Quantifying integration quality using feedback on mapping results

Serrano, Fernando R. S.

iiWAS2017

http://hdl.handle.net/11728/10709

Downloaded from HEPHAESTUS Repository, Neapolis University institutional repository
Quantifying integration quality using feedback on mapping results

Abstract
Traditional data integration delivers high integration quality but requires significant upfront effort because of the need for expensive experts to be involved. The pay-as-you-go approach to data integration aims to reduce this effort by relying on a bootstrap phase where algorithms replace experts in identifying or validating source-to-target semantic correspondences and executable mappings. Since the results of this phase are expected to be of lower quality, a continuous improvement phase is then launched where user feedback is collected and assimilated in order to improve the integration. It is crucial, therefore, to quantify integration quality. This paper presents a solution to this problem using feedback on mapping results as evidence. We contribute a methodology for quantifying integration quality while taking into account the inherent uncertainty of user feedback. The approach is evaluated in synthetic and real-world integration scenarios and shown to accurately and cost-effectively quantify their quality as a conditional probability.