Augmenting Under-Researched Medical Domains with an Intuitive Querying System on Resourceful Crowd-Collected Data

Vizitiv, Ana (School: Da Vinci School)
Griza, Madalina (School: Theoretical High School " Orizont, Durlesti ")

During the early months of COVID-19, neither doctors nor patients knew how dangerous the virus was and how to best treat it. Because formal research took several months to complete, people didn’t have many sources of knowledge to turn to. This lack of information caused panic, wrong treatments to take place and many deaths that could’ve been avoided. When people couldn’t find official information, they resorted to blogs, forums, discussion threads and internet articles. Even though such sources carried important insights driven primarily by individual experiences and amateur home trials, the research community disregarded this data because of its unofficial nature. We believe that there is value in this grassroots data, however. Especially in times of need when official research is lacking, such data could be really helpful and comforting to people. We propose a method to efficiently collect crowdsource data, and a querying tool that aims to solve the main challenges of such community-derived data, namely incomprehensibility, subjective validation, and misinformation. Our model won’t fully replace formal medical research, but can serve as a supplementary and personalized tool for patients and doctors. As the final product, we built a search engine that, given a situation, returns the results of people reporting similar experiences and what they did. We tested the search engine with multiple common COVID-19 related questions, and asked a doctor to assess the accuracy of the outputs. Not only did we obtain an estimated accuracy of 83%, but the engine also proved to be significantly more targeted and resourceful than a classic google search.

Awards Won:
Oracle Academy: Award of $5,000 for outstanding project in the systems software category.