

Ceník obvyklých cen IT prací

Vysoká škola ekonomická v Praze zpracovává jako zakázku pro sdružení třinácti státních institucí Ceník obvyklých cen ICT prací na základě skutečně sjednaných cen v ICT smlouvách. Na zakázce pracuje tým na katedře informačních technologií Fakulty informatiky a statistiky.

Zdrojem dat pro ceník jsou smlouvy zveřejněné v registru smluv <http://smlouvy.gov.cz>. Účelem ceníku je poskytnout přehled zadavatelům zakázek v oblasti ICT, jako podklad pro odhad očekávané ceny zakázky. Ceník se zaměřuje výhradně na ceny explicitně ve smlouvách uvedené jako ceny za člověkodenní nebo člověkohodinu práce. Ceník tedy poskytuje informaci o tom, jaké hodinové či denní sazby byly v oblasti ICT skutečně sjednány. Pozor však na dezinterpretaci: Z principu nemohou být nalezeny všechny sjednané ceny, a tedy z cen uvedených v ceníku nelze generalizovat na celý trh. Ceník neobsahuje ceny z trhu mimo veřejnou správu, mnohé ceny byly ve smlouvách zaslepeny, mnohé ceny mohly být sjednány v jiné výši z důvodu podmínek výběrových řízení nebo ve vazbě na jiné, např. paušální složky ceny v dané smlouvě.

Ceny jsou extrahovány procesem s vysokou mírou automatizace, kdy nejprve jsou v registru ekonomických subjektů pomocí zveřejněných ekonomických činností identifikovány subjekty – potenciální dodavatelé ICT prací, a smlouvy s těmito subjekty jsou pak staženy z registru smluv a obsahově analyzovány. Z plného textu relevantních smluv jsou pak pomocí regulárních výrazů extrahovány kandidátní cenové údaje pro jednotlivé typové práce. Každá takto automatizovaně extrahovaná cena prochází ruční validací, kdy expert znalý v oblasti ICT smluv zhodnotí, zda je cena extrahovaná správně, případně z dané smlouvy doplní další cenové údaje a tagy, které cenový údaj blíže specifikují.

Ceník je aktualizován v půlročních intervalech na základě nově přidaných smluv do registru smluv. Do ceníku zahrnujeme údaje platné klouzavě za poslední dva roky.

V první, pilotní verzi, bylo cca z 200000 relevantních smluv identifikováno cca 300 smluv se sjednanými a nezaslepenými jednotkovými cenami ICT prací, ze kterých se povedlo extrahovat cca 1000 cenových údajů. Pokud bude situace v registru smluv obdobná, očekáváme podobné výsledky i v dalších intervalech.

Ceník popisuje distribuci cenových údajů pro jednotlivé typy prací a popisuje je číselně, pomocí histogramů a krabíčkových grafů. Sdružení zadavatelů zastupuje Ministerstvo vnitra a to také ceník zveřejňuje na adrese <http://www.mvcr.cz/clanek/cenik-obvyklych-cen-ict-praci.aspx>.

Výsledkem pilotního projektu je kromě samotného ceníku také metodika tvorby ceníku a prototyp softwarových skriptů a technologického prostředí pro tvorbu ceníku s využitím výhradně opensource technologií. Byly použity zejména technologie Debian, ElasticSearch, Kibana, node.js, ethercalc a R studio.