

Pojmenované entity jsou slovní spojení, která v textu označují objekty reálného světa. Normalizací entit nazveme jejich převod do základního tvaru. Práce se zabývá vytvořením pravidlové procedury určené k normalizaci pojmenovaných entit v českých textech. Proces návrhu jednotlivých pravidel této procedury je důkladně zmapován. Důraz je kladen na to, aby každé pravidlo bylo motivováno příklady reálných entit. Za účelem dosažení co největší úspěšnosti jsou také analyzovány některé aspekty syntaxe českého jazyka. Na základě teoretického popisu procedury je dále implementován normalizační program a jeho úspěšnost je vyhodnocena srovnáním s ručně normalizovanými entitami. Ve spojení s již existujícími nástroji pro automatické rozpoznávání pojmenovaných entit v textu je možné tento normalizátor využít v jiných procesech strojového zpracování textu, například překladu do jiného jazyka, vyhledávání a kategorizaci apod.