

Computer software is playing an increasingly important role in the modern world. In particular, data-centric applications that help collect, store, share and analyze information are of critical importance to society. Unfortunately, the development of software is challenging: it is time-consuming, error-prone and requires highly-skilled personnel. To make things worse, the fast evolution of the world requires developers to capture in computer code increasingly complex knowledge about problem domains, as well as to integrate an increasing number of diverse sources of information. Due to this complexity, reliable software often takes years to develop.

Defiantly, today's world needs technology to rapidly create highly reliable yet very customized applications. For example, when coping with a natural disaster or some other unexpected crisis, we would like to have reliable and situation-specific tools for the responders and the general public to report and analyze the crisis-related information. Another example are one-off clinical trials, in which researchers need reliable tools to collect and analyze data reported by trial participants, all in a very knowledge-intensive and highly regulated domain. To ensure the quality of data, the applications for data reporting should be tailored to the specific trial requirements, and the medical conditions and needs of the trial participants. These examples share a challenging combination of requirements, which current technologies do not allow to fulfill without very costly development efforts: (i) a very complex problem domain, (ii) the need for customization, and (iii) the need for accessibility to users with little or no training.

This challenge calls for a major breakthrough in automating software creation, which is now possible due to the recent advances in Knowledge Representation and other areas of Artificial Intelligence. The main goal of this project is to obtain automated-reasoning techniques that exploit formally represented domain knowledge in order to raise the quality of data-centric applications, and to dramatically reduce their development efforts. Specifically, we will develop techniques that exploit domain knowledge given as two components: (1) a knowledge base that captures in a machine-readable language the general knowledge about a given problem domain, and (2) a focus specification that captures the requirements of a specific application. Using automated reasoning, these two components will be compiled into a reliable yet very accessible application, tailored for collecting and analyzing data about the entities described in the focus specification. The objectives of this project will involve work that connects several areas of computer science: Knowledge Representation, Artificial Intelligence, Database Theory and Software Verification.