

V posledních letech se využití textové analytiky v komerční sféře postupně stává významným tématem pro vědecké a praktické aplikace. Zaměřili jsme se na určování vztahů mezi entitami z dat dodaných partnerskou společností. Analýza textu z této sféry ale vyžaduje jiný přístup: počítání s nepřesnostmi a specifickými atributy. V této práci jsme se rozhodli ukázat využití dvou metod pro určování vztahů: tzv. Snowball systém a Metodu vzdáleného dohledu (z angl. Distant Supervision), které jsme přizpůsobili pro dodaná data. Dané metody byly implementovány pro využití strukturovaných a nestruturovaných dat z firemní databáze.

Klíčová slova: Získávání informací, Určování vztahů mezi entitami, Textová analytika, Distant Supervision, Snowball