Summary of the project for the general public

Reasoning can be defined as a process of inferring conclusions from accepted premises, or of verifying hypotheses based on already accepted premises. Traditionally, the area of science devoted to investigating the nature of reasoning is logic, which employs a formal notion of language, and inference rules. On the other hand, real human reasonings are not limited to applying general rules (for example: if I know that p implies q, and I know that p, then I know that q) but they also benefit from memory, associations, experience, and other mental tools that go beyond methods of logic. A question arises how to model these non-logical aspects of reasoning – and how to automatize them.

In recent years, there has been a rapid progress in techniques of machine learning based on deep neural networks, where the learning process does not rely on any abstract or general rules. One can say that deep neural networks learn only from the provided data. In some sense, they make 'associations' between the data, by optimizing weights of connections during training. In this way they gather necessary 'experience' needed to solve the required tasks.

These developments have led to new, and very efficient tools used in natural language processing, for example, Google Translate. There were hopes that in a similar way, computers can be taught how to reason in natural language – without resorting to logic at all. However, these hopes have not come true so far, and scientists start to wonder were they not premature.

In this project, we will combine methods of logic and deep learning to study the so called hybrid models of reasoning. In hybrid models, reasoning is ultimately based on rules of logic, but neural networks are involved in the process of premise selection and construction of consecutive reasoning steps. Our research will give insight into mechanisms behinds reasoning in natural language, as well as tools helping automatize it.