





Objective of the project

Potamoi wants to optimize the process of cleaning hydrological data in order to allow the improvement of flood forecasts and, consequently, reducing the number of deaths from floods and damaged properties.



Current flood forecasting systems are highly based on machine learning algorithms, consequently, the high quality and the large quantity of data available become indispensable for the proper execution of a flood forecast. However, the required cleaning process and data integration are costly and laborious, mainly for flood forecasting, where we have a real-time hydrological data collection system feeding the forecast models. The lack of infrastructure in poorly developed regions makes it impossible to access modern equipment necessary for data integration, stocking, and collection. Besides the lack of investments in personnel and the absence of dedicated permanent staff. There are few experts employed in the sector, and there's an even higher workload for rescue and post-disaster activities during major flood events. According to a survey led in 2020, 74% of the forecasting personnel confirm that their centers "do not have the experts and staff capable of integrating data, performing forecasts, and disseminating information." Unfortunately, these limitations delay the flow of data for forecasting and storage in historical databases. The results of this are detrimental, the accuracy of flood forecasts is reduced, the models are limited to short-range forecasts and the development of forecasting techniques is limited due to gaps in historical databases. In other words, more lives are at risk, and less time is available to prepare for a flood.



Potamoi is a service that provides data quality control, it manages the data flow between data collectors and machine learning models. Our service focuses on providing an easy-to-use and highly customized data cleaning and data integration pipeline, it ensures data validity and completeness, statistical quality and missing values restoration. Potamoi wants to facilitate access and customization of data cleaning pipelines by outsourcing the data quality control process. In a way that the process is cheaper, because it does not require internal development, and fast, because it has an availability that supplies the collection and processing service of the forecast center in which the latter alone cannot handle. Potamoi also provides an interactive user interface that allows the user to customize their data cleaning pipeline according to their needs, without requiring a high level of knowledge on the subject. In addition to optimizing the data reporting cycle and analysis.



The flood forecasting centers will not have to increase their staff, hire new specialists or relocate professionals who are working in other emergencies during the flood seasons. The use of an outsourced service will reduce costs related to employees, hours of work and infrastructure. The local system does not need to have the necessary infrastructure to maintain the flow without gaps, Potamoi's outsourced system will contribute with richer data and will not have gaps in sending clean and processed data. When improving data entry after adequate quality control, forecasting models are expected to perform better, therefore, more lives could be protected.



Current state of development and What is the project looking for?

We're currently working on Potamoi CS, as our prototype, which is an intuitive and easy-to-use data cleaning API for citizen scientists that wish to prepare their datasets for model training, exploratory analysis and research. This intends to be a service, where users can prepare their datasets for a range of different applications with the click of a button. Potamoi CS also allows the user to store the cleaned dataset at our databases, so that other users could have access to a catalogue of cleaned datasets for free, empowering the pillars of citizen science. Note: Fundamentally Potamoi CS is a simpler version with a few limitations of Potamoi, without the complexity and personalization that Potamoi provides for each forecast center contract service.

Contacts

We are a team of 4 students with experience in complementary areas, such as data science, econometrics, computer science and sustainable development. If you are interested in knowing more about the project, check our documentation (in progress) and our animation. - Paris Mollo parismollochristondis@gmail.com













