

## **Collaborative Data Science at the Swiss Data Science Center**

Rok Roškar, Guillaume Obozinski

Swiss Data Science Center, ETH Zürich  
rok.roskar@sdsc.ethz.ch

Collaborating efficiently on data science projects is a difficult task that requires a good knowledge of many different technologies and practices. At the Swiss Data Science Center (SDSC), we are building Renku, a platform for sustainable data science, whose aim is to make collaboration easier and future-proof. Renku strives to serve as the connector between the different needs of data-centric projects by brokering access to code, sources of data, and computational environments. The relationships between these entities are preserved in a knowledge graph which can be used for building an understanding about the propagation of data and results within an individual project, among a research group, or across institutions and communities. In this contribution, we will present Renku, its current features and roadmap. In addition, we will share a few examples of projects carried out in collaboration with research groups in the ETH domain, where SDSC machine learning expertise has been applied to help answer scientific questions.