50x Faster: Speeding up an SQL-based legacy system with few changes

Lei Shi, Lenneke Jong, Wolfgang Mueller, Enkhjargal Algaa

Lei.Shi@h-its.org
Lenneke.jong@h-its.org
Wolfgang.Mueller@h-its.org
Enkhjargal.Algaa@h-its.org

Abstract: SABIO-RK http://sabio.h-its.org/ is a database for reaction kinetic data. Its complex schema led to slow querying. The challenge described in this paper is how to improve the speed with as few as possible changes in the time tested SABIORK code. The dilemma of whether to rewrite a whole system or to best reuse the expertise embodied in existing code appears often in existing, grown academic bio software solutions. We describe our solution that uses the existing SQL database for consistent data storage, and uses inverted indexes for fast retrieval. We further outline how we migrated to a new interface and RESTful web services in order to improve SABIO-RK code reuse and standards compliance.

The work presented in this paper has been partly funded by the DFG project „Integrierte webbasierte Forschungsinfrastruktur zur Speicherung und Bereitstellung quantitativier zeitaufgeloster Daten im Laborkontext und fuer die Wissenschaftschaftlichkeit”.