

## Závěrečné hodnocení průběhu realizace projektu programu LN oponentem

<i>Příjemce dotace:</i>	Univerzita Karlova v Praze
<i>Řešitel:</i>	Prof. PhDr. Eva Hajičová, DrSc.
<i>Název projektu:</i>	Centrum počítačnické lingvistiky
<i>Identifikační číslo projektu:</i>	<b>LN00A063</b>

### Závazná kritéria hodnocení projektu:

#### 1. Zhodnocení průběhu, dosažených výsledků a splnění cílů projektu stanovených smlouvou:

- a) *Průběh realizace projektu; způsob řešení projektu ve vazbě na jeho koncepci, věcná i časová přiměřenost postupu řešení, zabezpečení projektu ze strany příjemce a jeho podíl na dosažených výsledcích v porovnání s obsahem zakladatelské smlouvy výzkumného centra:*

Práce a koordinace výzkumu v Centru vycházela z promyšlené koncepce a byla ze strany příjemce příkladným způsobem zabezpečována. Časový postup řešení stanovených byl dodržen, až na drobné úpravy, které byly vždy specifikovány v upřesnění dílčích cílů projektu pro následující rok řešení projektu. Dosažené výsledky plně odpovídají obsahu zakladatelské smlouvy výzkumného centra.

- b) *Dosažené výsledky projektu a jejich zhodnocení v porovnání s domácí a světovou špičkou:*

Centrum dosáhlo celou řadu vynikajících výsledků, jejichž plný výčet by zde zabral příliš mnoho místa, jmenujme zde aspoň v hrubých rysech základní oblasti:

- Vybudování Pražského závislostního korpusu (Prague Dependency Treebank, PDT), morfologicky a analyticky anotovaného korpusu, obohaceného o značkování na tektogramatické rovině, která zachycuje význam vět. Jedná se o zásadní výsledek, bez něhož by další kvalitní výzkum v oblasti syntaxe a sémantiky českého jazyka nebyl možný. Jak konstatují přední světoví odborníci v rozpravě k výsledkům Centra, PDT obsahuje více informace a je lépe strukturovaný než Penn Treebank, což je v současnosti nejužívanější světový analogický korpus.

- Výsledky v oblasti matematické lingvistiky a teoretický výzkum. Byl vyvinut nástroj pro systém strojového učení na základě průměrovaného perceptronu a dále byly testovány na dalších jazycích nástroje založené na exponenciálním statistickém modelu. Byl rovněž prováděn výzkum v oblasti využití konečných automatů a převodníků pro jazykové modelování a morfologické značkování. Teoretický výzkum byl vhodně motivován prací v ostatních oblastech činnosti Centra.

- Výsledky v oblasti rozpoznávání řeči. Zvláštní pozornost byla věnována studiu tzv. suprasegmentálních jevů, jako je význam větné prozodie, pro který je výhodným východiskem analýza aktuálního členění věty a studiu modelování jazyka, se zřetelem na flexivní povahu češtiny. Výsledky výzkumu v této oblasti byly aplikovány v mezinárodním projektu MALACH (Multilingual Access to Large Spoken Archives), jehož cílem je vývoj systémů pro automatický předpis svědeckých výpovědí lidí, kteří přežili holocaust.

- Výsledky v oblasti vytváření a využívání vícejazyčných zdrojů, zvláště ve studiu a uplatnění paralelních korpusů se zaměřením na strojový a strojem podporovaný překlad a další aplikace, jako vyhledávání informací a data mining ve vícejazyčných textech.

- a řada dalších výsledků, tak jak jsou uvedeny ve zprávě.

- c) *Splnění cílů projektu stanovených smlouvou Využitelnost a využití výsledků projektu (způsob využití a uživatelé výsledků v aplikační sféře, v rámci regionu a státu, aktuální stav a s ohledem na vývojové trendy v daném vědním oboru):*

Cíle stanovené smlouvou byly splněny a v mnoha ohledech překročeny. Činnost Centra vedla k navázání těsných kontaktů s českým a mezinárodním průmyslem využívajícím počítače, o čemž svědčí zájem partnerů a uživatelů z oblasti aplikační sféry.

- d) *Přínos projektu (pro aplikační sféru, region a další výzkumnou práci):*

Výsledky projektu jsou zcela zásadním přínosem jak pro základní a aplikovaný výzkum v oblasti počítačového zpracování řeči, tak i pro potencionální aplikace v oblasti informačních technologií, jako jsou např. syntéza a rozpoznávání řeči, dialogové systémy, strojový překlad a pod.

- e) *Dosažená odborná úroveň týmu; zhodnocení vlivu personálního zabezpečení na průběh řešení projektu (kmenoví pracovníci, nová pracovní místa):*

Odborná úroveň týmu je velmi vysoká a dobře vyvážená. Bylo by velmi žádoucí, aby tento tým mohl dále produkovat špičkové vědecké výsledky, jako dosud.

- f) *Srovnatelnost projektu s obdobnými projekty na analogické úrovni v ČR i v zahraničí:*

Projekt je unikátní v rámci ČR a srovnatelný se špičkovými zahraničními projekty v dané oblasti.

- g) *Zhodnocení kvality dosažených výsledků (dle seznamu – viz příloha) řešitelem a jeho pracovním týmem ve výzkumu a vývoji za dobu řešení projektu:*

Jak již bylo řečeno, jde o velmi kvalitní, zcela zásadní výsledky, jak v teoretické, tak i aplikační oblasti.

- h) *Vyhodnocení naplnění cíle programu v oblasti odborného růstu mladých vědeckých pracovníků (celkový počet zapojených studentů doktorandského studia, šíře jejich vědeckého působení ve výzkumném centru, počet mladých vědeckých pracovníků zapojených do řešení projektu po celou dobu trvání):*

Cíle v oblasti odborného růstu mladých vědeckých pracovníků a jejich vědeckého působení ve výzkumném centru byly splněny.

- i) *Zhodnocení spolupráce výzkumných týmů sloučených ve výzkumném centru:*

Spolupráce výzkumných týmů sloučených ve výzkumném centru byla na vysoké úrovni a výrazně přispěla k vysoké kvalitě dosažených vědeckých výsledků.

- j) *Zhodnocení přístupu vědecké obce k výsledkům a výstupům centra pro vědeckou a širokou veřejnost:*

Výsledky byly prezentovány na celé řadě mezinárodních konferencí a odborných akcí, v 259 publikacích a 17 technických zprávách. Byly vytvořeny www stránky CKL, <http://ckl.mff.cuni.cz/>, které podávají komplexní informaci o činnosti Centra. Centrum přispívalo k informovanosti o svých výsledcích rovněž jako člen mezinárodní sítě ENABLER Network. Každoročně v rámci Dnu otevřených dveří na MFF UK představovalo CKL svou činnost širší odborné veřejnosti. Celkově se dá říci, že CKL velmi dobře zabezpečovalo přístup vědecké obce k produkováným výsledkům.

## **2. Výše finančních prostředků a účelnost jejich využití:** (Posouzení výše a struktury vynaložených prostředků na realizaci projektu ve vztahu k dosaženým výsledkům ve sledovaném období a k efektivnosti jejich využití)

Vzhledem k získaným výsledkům a údajům uvedeným ve zprávě lze s jistotou říci, že finanční prostředky byly využity velmi účelně.

**3. Připomínky, výhrady a doporučení oponenta:**

Bylo by velmi žádoucí zajistit financování další činnosti Centra.

**4. Celkové hodnocení projektu: (slovně i klasifikací)**

*V – vynikající výsledky (mezinárodního významu)*

**5. Závěr:**

Souhlasím s předloženou závěrečnou zprávou bez připomínek.

Příjmení, jméno, titul oponenta: Kopeček Ivan, Doc. RNDr., CSc.

Pracoviště : Fakulta informatiky MU Brno

Popište Váš případný vztah k příjemci dotace (organizaci), popř. spolupříjemci; k řešitelskému týmu:

**Prohlášení oponenta:**

Prohlašuji, že mi byly příjemcem včas poskytnuty všechny potřebné informace a předloženy mnou požadované materiály a podklady pro posuzování (zpráva, smlouva, rozhodnutí, popř. na vyžádání účetní doklady, seznamy výsledků, abstrakta publikací....). Dále prohlašuji, že při posuzování tohoto projektu jsem se řídil(a) výhradně objektivními hledisky.

Datum: 17. prosince 2004

Podpis : \_\_\_\_\_