

Posudek oponenta habilitační práce

Masarykova univerzita

Fakulta Fakulta informatiky MU

Habilitační obor Informatika

Uchazeč Pavel Rychlý, Ph.D.

Pracoviště Fakulta informatiky MU

Habilitační práce Statistical Processing of Text Corpora

Oponent Prof. Ing. Václav Matoušek, CSc.

Pracoviště Fakulta aplikovaných věd, Západočeská univerzita v Plzni

Předložená habilitační práce obsahuje výběr významných publikací vytvořených v rámci vědeckovýzkumné aktivity habilitanta zhruba za posledních deset let, přičemž publikace pokrývají poměrně široký rozsah problematiky z oblasti zpracování přirozeného jazyka. Tyto publikace tvoří věcné jádro předložené habilitační práce a jsou doplněny velmi stručnou (osmistránkovou) úvodní částí práce. Předložené články prezentují tři oblasti odborné činnosti habilitanta (kolokace, distributivní sémantika a jazykové modelování) a dostatečně podrobně ilustrují jeho významné vědeckovýzkumné výsledky, které byly dosaženy na pracovišti Fakulty informatiky Masarykovy univerzity. Práce je koncipována přehledně, obsahuje řadu původních výsledků ilustrovaných realizovanými systémy a výsledky provedených experimentů. Konstatuji, že uvedené tři oblasti habilitační práce jsou aktuální, originální a pro počítačové zpracování přirozeného jazyka značně potřebné. Předložené články obsahují několik hodnotných výsledků, které svědčí o dobrých znalostech a schopnostech habilitanta řešit i náročné vědeckovýzkumné úlohy.

Náplň habilitační práce vybranými publikacemi je dobře logicky uspořádána, přičemž je velmi obtížné říci, která ze tří v práci prezentovaných oblastí je nejhodnotnější; všechny předložené články byly prezentovány na významných světových konferencích nebo ve známých mezinárodních časopisech s impact faktorem. Především delší zahrnuté články (kapitoly 4, 5 a 11) podrobněji prezentují významnost dosažených výsledků a určitě by bývalo vhodné uvést také procentní podíl habilitanta na autorství jednotlivých publikací (4 – 7). Z některých článků dále vysvítá, že se nejedná o uzavřené výsledky vědeckovýzkumné práce, neboť řešení některých problémů se nalézají teprve v počátcích a budou v následujících letech dále rozvíjena. Nicméně všechny uvedené výsledky jsou velmi hodnotné a budou zcela jistě využity autorem i dalšími (mladšími) řešiteli k dalšímu výzkumu.

Každá z uvedených tří oblastí vědeckovýzkumné práce – viz výše – představuje jednak poměrně rozsáhlou oblast čítající množství dosud neřešených problémů a jednak jde o relativně uzavřené oblasti, které spolu souvisejí jen zčásti. Práce tak dokládá schopnost habilitanta orientovat se při řešení vědeckovýzkumných úkolů na několik širších problémových oblastí a také jeho schopnost pokusit se aplikovat významné výsledky získané v jedné oblasti i na oblasti další.

Jako jeden z nejvýznamnějších výsledků popisovaných v práci považuji články pojednávající o originálním systému pro lingvisty a lexikografy (the sketch engine), který je dnes známý po celém světě, je považován za standard mezi nástroji pro lexikografy a jehož původním a dlouho jediným autorem je právě Pavel Rychlý. Lze tak konstatovat, že habilitant dosáhl výsledku na špičkové světové úrovni.

K obsahu práce nemám žádné podstatné připomínky, neboť články, jak již bylo uvedeno, byly prezentovány ve významných časopisech a na konferencích a je jen škoda, že jejich ohlas je uveden pouze v příloženém materiálu (dokladech k návrhu na zahájení habilitačního řízení). V závěru práce (zcela chybí) pak mělo být uvedeno zaměření příštích výzkumných aktivit habilitanta v uvedených oblastech, popř. jeho obecnější záměry.

Práce neobsahuje formální chyby, úvodní kapitola práce je napsána zcela srozumitelně a pěknou angličtinou, která čtenáři nečiní potíže s jednoznačným porozuměním textu. Celkově jsem z práce nabyl dojmu, že habilitant věnoval značnou pozornost konečné podobě práce a nenechal se ovlivnit např. tlakem na urychlené dokončení práce.

Shrnu-li význam práce pro rozvoj vědního oboru, pak musím konstatovat, že v práci uváděné výsledky jsou pro rozvoj oboru velmi významné a jsou dokumentovány několika úspěšně realizovanými systémy publikovanými v jednotlivých člancích. Předložená habilitační práce má jasnou strukturu a mohu konstatovat, že obsahuje přínos k rozvoji vědního oboru a její obsah je hodnotný. O praktické využitelnosti prezentovaných postupů není pochyb, uvedené výsledky jasně dokladují cílevědomost při realizaci jednotlivých experimentů a systémů a také užitečnost získaných výsledků.

Dotazy k obhajobě:

1. Jak (v čem) je Sketch Engine implementován a jak je dodáván zájemcům o jeho instalaci ?
2. Jak (kdy) bude možno získat Sketch Engine pro český jazyk a s jak rozsáhlými korpusy ?

Závěr:

Za rozhodující ve svém hodnocení považuji, že autor předložené práce prokázal svoji erudici v dané problémové oblasti vypracováním řady článků publikovaných na významných konferencích a v impaktovaných časopisech, které navíc mají rozsáhlý citační ohlas. Ve světle tohoto faktu považuji své výše uvedené drobné připomínky pouze za formální a nepodstatné, předložená habilitační práce splňuje podmínky dle §72 zákona 111/98 Sb. o vysokých školách a doporučuji habilitační komisi, aby předloženou práci jako práci habilitační akceptovala.

V Plzni dne 1. října 2015

Prof. Ing. Václav Matoušek, CSc.

katedra informatiky a výpočetní techniky
Fakulty aplikovaných věd
Západočeské univerzity v Plzni