

V předložené práci studujeme možnosti automatického třídění příchozí emailové komunikace. Naším hlavním cílem je rozpoznání informací o nadcházejících workshopech a konferencích, nabídkách práce a vydávaných knihách. Snažíme se vyvinout nástroj, který informace vydoluje z dat získaných z oborových konferencí. Nabídky v konferencích přicházejí ve formě html, rtf, nebo prostého textu, ale informace v nich je zapsána v běžném jazyce. Text{miningovými metodami získáváme informace z běžného textu a ukládáme je ve strukturované formě, kterou je možné jednoduše strojově zpracovávat. Zkoumáme způsob zpracování pošty člověkem a následně tyto poznatky aplikujeme při tvorbě systému. V průběhu práce řešíme problémy se samotným získáním zpráv, rozpoznáním jazyka a kódování a rozpoznáním typu zprávy. Informace, kterou ze zprávy potřebujeme získat se různí v závislosti na typu zprávy a události, které se týká. Teprve po rozpoznání nosné informace ve zprávě jsme schopni vydolovat data pro zjištěný typ události. Na závěr ukládáme získané znalosti do databáze, která umožňuje rychlou interakci s uživatelem.