Disobedient Democracy – Data Management Plan 2017 (developed according to SNF Guidelines)

| **Data Collection and Documentation**

What data will you collect, observe, generate or re-use? How will the data be collected, observed or generated? What documentation and metadata will you provide with the data? | The project has two phases of data collection. In the first phase (2017-2018), we collect quantitative data by applying Protest Event Anaysis. In the second phase (2018-2020), we collect qualitative, ethnographic data on urban social movements. This DMP currently contains only information about data for the first phase of data collection.
Protest Event Analysis data is collected from print and digital archives of national newspapers. This project collects data on all types of protest events that happened between 2000 and 2017 in four countries: Croatia, Serbia, Spain and Portugal.
For each country we use two national daily newspapers:
• Croatia – Večernji List and Jutarnji List
• Serbia – Politika and Danas
• Portugal – Diario de Noticias and Publico
• Spain – El Pais and El Mundo
Coders go through each newspaper issue in the given period, from front to back, identify protest events, and then input data which is required by the Data Protocol.
The Data Protocol is web-based, hosted on the project website disdem.org. The web form requires coders to insert various types of data on each protest event, including uploading the scanned newspaper article, precise information on the source, and as many features of the protest event as possible.
Each coder is assigned a reviewer who helps resolve ambiguous cases and generally supervise coders.
The online protocol enables simultaneous input of data from multiple locations and by many coders. Each input into the protocol is saved, while the coders and the reviewers have options of deciding when the input is moved from "draft" status to completed.
The online protocol enables easy export of all inputted data into CVS format, which is compatible with various software for data analysis.
The Codebook accompanies the Data Protocol. The Codebook explains the main features of the method of data collection, key terms as well as provides detailed instructions that accompany each item in the Data Protocol.
After the data collection is completed, we will make available the Dataset, the Data Protocol, the Codebook and this Data Management Plan. |

Ethics, Legal and Security Issues *How will ethical issues be addressed and handled? How will data access and security be managed? How will you handle copyright and Intellectual Property Rights issues?*	This method of data collection relies on publicly available data and in therefore does not require consent from subjects nor does it raise direct ethical concerns. With respect to making the data openly available, there is the restriction of not being able to publish data sources, since these are national newspapers with copyrights. Though the project scans all archival newspaper articles in order to be able to check reliability of coders and correct for mistakes in data collection, due to copyright restrictions we will not make publicly available the scanned newspaper items, but only the dataset containing information collected from these sources.
Data Storage and Preservation *How will your data be stored and backed-up during the research? What is your data preservation plan?*	Data is stored on a virtual server hosted by Linode, LLC (https://www.linode.com/) in a data center. Data center is located in Frankfurt, Germany. Data is backed up by using Linode backup service, which stores a backup of the entire server once per day. Server can be restored from daily backup (less than 24 hours old), weekly backup(less than 7 days old), or bi-weekly backup (between 8 and 14 days old). Backups are stored on a separate system in the same data center as the server. Fundraising efforts will be made so that after the end of the project, the cost of hosting the data is secured from other sources.
Data Sharing and Reuse *How and where will the data be shared? Are there any necessary limitations to protect sensitive data? Will you choose digital repositories that conform to the FAIR Data Principles? Will you choose digital repositories maintained by a non-profit organisation?*	After the completion of data collection, in the spring of 2018 the data will be available on the project website, and shared under the CC BY-NC-SA Creative Commons licence. This means that users will be free to copy, redistribute, transform and build upon the dataset as long as they give appropriate credit, provide a link to the license, and indicate if changes were made. It also implies that the data is not used for commercial purposes. Apart from making it available on the project website, we will make the data available on digital repositories that conform to the FAIR Data principles[1]. We will reposit the dataset on the Croatian Data Science Archive and the Harvard Dataverse.

---

[1] See http://www.snf.ch/SiteCollectionDocuments/FAIR_principles_translation_SNSF_logo.pdf