

*Kaser A., Starke M., Sachse P.*

**Projekt:**

**Natürliche Datensätze vs. künstliche Datensätze – Internetsuchmaschinen im Vergleich**

Ausgangspunkt für unser Projekt war eine kulturhistorische Dokumentation (Kartei), welche von Herrn Dr. Peter Sommer (CH) über vier Jahrzehnte erarbeitet wurde. Diese einzigartige Dokumentation bildet ein semantisches Netzwerk, welches ca. 3000 Begriffe korrekt miteinander verknüpft. Das Netzwerk beruht damit auf einem natürlichen Datensatz.

Im Gegensatz dazu können Datenbestände, die Internet-Suchmaschinen zu Grunde liegen (beispielsweise Google, Yahoo, Microsoft Network, Wikipedia), als „künstliche“ Datensätze betrachtet werden, da sie von Computerprogrammen automatisch erstellt werden.

In einer aktuellen Studie wird u.a. ermittelt, ob „künstlich“ erstellte Datensätze in ihrer Struktur vergleichbar aufgebaut sind wie jene, die unmittelbar von Menschen erzeugt und genutzt werden.