Crowd-Sourced Buildings Data Collection and Remote Training: New Opportunities to Engage Students in Seismic Risk Reduction

Antonella Peresan*, Chiara Scaini and Carla Barnaba

National Institute of Oceanography and Applied Geophysics—OGS, Trieste, Italy

Young generations are increasingly committed to understanding disasters, and are a key player in current and future disaster risk reduction activities. The availability of online tools opened new perspectives in the organization of risk-related educational activities, in particular in earthquake-prone areas. This is the case of CEDAS (building CEnsus for seismic Damage Assessment), a pilot training activity aimed at collecting risk-related information while educating high-school students about seismic risk. During this experimental activity, students collected and elaborated crowdsourced data on the main building typologies in the proximity of their homes. In a few months, students created a dataset of valuable risk-related information, while getting familiar with the area where they live. Data collection was performed both on-site, using smartphones, and online, based on remote sensing images provided by multiple sources (e.g., Google maps and street view). This allowed all students, including those with limited mobility, to perform the activity. The CEDAS experience pointed out the potential of online tools and remote sensing images, combined with practical activities and basic training in exploratory data analysis, to engage students in an inclusive way. The proposed approach can be naturally expanded in a multi-risk perspective, and can be adjusted, eventually increasing the technical content of collected information, to the specific training and expertise of the involved students, from high-school to university level.

Keywords: exposure data collection, online education, crowdsourcing, citizen science, seismic risk awareness

THE CEDAS APPROACH

Citizens' perception of disaster risk is crucial to improve societal response and mitigate impacts (UNDRR, 2022). For this reason, there is a growing attention to risk-related educational activities for younger generations (UNISDR, 2017; UNICEF Education Section, 2019). However, during the pandemic, most educational activities were disrupted due to the COVID-related limitations.

The pilot project "CEDAS: building CEnsus for seismic Damage ASsessment" (Scaini et al., 2021) started in 2021 in northeastern Italy, a seismic-prone area where the last destructive event happened in 1976. Despite a number of didactic activities being developed in the area (Saraò et al., 2013; Barnaba et al., 2018), seismic risk perception is low among young generations (Peruzza et al., 2018). CEDAS was originally aimed at engaging young students during the lockdown, yet it turned out to be a flexible tool, suitable for a variety of training activities. Students

FIGURE 1 | Example of use of satellite images (extracted from Google Maps) for identifying the building shape. Similar images were also used to exemplify roof types. The questionnaire could be visualized on computer, tablet or smartphone. Smartphone icon created by Pedro Santos from Noun Project.

FIGURE 2 | Examples of training material based on Google Maps and Google Street View images for identification of homogeneous town compartments. Training was delivered online and recorded to deliver thematic videos. Icon of Computer screen by Philipp Petzka from Noun Project.

collected data on the main building characteristics (i.e., the buildings exposure) near to their homes, by compiling an online form using their phone or personal computer. The potential of this activity soon became evident: in less than 4 months, the project involved more than 170 students collecting a large amount of reliable data on common building typologies (Scaini et al., 2021). The experiment was then repeated in 2022 with another 150 students, reaching up to 6,000 building filled forms. This is the first example of exposure data collection performed by students, while similar activities were conducted by adults (e.g., Grigoratos et al., 2018), or by students to recognize damages (e.g., Davis, 2021). But CEDAS does not only consist of mere data collection. Students were trained on how to assess the characteristics of exposed assets and to statistically analyze the collected data, performing citizen science activities (Lee et al., 2020). At the end of each CEDAS edition, students shared and discussed their findings in a final public event. Preliminary exploratory data analysis carried out by the students, as well as specific cross-checking and validation performed by expert researchers (including independent expert evaluation of a subsample of representative buildings surveyed by the students), allowed estimating the quality of the collected data and their suitability for exposure assessment (Scaini et al., 2021). Hence the relevance of the CEDAS experience is twofold: educational,
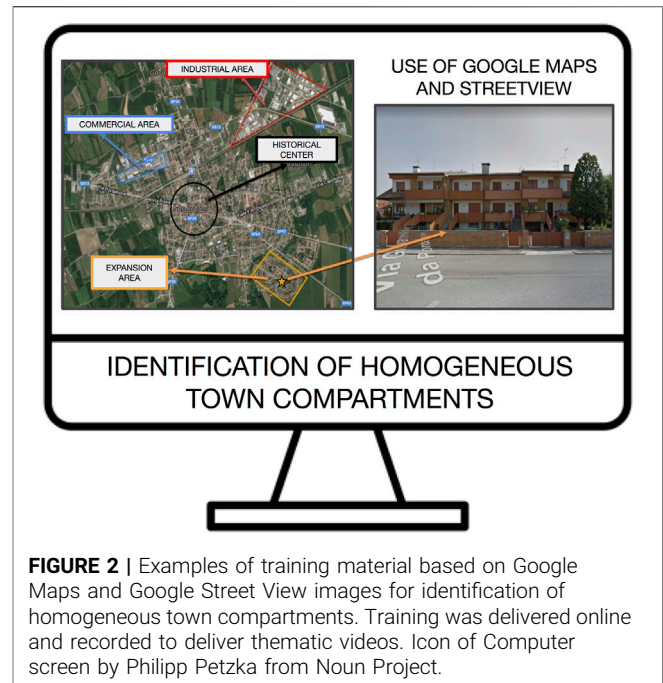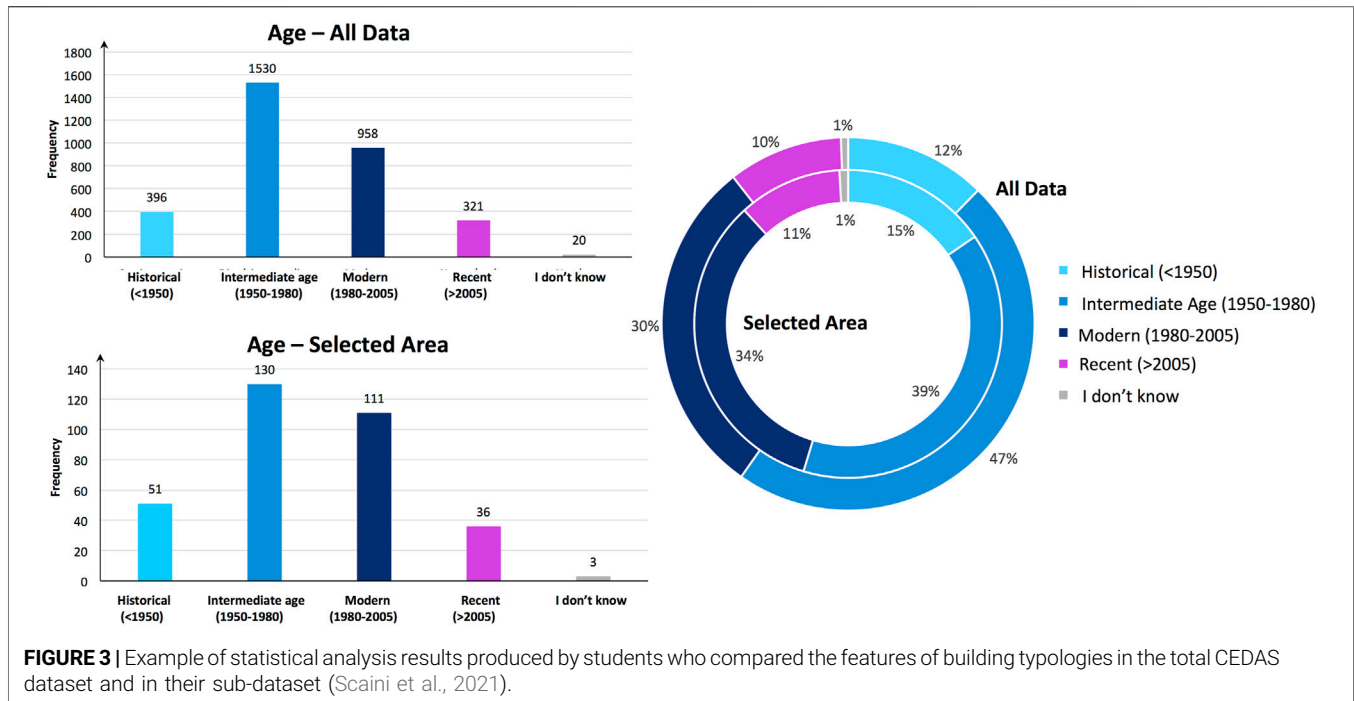
increasing knowledge and risk awareness level of involved students, and scientific, allowing not only for extensive data collection/update, but also providing useful methodological hints. The comparative analysis and critical discussion with data collected using remote sensing images and online data provided by multiple sources (including Google maps and Street View), in fact, allowed us to explore the limits and potential of such resources, which can be used to guide automated data analysis by artificial intelligence tools (e.g., Pelizari et al., 2021).

TRAINING METHODS AND DATASETS

All students attended two online training sessions. In the first one, students were familiarized with a form designed with a free online tool (Google Forms) to gather building data. Responders need a Google account only if they upload building images, which was not mandatory. Students were trained to answer the form questions on the main building characteristics (e.g., age, material and storey number). The training material included images collected in the specific study area, e.g., Google maps aerial images (**Figure 1**) and street view screenshots. Google maps was suggested as a tool to investigate features not directly visible from the street (e.g., roof type, building shape).

During the training, particular attention was devoted to explaining how buildings are geographically distributed in different areas (e.g., historical centers, commercial/industrial areas, suburbs). This helped students in recognizing building typologies from the urban context. To this aim, Google Maps and Street View images were particularly useful (**Figure 2**). The

FIGURE 3 | Example of statistical analysis results produced by students who compared the features of building typologies in the total CEDAS dataset and in their sub-dataset (Scaini et al., 2021).

possibility of using different data sources, including satellite images and/or land use data, was mentioned and might be included in the activity in future. Based on our experience, we suggest that training material is adapted to the characteristics of the different areas, using images taken from the specific context, as this facilitates recognition of the building features (e.g., age or maintenance).

The second training session was focused on data analysis and interpretation. Students received a subset of the collected data [see Scaini et al. (2021) for details] together with the entire dataset. **Figure 3** shows an example of the analyses performed by a group of students comparing the two datasets. All analyses were done using electronic spreadsheets (Microsoft Excel or Libreoffice) and online open access tools [e.g., Google spreadsheets, Easystat[1]]. All training material was made available online, so that students were able to access it at every time.
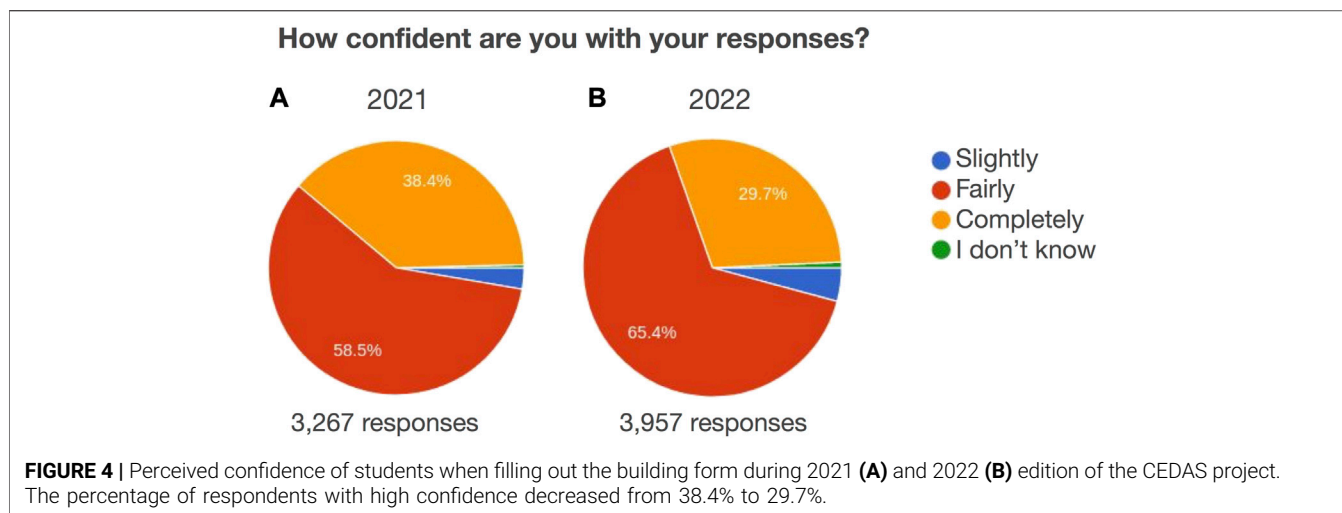
During the first edition of the CEDAS experiment in 2021, students were requested to collect the relevant information based on a direct on-site observation of the buildings. In 2022 the experiment was carried out in fully virtual mode, based on indirect observation of buildings, through online images and maps, with the aim to explore the possibilities provided by satellite images and other online data resources. The comparative analysis between results obtained from the two experiments, provided new insights on the limits and potential of such resources, which can be used to guide automated data analysis by artificial intelligence tools (e.g., Pelizari et al., 2021). At the stage

of final data interpretation and discussion, in fact, students were requested to comment on the building features that could be easily defined based on online information (e.g., roof type and shape), and those for which limited or no information could be retrieved (e.g., presence of a basement). Accordingly, the acquired information can be used either to provide the learning set for specific well defined features, or as a constraint towards deep learning analysis of satellite images.

STUDENTS' FEEDBACK AND LESSONS LEARNED

In 2021, after the first edition of the project, we collected students' feedback with an online questionnaire (Scaini et al., 2021). Students enjoyed the activity both in remote and on-site mode, but more enthusiasm was expressed for the active part (i.e., collecting buildings data). 45% of students preferred to survey buildings *in situ*, while 35% favored a hybrid mode (online and *in situ*). In 2022 the activity was performed only in virtual mode, with the students compiling the survey only based on images available online (i.e., Google maps and StreetView). The overall students' perceived confidence decreased from 2021 to 2022 (**Figures 4A,B**, respectively), despite having received the same training. This is likely due to the limitations of carrying out the survey based on online information, without direct observation of the building. During discussions, students pointed out advantages and disadvantages of carrying out the activity online. Online images might be out of date or might not allow them to

---

[1]https://easystat.com/

FIGURE 4 | Perceived confidence of students when filling out the building form during 2021 **(A)** and 2022 **(B)** edition of the CEDAS project. The percentage of respondents with high confidence decreased from 38.4% to 29.7%.

recognize building characteristics, increasing the perceived uncertainty. However, online activity enhances their abilities with online tools. According to students, the most difficult aspect to be recognized is the material, while in the previous edition age was found to be the most difficult one, followed by material (Scaini et al., 2021). Both during practical and online activities, students declared to benefit from direct interaction with their schoolmates. However, only the activity *in situ* supported the direct interaction with building residents (to whom questions were seldom posed). During the discussion, students' declared that the interaction with others increased their confidence in the responses.

Students' feedback suggests that practical activities increase both their observation ability and their personal engagement. Teachers also expressed a high interest in the activity, but pointed out two main difficulties associated with remote mode: communication with the students and statistical analysis. According to students, these difficulties might be mitigated by personal interaction, and in particular team-working. We therefore conclude that the training on statistical analysis and results interpretation, might be more effective in person.

Most students perceived that the activity is useful for them given that they learned new tools (e.g., advanced use of excel) and developed skills (active observation, self-organization, team-working) demanded in most workplaces. They also perceived the importance of their contribution for the scientific community and were highly motivated to contribute to seismic risk reduction in their homeplace. The importance of building performance emerged multiple times during the discussion, in particular in areas where buildings were partially reconstructed after the 1976 destructive Friuli earthquake, which was explicitly mentioned by the students. Students coming from technical schools also devoted special attention to the building's state of conservation which was found to be moderate or poor in some cases.

PRELIMINARY RESULTS AND POTENTIAL DEVELOPMENTS OF THE CEDAS APPROACH

The CEDAS experience aims to go beyond the limits of traditional school programs and citizen science activities. The activity is not only a mere data collection and building inventory. Students learn to observe and classify buildings, based on visual inspection and remote sensing images. They also learn to use processing software, to support their considerations with data, as well as to improve their critical attitude and confidence in discussion of the results. Results demonstrate that trained citizens and scientists can fruitfully collaborate in increasing risk-related knowledge and, subsequently, societal resilience in seismic-prone areas. Citizens, in this case students, are rewarded by contributing to scientific knowledge: most respondents would like to repeat the activity in future. Scientists, in turn, can benefit from the exposure data provided by trained citizens, which enriches their databases.

All phases of the proposed activity, from training, to data collection and validation, benefitted from remote sensing images (e.g., Google maps satellite images). Online tools (e.g., Google Street View) successfully enabled the project in virtual mode, reaching a large audience of citizens with different mobility needs. However, some activities (e.g., statistical analysis and interpretation of collected data) may significantly benefit from in presence interaction. Despite demonstrating the full potential of online tools, we also stress the importance of in presence activities. Future editions of CEDAS should therefore aim at an optimal balance between virtual and in presence activities, depending on the specific context.

The CEDAS pilot experiment demonstrated the validity of the proposed approach, namely, the use of a simple and specific tool (the online form) that is easy to use (through the smartphone) and may use readily available data (Google Maps and images). The CEDAS experience also highlighted its double relevance: the first is the educational value, aimed at

increasing the level of knowledge and resilience of new generations. The second is scientific, as it demonstrated that the data collected are of high quality, suitable for quantifying the population's exposure to risk, and can provide useful insights towards automated analysis by machine learning and artificial intelligence tools. Artificial intelligence and machine learning methods are already used for damage recognition (Xie et al., 2020), but their potential for exposure assessment is still poorly explored despite the increasing availability of potential source data (e.g., from remote sensing). Data collected here and their associated exposure features might serve as a training dataset for artificial intelligence applications, towards a data driven exposure development. The approach might as well be integrated with other existing platforms that collect crowdsourced data (e.g., Gomez Zapata et al., 2021) and/or street view images (e.g., Mapillary[2]) relevant for the exposure assessment.

The proposed approach can be adopted in different seismic-prone areas worldwide, and even for multiple risks, reaching a wide audience of citizens with different backgrounds and mobility limitations and contributing to mitigate disaster risk. CEDAS can be especially useful in developing countries, where information about exposed assets is still limited and rapidly evolving. The success of CEDAS relies on the involvement of local communities, in particular school students, to target the activity and define training materials depending on the educational settings of the country/region at stake. In addition, local research groups and other stakeholders (e.g., from the local governments) should be involved and support the definition of priorities (e.g., building relevant features, most relevant areas at stake).

THE CEDAS PERSPECTIVE FOR DISASTER RISK REDUCTION

Natural disasters may seriously affect the achievement of Sustainable Development Goals (SDGs[3]). Earthquakes, in particular, have serious societal impacts such as the increase of poverty and inequalities. Schools play a central role in disaster risk reduction (UNICEF Education Section, 2019) as demonstrated by the impact of educational activities to effectively increase community awareness and preparedness (Parham et al., 2021). However, DRR concepts should be increasingly integrated in teaching programs and encompass both prevention and preparedness, with specific attention to practical activities (Apronti et al., 2015). CEDAS contributes in this sense by proposing a practical activity that involves both students and teachers, and can be adapted to different contexts (e.g., limited mobility, lack of data). By engaging and educating young students from local communities, CEDAS increases citizens' risk awareness and

preparedness, both identified as uppermost aspects for disaster risk reduction (UNDRR, 2022).

Besides contributing to increase risk-related knowledge in citizens, the CEDAS approach can also support scientific advances, by providing up-to-date datasets to be used to enhance existing exposure layers in combination with other data sources (e.g., remote sensing data, ancillary data such as building Census). This could be achieved by developing Artificial Intelligence (AI) tools, taking advantage from the experience provided by human analysis (students, citizens and experts), and using Machine Learning (ML) algorithms to classify and make inferences about building typologies, based on a heterogeneous set of available information. Such an approach should be tailored depending on the specific context and on the amount of data available for the AI learning and classification phases. The use of AI and large amounts of data may provide important insights about the uncertainties related with crowd sourced exposure data collection, as well as about the uncertainty in the classification of buildings. In the future, this procedure might be implemented dynamically, in order to capture rapid changes in building typologies and urban expansion, and might contribute developing the "digital twin" approach in the Earth system sciences (e.g., the DTGEO project[4]).

The development of up-to-date exposure datasets is of paramount importance to define medium-to-long-term disaster risk reduction strategies, as envisaged by the Sendai framework objectives for disaster risk reduction (UNISDR, 2017). CEDAS could be extended, with minimum changes, to the collection of exposure data for other natural phenomena (e.g., flood, tsunami, landslides) with the aim of developing multi-hazard and multi-risk strategies and foster disaster risk reduction practices (UNDRR, 2022).

With the involvement of local communities, CEDAS is potentially adaptable to different contexts and might be integrated in ongoing global disaster risk reduction efforts by engaging young citizens, increasing risk awareness and collecting reliable and up-to-date exposure data, especially needed in less developed countries.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because Data can be provided only in aggregated form by request. Requests to access the datasets should be directed to the corresponding author.

ETHICS STATEMENT

Humans contributed only to data collection and analysis, and they were not the object of investigation and study. For this reason ethical approval was not required for the study in accordance with the local legislation and institutional requirements. Written

---

[2]https://www.mapillary.com/datasets
[3]https://sdgs.un.org/goals

[4]https://dtgeo.eu/

informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

AUTHOR CONTRIBUTIONS

Specifically: AP and CS drafted the manuscript; AP, CS, and CB revised the manuscript. The fields of prevalent contribution by each author are as follows: CS exposure data collection; AP statistical data analysis; CB education and organisation. All authors contributed to the article and approved the submitted version.

FUNDING

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ACKNOWLEDGMENTS

REFERENCES

Apronti, P. T., Osamu, S., Otsuki, K., and Kranjac-Berisavljevic, G. (2015). Education for Disaster Risk Reduction (DRR): Linking Theory With Practice in Ghana's Basic Schools. *Sustainability* 7 (7), 9160–9186. doi:10.3390/su7079160

Barnaba, C., Contessi, E., and Girardi, M. R. (2018). A High School Students' Geophysical Survey in a Seismically Active Area: The PRESS40 Project. *Seismol. Res. Lett.* 89 (4), 1539–1545. doi:10.1785/0220170219

Davis, W. (2021). Google Street View as a Teaching Tool for Assessing Earthquake Building Damage. *Seismol. Res. Lett.* 93, 474–476. doi:10.1785/0220210148

Gomez-Zapata, J. C., Brinckmann, N., Harig, S., Zafrir, R., Pittore, M., Cotton, F., et al. (2021). Variable-Resolution Building Exposure Modelling for Earthquake and Tsunami Scenario-Based Risk Assessment: An Application Case in Lima, Peru. *NHESS* 21 (11), 3599–3628. doi:10.5194/nhess-21-3599-2021

Grigoratos, I., Monteiro, R., Ceresa, P., Di Meo, A., Faravelli, M., and Borzi, B. (2018). Crowdsourcing Exposure Data for Seismic Vulnerability Assessment in Developing Countries. *J. Earthq. Eng.* 25, 835–852. doi:10.1080/13632469.2018.1537901

Lee, K. A., Lee, J. R., and Bell, J. (2020) A Review of Citizen Science Within the Earth Sciences: Potential Benefits and Obstacles, *Proc. Geologists' Assoc.* 131 (6) 605–617. doi:10.1016/j.pgeola.2020.07.010

Parham, M., Teeuw, R., Solana, C., and Day, S. (2021). Quantifying the Impact of Educational Methods for Disaster Risk Reduction: A Longitudinal Study Assessing the Impact of Teaching Methods on Student Hazard Perceptions. *Int. J. Disaster Risk Reduct.* 52, 101978. doi:10.1016/j.ijdrr.2020.101978

Pelizari, A. P., Geiß, C., Aguirre, P., Santa María, H., Merino Peña, Y., and Taubenböck, H. (2021). Automated Building Characterization for Seismic Risk Assessment Using Street-Level Imagery and Deep Learning. *ISPRS J. Photogrammetry Remote Sens.* 180, 370–386. doi:10.1016/j.isprsjprs.2021.07.004

Peruzza, L., Sarò, A., Barnaba, C., and Massolino, G. (2018). Elapsed Time: 40 Years What Young People of Friuli Venezia Giulia Know About the 1976 Earthquakes, Natural Hazard and Seismic Safety. *Boll. Geof. Teor. Appl.* 59, 575. doi:10.4430/bgta0227

Saraò, A., Barnaba, C., and Peruzza, L. (2013). Keeping Focus on Earthquakes at School for Seismic Risk Mitigation of the Next Generations. *Geophys. Res. Abstr.* 15, EGU2013–10560.

Scaini, C., Peresan, A., Tamaro, A., Poggi, V., and Barnaba, C. (2021). Can High-School Students Contribute to Seismic Risk Mitigation? Lessons Learned From the Development of a Crowd-Sourced Exposure Database. *Int. J. Disaster Risk Reduct.* 69 (2022), 102755. doi:10.1016/j.ijdrr.2021.102755

UNDRR (2022). Global Assessment Report on Disaster Risk Reduction. Available at: https://www.undrr.org/gar2022-our-world-risk (Accessed July 1, 2022).

UNICEF Education Section (2019) *Risk-Informed Education Programming for Resilience Guidance Note*. New York: UNICEF.

UNISDR (2017). Sendai Framework for Disaster Risk Reduction - UNISDR, Available at: http://www.unisdr.org/we/coordinate/sendai-framework (Accessed July, 1 2022).

Xie, Y., Ebad Sichani, M., Padgett, J. E., and DesRoches, R. (2020). The Promise of Implementing Machine Learning in Earthquake Engineering: A State-Of-The-Art Review. *Earthq. Spectra* 36 (4), 1769–1801. doi:10.1177/8755293020919419