible for providing institutional (contextual) data for the different data sets that may be used for scientific research. It will be central to the aim of linking different types of election studies to actual electoral outcomes. The proposed host for the CID is the Center for Socio-Political Data at Sciences Po Paris.

This is a component currently contributed to separately by many of the individual projects, although the contributions of different projects differ in terms of standards and coverage.

MEDem will attempt to simplify and standardize the collection and distribution of institutional and contextual data to its various projects. It will also link to existing projects by gathering more fine-grained comparative data on matters such as campaign conduct, electoral systems and districting principles, or media regulations. Providing a separate Competence Centre to undertake this task will relieve other projects of the need to do so while ensuring the standardization of data used in virtually all of the research undertaken with MEDem data.

The centre will:

- Consult with comparative projects involved in the collection of survey and textual data, as set out above;
- Promote the usability of contextual data across datasets collected in the different projects, especially for data that exist in different languages;
- Document the methodological and technical approaches for collecting contextual data from a variety of sources;
- Ensure that all the contextual data is well documented and updated following a joint metadata and documentation standard, and archived in established data archives if not archived at the competence centre itself;
- Contribute to the development of post- and pre-harmonization procedures and standards for previous and future contextual data collections;
- The CID may also be involved in the preparation of the European Election Study (EES).

The CID is the unit responsible for providing contextual data on electoral democracies for the different projects under the MEDem umbrella. It involves consultation with projects within the MEDem scope and outside of it, promote the use of contextual data, document the data and the approaches for collecting, documenting and using contextual data, develop post and pre-harmonization procedures and standards.

There are indeed different types of data reflected by the category of Institutional/contextual data:

- Electoral outcomes (down the more precise geographical unit possible, including by-elections)
- Macro socio, economic and political indicators (from World bank, OECD, Eurostat,…)
- Electoral regulations (from electoral systems, districting principles to media and campaign regulations)
- Behavioral measurements (campaign conduct for instance is cited).

The CID suggests to extend this definition, by including, one the one hand, "parliamentary behavior" (from roll call records of individual MPs to a number of indicators of legislative activities (such as questions, amendments, participation and activities of committees,…) and, on the other hand, the sociology of the elite (starting with at least gender and profession of MPs and MEPs).

The reason for these proposed extensions are substantial (because both could be significant contributions to the actual monitoring of electoral democracy) but also methodological (as the rationale here is the collection of comparative of "natural" data (i.e. neither experimental nor based on interviews). Of course, this leaves open for discussion the role of expert (political elite) surveys in the domain of interest of this centre. It would probably mean close collaboration with the CSD in this regard.

As for the other centres, the CID is not aiming at producing data by itself. The main aim is to gather and coordinate various efforts of data collection which are pre-existing. Most of the projects currently participating in the MEDem project already collect their own contextual data (for instance, CSES has already a wealth of contextual variables from district level variables to "macro level variables"). A number of projects outside of MEDem at the present stage have also already provided amazing efforts for combining existing indicators (for instance the QoG dataset or, in more specialized area, the coding of electoral systems by M. Golder or D. Caramani).

- In this sense, the main task of the CID lays in the **networking and organizational efforts** to lay out the map of comparative project dealing with contextual data relevant for MEDem, reaching out to them, and trying either to convince them either to participate directly in MEDem or at least to allow the utilization of the data they produce.

This has for course to be done in close collaboration with headquarters and also the other centres. Beyond contact with existing comparative project, it is likely that a network of national collaborators would be needed. In this regard, working with existing projects on a list of national contact points and their relevant field of expertise would be an important step towards more integration.

> ➢ The second broad task of the CID is working on the data and data format. As a national social science archive, CDSP is quite used to this vast enterprise of integration. Once again in collaboration with CAD**, standards for data documentation and data formats** have to be set up.

If DDI 3 is a likely target for documentation, an important task would be establishing the strategy for missing data imputation and trend estimation. It is indeed very likely that the main output of the CID be a type of database with country-year characteristics as the main unit of analysis. The CID proposes to add a database at the individual candidate level as well; party characteristics could be in the same way be considered for future work of integration. The ParlGov project (a database comprising parties, elections, and cabinets as main entries) has clearly set the path in this direction.

> ➢ The third broad task for CID is the **promotion of contextual data**, in close collaboration with the headquarters and the CAD. A number of projects have indeed shown that contextual data does not always meet the expectations in terms of use by the academic community. Easier access, better country coverage and expanded geographical area are likely to foster scientific interest.

Yet, it may well be insufficient to justify the effort if nothing else is done. That is why data promotion should also be at the centre of the strategy of the CID. Up to a certain extent, networking with existing comparative projects and national country experts is likely to provide a first solid audience for the CID. If a specific academic conference may not seem to the most efficient strategy, specific workshop targeting methods for contextual data use would of course be key. This kind of workshop can also be easily integrated in summer school projects. A second strategy could consist in dedicated help for research groups aiming to use contextual data. A pool of experts in support of the centre could be constituted so as to provide up- to-date advice (either for specific question or more long term developments). CID would coordinate these experts to ensure efficient response. This of course does not preclude the use of other traditional means for data promotion, from actual research to organization of events and social media animation, depending on the future communication strategy of MEDem.

3.3.3   The General Assembly

The General Assembly consists of representatives from member countries who shall have overall responsibility for ensuring the financial and operative viability and sustainability of MEDem. Its responsibilities shall also include, without being limited to:

> ➢ Appointing, replacing or removing the MEDem Director, and the directors of centres in collaboration with the host institution;
> ➢ Appointing, replacing or removing the Main Office and centres;
> ➢ Receiving periodic reports from the Director on the exercise of his/her duties;
> ➢ Reviewing and approving accounts and rolling work programs;
> ➢ Appointing, replacing or removing the members of the MEDem Scientific Board and its sub-committees;
> ➢ Monitoring and ensuring that members fulfil their obligations towards MEDem;
> ➢ Approve the overall strategic planning of MEDem.

3.3.4   The Scientific Board (SB)[2]

The Scientific Board has the ultimate scientific responsibility and decides on all scientific issues related to MEDem. It is important that existing national and comparative projects retain ownership of MEDem and that they are represented in the SB. This is assured through their right to nominate members to the SB. The members are expected to include well-known scholars in the fields covered by MEDem.

The Scientific Board,

---

[2] MEDem may consider putting also an international scientific advisory board with highly experienced scholars in place.

➢ Ensures the overall scientific quality of all MEDem operations at large and decides on the future scientific direction of MEDem;

➢ Decides on standards of data integration, data collection and data comparability (methodological standards, measurement and coding standards);

➢ Defines in the future on procedures for innovation and the inclusion of new instruments across different sub-fields of MEDem, which are to be included in one or several MEDem components.

The Scientific Board can create or delegate to sub-committees for permanent or specific tasks. Already envisaged are sub-committees on methodology, on translation and on harmonization.

3.4 Funding

A draft budget will follow some key principles:

➢ The main office and the competence centres, including the Centre for data archiving and dissemination, are funded by contributions from the member and observer countries as well as by substantial funds from the host countries of the main office and the respective centres;

➢ Actual data collection is funded through existing projects and/or national nodes, thus its collection costs are usually not part of the MEDem budget;

➢ Special funds will be required for the implementation and development of MEDem. Some of these financial needs may be covered through the funds especially provided for research infrastructures in the Horizon 2020 program. For that purpose, and especially for the build-up phase of MEDem, an application to the H2020-INFRAIA funding scheme will be submitted in early 2020. Meanwhile, additional funds will very likely also be needed to establish MEDem.

Note that the largest costs associated with MEDem are those of colleting survey data from mass publics – costs that in many countries have already been covered from national sources, in some cases for decades already. The major purpose of MEDem is to get greater returns from these expenditures by releasing added value inherent in the data already being collected (and already being paid for) through linking with additional data that, for the most part, is much less costly to acquire.

4 Annex: MEDem Preparation and Implementation Phase Working Principles (Adopted at Gothenburg MEDem meeting, June 2018)

Purpose, Scope, and General

1. MEDem (for „Monitoring Electoral Democracy") is an initiative of scholars and institutions involved in research on modern democracies. Its purpose is to establish a permanent organisation that serves as an infrastructure supporting comparative research of the highest possible quality, by enabling connections between, and providing access to, data of the types most relevant for understanding the functioning of contemporary democracies in Europe and beyond.

2. The present MEDem initiative was established in April 2017 at a Special Meeting of the Consortium for European Research with Election Studies (CERES) in Vienna. Austria, with the objective of making a trial application for ESFRI roadmap status in the then current round of applications. This trial was unsuccessful in overcoming bureaucratic hurdles in the way of a new organization, and our revised goal is to make MEDem formally part of the ESFRI roadmap 2020. An initial Steering Committee was elected at a follow-up meeting in Mannheim, Germany, in November 2017 and was tasked with developing these Working Principles and to further develop the infrastructure initiative.

3. The present Working Principles regulate the cooperation among members of the MEDem initiative until that permanent organisation has been installed. They come into force immediately after having been initially confirmed by supporters of the initiative at a meeting held in Gothenburg in June 2018. They can be modified by a normal vote of the MEDem plenary.

4. The MEDem initiative will cease to exist once the permanent organisation has been formally installed, and these Working Principles will be fully replaced by the regulations of that future organisation.

Membership

5. Membership to the MEDem initiative is by invitation and can be acquired by individuals representing national data collections, by relevant institutions, and by existing comparative data collection initiatives.

6. Members commit to fostering the objectives of MEDem.

7. Membership rights include participation and voting at MEDem plenary meetings, and using the name and logo of MEDem when working for its stated objectives.

The MEDem Plenary

8. The Plenary is the ultimate decision-making body of the MEDem initiative and is constituted by all members of MEDem.

9. Plenary meetings can be called by the Steering Committee, or by a quorum of at least 25% of the membership. Outside meetings, members can demand a vote on a particular issue with a quorum of 25%.

10. Decisions in the Plenary are made by simple majority of all valid votes cast. Decision-making outside of plenary meetings can be conducted by electronic communication.

11. The Plenary elects the Steering Committee and assigns one of its members to be its speaker.

12. The Plenary decides about these Working Principles by simple majority.

13. The speaker of the Steering Committee chairs the meetings of the Plenary.

The Steering Committee

14. The Steering Committee can have between 3 and 7 members. Members do not have a fixed term of office.

15. The Steering Committee conducts the day to day business of MEDem. Specifically, this includes:

   a. Organisational and administrative work to run the MEDem initiative, including the organisation of plenary assemblies and voting processes.
   b. Designing roadmaps for establishing the permanent infrastructure, in interaction with the wider community and the MEDem membership.
   c. Initiating the development of applications for relevant grants.
   d. Supporting members in their national efforts to obtain the required funding for participation in the permanent organization.
   e. Reaching out to existing and potential members in order to collect information about their current status and to encourage their membership in MEDem.

16. The Steering Committee is obliged to inform the membership in an at least semi-annual schedule about its activities and current developments. The plenary will be consulted outside this schedule on issues of foundational importance to MEDem.