

# Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

**Autor práce** Bc. Rastislav kadleček  
**Název práce** Converting HTML product data to Linked Data  
**Rok odevzdání** 2017  
**Studijní program** Informatika      **Studijní obor** Softwarové systémy

**Autor posudku** Doc. Mgr. Martin Nečaský, Ph.D.      **Role** Vedoucí  
**Pracoviště** KSI

## Text posudku:

Práce se zabývá analýzou, návrhem a implementací softwarového nástroje pro konverzi tabulkových HTML dat do podoby sémantických Linked Data. Konkrétně se zaměřuje na data o produktech (např. z elektronických obchodů). Cílem práce nebylo vytvořit zcela nový softwarový nástroj ale rozšířit existující nástroj, na kterém student pracoval v rámci svého studentského softwarového projektu.

Student v práci rozšířil existující algoritmus TableMiner+, kdy aplikoval metody strojového učení (Machine Learning - ML) na oblast rozpoznávání produktů v HTML tabulkách. Rozšíření úspěšně implementoval a provedl řadu testů s různými ML technikami a srovnal tyto techniky pro jím řešený případ.

Samotné řešení je dobře zpracováno. Řešitel v průběhu řešení práce zapracovával mé připomínky jako vedoucího práce a průběžně demonstroval na živých ukázkách dosažené výsledky. Výhrady mám k návrhové části práce, kdy funkční a kvalitativní požadavky nejsou zpracovány příliš detailně a autor se tak velmi brzy věnuje implementačním detailům. Též mám výhrady k evaluační části práce, kdy významná část evaluace je zahrnuta v příloze bez detailního vysvětlení, což práci zbytečně ochuzuje. Též není zřejmé, jak vytvořený systém obstojí ve srovnání s jinými existujícími řešeními, pokud nějaká existují. Řešitel měl však lehce ztížené podmínky tím, že v průběhu řešení práce se bez jeho zavinění změnil vedoucí.

**Práci doporučuji k obhajobě.**

**Práci nenavrhuji na zvláštní ocenění.**

*Pokud práci navrhuje na zvláštní ocenění (cena děkana apod.), prosím uveďte zde stručné zdůvodnění (vzniklé publikace, významnost tématu, inovativnost práce apod.).*

**Datum** 24. srpna 2018

**Podpis**