

Local Media for Democracy (LM4D) // Methodology of the research

The research task of the “Local Media for Democracy” (LM4D) project is aimed at assessing the state of play for local and community media in Europe, identifying news deserts in the areas and/or communities where these outlets are at risk, impairing a smooth dissemination of diverse and quality information.

One of the final aims of this study is outlining which are the relevant dimensions for studying the phenomenon of news deserts in a European context, while informing about local and community media outlets and journalists in the EU countries and offering examples of best practices in the management and innovation of local newsrooms throughout Europe.

The theoretical part of the project, concerning the main definitions used across the research, such as ‘news deserts’, ‘local media’ and ‘community media’ applied to the European Union context, has been developed through bibliographical and desk research (please see the [Glossary](#)). It moves from the concept of ‘news deserts’ formulated by Anglo Saxon authors, and then contrasts this concept to the actual situation of local and community media in the European Union, to finally build a ‘news deserts’ definition that is context based.¹

For the mapping of ‘news deserts’ in the European Union, LM4D counts on the collaboration of country-based researchers who analyse the situation of local and community media in their respective countries, indicating the level of risk for local and community media outlets, and consequently signalling the existence or potential emergence of ‘news deserts’ across the 27 EU member states. For this part of the research, the main method used is the collection of data and information on a country basis through the answering of a structured questionnaire by the researchers (please see the [Questionnaire](#)).

The LM4D project is a pilot, thus the questionnaire here developed must be considered as first test throughout its first implementation and might need fine-tuning in the future. Nonetheless, it has solid foundations based on the long-tested Media Pluralism Monitor project, implemented from 10 years by CMPF: the current research builds on its methodology.

The questionnaire consists of **55 questions (variables)** of legal, economic and socio-political nature, clustered into 6 indicators. It is built in a way that allows for the gathering of both quantitative and qualitative data, and for a cross-comparative analyses between EU member states. The data gathered with the questionnaire will be reviewed by CMPF researchers and will be compiled in the form of a final report divided into country sections. Before beginning the data collection, the questionnaire was reviewed by the roster of local researchers, and by the LM4D Consortium’s Research Committee.

The questionnaire is composed by **12 contextual variables**, to be considered as independent variables for the purpose of this research (please see the paragraph

¹ [What are “news deserts” in Europe? - CMPF Blog \(eui.eu\)](#)

below: “Independent variables: ex-post analysis”). They cover general questions about the country size, the population, and the general state of play in the local and community media sector. No risk score is associated to these questions.

Following the above-mentioned section, the questionnaire is composed by 5 indicators aimed at evaluating the risks for local and community media, in the countries under study. These indicators are:

- **Granularity of the infrastructure of local media:** This indicator assesses the presence and offer of local and community media services in a country, as well as of local journalists. It is composed of 6 variables.
- **Market and reach:** This indicator assesses the economic conditions, the viability and sustainability for local and community media in the country. It investigates revenues, supply distribution levels, subsidies, state advertising and market shares. It is composed of 13 variables.
- **Safety of local journalists:** This indicator assesses the situation for local journalists when it comes to their working and physical safety, assessing also the presence of SLAPPs or other forms of harassment. It is composed of 6 variables.
- **Editorial independence:** This indicator assesses the independence of local and media from political and commercial pressures. It investigates the risks of conflict of interest, the fairness and transparency in the allocation of state subsidies and state advertising, the diversity of news content, to name a few. It is composed of 8 variables.
- **Social inclusiveness:** This indicator assesses the extent and quality of news offered for and about minorities and marginalised communities, and whether local and community media meet the critical information needs of community they serve, and if they offer public interest news. It is composed of 7 variables.

In addition, the questionnaire poses two questions investigating the existing “**Best practices and open public sphere**”: these variables assess the actual existence of innovative practices for enhancing an open and thriving public sphere in specific communities, not only by professional media services but also, for example, through citizens’ initiatives and social media (“other media actors”, see the Glossary attached). No risk score is associated to these questions.

Finally, the questionnaire asks to the local researcher to draw, building on the data and information collected throughout the filling in of the questionnaire, the areas that could be considered as news deserts as from the definition taken under consideration for this research, namely a geographic or administrative area, or a social community, where it is difficult or impossible to access sufficient, reliable, diverse and independent local, regional and community media and information.

Standardization and calibration of data

All the questions in this questionnaire are classified as variables. Variables are grouped into indicators (see above). To determine the risk for each variable, and indicator, a standardised formula has been applied to the LM4D questionnaire. This

method builds on CMPF's Media Pluralism Monitor project, as well as on previous studies, in which the indices were based on a list of questions/ indicators for which the answers were calibrated on a scale from 0 to 1 (e.g., Gilardi, 2002; Hanretty, 2009). In other words, each variable is assessed by a question and receives a score from 0 to 1, according to the specific answer. The variables (five-option replies, in the form of risk level or Likert scale) are rated 0/0.25/0.5/0.75/1 according to the band into which the reply falls. The overall risk level per country is then obtained by calculating the mean of the mean of the variables according to indicator. Scores closer to 0 indicate a low risk assessment, while those closer to 1 indicate a high-risk assessment, as outlined in Table 1.

Table 1: Risk scores for variables 13 to 52 in the LM4D questionnaire

Risk level	Likert-scale equivalent	Numerical value
Very High Risk	Strongly Disagree	1
High Risk	Disagree	0.75
Medium Risk	Neither Agree nor Disagree	0.5
Low Risk	Agree	0.25
Very Low Risk	Strongly Agree	0
Not applicable	Not applicable	n.a
No Data	No Data	low risk (0.3)/ high risk (0,7)/missing value

No data and Not applicable answers. We acknowledge that for the scope of this research, it might be difficult for local researchers to access certain datasets and/or to answer to some qualitative questions due to a generalized lack of data (e.g. on revenues, media reach, audience measurements) or of specific research on local media and community media in given EU countries. Therefore, this questionnaire allows for the possibility to answer by using the option "no data". Following the choice of this answer, the country teams are asked to evaluate whether the lack of data represents a systemic problem within their national context (i.e., to evaluate whether the lack of data should be assessed as a lack of transparency, thus problematic in their country). In this way, the specific characteristics of the national context are accounted for, since there may be a variety of reasons why certain data are not available/accessible across EU Member States, and not all these reasons are necessarily always causes for concern. In other words, the country teams should provide the 'no data' option with comments, specifying if the lack of data is problematic or not and why, as well as suggesting the level of risk they would assign to the unavailability of data on specific matters. All answers that are marked as "no data" are double-checked by CMPF, meaning that one independent coder will eventually assign

a risk score to each “no data” answer. The risk scores for “no data” assigned by the CMPF are: low risk (0.3) or high risk (0.7). The no data value will instead not be considered in the calculation of the risk (“missing value”), in cases when the absence of data in that specific case is not relevant.

Generally, the following procedure is applied: if the country team indicates that the lack of data was not problematic, then the CMPF follows this suggestion, and codes it accordingly, as missing value; if the country team considers it a risk but not a severe one (e.g. lack of monitoring of the representation of minorities on broadcasting channels, but in a general context where the presence of minorities is minimal), then the CMPF team scores it as “low risk”. Finally, if the country team considers the lack of data as indicative of a transparency problem, violating existing EU or domestic laws, or impairing the implementation of necessary policies, then a high-risk score is coded.

Finally, each variable provides also for the possibility to choose a “not applicable” answer: this must only be used as a last resort. It must only be selected when the question is not relevant to the specific context of the country (e.g. because of the small size of the country). Crucially, when opting for the “Not Applicable” option, country teams must provide a substantive explanation justifying their decision, possibly supported by data and/or sources. All the questions coded as “not applicable” are excluded from the final calculation.

Aggregation method. Inspired by the MPM methodology, an aggregation method for the risk scores by indicator has been developed, so that the indicators’ trends in each country would be visible, and to eventually better grasp differences and similarities among the member states as reflected in the risk scores. The method is based on the mean of the item scores, used as the most common aggregation method to calculate indices.

At the indicator level, scores between 0 and 20% are considered to be very low risk, above 20 to 40% are considered to be low risk, above 40 to 60% medium risk, above 60 to 80% high risk, above 80 to 100% very high risk. On the level of indicators, scores of 0 were rated as 3%, and scores of 100 were rated as 97%, by default, in order to avoid an assessment of a total absence or a total certainty of a given risk, as these concepts contrast with real-world situation. It should be noted that all values were presented as percentages, for ease of use and interpretation (e.g., a score of 0.46 is presented as a risk of 46%). The results for each area and indicator are presented on a scale from 0% to 100%. Each score is rounded to hundredths. If the digit at the hundredth place is equal or superior to 5, the risk level is rounded up. If the digit of the hundredth place is inferior to 5, the risk level is rounded down.

Independent variables: ex- post analysis

The questionnaire, moreover, contains 12 contextual/independent variables which we have included for two main reasons. Firstly, these questions assist us in understanding the contextual differences between countries. For instance, in relatively small countries in terms of population and geographical size, we acknowledge that some of the questions might not be applicable, or that the researchers might struggle to access

relevant datasets or specific academic literature, as well as other kind of reports². Secondly, the inclusion of these independent variables provides an opportunity for us to conduct a deeper statistical analysis of the news deserts' phenomenon. Specifically, we wish to carry out a multiple regression analysis to ascertain which country-level features are strongly correlated with the indicators of this questionnaire aimed at identifying a risk for existing or emerging "local news deserts". To this end, we will assign dummy code variables to Boolean and categorical data while leaving the numerical variables unaltered. For example, in the question that asks whether a legal definition exists for local and community media, 'Yes' responses will be assigned the Boolean value of 1, and 'No' values will be assigned the Boolean value of 0.

When dealing with categorical variables, we will adopt one-hot encoding (also known as dummy variables) in order to incorporate them into the regression model. One-hot encoding involves creating a binary variable for each category, where a value of 1 indicates the presence of the category and a value of 0 indicates its absence. We will use a technique known as "one-hot coding" to handle the presence or absence of categories while working with polytomous categorical data (categorical variables with more than two levels)³. This method will allow us to incorporate these variables into our regression model as independent variables along with other continuous variables we have identified. Furthermore, for each domain of the dependent variable we wish to explore which country-level independent variables are correlated with the aggregate risk score. The advantage of applying the regression model is that if the model fits the data well, we could predict or anticipate other areas at risk of news deserts beyond the countries participating in the study.

Some of the data for these variables will be retrieved from the questionnaire (primary data), while some will be retrieved from pre-existing secondary data. As part of a non-exhaustive list, the independent variables include the following: *Land area (sq. km)*, *Population size*, *Population density (per km2)*, *Rural population (% of total)*, *Immigrant population (%)*, *LMD debate (yes/no)*, *OMRs (yes/no)*, *Type of government system*, *Number of levels of government*, *Minority groups legally recognized (yes/no)*, *Legal definition 4LM (yes/no)*, *NUTS2 GDP per capita (range)*, *Net GDP growth in the last 10 years (%)*, *Median household income*, *Percentage of government spending earmarked for media, culture, etc.*, *Internet penetration (%)*, *Levels of trust in public*

² For example, in member states such as Cyprus, Luxembourg, and Malta, questions pertaining to the existence of a 'stable network of local correspondents' may not be applicable. This is because there are no local branches in these countries given their relatively small size. Moreover, there may not be data pertaining to questions asking for a breakdown of urban/rural 'demographics in terms of the residence of journalists' in, for example, Malta (which is the most densely populated country in Europe at circa 95%) where there are very few rural areas left. In this case, we would take this into account when considering whether 'no data' was considered a risk factor.

³ For example, in Question 4 of the LM4D Questionnaire which has at least three categories (the taxonomy is still to be decided inductively based on the responses in the questionnaire), we would create two binary variables, for example, 'is_parliamentary' and 'is_presidential' wherein the former would have a value of 1 if the governance structure is parliamentary, and 0 otherwise, and the latter would have a value of 1 if the governance structure is presidential, and 0 otherwise. If both binary variables are 0, then the governance structure would be 'semi-presidential' (or another type of governance structure, TBC).

institutions (% - source: Eurobarometer), Levels of trust in news media (%), Education level, Media literacy level, and Age demographics.

By utilizing these independent variables, we hope to gain a better understanding of the factors that contribute to the risk of a news desert and to develop effective strategies to combat this phenomenon.