

# Statistical Machine Learning with Linked Data

## Talk Abstract

Volker Tresp<sup>1</sup>

Siemens CT,  
`volker.tresp@siemens.com`,

The size of the Linked Open Data (LOD) cloud is constantly increasing where the term Linked Data is used to describe a method of exposing, sharing, and connecting data via dereferenceable URIs on the Web. In this talk we explore the usefulness of statistical machine learning for LOD. Statistical machine learning has the chance of exploiting statistical regularities in the data that cannot easily be captured by logical statements and can handle contradictory, uncertain and missing data. In general, the data quality on LOD is varying: whereas LOD for the life sciences has reasonably good quality, other portions of the LOD cloud are not maintained as well and are still quite noisy. We present existing machine learning approaches to learning with LOD. We conclude that machine learning can be quite effective on LOD if the data quality fulfils some minimal quality requirements.