

## Eigennamenannotation

Beim Vergleich der populärwissenschaftlichen Texte mit dem Zeit-Korpus hat sich gezeigt, dass bei der Frequenz von Eigennamen relativ große Unterschiede bestehen. Im Schnitt kommen im Zeit-Korpus Eigennamen fast doppelt so häufig vor, wie im PopSci-Korpus. Deshalb werden Eigennamen unabhängig von der Wortartannotation - bei der Eigennamen ebenfalls eine eigene Kategorie bilden - behandelt und weiter ausdifferenziert.

## Annotierte Kategorien

Eigennamen sind prinzipiell nur solche Bezeichnungen, die eine räumlich oder zeitlich klar umgrenzte Entität bezeichnen. D.h. alle Namen, die eine Kategorie oder eine Gattung benennen werden in dieser Annotationsebene nicht oder nur indirekt (siehe *MISC*) erfasst. Insbesondere bei Benennungen im wissenschaftlichen Bereich fällt es manchmal schwer, zwischen Eigennamen und Gattungsbezeichnungen zu unterscheiden. So handelt es sich z.B. beim *Aids-Virus* nicht um einen Eigennamen, da mit dieser Bezeichnung keine einzelne Entität sondern eine Klasse von Viren bezeichnet wird.

Bei der Annotation werden folgende Eigennamen unterschieden:

- **Person** (*PER*): Namen von lebenden oder toten, fiktiven oder realen Personen.
- **Organisation** (*ORG*): Organisationen sind soziale Strukturen, die als solche in der Öffentlichkeit auftreten und dort bestimmte Interessen verfolgen. Darunter fallen unter anderem auch: große Tagungen, Länder, die stellvertretend für ihre Regierung genannt werden, Zeitungsverlage oder eigenständige Forschungsprojekte.
- **Ort** (*LOC*): Diese Kategorie benennt alle geographischen Bezeichnungen wie z.B. Länder, Städte, Meere, Seen, Flüsse oder Berge. Auch Orte die außerhalb der Erde liegen wie die Milchstraße oder Bezeichnungen für einzelne Sterne fallen in diese Kategorie.
- **Objekt** (*OBJ*): Von Menschenhand geschaffene Objekte, denen ein eindeutiger Namen zugewiesen wurde. Diese Kategorie wird z.B. für Satellitennamen oder Bezeichnungen für andere technische Geräte verwendet.
- **Ereignis** (*EVT*): Historische Ereignisse, die einen eigenen Namen tragen wie beispielsweise der Zweite Weltkrieg oder der Kalte Krieg.
- **Sonstige** (*MISC*): In diese Kategorie fallen vor allem Komposita oder Ableitungen die einen Eigennamen enthalten, selbst aber keinen Eigennamen bilden. Zu den Ableitungen zählen z.B. adjektivierte Länder- oder Städtebezeichnungen. Beispiele für Komposita wären dagegen: US-Forscher, Hawking-Strahlung, Mandelbrot-Menge oder Nobelpreis. Bezeichnungen werden allerdings nur dann in die Kategorie *MISC* eingeordnet, wenn die Bezeichnungen als Ganzes nicht in eine der andern Kategorien fällt. Ein Beispiel hierfür wäre das Hubble-Teleskop, das zwar eine Kompositum mit einer Personenbezeichnung darstellt, als ganzes aber ein spezielles Objekt benennt und deshalb in die Kategorie *OBJ* eingeordnet wird.

Die Eigennamen werden immer im Kontext interpretiert. D.h. es kann sein, dass der

selbe Name in verschiedenen Kontexten verschiedenen Kategorien zugeordnet wird. Zum Beispiel ist eine Länderbezeichnung im geographischen Sinne ein Ort, wogegen sie im institutionellen Sinne - Deutschland schließ ein Vertrag ab - der Kategorie *ORG* zugeordnet wird.