StackOverflow Communities Analysis
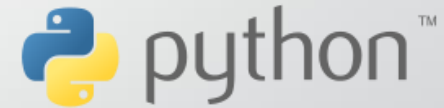
Dmitry Kalashnikov

The Project

- **Goal**: Come up with model for calculation of community size and user flow among the communities on StackOverflow

- **Objectives:**
  - Define a set of rules or a model to determine if a user is a representative one
  - Define features which describe community
  - Group tags which are related to each other
  - Estimate community size and movement of users across the communities

The Requirements

- **Necessary**
  - You can query DB with **SQL**

  - You develop with **R/Python**

  - You know **statistics** and **machine learning**

The Format

- **Internship may be continued in Spring**
- **2 students**
- Daily weekly meetings or calls
- Code review
- 10 hours per week

Thank you for your attention