

Abstrakt

Cílem mé práce bylo zjistit specifické vlastnosti českých lékařských zpráv z hlediska možnosti extrahovat z nich konkrétní informace.

Pro svoji práci jsem měl k dispozici celkem 268 anonymizovaných narativních lékařských zpráv ze dvou ambulantních pracovišť. Studoval jsem standardy pro uchování elektronické zdravotnické dokumentace i pro přenos klinických informací mezi informačními systémy ve zdravotnictví. Věnoval jsem se také implementování elektronického zdravotního záznamu v zubním lékařství.

Nejprve jsem se narativní lékařské zprávy snažil zpracovat pomocí nástrojů pro zpracování přirozeného jazyka (Natural Language Processing, NLP). Dospěl jsem k závěru, že narativní lékařské zprávy v českém jazyce jsou typickému českému textu velmi vzdálené zejména pro svoji heslovitost a absenci české větné stavby. Obsahují také velké množství překlepů, zkratk a zkrácených slov. Vzhledem k nedostupnosti hlavních mezinárodních klasifikačních systémů v českém jazyce jsem se rozhodl pokračovat ve výzkumu vývojem metody pro přípravu vstupního textu pro překlad a jeho sémantickou anotaci.

Hlavním cílem této části výzkumu bylo navrhnout metodu a podpůrný software pro interaktivní korekci a sémantickou anotaci narativních lékařských zpráv, které by umožnily jejich snadnější použití, s menším množstvím chyb i mimo jejich původní kontext.

Vyvinul jsem třífázovou metodu předběžného zpracování s cílem podpořit druhotné využití lékařských zpráv. Metoda třífázového předzpracování narativních klinických zpráv byla ověřena na 49 anonymních českých lékařských zprávách z oblasti kardiologie.

Klíčová slova

Lékařské zprávy, zpracování přirozeného textu, extrakce informací, elektronický zdravotní záznam.