A controlled vocabulary of relations between variables

Claus-Peter Klas^{*1}, Benjamin Zapiloko¹, and Janete Saldanha Bach^{*1}

¹GESIS – Leibniz Institute for the Social Sciences – Germany

Abstract

Research outputs are increasingly interdependent, interconnecting many entities from different levels into relationship maps such as the knowledge graphs. Within DDI LC implicit and explicit relations are modelled within and across studies on all entities, such as studyunits, instruments, questions, response scales, and variables. As the variable is one of the most relevant entities, particularly for re-use purposes, we present a Social Sciences framework design for the relations and their descriptions to build a Social Science knowledge graph. These relations explicit between variables will describe links across different waves and studies in conjunction with questions. In addition, relations such as "derived by", "aggregated by" or variables annotated with controlled vocabularies are also defined. We provide a brief textual identification of the relation type, supported by a controlled vocabulary (CV) and an extended description of the relationship. This will be published as a controlled vocabulary for variable relations via the CESSDA vocabulary manager to foster vocabulary re-use. Documenting and describing all these relations will enrich the data reuse by supporting search and browse functionality. By re-using the CV, a general Social Science research knowledge graph can be created across institutes, in line with the FAIR principles. As next step we will extend the descriptions of relations to further entities within DDI also published as CV.

^{*}Speaker