Engineering Structures 136 (2017) 535–546

Contents lists available at ScienceDirect

Engineering Structures

journal homepage: www.elsevier.com/locate/engstruct

Next directions in experimental data for seismic hazard mitigation

Ignacio Lamata Martínez [a], Martin S Williams [a], Shirley Dyke [b,*], Markus Krötzsch [d], Pierre Pegon [c]

[a] University of Oxford, Department of Engineering Science, United Kingdom
[b] Purdue University, Lyles School of Civil Engineering, United States
[c] European Commission, Joint Research Centre (JRC), Directorate for Space, Security and Migration, Safety & Security of Buildings Unit, Italy
[d] Center for Advancing Electronics Dresden (cfaed), TU Dresden, Germany

ARTICLE INFO

Article history:
Received 24 August 2016
Revised 4 October 2016
Accepted 10 December 2016

Keywords:
Data integration
Experimental data
Distributed systems
Earthquake engineering
Semantic web
Interoperability

ABSTRACT

Data are one of the main assets of earthquake engineering. Laboratory experiments can be extremely expensive and time consuming to replicate and, therefore, long-term preservation of experimental data and sharing the data with users has become one of the disciplinary priorities. There is a growing demand for international partnerships, which creates a need for data sharing, in an attempt to maximise research impact and to tackle experimental set-ups that could not be realised otherwise. However, there is a patent lack of interoperability between the institutions forming the earthquake engineering community, which inhibits efficient collaboration between them. In this paper, we discuss a vision about the directions that experimental data should take in the coming years, focusing on two aspects: enhanced international collaborations and implementation of open data access. We also describe the progress that has been made towards this vision, by establishing an open platform for the integration of earthquake hazard mitigation resources called Celestina. Celestina is supported by Semantic Web technologies, and uses an ontology as its integration data model. A prototype of the platform has been developed and tested between NEES (Purdue University, in the US), the University of Oxford (in UK) and EUCENTRE (in Italy), and a small proof of concept has enabled integrated experimental data from Oxford and EUCENTRE through the NEES cyberenvironment. This demonstration provides an example which has the potential to catalyze a new generation of research progress enabled by international data sharing.

1. Introduction

Earthquake engineering is the scientific discipline that studies the nature of earthquakes and their effects on structures, such as buildings or bridges, in order to diminish their damage and with the ultimate goals of saving human life and minimising economic loss. The study of how an earthquake affects different structures and their materials is either conducted by observing the behaviour of structures that have been subjected to a real earthquake, or conducted by means of experiments, where part or a full structure, soil or devices are placed under a seismic stimulus in a laboratory and their behaviour is recorded for subsequent analysis. Experiments provide valuable data for the development of systems to minimise the impact of earthquakes, but their execution can be very costly. Computing advances have enabled new methods of evaluating seismic performance, for example by simulating the behaviour of structures numerically. However, uncertainty over physical parameters and limitations on the understanding of element behaviour can limit the applicability of computer models, requiring the actual physical testing of the structure.

Data collected from real earthquakes and experiments are a highly valued resource in earthquake engineering, since laboratory experiments can be extremely expensive and time consuming to replicate, and the effects of real earthquakes are impossible to record a second time. Besides the need of a formalised way to store experimental data, the expanding demand for international collaborations creates a need for data sharing in an attempt to maximise research efforts and to tackle experiment set-ups that could not be realised otherwise. However, there is a patent lack of interoperability between the institutions forming the earthquake engineering community, which introduces barriers to efficient collaboration between them. Moreover, even though there exist some efforts to share experimental data (such as those of NEES[1] in the US and SERIES[2] in Europe), this only addresses the consumption of data by

* Corresponding author.
E-mail addresses: ignacio.lamata@oxfordalumni.org (I. Lamata Martínez), martin.williams@eng.ox.ac.uk (M.S Williams), sdyke@purdue.edu (S. Dyke), markus.kroetzsch@tu-dresden.de (M. Krötzsch), pierre.pegon@jrc.ec.europa.eu (P. Pegon).

[1] https://nees.org.
[2] http://www.series.upatras.gr.

http://dx.doi.org/10.1016/j.engstruct.2016.12.012
0141-0296/

humans, since the lack of interoperability prevents automated data consumption.

In this paper, we present a vision about the directions that experimental data should take in the coming years that builds upon the past successes in the earthquake engineering community (e.g. [1,2]). We focus the discussion on two aspects: enhanced international collaborations and implementation of open data policies. We also describe the work we have done towards this vision, by creating a platform for the integration of earthquake hazard mitigation resources.

2. State of the art

2.1. Current status of experimental data in earthquake engineering

The earthquake engineering community has been a leader in the effort to provide open data to accelerate scientific advances needed to build more resilient communities around the world. The most noteworthy examples of experimental data repositories are those of NEES (Network for Earthquake Engineering Simulation) and SERIES (Seismic Engineering Research Infrastructures for European Synergies). The NEES infrastructure [3] has a centralised data repository to store and share experimental data, whose data model is described by [4,5]. The quality and completeness of the data are verified and improved by means of a process called data curation, in which the data are monitored and approved for publication [6]. NEES has commitments in terms of digital data preservation, to ensure that data remain accessible and usable [7], and provides open licenses and digital identifiers for both documents and data resulting from experimental activities. The NEES database provided data integration to the earthquake engineering institutions in the USA. The database is supported by a number of tools such as inDEED [8], and PEN [9]. NEES has been superseded by NHERI (Natural Hazards Engineering Research Infrastructure), whose cyberinfrastructure component[3] is in charge of NEES legacy of data. Changes to the data management model are expected in the next few years.

In Europe, SERIES created the SERIES Virtual Database [10], an infrastructure of distributed data sources that allows access to the SERIES European laboratories' data through a single, centralised user interface, and that enables data integration between the different SERIES institutions. The main reasons for a decentralised approach (instead of centralised) are that European institutions need to be autonomous and have control of their own data; also, a decentralised solution contributes to reduce the technological gap between each laboratory. The distributed nature of the virtual database is intended to be invisible to the end user, whose experience should be similar to accessing a single data repository.

Other databases have been internally created at different institutions around the world, such as at the Joint Research Centre and EUCENTRE in Italy, and E-defense in Japan, but they have not been documented publicly. EPOS[4] (European Plate Observing System) also provides public access to integrated multidisciplinary Earth science data at a European level, such as data from volcano observatories, seismic waveform data from ORFEUS[5] and laboratory experimental data from other European institutions.

A digital archive of video, audio, images and documents related to seismic events across New Zealand has also been established within the CEISMIC center[6]. The number and variety of repositories is evidence of the continued importance of accessing data from experiments to enable research and scientific advances in this field.

2.2. Data policies

Globally, data policies on publicly funded research are increasingly moving to an Open Access schema, in which research output is published free to access, redistribute and, in many cases, reuse. Examples of funding bodies requiring or encouraging to publish research outcomes on an Open Access basis are the Research Councils UK [11], Wellcome Trust [12], the US National Science Foundation [13], the US Department of Energy [14] and the US National Institutes of Health [15]. In the UK, HEFCE (Higher Education Funding Council for England, the government agency responsible for university core funding) will require open access publications for the next research assessment process [16]. In the US, the National Science Foundation, which supports about one-quarter of all federally-funded research in the US, has required a data management plan for all research projects since 2013, and the US government recently adopted the Open Government Plan 3.5 [17]. It will not be long before public funding bodies strictly require all their funded publications to be disseminated under an Open Access schema. Open Access is often supported by publishers (e.g. BioMed Central [18], Plos [19], etc.) and repositories (e.g. PubMed Central [20], Europe PMC [21], repositories indexed by OpenAire [22], etc.).

The next step after the publishing industry adopts Open Access is to embrace Open Data. Sources of public data are also moving in the Open Data direction, and good examples are the governments of the USA [23], Europe [24] and Asia [25] amongst others. In Earthquake Engineering, the NEES repository has encouraged the use of Open Data, and nearly all of their experimental outcomes have been published under such schema [26]. Also, other seismic related repositories such as the ones at EPOS and ORFEUS are encouraging the adoption of open data [27,28].

Open Data creates new opportunities for data sharing and reuse, but at the same time imposes additional requirements on infrastructures to enable such interoperability. Data management and publishing systems need to shift their focus accordingly: new tasks of data exchange and information integration are gaining prominence, while traditional concerns such as user rights management and access control are less critical. New types of Open Data management systems should therefore be tailored towards this new form of publishing.

2.3. Interoperability efforts in other sciences

Science Gateways have been developed within several scientific domains in an effort to provide essential, often domain-specific, cyberinfrastructure services to researchers to accelerate research and knowledge generation [29]. Some other sciences have also attempted to create a global interoperability environment. For example, there has been an effort in bioinformatics to integrate different data sources that already expose their data in RDF, a Semantic Web technology [30]. This effort primarily focused on the integration of three existing datasets by using SPARQL, another Semantic Web technology that is briefly described later, without alteration of the existing data. By using SPARQL, an integrated layer is created and the three repositories can be queried in a unified, uniform manner. One of the challenges to integration that the authors had to resolve was the use of a common identifier when a same object is referenced in different repositories. Such problematic is not currently found in seismic data, since disparate repositories do not store information for the same objects.

Bio2RDF is another effort from the domain of bioinformatics that aims for the integration of data in the life sciences [31]. It is also based on RDF and provides a simple convention guideline

[3] https://www.designsafe-ci.org.
[4] https://www.epos-ip.org/.
[5] http://www.orfeus-eu.org/.
[6] http://www.ceismic.org.nz.