

Managing Data of Historical Earthquakes at Italian, European and Global Scale: Insights into a ten-year Experience

Mario Locati, Istituto Nazionale di Geofisica e Vulcanologia (INGV), Milan, Italy

Improving the knowledge on past earthquakes is a key-activity providing useful information for multiple purposes, for the society in general, as well as for the scientific community. A sound picture of the seismicity may help adjusting how everyone perceive the potential impact of an often-damaging natural phenomenon. In the scientific field of seismology, the knowledge of the past seismicity is at the base of a variety of analyses, ranging from the study of seismogenic sources, the seismic history of a place, to the input for assessing the seismic hazard.

Over the centuries, the European literate society consolidated a tradition of leaving written records reporting effects of natural phenomena that had an impact on the people's life. Many of these important pieces of information survived the passing of time, bypassing wars, fires, floods and an enormous amount of changes of ownerships and locations. Until scientific discoveries in the late nineteenth century on how earthquakes generate and propagate, the content of the reported information was strongly affected by who wrote (or drawn, or carved) the historical record. The reported information is also affected by the reason why the author decided to leave a record, and it is closely combined with the overall historical and geographical context in which the author was living. Understanding from these historical records what happened, where and when in a reliable way, is the challenging task of modern historical seismologists, a few people who developed a cross-disciplinary experience and who can unravel this closely interconnected and intricate context. The more pieces of resulting interpreted information can be put together about an earthquake, the more complete and reliable is the overall resulting picture of its effects.

In the seventies, eighties and nineties of the last century there was a drastic change in the way the past seismicity was investigated due to the need of finding locations for nuclear power plants. All of a sudden, previously scattered, personal, and voluntary based initiatives were replaced by more focused, well-organized and well-financed research projects. Tens of historians were hired in the most developed countries (i.e. Italy, France, United Kingdom, Germany) to go and search all kinds of archives where information could be potentially found. The huge effort, unthinkable today due to the limited funding, generated an amazing amount of earthquake related data. However, not all the resulting data was made publicly accessible, as only the most important outcome took the form of hundreds of studies (scientific papers, reports, and books). Private companies were keeping some of the data underlying material back, and some collected material was simply dumped after a couple of decades as it was not considered important. In fact, only earthquake catalogues were considered worth the effort of being preserved, whereas single pieces of information supporting the earthquake parameters were often overlooked. It should be stressed that the digital revolution had just started and the early attempts of building databases had a very limited complexity and storage capability.

Today's available knowledge on past European seismicity still relies heavily on the seismological interpretations made in the last decades of the twentieth century, when the European historical seismology research was at its top. Recently, there has been a renewed interest in questioning the reliability of earthquake parameters, and the only way to perform a robust check is to go back to the underlying interpretation work performed by the historical seismologists, and, in the most doubtful scenarios, re-creating and re-checking the link between the interpretation and the original historical sources.

The online European Archive of Historical Earthquake Data (AHEAD, <https://doi.org/10.6092/ingv.it-ahead>) aims at fulfilling the task of preserving the interpretation work performed by historical seismologists. First publicly opened in 2013, it was built thanks to the support of a project funded by the European Commission. AHEAD was -and still is- coordinated by a research group working at INGV, the Italian National Institute of Geophysics and Volcanology, and it is conceived as a pan-European, common, and open platform to support the research on historical earthquake data. To achieve this goal, AHEAD archives published papers on earthquakes in the time window 1000-1899 reporting the seismological interpretations based on historical records. In particular, publications are preferred with earthquake effects expressed in terms of macroseismic intensity. The AHEAD data curators extract from the descriptive interpretations the relevant data, such as the MDPs (Macroseismic Data Points) and earthquake parameters and associate the extracted data to the corresponding earthquake. By collecting and making available all these published interpretation works, the archive ends up with multiple research results for the same earthquake, providing users a convenient tool to access the available seismological knowledge.

The collection process is complicated by the heterogeneity of the studies, because each historical seismologist structures his/her own research outcomes adopting different approaches, either providing ready-to-use data tables, or maps only, or lengthy descriptions that must be manually and carefully parsed in order to preserve the complexities of the information. Along the years, each European country has chosen different ways to perform new research on the past seismicity and to make the results available. Taking into account that there are around 50 sovereign states with 24 different languages, the resulting picture is quite complex. AHEAD deals with these difficulties by adopting regional nodes, each responsible for his own geographical area. AHEAD so far managed to involve 12 institutions that covers 9 countries. Designed and developed in 2010 and later on updated, an in-house developed software tool called MIDOP (Macroseismic Intensity Data Online Publisher) was established in order to support the growth of regionally managed archives. The same INGV research team that coordinates AHEAD manages also the Italian regional node called ASMI, the "Italian Archive of Historical Earthquake Data". ASMI was publicly opened in 2017 thanks to the financial support of the Italian Civil Protection.

In the last decade the European Commission strongly committed herself to adopt the Open Science paradigm in all areas of ERA (European Research Area), a new approach that is trying to replace the slow habit of sharing knowledge by means of the publications, with much faster, and more flexible solutions based on modern communication solutions (i.e. web services). In this framework, the European Commission supported the establishing of disciplinary-specific ERICs, European Research Infrastructures, legal entities aiming at introducing Open Science in their own scientific field. In the field of Geosciences, the reference ERIC is EPOS (European Plate Observing System). In 2015 AHEAD applied to be the EPOS data provider for data on historical earthquakes. For the application to be validated, AHEAD had to setup a series of web services to enable the automatic exchange of machine-readable data, maintaining in parallel the traditional human driven of flowing of data among researchers.

The very same approach of the archiving methodologies adopted by AHEAD, was replicated at a worldwide scale in the Global Historical Earthquake Archive (GHEA), a project financed between 2010 and 2013 by the Global Earthquake Model (GEM) Foundation.

This presentation will guide you through the above-described scenario, focusing on the last decade and explaining the adopted solutions, to provide you with details on how historical seismology in Europe is striving to contribute to the available seismological knowledge.