

# Measuring Similarity between Ontologies

Alexander Maedche<sup>1</sup> and Steffen Staab<sup>2,3</sup>

<sup>1</sup> FZI - Research Center for Information Technologies at the University of Karlsruhe,

Haid-und-Neu-Str. 10-14, D-76131 Karlsruhe, Germany

<http://www.fzi.de/WIM>

<sup>2</sup> Institute AIFB, Univ. Karlsruhe,

D-76128 Karlsruhe, Germany

<http://www.aifb.uni-karlsruhe.de/WBS>

<sup>3</sup> Ontoprise GmbH,

76131 Karlsruhe, Germany

<http://www.ontoprise.de>

**Abstract.** Ontologies now play an important role for many knowledge-intensive applications for which they provide a source of precisely defined terms. However, with their wide-spread usage there come problems concerning their proliferation. Ontology engineers or users frequently have a core ontology that they use, e.g., for browsing or querying data, but they need to extend it with, adapt it to, or compare it with the large set of other ontologies. For the task of detecting and retrieving relevant ontologies, one needs means for measuring the similarity between ontologies. We present a set of ontology similarity measures and a multiple-phase empirical evaluation.

## 1 Introduction

A core purpose for the use of ontologies is the exchange of data not only at a common syntactic, but also at a shared semantic level. Especially on the WWW more and more ontologies are constructed and used, beginning to replace the old-fashioned ways of exchanging business data via standardized comma-separated formats by standards that adhere to semantic specifications given through ontologies. Thus, in the near future more and more ontologies will be made available on the WWW. With this upswing and beginning widespread usage of ontologies, however, new problems are incurred. Ontology engineers or users frequently have a core ontology that they use, e.g., for browsing or querying data, but they need to extend it with, adapt it to, or compare it with the large set of other ontologies. For the task of detecting and retrieving relevant ontologies, one needs means for measuring the similarity between ontologies on a canonical scale (e.g., the reals in  $[0, 1]$ ).

So, how may we measure the similarity of ontologies or of ontology parts? One could make use of the formal structures of ontologies and try at the unification of ontologies or ontology parts (which is essentially subgraph matching). The drawback here would be that all real-world ontologies that we know of do not only specify its conceptualization by logical structures, but to a large extent also by reference to terms that are grounded through human natural language use. For instance, modeling that MAN