



reference to a citation of individual data items. Ignoring the above mentioned methodological questions and challenges, results in a shortage of clear and unified guidelines for citing data, mainly from online databases.

Unified format of citation of references to data in a large collection of publications can enable an automatic analysis of data citation patterns. For example, as in the case of papers' citation analysis, one can infer the relative importance of a certain data source compared to other sources, or can study the change in authors' preferences of data sources over time. The absence of such citation guidelines prevents exploiting this potential.

This thesis studies how data citation should be accommodated in order to cope with the described challenges. We propose a conceptual model that describes both how data should be published and how it should be cited. The core element of the model is a simple datum defined by its subject, its property, possibly by its time reference, and its publication time. The proposed model treats the datum as an independently citable element which is uniquely defined using a formal definition. The publication time component of the datum's identification is used to differentiate between two revisions of the same measured phenomena, published in different times. The model also defines an abstract element called a publication framework. A publication framework bundles data together and provides them a context beyond the definition and interpretation of each datum separately. In a concrete implementation, a publication framework can be a single table, a multi-table digital document or an entire database.

In order to demonstrate the feasibility of the proposed model and part of its validation process, we present in this thesis an outline for building a system for data publication and citation based on the proposed model. Using this outline we have implemented a prototype for such a system that contains two numerical databases simulating two different data publishers. We report the result of using this prototype in which the users acted both as readers of texts containing data citations, as well as authors that use the system to create such data citations

System No.

1131148