Date of Submission	15 July 2025

IPL Project Proposal Form 2025
(MAXIMUM: 3 PAGES IN LENGTH)

1. Project Title: (2 lines maximum): **AI-Powered Extraction of Rainfall-Induced Landslide Information (AI-PERIL)**

   Select one of two below.

   (1) New project

2. Main Project Fields

   Select the suitable topics. If no suitable one, you may add new field.

   (1) Technology Development

   A. Monitoring and Early Warning

   (3) Capacity Building

   B. Collating and Disseminating Information/ Knowledge

   (4) Mitigation, Preparedness and Recovery

   B. Mitigation

3. Name of Project leader: **Maria Teresa Brunetti**

   Affiliation: **CNR IRPI – Senior Researcher**

   Contact: (postal address, fax, phone, email): **Via Madonna Alta, 126, 06128 Perugia, phone number: +39 075 5014410, email: mariateresa.brunetti@cnr.it**

   Core members of the Project

   Names/Affiliations: **Stefano Luigi Gariano (CNR IRPI – Senior Researcher)**, **Massimo Melillo (CNR IRPI –Researcher)**, **Elisabetta Napolitano (CNR IRPI – Associated Researcher)**, **Silvia Peruccacci (CNR IRPI – Senior Researcher)**

4. Objectives: (5 lines maximum; what you expect to accomplish?):

   The objective of the project is to develop AI-based procedures to automatically extract accurate information on the occurrence of rainfall-induced landslides from textual sources available on the internet, including web pages and PDF documents. The information includes the location of the landslide and the date/time of its occurrence, as well as the spatial and temporal accuracy of the extracted information.

5. Background Justification: (10 lines maximum)

   A reliable landslide forecast requires the collection of information on historical landslides in a structured catalogue. Using landslide catalogues that contain inaccurate spatial and temporal information leads to unreliable and uncertain operational landslide forecasting. Therefore, the availability of accurate and detailed catalogues is essential to reduce the uncertainties, which are to some extent unavoidable. To this end, for the last 15 years, many researchers from CNR IRPI have been actively involved in the compilation of ITALICA (the ITAlian rainfall-induced LandslIdes CAtalogue), which currently contains 6312 records with information on rainfall-induced landslides that have occurred on the Italian territory between

January 1996 and December 2021 (Peruccacci et al., 2023; Brunetti et al., 2025). Extracting data on the occurrence of rainfall-induced landslides manually from the main information sources (online newspapers, technical reports, blogs and social media) is extremely time-consuming and requires considerable human and financial resources. The use of AI could solve this critical issue.

6. Study Area: **World**

7. Project Duration: **Two years**

8. Resources necessary for the Project and their mobilization
   Personnel, Facilities, and Budgets

Researchers from CNR IRPI (permanent staff and research associates);

Computers and servers to manage and store data at CNR IRPI;

Budget: A financial contribution may be needed to buy subscriptions to AI tools, to optimize and speed up information extraction operations. Additional expenses are foreseen for the publication of the results.

9. Project Description: (30 lines maximum)

Rainfall-induced landslides pose a serious geo-hydrological hazard, causing extensive damage to infrastructure and loss of life. Understanding the rainfall conditions that cause their occurrence is critical to their prediction. For large and diverse territories, the prediction of landslides is based on the definition of rainfall thresholds, which require the reconstruction of the rainfall conditions that caused the landslides. Therefore, the collection of accurate information on the spatial and temporal occurrence of past landslides is necessary so that the reconstruction of the rainfall that triggered them can be objective and reliable. The search for and selection of news from information sources is a long, complex, and repetitive procedure. CNR IRPI has invested more than a decade in information gathering and built ITALICA, the ITAlian rainfall-induced LandslIdes CAtalogue (Peruccacci et al., 2023; Brunetti et al., 2025).

Traditionally, researching and collecting accurate spatial and temporal information on rainfall-induced landslides has been time-consuming. This involves searching for news reports in newspapers and online blogs, as well as technical and scientific articles. In addition, this activity must be preferably performed by an experienced operator, since it can be prone to human errors and limited by individual processing capacity. Artificial intelligence (AI) can be a valuable support at the stage of collecting and selecting data for analysis. In fact, AI, properly trained by an experienced operator, can process and interpret large volumes of information in significantly shorter times than those required by a human operator. This, therefore, allows a greater amount of data to be collected in reasonable time, thereby reducing the uncertainty of the parameters that define rainfall thresholds.

The project aims to implement AI-based procedures for accurate reconstruction of the location and time of occurrence of landslides from online sources.

The AI-based procedures will be obtained using an iterative refinement method and natural language process (NLP) to interact with the large language model (LLM). This will enable the design of prompts that can generate structured and standardized datasets containing accurate, objective, reliable, and reproducible spatiotemporal information on rainfall-induced landslide occurrences in several areas of the world.

Language issues will also be considered. The outputs provided by AI will be validated with those of the experienced operator to optimize the prompt for information retrieval. The results, i.e. the creation of new landslide catalogues in various regions of the world and the updating of Italian rainfall-induced landslide catalogues, will be published in open-access repositories. This will allow the catalogues to be used to define, validate, and improve methods and tools for predicting landslides in space and time.

10. Work Plan/Expected Results: (30 lines maximum; work phases, milestones and publication) including the contribution plan of articles on the IPL project (progress/result) to the Open Access Book Series P-LRT in the coming few years.

The main phases of the project workflow are listed below:

1) retrieving and collecting landslide news from the internet;

2) designing, implementing and applying prompts (i.e., textual inputs provided to the AI tool) for the automatic extraction of spatiotemporal information from the collected sources;

3) verifying and validating the extracted information through manual cross-checking and

4) compiling or updating the landslide catalogues.

The workflow is scalable and adaptable: it can be applied in Italy, using textual inputs and parameters in Italian, as well as in other countries, by adapting the textual inputs to English or other local languages in which the AI has a high degree of linguistic expertise and skills.

Milestones:

M1) Collecting a sample of news articles about rainfall-induced landslides in pre-defined areas.

M2) Extracting accurate spatial and temporal information from news using the designed prompt.

The main outcomes of the project, as well as the main methodological innovations, will be summarized in an article to be submitted to the Open Access Book Series P-LRT in the coming few years.

11. Deliverables/Time Frame: **One report at the end of each year**

12. Project Beneficiaries: (5 lines maximum; who directly benefits from the work?): **The beneficiaries of the project are research institutions involved in forecasting rainfall-induced landslides and ultimately civil protection authorities.**

13. References (Optional): (6 lines maximum; i.e. relevant publications)
- Brunetti, M.T., Gariano, S.L., Melillo, M., Rossi, M., Perucacci S. (2025) An enhanced rainfall-induced landslide catalogue in Italy. Scientific Data 12, 216, https://doi.org/10.1038/s41597-025-04551-6
- Peruccacci, S., Gariano, S. L., Melillo, M., Solimano, M., Guzzetti, F., Brunetti, M. T. (2023) The ITAlian rainfall-induced LandslIdes CAtalogue, an extensive and accurate spatio-temporal catalogue of rainfall-induced landslides in Italy, Earth System Sciences Data, 15, 2863–2877, https://doi.org/10.5194/essd-15-2863-2023

Note: Please fill and submit this form by 15 August 2025 to:
KLC secretariat <klc2020@landslides.org> and ICL Network <icl-network@landslides.org>