

SMLOUVA
o poskytnutí účelové podpory
na řešení projektu velké výzkumné infrastruktury
s názvem
Český národní korpus
č. j.: MSMT-44/2023

Ministerstvo školství, mládeže a tělovýchovy

IČO: 00022985

se sídlem: Karmelitská 529/5, 118 12 Praha 1,
jednající PhDr. Lukášem Levákem, ředitelem odboru výzkumu a vývoje,
(dále jen „Poskytovatel“)

a

Univerzita Karlova

IČO: 00216208

právní forma: veřejná vysoká škola
se sídlem: Ovocný trh 560/5, 116 36 Praha 1
číslo účtu: ██████████

zastoupena prof. MUDr. Milenou Králíčkovou, Ph.D., rektorkou,
(dále jen „Příjemce“)

(společně dále také jako „smluvní strany“)

uzavírají

podle § 3 odst. 2 písm. d), § 4 odst. 1 písm. e) a § 9 odst. 1, 2 a 3 zákona č. 130/2002 Sb., o podpoře výzkumu, experimentálního vývoje a inovací z veřejných prostředků a o změně některých souvisejících zákonů (zákon o podpoře výzkumu, experimentálního vývoje a inovací), ve znění pozdějších předpisů, zákona č. 218/2000 Sb., o rozpočtových pravidlech a o změně některých souvisejících zákonů (rozpočtová pravidla), ve znění pozdějších předpisů, a subsidiárně podle zákona č. 89/2012 Sb., občanský zákoník, ve znění pozdějších předpisů, a zákona č. 500/2004 Sb., správní řád, ve znění pozdějších předpisů, tuto **smlouvu o poskytnutí účelové podpory na řešení projektu velké výzkumné infrastruktury (dále jen „Smlouva“)**:

Článek 1

Předmět Smlouvy

- 1) Předmětem Smlouvy je poskytnutí účelové podpory podle § 3 odst. 2 písm. d) zákona č. 130/2002 Sb. (dále též „dotace“) Poskytovatelem Příjemci na řešení projektu velké výzkumné infrastruktury schváleného usnesením vlády České republiky ze dne 14. prosince 2022 č. 1043 a identifikovaného názvem **Český národní korpus** (akronym: **CNC**) a identifikačním kódem **LM2023044** (dále jen „Projekt“). Předmětem řešení projektu je zajištění realizace výzkumných kapacit Projektu a jejich zpřístupnění v režimu otevřeného přístupu v rozsahu uvedeném v Příloze I. Smlouvy.
- 2) **Přílohou I.** Smlouvy je popis projektu velké výzkumné infrastruktury, který obsahuje cíle Projektu a jeho předpokládané výsledky. **Přílohou II.** Smlouvy je výše celkových uznaných nákladů Projektu

a jejich členění časové (náklady v jednotlivých letech řešení Projektu) i účelové (podle druhu výdajů) a celková výše podpory (dotace) a její členění. Pokud se na Projektu podílí další účastník/účastníci, výše podpory je vyčíslena celkově i pro příjemce a každého dalšího účastníka zvlášť.

- 3) Osobou odpovědnou příjemci za odbornou úroveň Projektu, tzv. řešitel, je [REDACTED]
Řešitel je příjemcem určen jako kontaktní osoba pro komunikaci s poskytovatelem v záležitostech týkajících se projektu.
- 4) Příjemce je povinen:
 - a) zahájit řešení Projektu v souladu s Přílohou I., nejdříve však dne **1. ledna 2023** a nejpozději do 60 kalendářních dnů ode dne nabytí účinnosti Smlouvy,
 - b) ukončit řešení Projektu, tj. ukončit věcně zaměřené projektové aktivity a čerpání poskytnuté podpory nejpozději do dne **31. prosince 2026**.
- 5) Příjemce je povinen realizovat Projekt v rozsahu a za podmínek vyplývajících ze Smlouvy a dotaci použít výlučně na úhradu uznaných nákladů Projektu.
- 6) Příjemce prohlašuje, že je organizací pro výzkum a šíření znalostí a splňuje její definiční znaky stanovené v části 1.3 písm. (ff) Rámce pro státní podporu výzkumu, vývoje a inovací (Sdělení Evropské komise č. 2022/C 414/01 – dále jen „Rámec“).
- 7) Příjemce souhlasí se zveřejněním svého názvu, sídla, dotačního titulu, výše poskytnuté dotace a závěrečné zprávy o řešení Projektu.

Článek 2

Poskytnutí podpory, její výše a podmínky jejího čerpání

- 1) Celková výše uznaných nákladů Projektu je
54 462 000 Kč
(slovy padesát čtyři milionů čtyři sta šedesát dva tisíc korun českých).
- 2) Poskytovatel poskytne Příjemci dotaci na řešení Projektu ve formě finančních prostředků převedených na účet Příjemce uvedený ve Smlouvě. Poskytovatel stanovuje celkovou výši dotace přidělenou na celé období řešení Projektu na
54 462 000 Kč
(slovy padesát čtyři milionů čtyři sta šedesát dva tisíc korun českých).
- 3) Dotace bude vyplácena v každoročních splátkách ve výši stanovené v Příloze II smlouvy v termínech podle § 10 odst. 1 zákona č. 130/2002 Sb., nedojde-li v důsledku rozpočtového provizoria podle rozpočtových pravidel k regulaci čerpání výdajů státního rozpočtu České republiky, jsou-li povinné údaje o Projektu zařazeny do Informačního systému výzkumu, vývoje a inovací (dále jen „IS VaVal“) v souladu se zákonem č. 130/2002 Sb. a jsou-li zároveň splněny všechny relevantní podmínky a dodrženy ostatní povinnosti Příjemce vyplývající ze Smlouvy a právních předpisů. V případě rozpočtového provizoria bude nevyplacená část dotace vyplacena do 60 kalendářních dnů po jeho skončení.

Článek 3 **Způsobilé a uznané náklady Projektu, účetní evidence**

- 1) Způsobilými náklady Projektu ve smyslu § 2 odst. 2 písm. m) zákona č. 130/2002 Sb. mohou být pouze takové náklady, které jsou hrazeny výlučně v souvislosti s Projektem. Náklady musí být vynaloženy v období řešení Projektu stanoveném v čl. 1 odst. 4 Smlouvy; při splnění této podmínky jsou za způsobilé považovány i náklady vynaložené před účinností Smlouvy. Uznanými náklady Projektu ve smyslu § 2 odst. 2 písm. n) zákona č. 130/2002 Sb. jsou způsobilé náklady, které jsou vynaloženy za účelem dosažení cílů Projektu, jsou vynaloženy v souladu se Smlouvou, Příjemce jejich vynaložení přesvědčivě zdůvodnil a byly schváleny Poskytovatelem.
- 2) Podpora poskytnutá podle Smlouvy směřuje na úhradu nehopodářských činností vykonávaných v rámci Projektu ve smyslu části 2.1 Rámce. Podíl využití celkové kapacity velké výzkumné infrastruktury pro hospodářské činnosti musí splňovat podmínky stanovené zejména v odst. 21 Rámce.
- 3) Příjemce je povinen vést v souladu se zákonem č. 563/1991 Sb., o účetnictví, ve znění pozdějších předpisů, oddělenou evidenci o všech nákladech a výdajích Projektu a v jejím rámci sledovat náklady nebo výdaje hrazené z podpory. Tato evidence může být kdykoliv v průběhu řešení Projektu i po jeho ukončení, a to po dobu stanovenou pro uchovávání účetních dokladů zákonem, předmětem kontroly ze strany Poskytovatele, místně příslušného Finančního úřadu a případně i dalších orgánů zmocněných ke kontrole platnou legislativou. Oddělenou účetní evidenci je Příjemce povinen vést také pro hospodářské (ekonomické) činnosti využívající kapacitu Projektu; tuto evidenci je Příjemce povinen uchovávat po dobu 10 let od konce účetního období, v němž bylo řešení Projektu ukončeno.

Článek 4 **Změny uznaných nákladů a výše poskytnuté podpory**

- 1) Změnu celkové výše uznaných nákladů Projektu nebo celkové výše dotace lze provést jen na základě předchozí písemné žádosti Příjemce, s odůvodněním, které je v souladu s plněním cílů Projektu, a lze ji provést jen uzavřením písemného dodatku ke Smlouvě. Uznané náklady a s nimi související výše podpory nemůže být v průběhu řešení Projektu změněna více, než jak připouští § 9 odst. 7 zákona č. 130/2002 Sb., které se jinak uplatňuje v případě podpory udělené na základě veřejné soutěže.
- 2) Změny finančních objemů v položkovém členění podle věcné specifikace uznaných nákladů Projektu podle Přílohy II. nebo změna rozdělení podpory mezi účastníky Projektu, které nemají vliv ani na celkovou výši uznaných nákladů Projektu, ani na celkovou výši dotace, Poskytovatel schvaluje na žádost Příjemce písemným souhlasem, bez nutnosti uzavírání dodatku Smlouvy. Při změně nesmí přesunutá částka přesáhnout 20 % celkových uznaných nákladů pro daný kalendářní rok, přičemž její maximální výše je 20 milionů Kč.
- 3) O změnu výše uznaných nákladů nebo poskytnuté podpory Projektu podle odst. 1 nebo o změnu v položkovém členění podle věcné specifikace uznaných nákladů Projektu podle odst. 2 může Příjemce požádat do dne 31. října daného kalendářního roku, nejpozději však 90 kalendářních dnů před datem ukončení řešení Projektu. Poskytovatel může vyhovět žádosti podané i po uplynutí uvedených termínů, ale nedodržení termínu může být důvodem pro nevyhovění žádosti.
- 4) Na souhlas Poskytovatele se změnou uznaných nákladů Projektu nebo změnou výše podpory podle tohoto článku nemá Příjemce právní nárok.

Článek 5 **Finanční vypořádání poskytnuté podpory**

- 1) Příjemce je povinen dotaci finančně vypořádat a nepoužité prostředky dotace vrátit do státního rozpočtu na depozitní účet Poskytovatele č. 6015-821001/0710 podle pravidel obsažených ve vyhlášce č. 367/2015 Sb., o zásadách a lhůtách finančního vypořádání vztahů se státním rozpočtem, státními finančními aktivy a Národním fondem (vyhláška o finančním vypořádání), ve znění pozdějších předpisů, a to předepsaným způsobem, zveřejněným každoročně na internetových stránkách Poskytovatele www.msmt.cz.
- 2) V případě, že Příjemce prostředky poskytnuté z dotace v daném kalendářním roce nedočerpá do dne 31. prosince daného kalendářního roku, lze tyto prostředky vrátit zpět na výdajový účet Poskytovatele č. 821001/0710, ze kterého mu byly poskytnuty, a to nejpozději do konce daného kalendářního roku. V případě předložení žádosti o změnu časového plánu čerpání dotace musí vrácení prostředků této žádosti předcházet, přičemž je nutné dodržet termíny podle čl. 4 odst. 3 Smlouvy.
- 3) V případě ukončení Projektu před původně plánovaným termínem je Příjemce povinen vrátit nevyčerpanou část dotace do 30 kalendářních dnů ode dne ukončení Projektu.
- 4) Příjemce je povinen vyrozumět o vrácení finančních prostředků souvisejících s poskytnutou podporou avízem Poskytovatele, a to v elektronické podobě na adresu elektronické korespondence aviza@msmt.cz a rovněž informovat ve stejné lhůtě o této skutečnosti odbor výzkumu a vývoje MŠMT (vyzkumneinfrastruktury@msmt.cz). Poskytovatel musí avízo obdržet nejpozději v den připsání vratky na účet.
- 5) V případě, že zvláštní zákon umožňuje Příjemci převádět část nespotřebovaných prostředků podpory do Fondu účelově určených prostředků (dále jen „FÚUP“), je povinen tu část dotace, která byla převedena do FÚUP, spotřebovat v následujícím roce řešení Projektu, a to pouze na úhradu uznávaných nákladů, na které byla původně určena podle Přílohy II.

Článek 6 **Poskytování informací a údajů o Projektu a jeho výsledcích**

- 1) Příjemce je povinen předkládat Poskytovateli za jednotlivé kalendářní roky trvání řešení Projektu průběžnou zprávu o plnění Projektu vždy **do dne 30. ledna** následujícího kalendářního roku, nebude-li Poskytovatelem stanoven jiný termín, a to včetně výkazu výdajů vynaložených v zúčtovacím období a seznamu členů řešitelského týmu, který je závazný ve vztahu k uznatelným nákladům Projektu.
- 2) Souhrnný výkaz výdajů Projektu je součástí závěrečné zprávy o plnění Projektu, kterou je Příjemce povinen předložit **do 30 kalendářních dnů** po ukončení řešení Projektu. Tato lhůta platí i v případě ukončení řešení Projektu před termínem uvedeným v čl. 1 odst. 4 Smlouvy.
- 3) Příjemce je povinen předávat Poskytovateli úplné, pravdivé a včasné informace o Projektu a získaných poznatcích a jiných výsledcích Projektu, přitom je povinen postupovat podle pokynů Poskytovatele. Příjemce souhlasí se zveřejňováním těchto požadovaných údajů a se zpřístupněním redakčně upravené závěrečné zprávy Projektu veřejnosti Poskytovatelem. Poskytovatel předává údaje o Projektu do IS VaVal a případně dalších informačních systémů dle platné legislativy.
- 4) Příjemce je povinen spravovat výzkumná data v souladu s FAIR principy a zajistit jejich dostupnost a šíření dle obvyklých zvyklostí daného oboru, jak je uvedeno v Příloze I. Pokud je předmět řešení

Projektu předmětem obchodního tajemství, je Příjemce povinen poskytnout konkrétní informace o Projektu a poznatcích a jiných výsledcích Projektu v takovém rozsahu a formě, aby byly zveřejnitelné. Pokud předmět řešení Projektu nebo jiné aktivity výzkumu, vývoje a inovací podléhají mlčenlivosti stanovené příslušným zvláštním právním předpisem, Poskytovatel a Příjemce poskytují informace o prováděném výzkumu, vývoji a inovacích a jejich výsledcích s vyloučením těch informací, o nichž to stanoví příslušný zvláštní právní předpis.

Článek 7 **Povinnosti Příjemce**

Příjemce je povinen:

- a) vyvíjet veškeré úsilí k dosažení cílů uvedených v Projektu a splnění veškerých závazků vůči Poskytovateli;
- b) po celou dobu řešení Projektu nakládat s prostředky z dotace i s veškerým majetkem získaným z těchto prostředků hospodárně, efektivně a účelně v souladu se zákonem č. 320/2001 Sb., o finanční kontrole, ve veřejné správě a o změně některých zákonů (zákon o finanční kontrole), ve znění pozdějších předpisů, zejména jej zabezpečit proti poškození, ztrátě nebo odcizení; vynakládané prostředky musí být přiměřené k cenám v místě a čase obvyklým;
- c) ve lhůtách uvedených v čl. 6 předkládat Poskytovateli průběžné zprávy a závěrečnou zprávu o plnění Projektu a respektovat pokyny Poskytovatele týkající se obsahu a struktury podávaných zpráv a termínů a lhůt pro jejich odevzdání;
- d) zamezit dvojímu financování uznaných nákladů Projektu a způsobilých výdajů vykazovaných ve stejném účetním období v dalších dotačních titulech Poskytovatele a zároveň je povinen zabránit v případě vícezdrojového financování nedovolenému křížovému financování;
- e) písemně informovat Poskytovatele o všech změnách, které nastaly v době účinnosti Smlouvy a týkají se údajů uvedených ve Smlouvě, právní osobnosti Příjemce nebo dalších účastníků Projektu, údajů požadovaných pro prokázání způsobilosti nebo které mohou mít vliv na řešení Projektu nebo jeho rozpočet, a to nejpozději do 7 kalendářních dnů ode dne, kdy tato skutečnost nastala nebo se o ní dozvěděl; výslovně se tato povinnost vztahuje také na prohlášení podle čl. 1 odst. 6 Smlouvy;
- f) v případě změny řešitele o tuto změnu Poskytovatele písemně požádat s nutností následného uzavření dodatku ke Smlouvě; novým řešitelem může být jmenována jen osoba plně odborně způsobilá, která se na řešení Projektu účastní v rozsahu potřebném k dosažení účelu Projektu a má o své účasti na Projektu s Příjemcem uzavřenou písemnou smlouvu; v případě změn ostatních členů řešitelského týmu, které neovlivní předmět, cíl a rozpočet Projektu, Příjemce informuje Poskytovatele prostřednictvím průběžné nebo závěrečné zprávy o plnění Projektu;
- g) v případě potřeby změn v položkovém členění prostředků podpory Projektu nebo v rozdělení prostředků podpory mezi účastníky Projektu o tyto změny požádat Poskytovatele s dostatečným předstihem;
- h) písemně a bezodkladně informovat Poskytovatele o podezření na nesrovnalosti zjištěné při řešení Projektu; nesrovnalostí se rozumí porušení ustanovení právních předpisů EU, právních předpisů ČR nebo ustanovení Smlouvy;
- i) řádně uchovávat originály všech rozhodnutí, smluv a dalších dokumentů týkajících se řešení Projektu v souladu s právními předpisy po dobu 10 let od data ukončení Projektu;

- j) zajišťovat kontakt Poskytovatele s řešitelem, čímž se rozumí např. předávání pokynů a dalších informací Poskytovatele řešiteli;
- k) umožnit kontrolu podle čl. 10 Smlouvy, sledování a hodnocení Projektu a účastnit se jednání, která byla svolána za tímto účelem;
- l) mít vnitřní předpis (metodiku) k vykazování režijních nákladů a vnitřní předpis pro stanovení výše osobních nákladů, včetně podmínek pro stanovení výše odměn, tyto vnitřní předpisy po celou dobu řešení Projektu dodržovat a Poskytovateli kdykoliv na vyžádání předložit jejich aktuální znění;
- m) vést internetovou stránku Projektu v anglickém znění a zveřejňovat na ní příležitosti pro využití výzkumných kapacit zajišťovaných Projektem uživateli v režimu otevřeného přístupu;
- n) uvádět v souvislosti s Projektem ve všech zveřejňovaných informacích identifikační kód Projektu podle čl. 1 odst. 1 Smlouvy a skutečnost, že na řešení Projektu byla poskytovatelem poskytnuta dotace z prostředků účelové podpory velkých výzkumných infrastruktur, přičemž v této souvislosti vždy uvádět i oficiální logo Poskytovatele v souladu s pravidly, která jsou zveřejněna na internetových stránkách Poskytovatele www.msmt.cz;

Článek 8 **Další účastníci Projektu**

- 1) Projekt nemá další účastníky.
- 2) Dalším účastníkem může být pouze subjekt, který splňuje podmínku uvedenou v čl. 1. odst. 6 Smlouvy.
- 3) Další účastníci Projektu (viz § 2 odst. 2 písm. j) zákona č. 130/2002 Sb.) se mohou podílet na využití poskytnuté dotace, pouze pokud je jejich výzkumný přínos nezbytný k řešení Projektu v souladu s Přílohou I. Příjemce je povinen koordinovat činnost všech účastníků Projektu a uzavřít s nimi písemnou smlouvu o účasti na řešení Projektu, která obsahuje zejména rozdělení jednotlivých činností mezi účastníky, rozdělení dotace mezi Příjemce a další účastníky Projektu (včetně termínů a způsobů jejího poskytování a kontroly) a úpravu práv k výsledkům dosaženým účastí jednotlivých účastníků Projektu. Úprava sjednaná ve smlouvě o účasti na řešení Projektu musí Příjemci umožnit zveřejňovat úplné, pravdivé a včasné informace o Projektu a jeho výsledcích. Příjemce odpovídá za to, že jím uzavřené smlouvy o účasti na řešení Projektu budou obsahovat ustanovení opravňující Poskytovatele provádět u dalších účastníků Projektu kontrolu ve stejném rozsahu, v jakém je Poskytovatel oprávněn kontrolovat Příjemce.
- 4) Smlouva o účasti na řešení Projektu je mezi Příjemcem a dalším účastníkem sjednána do 60 dnů od podpisu Smlouvy a přistoupí-li další účastník v průběhu řešení Projektu, je sjednána do 60 dnů od uzavření dodatku Smlouvy, který přítomnost dalšího účastníka reflektuje. Příjemce předloží smlouvy o účasti na řešení projektu Poskytovateli na vyzvání.
- 5) Příjemce je povinen poskytnout část podpory připadající na další účastníky Projektu těmto účastníkům nejpozději vždy do 30 kalendářních dnů ode dne, kdy ji obdržel od Poskytovatele. Výše prostředků, které z dotace získávají další účastníci Projektu, a jejich rozdělení v jednotlivých letech je uvedeno v Příloze II. Smlouvy.

Článek 9 Dodavatelé

Dodavatelé, jejichž plnění je potřebné k řešení Projektu, musí být Příjemcem vybráni v souladu s režimem stanoveným v zákoně č. 134/2016 Sb., o zadávání veřejných zakázek, ve znění pozdějších předpisů. Cena jakékoliv dodávky nesmí přesáhnout cenu v místě a čase obvyklou se zohledněním charakteru dodávky.

Článek 10 Kontrola řešení Projektu

- 1) Poskytovatel je v souladu s platnými právními předpisy (především podle § 13 zákona č. 130/2002 Sb., podle zákona č. 255/2012 Sb., o kontrole (kontrolní řád), ve znění zákona č. 183/2017 Sb., a podle zákona č. 320/2001 Sb., o finanční kontrole,) oprávněn provádět u Příjemce kontrolu řešení Projektu, plnění cílů Projektu, personálního a finančního řízení Projektu, čerpání a využívání dotace, včetně zhodnocení účelnosti vynaložených výdajů, dosažených výsledků a jejich právní ochrany, v průběhu řešení Projektu a následně i po dobu až 10 let od ukončení řešení Projektu. Využívá k tomu předložených průběžných zpráv o realizaci Projektu a dalších informací, které si za tímto účelem od Příjemce vyžádá. Kontrola podle tohoto odstavce se provádí také vždy po ukončení řešení Projektu, a to na základě předložené závěrečné zprávy o realizaci Projektu.
- 2) Příjemce je povinen poskytnout osobám provádějícím kontrolu přístup na svá pracoviště a k osobám podílejícím se na řešení Projektu, stejně jako ke všem účetním a dalším dokumentům, datovým záznamům a zařízením, která byla za prostředky z dotace pořízena nebo která s Projektem souvisejí.
- 3) Poskytovatel je oprávněn pozastavit poskytování prostředků dotace, pokud mu nebyly Příjemcem předloženy doklady k prokázání uznaných nákladů Projektu, průběžná zpráva o realizaci Projektu nebo ostatní podklady ve lhůtách stanovených Smlouvou.
- 4) Příjemce je povinen informovat Poskytovatele o kontrolách, které u něj byly v souvislosti s poskytnutou podporou provedeny externími kontrolními orgány, včetně závěrů těchto kontrol, a to bezprostředně po jejich ukončení.

Článek 11 Zrušení Smlouvy, sankce za porušení Smlouvy

- 1) Smluvní strana je oprávněna podat písemný návrh na zrušení této Smlouvy podle § 167 zákona č. 500/2004 Sb., správní řád, ve znění pozdějších předpisů. Návrh na zrušení Smlouvy lze podat také v případě závažného porušení povinností souvisejících s poskytnutím dotace podle této Smlouvy stanovených právním předpisem či Smlouvou.
- 2) V případě nesplnění povinností Příjemce podle čl. 7 písm. c), e), f) h), i), j) k), l), m), n) nebo čl. 8 odst. 4 vzniká Poskytovateli nárok na smluvní pokutu ve výši 50 tisíc Kč. Jestliže v přiměřené lhůtě od oznámení o uplatnění nároku na smluvní pokutu dle předchozí věty Příjemci nedojde k nápravě, nejdříve však po marném uplynutí 15 dnů od tohoto oznámení, může být smluvní pokuta udělena opakovaně. Smluvní pokuta je splatná do 30 kalendářních dnů ode dne doručení výzvy Poskytovatele Příjemci k jejímu uhrazení.

- 3) Odpovědnost za plnění Smlouvy vůči Poskytovateli nese Příjemce. Proto v případech, kdy porušení smluvní povinnosti zavinil případný další účastník Projektu, povinnost úhrady smluvní pokuty podle tohoto článku nese Příjemce. Povinnost k náhradě takto Příjemci vzniklé škody je upravena ve Smlouvě o účasti na řešení Projektu.
- 4) Za podmínek uvedených v zákoně č. 218/2000 Sb., o rozpočtových pravidlech a o změně některých souvisejících zákonů (rozpočtová pravidla), je Poskytovatel oprávněn podporu (dotaci) nebo její část nevyplatit, nebo žádat vrácení prostředků, které na základě Smlouvy již byly Příjemci vyplaceny, či jejich části.

Článek 12

Práva k výsledkům Projektu

- 1) Všechna vlastnická a užívací práva a práva duševního vlastnictví k výsledkům Projektu, jejichž využívání je upraveno zvláštními právními předpisy, náleží Příjemci. Jsou-li v Projektu zapojeni kromě Příjemce další účastníci, jsou uvedená práva mezi nimi rozdělena v poměru vyplývajícím ze smlouvy o účasti na řešení Projektu podle článku 8 Smlouvy, resp. v poměru, v jakém se na dosažení výsledku podíleli.
- 2) Příjemce a další účastníci Projektu, kteří uplatňují práva k výsledkům Projektu, jsou povinni zajistit, aby výsledky, k nimž mají vlastnická práva a které mohou být využity, byly přiměřeně a účinně chráněny a využít je nebo umožnit jejich využití při respektování nezbytné ochrany vlastnických a uživatelských práv k výsledkům a mlčenlivosti podle zvláštních právních předpisů.
- 3) Výsledky, které nepodléhají ochraně podle zvláštních právních předpisů nebo nejsou předmětem obchodního tajemství, jiného tajemství nebo utajovanou informací podle zvláštního právního předpisu, je Příjemce povinen aktivně veřejně šířit.

Článek 13

Práva k majetku

Vlastníkem hmotného majetku, potřebného k řešení Projektu a pořízeného z poskytnuté dotace, je Příjemce či další účastník Projektu, který si uvedený majetek pořídil nebo ho při řešení Projektu vytvořil. Po dobu realizace Projektu Příjemce ani další účastníci nejsou oprávněni bez souhlasu Poskytovatele s tímto majetkem nakládat ve prospěch třetí osoby, tj. například tento majetek zcizit, pronajmout, půjčit, zapůjčit či zastavit.

Článek 14

Odpovědnost za škodu

Poskytovatel nenese odpovědnost za jednání nebo naopak nečinnost Příjemce. Poskytovatel žádným způsobem neodpovídá za nedostatky výrobků nebo služeb, které spočívají v poznacích dosažených v rámci řešení Projektu.

Článek 15 **Spory smluvních stran**

Spory smluvních stran vznikající ze Smlouvy a v souvislosti s ní budou řešeny podle právních předpisů České republiky.

Článek 16 **Vyhodnocení výsledků Projektu**

Projekt je průběžně vyhodnocován Příjemcem na základě průběžných zpráv o řešení Projektu. Konečné vyhodnocení z hlediska vytýčených a dosažených cílů je předmětem závěrečné zprávy o řešení Projektu. Poskytovatel výsledky Projektu vyhodnocuje průběžně, přičemž průběžné zprávy a závěrečná zpráva o řešení Projektu jsou podkladem pro komplexní hodnocení velkých výzkumných infrastruktur, které Poskytovatel provádí prostřednictvím zahraničních hodnotitelů.

Článek 17 **Závěrečná ustanovení**

- 1) Smlouva nabývá platnosti dnem podpisu poslední ze smluvních stran a účinnosti dnem jejího zveřejnění v registru smluv podle zákona č. 340/2015 Sb., o zvláštních podmínkách účinnosti některých smluv, uveřejňování těchto smluv a o registru smluv (zákon o registru smluv), ve znění pozdějších předpisů. Účinnost Smlouvy končí ke 180. dni po ukončení Projektu.
- 2) Jakmile Smlouva nabude účinnosti, Poskytovatel bude považovat za způsobilé i ty náklady, které vznikly Příjemci, popřípadě dalším účastníkům Projektu, v době řešení Projektu podle článku 1 odst. 4 Smlouvy před datem účinnosti Smlouvy.
- 3) Změny Smlouvy, není-li ve Smlouvě výslovně uvedeno jinak, mohou být prováděny pouze dohodou smluvních stran formou písemných vzestupně číslovaných dodatků, podepsaných oprávněnými zástupci smluvních stran.
- 4) Smlouva je uzavírána v elektronické formě a podepisována digitálním podpisem osob oprávněných jednat jménem smluvních stran.
- 5) Poskytovatel zajistí uveřejnění Smlouvy a metadat Smlouvy v registru smluv včetně případných oprav uveřejnění. Nedodrží-li tento svůj závazek ve lhůtě 30 kalendářních dnů ode dne uzavření Smlouvy, je oprávněn zajistit uveřejnění Příjemce. Příjemce souhlasí s uveřejněním celého obsahu Smlouvy vyjma případných osobních údajů.
- 6) Smluvní strany souhlasně prohlašují, že si Smlouvu řádně přečetly, jejímu obsahu porozuměly, nejsou jim známy žádné důvody, pro které by Smlouva nemohla být řádně plněna nebo které by způsobovaly její neplatnost, a že Smlouva je projevem jejich vážné vůle, což stvrzují svými podpisy:

Za Poskytovatele:

Za Příjemce:

V Praze dne:

V Praze dne:

PhDr. Lukáš Levák
ředitel odboru výzkumu a vývoje

prof. MUDr. Milena Králíčková, Ph.D.
rektorka

Ministerstvo školství, mládeže a tělovýchovy

Univerzita Karlova

PŘÍLOHA I – POPIS PROJEKTU VELKÉ VÝZKUMNÉ INFRASTRUKTURY

CNC

Název: Český národní korpus

Akronym: CNC

Vědní oblast: Společenské a humanitní vědy

Příjemce: Univerzita Karlova

Statutární orgán: Prof. MUDr. Milena Králíčková, Ph.D. (rektorka)

Odpovědná osoba: ██████████

Webové stránky: <http://www.korpus.cz>

1. ZAMĚŘENÍ A VÝZNAM VELKÉ VÝZKUMNÉ INFRASTRUKTURY

Velká výzkumná infrastruktura Český národní korpus / Czech National Corpus (CNC) se soustředí na kontinuální mapování českého jazyka prostřednictvím vytváření a zpřístupňování rozsáhlých databází autentických textů (jazykových korpusů), které primárně slouží jazykově orientovanému empirickému výzkumu v oblasti společenských a humanitních věd. CNC plní úlohu analogickou k zahraničním pracovištím rozvíjejícím národní korpusy svých jazyků a jako jediná výzkumná infrastruktura svého druhu tak zastává obdobnou funkci jako Národní knihovna či Národní archiv při uchovávání národního slovesného a historického dědictví.

Rozsáhlé a kvalitní jazykové korpusy jsou dnes nezbytným předpokladem nejenom pro moderní lingvistický výzkum, ale také pro rozvoj jazykových aplikací a technologií. Při budování korpusů se hlavní pozornost CNC soustředí na jejich pestré a reprezentativní složení, vysokou kvalitu zpracování a bohatou anotaci. Díky rozsahu a různorodosti nabízeného jazykového materiálu se čeština řadí mezi datově nejlépe pokryté jazyky na světě. Neméně důležitý je však také způsob zpřístupnění jazykových korpusů prostřednictvím inovativních a uživatelsky přívětivých uživatelských rozhraní. CNC se proto kromě sběru a zpracování jazykových dat soustřeďuje také na vývoj specializovaných webových aplikací, které jsou zpravidla jazykově nezávislé a které jsou využívány také podobnými projekty v zahraničí.

CNC představuje etablovanou infrastrukturu, která se objemem dat, mírou využití (průměrný počet uživatelských dotazů za den se pohybuje okolo 4 000) a rozsahem uživatelské komunity (více než 8 000 registrovaných aktivních uživatelů) řadí mezi největší centra svého druhu na světě.

CNC pro svoji činnost disponuje rozvinutým technickým a odborným zázemím, které lze rozdělit na dvě základní části: jednak infrastrukturu související se sběrem a zpracováním dat spolu s technickým zázemím pro jejich archivaci, a jednak infrastrukturu související s poskytováním služeb koncovým uživatelům, která zahrnuje uživatelská rozhraní pro přístup k datům (včetně související techniky, zejména počítačových serverů spolu s databázemi a dedikovanými aplikacemi) a dále nástroje uživatelské podpory.

CNC bude dále rozvíjet následující 3 hlavní oblasti své činnosti: a) rozvoj a anotace datové základny, b) vývoj nástrojů pro práci s korpusy, c) uživatelská podpora.

a) Rozvoj datové základny se v CNC zaměřuje především na tyto směry:

- současný psaný jazyk představovaný pravidelně aktualizovanými korpusy řady SYN, které již po 20 let mapují českou tištěnou produkci; řada SYN zahrnuje také reprezentativní

stomilionové korpusy obsahující beletrii, odbornou literaturu a publicistiku, vydávané v pětiletých intervalech (celkový objem řady SYN dosahuje 5 mld. slov);

- pravidelně aktualizovaný monitorovací korpus ONLINE pokrývající primárně internetovou žurnalistiku (přes 6 mld. slov);
- korpusy spontánní mluvené češtiny řady ORTOFON s přepisy a zvukovými nahrávkami výhradně autentického, nepřipraveného a neformálního jazyka pokrývající celé území ČR (2 mil. slov);
- mnohojazyčný paralelní korpus InterCorp zahrnující texty spolu s jejich překlady mezi češtinou a přibližně 40 dalšími jazyky pro kontrastivní a translátologický výzkum (cca 1,8 mld. slov; nadále pouze automatické zpracování);
- diachronní korpus DIA se zaměřením na 19. století;
- specializované korpusy vznikající pro výzkum specifických témat, např. nářeční korpus DIALEKT.

Rozvoj datové základny je doplněn rozvojem morfologického a syntaktického značkování směřujícího k pokrytí různých variet češtiny jedním univerzálním a lingvisticky adekvátním schématem s vysokou úspěšností anotace. Pro zvýšení mezijazykové kompatibility bude CNC i nadále věnovat pozornost mezinárodnímu anotačnímu standardu Universal Dependencies.

b) Důraz na empirii v lingvistice spolu s objemem a pestrostí dat CNC přináší nutnost vyhodnocovat výsledky statistickými metodami. To klade nemalé nároky na rozvoj infrastruktury v oblasti uživatelských rozhraní a nástrojů, které musejí být přístupné pro badatele z humanitních a společenských věd, kteří často nedisponují expertní znalostí statistiky či automatického počítačového zpracování rozsáhlých dat. CNC je v této oblasti aktivní jak na poli rozvoje metodologie vytěžování korpusových dat, tak v oblasti vývoje vlastních aplikací pro inovativní práci s nimi. CNC momentálně nabízí na portále <http://www.korpus.cz> celkem 14 uživatelských webových aplikací s různým odborným zaměřením, některé z nich jsou přístupné také pomocí API.

c) Pro přístupnost infrastruktury a pro její optimální využití je klíčová rozsáhlá a všestranná uživatelská podpora zajišťující vstřícnost vůči uživatelské komunitě. Uživatelská podpora CNC zahrnuje následující prvky:

- rozsáhlá dokumentace v podobě wiki, která obsahuje manuály k uživatelským aplikacím, popis dostupných korpusů, on-line kurz pro práci s CNC, slovníček pojmů z oblasti korpusové lingvistiky aj.;
- on-line poradna sloužící registrovaným uživatelům pro konzultaci výzkumných i technických obtíží při práci s CNC; poradna je spojená se systémem hlášení chyb a požadavků na implementaci další funkcionality uživatelských rozhraní, který mj. pomáhá získávat od uživatelů zpětnou vazbu;
- repozitář vědeckých výstupů založených na CNC, který zároveň představuje bibliografii české korpusové lingvistiky;
- repozitář korpusových cvičení pro využití ve výuce češtiny v podobě vzorových úloh a pracovních listů;
- hostování korpusů: služba expertního zpracování dat vytvořených třetími subjekty do podoby jazykových korpusů a jejich zveřejnění prostřednictvím CNC;
- datové balíčky: příprava jazykových dat odvozených z korpusů CNC podle individuálních požadavků uživatelů;

- mezi prvky uživatelské podpory můžeme zařadit také školení, popularizaci v médiích a workshopy pro uživatele konané pravidelně dvakrát ročně s kapacitou okolo 50 osob.

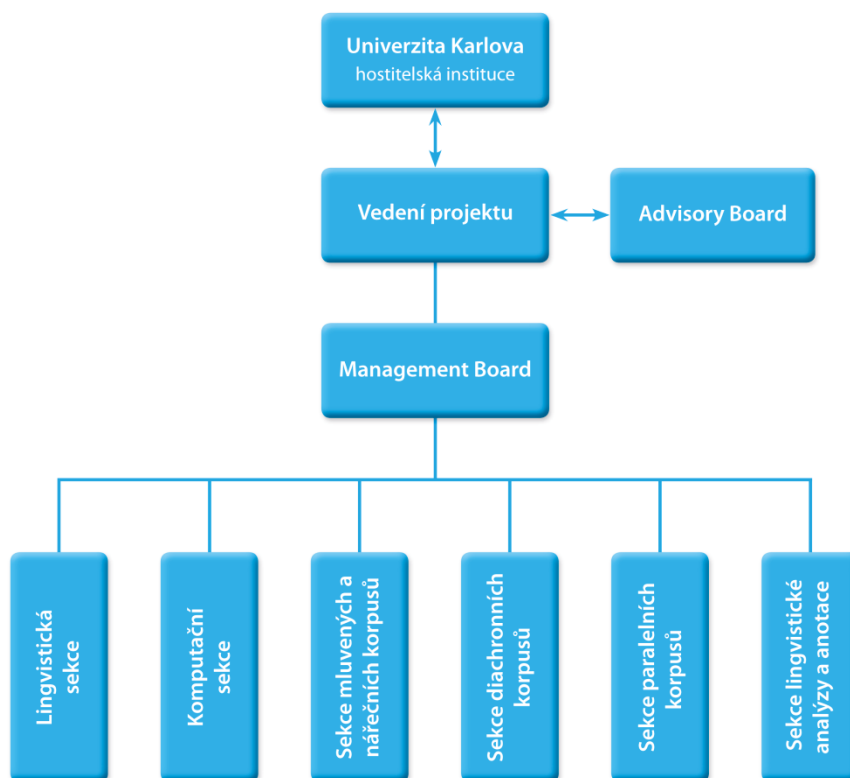
Pro CNC je prioritou udržení současné vysoké úrovně uživatelské podpory, která je jedním z klíčových faktorů efektivního využívání infrastruktury. Vedle faktického zpřístupnění dat a nástrojů proto bude CNC i nadále poskytovat metodologickou podporu pro jejich uplatnění ve vědecké a aplikační praxi.

Ve všech výše popsaných oblastech se infrastruktura rozvíjí a bude rozvíjet jak v závislosti na světových trendech v oblasti korpusové lingvistiky, počítačnické lingvistiky a digital humanities, tak na uživatelské zpětné vazbě a na vlastním infrastrukturním výzkumu, který se zaměřuje především na metodologické otázky vytěžování jazykových dat.

2. MANAGEMENT VELKÉ VÝZKUMNÉ INFRASTRUKTURY

Hostitelskou institucí VVI CNC je Univerzita Karlova (UK) reprezentovaná svojí Filozofickou fakultou (FF UK). Univerzita Karlova poskytuje infrastrukturu prostřednictvím FF UK především prostory spolu s potřebným administrativně-technickým zázemím (administrativní oddělení děkanátu, serverovna, připojení k internetu, prostory pro pořádání workshopů apod.). Díky tomuto institucionálnímu zakotvení v univerzitním prostředí je infrastruktura s uživateli z řad studentů a vědecko-pedagogických pracovníků v každodenním kontaktu.

Řízení infrastruktury CNC je rozděleno mezi tři orgány: vedení projektu, Management Board a Advisory Board (viz organigram).



Vedení projektu tvoří odpovědná osoba infrastruktury CNC a dále tajemník, kteří jsou zodpovědní za exekutivu, administrativu, koordinaci všech aktivit infrastruktury a implementaci strategických rozhodnutí.

Pracovníci infrastruktury jsou podle oblasti své činnosti v CNC rozděleni do 6 sekcí:

- lingvistická sekce (strategický infrastrukturní výzkum, rozvoj metodologie, design korpusů a aplikací, popularizace);
- počítačnická sekce (celková technická podpora projektu, vývoj aplikací);
- sekce mluvených a nářečních korpusů (sběr a zpracování dat pro mluvené a nářeční korpusy);
- sekce diachronních korpusů (sběr a zpracování starších českých textů);
- sekce paralelních korpusů (zpracování mnohojazyčných dat pro paralelní korpus InterCorp);
- sekce lingvistické analýzy a anotace (lingvistická anotace korpusových dat).

Vzhledem k rozpočtovým škrtnutím a z toho vyplývajícím omezení aktivit VVI dojde během roku 2023 k vnitřní reorganizaci CNC. Její konkrétní podoba je zatím předmětem jednání, velmi pravděpodobně se však dotkne pouze rozdělení na sekce (snížení počtu sekcí a aktualizace jejich zaměření).

Vedoucí těchto sekcí tvoří spolu s vedením projektu **Management Board**, který je klíčový pro koordinaci běžných provozních aktivit infrastruktury a který také přijímá středně- a dlouhodobá strategická rozhodnutí.

Advisory Board je mezinárodním kontrolním a poradním orgánem infrastruktury. Skládá se z předních světových odborníků na problematiku výstavby a využití jazykových korpusů, zpracování přirozeného jazyka, a také z expertů v oblasti gramatického a lexikálního popisu češtiny. V souladu se svým statutem zasedá Advisory Board pravidelně jednou ročně v prostorách CNC na FF UK. V rámci zasedání se jeho členové vyjadřují k výroční zprávě CNC, kterou předkládá vedení projektu a která pokrývá všechny oblasti činnosti infrastruktury. Advisory Board dále poskytuje zpětnou vazbu k prioritám dalšího rozvoje infrastruktury (rozvoj datové základny v jednotlivých korpusových řadách, její anotace, vývoj nových aplikací pro práci s korpusy apod.).

3. SPOLUPRÁCE VELKÉ VÝZKUMNÉ INFRASTRUKTURY

CNC je asociovaným členem konsorcia CLARIN-CZ a certifikovaným K-centrem ESFRI CLARIN se zaměřením na korpusovou lingvistiku, zdroje CNC jsou dále zapojeny do CLARIN Federated Content Search. CNC udržuje a bude dále rozvíjet úzkou a mnohostrannou spolupráci s VVI LINDAT/CLARIAH-CZ, národním uzlem CLARIN ERIC a DARIAH ERIC v ČR. Spolupráce s LINDAT/CLARIAH-CZ probíhá na mnoha úrovních: vzájemné zastoupení v poradních orgánech obou infrastruktur, zpřístupnění vybraných jazykových korpusů CNC výzkumné komunitě také prostřednictvím repozitáře LINDAT/CLARIAH-CZ, využívání nástrojů nabízených tímto repozitářem, společná koordinace rozvoje open-source nástroje pro uživatelskou práci s korpusy KonText vyvíjeného v CNC; výsledkem je mj. nasazení KonTextu také v několika dalších centrech CLARIN v Evropě.

Jako VVI zaměřující se především na kontinuální mapování češtiny je CNC ve světovém kontextu jedinečná, rozvíjí proto intenzivní spolupráci se zahraničními výzkumnými institucemi obdobného charakteru. Mimo napojení na ESFRI lze zmínit především Jazykovedný ústav Ľudovíta Štúra SAV, Slovensko; Institut für Deutsche Sprache, Mannheim, Německo; Trojina, Institute for Applied Slovene Studies, Ljubljana, Slovinsko a řadu dalších; u mnoha zahraničních institucí probíhá spolupráce na základě rámcových smluv. Důležitá je také kooperace CNC s týmy tvořícími národní korpusy jednotlivých jazyků, iniciativami SlaviCorp a International Comparable Corpus atd. V tomto typu spolupráce jde především o výměnu expertíz, zkušeností, dat a nástrojů, které jsou jazykově nespécifické, dále o hostování korpusů (CNC tak může uživatelům nabízet webové korpusy řady Aranea, korpusy horní a dolní lužické srbštiny aj.) a o vývoj uživatelských rozhraní CNC s potenciálem širokého mezinárodního využití (kromě KonTextu např. editor paralelních textů InterText, webové rozhraní Word at a Glance nebo statistická kalkulačka Calc).

Na infrastrukturní spolupráce navazují spolupráce na vědeckých projektech se zahraničními jazykovědnými pracovišti, jejichž výsledky mají pozitivní dopad na vylepšení služeb infrastruktury: v rámci společného projektu s Universitou v Tromsø například vzniká nová, zásadním způsobem přepracovaná verze webového rozhraní KWords pro analýzu klíčových slov. CNC se dále aktivně zapojuje do dění v oblasti digital humanities, má zastoupení v CzADH (Česká asociace pro digitální vědy) a jejím prostřednictvím také v EADH (European Association for Digital Humanities), účastní se projektu H2020 „Computational Literary Studies Infrastructure“ aj.

Zaměření na zahraniční badatele je důležité také při budování uživatelské základny. Díky inovativním uživatelským aplikacím a vysoké úrovni souvisejících služeb spolu s jejich snadnou dostupností má CNC i přes svůj národní charakter okolo 20 % registrovaných aktivních uživatelů ze zahraničí. Těmito uživateli jsou typicky jednak zahraniční studenti a bohemisté, jednak uživatelé mnohojazyčného paralelního korpusu InterCorp nebo hostovaných cizojazyčných korpusů. Tento fakt dokazuje, že CNC je významnou VVI i v mezinárodním měřítku, nezanedbatelnou přidanou hodnotou této činnosti je přitom také propagace českého jazyka a kultury.

V národním měřítku je klíčová úzká spolupráce s VVI LINDAT/CLARIAH-CZ, dále pobíhá infrastrukturní a vědecko-výzkumná spolupráce s řadou dalších národních institucí. Mezi uživateli CNC nalezneme zástupce prakticky všech akademických a vědeckých pracovišť ČR v oblasti humanitních věd a zejména lingvistiky, jejichž zpětná vazba je využita pro další rozvoj služeb VI: ve spolupráci s Ústavem pro českou literaturu AV ČR vznikl rozsáhlý Korpus současné poezie, v Ústavu pro jazyk český AV ČR vzniká na datech a rozhraních CNC Akademický slovník současné češtiny. Nad rámec pravidelných a dlouhodobých kooperací probíhá řada dílčích projektů s různými bohemistickými a jazykovědnými centry v ČR, typicky se jedná o výzkumné projekty, jejichž součástí je analýza a interpretace jazykových dat.

4. OTEVŘENÝ PŘÍSTUP A UŽIVATELÉ VELKÉ VÝZKUMNÉ INFRASTRUKTURY

CNC je otevřená virtuální infrastruktura přístupná na individuální bázi nepřetržitě, on-line (vzdálený přístup) a zdarma. Podmínkou využití některých aplikací je pouze registrace, která je nutná z licenčních důvodů a při níž uživatelé potvrzují svůj souhlas s nekomerčním využitím přístupu k CNC. Přístup není dále nijak omezen, mohou ho tedy používat jak vědečtí a akademičtí pracovníci, tak i široká veřejnost. Uživatelé s akademickou afiliací navíc mohou namísto registrace do CNC využít přihlášení účtem své domovské instituce (single sign-on Shibboleth).

Povaha infrastruktury a její technické vybavení umožňují plnohodnotně obsloužit simultánní požadavky mnoha uživatelů, není proto potřeba jakkoli regulovat přístup k infrastruktuře, zavádět opatření pro jeho prioritizaci či pro přednostní alokaci kapacity.

Centrálním přístupovým bodem ke službám infrastruktury CNC je portál <http://www.korpus.cz>, kde jsou k dispozici aplikace pro práci s korpusy, manuály, uživatelská podpora a veškerá dokumentace. Vedle toho jsou vybrané datové zdroje CNC přístupné i prostřednictvím repozitáře infrastruktury LINDAT/CLARIAH-CZ. Všechny aplikace vyvíjené v rámci infrastruktury jsou open-source.

Uživatelská komunita čítá více než 8 000 registrovaných aktivních uživatelů; z nich přibližně 70 % tvoří studenti, zbývající část je rozdělena mezi vědecké a výzkumné pracovníky (cca 15 %) a laickou veřejnost (překladaelé, redaktoři, pedagogové apod.; cca 15 %). Registrovaní uživatelé CNC pocházejí zhruba v 80 % případů z ČR, zbývajících 20 % tvoří výzkumníci a studenti z téměř 70 zemí světa. Z Evropy je nejvíce zastoupeno Německo, Polsko a Slovensko, ze světa pak USA a Kanada. Přibližně stejný poměr domácích a zahraničních uživatelů očekáváme i v nadcházejícím období. Pokud jde o afiliace, uživatelé z ČR pocházejí téměř ze všech filozofických a pedagogických fakult veřejných vysokých škol v ČR, nejvyšší podíl vědeckých pracovníků mezi uživateli pak má AV ČR.

Vzhledem k cílené otevřenosti infrastruktury pro všechny zájemce o jazykové zdroje je velká část služeb a aplikací CNC dostupná i bez registrace; počet ani původ neregistrovaných uživatelů však nelze uspokojivě odhadnout, sledován je pouze objem a druh jejich interakcí s infrastrukturou.

Průměrný počet uživatelských interakcí s infrastrukturou činí v průměru okolo 4 000 denně, celkem tedy téměř 1,5 milionu ročně. Za uživatelskou interakci v tomto smyslu považujeme pouze položení dotazu do některého z uživatelských rozhraní, ne tedy jakoukoli další práci s výsledky dotazu ani využívání informačních zdrojů portálu CNC. Vzhledem k tomu, že typické využití korpusových nástrojů předpokládá po položení dotazu následnou analýzu a vyhodnocení, je počet všech interakcí s infrastrukturou (zahrnující kromě dotazů i jejich další zpracování) ještě několikanásobně vyšší.

5. SOCIOEKONOMICKÉ DOPADY VELKÉ VÝZKUMNÉ INFRASTRUKTURY

Rozsáhlé jazykové korpusy se v posledních desetiletích staly nezbytným nástrojem pro empirický jazykovědný výzkum a pro vytváření slovníků, gramatik a dalších specializovaných popisů jazyka. V návaznosti na rostoucí možnosti zpracování rozsáhlých textových souborů se korpusy stále více uplatňují i v dalších oborech společenských a humanitních věd, což se odráží také ve využití zdrojů a služeb infrastruktury CNC; ty se tak používají nejenom v lingvistice, ale také v řadě oborů, v nichž hraje důležitou roli jazyk a interpretace textu: historie, sociologie, pedagogika, kognitivní vědy apod. Kontinuální záznam jazyka v korpusech pak slouží nejenom jako záznam způsobu komunikace, ale vydává svědectví i o společnosti, kultuře a historii, a lze ho tak chápat jako kulturní artefakt významný pro formování národní identity. CNC je v tomto smyslu unikátní a nezastupitelnou výzkumnou infrastrukturou, která pro tyto účely systematicky poskytuje data a nástroje.

Snadná dostupnost referenčních reprezentativních korpusů v CNC vytváří předpoklady pro replikovatelný empirický výzkum. Díky obecnému charakteru většiny korpusů navíc pro řadu výzkumných projektů odpadá nutnost věnovat úsilí a prostředky na sběr a zpracování vlastních dat. Centralizace těchto aktivit tak přispívá nejenom k zefektivnění podpory vědy v ČR, ale také ke zkvalitnění a vyrovnání výchozích podmínek výzkumu, protože veškeré služby CNC jsou poskytovány na základě rovného a otevřeného přístupu. Infrastruktura CNC tak v současnosti představuje pro mnoho jazykově orientovaných vědeckých projektů základní empirické východisko a nezbytnou podmínku realizace.

O hojném využívání CNC svědčí také počet záznamů v Repozitáři vědeckých výstupů založených na CNC, který je součástí webového portálu CNC a který v současné době eviduje více než 3 000 záznamů. Repozitář také ukazuje, že s využitím infrastruktury CNC vzniká na vysokých školách v ČR ročně více než 100 závěrečných prací (bakalářských, magisterských, doktorských), což dokládá mj. vliv CNC na formování nové generace výzkumníků a pedagogů v oblasti lingvistiky a obecněji humanitních věd směrem k práci s empirickými daty.

Vzhledem k tomu, že infrastruktura CNC je otevřeně přístupná široké veřejnosti, patří mezi její socioekonomické přínosy také rozšiřování povědomí o jazyce. V oblasti vzdělávání jde zejména o spolupráci při vytváření učebnic a výukových pomůcek pro výuku češtiny (jako mateřského i druhého jazyka) a dalších jazyků (zejména angličtiny, němčiny, francouzštiny a dalších evropských jazyků). Jak v případě mateřského jazyka, tak v případě češtiny jako jazyka cizího je efektivní výuka založená na empirických datech a umožňující brzké odstranění jazykové bariéry primárním předpokladem pro úspěšnou integraci do české společnosti (ať už ze sociálně vyloučeného, jazykově znevýhodněného prostředí nebo přímo z prostředí jiného jazyka). Ve výuce cizích jazyků obecně platí, že efektivní zacílení na komunikační kompetence předpokládá kontrastivní empirické poznání cílového jazyka na pozadí jazyka výchozího, které poskytují v unikátním rozsahu a pestrosti právě paralelní korpusy.

V neposlední řadě jsou důležitými oblastmi využití zdrojů CNC počítačové zpracování přirozeného jazyka a jazykové technologie, jejichž aplikační potenciál s rozvojem digitální ekonomiky neustále roste. Jde zejména o strojový překlad, rozpoznávání řeči, text mining atd., k jejichž vývoji infrastruktura

přispívá především poskytováním kvalitních a rozsáhlých dat pro trénování a testování jazykových modelů. Diskurzivní analýza využívající korpusové zdroje CNC, zejména korpus ONLINE, také pomáhá při identifikaci rizik, jakými jsou např. hybridní hrozby a dezinformační weby.

6. KOMUNIKAČNÍ STRATEGIE A PROPAGACE VELKÉ VÝZKUMNÉ INFRASTRUKTURY

Pro oslovení potenciálních uživatelů z řad veřejnosti CNC aktivně využívá sociální sítě (Facebook, Twitter, YouTube), na nichž pravidelně zveřejňuje příspěvky propagující nové nástroje, připravované workshopy či konference a také instruktážní videa. CNC se věnuje i intenzivní popularizaci v médiích. Členové týmu vystupují s rozhovory a popularizačními články např. v Lidových novinách, v Respektu, v Reflexu nebo v Deníku N, přijali řadu pozvání do ČT či ČRo; v případě ČRo a Deníku N jde navíc v současné době o pravidelné pořady/sérii článků.

Pro získávání nových uživatelů z řad odborné veřejnosti je klíčová především organizace workshopů práce s CNC (pravidelné uživatelské workshopy pořádané dvakrát ročně v sídle VVI v Praze, a dále zvané workshopy v jiných institucích v ČR i v zahraničí), zvané přednášky a účasti na konferencích; na nich členové týmu CNC prezentují jak výzkum směřující k vylepšení služeb CNC, tak také příkladové studie ukazující využití CNC v lingvistice a v dalších humanitních oborech. Příležitostně CNC konference také (spolu)pořádá či hostuje, v září 2022 to například byl 6. ročník mezinárodní konference Translation in Transition, kde jedním z hlavních témat bylo využití zdrojů CNC pro kontrastivní lingvistiku a translologii.

CNC dále pořádá pravidelný seminář, který se koná v hybridní podobě, a mohou ho tak navštěvovat zájemci z jiných vědeckých a výzkumných institucí z ČR i ze zahraničí. Cílem semináře je především seznámení odborné veřejnosti s nejnovějším vývojem v oblasti korpusové lingvistiky, jako přednášející jsou proto zvaní také renomovaní zahraniční odborníci.

7. UZNANÉ NÁKLADY VELKÉ VÝZKUMNÉ INFRASTRUKTURY

Výše uznaných nákladů VVI CNC je uvedena v tabulce, která tvoří Přílohu II smlouvy. Další text jednotlivé položky této tabulky podrobněji vysvětluje a komentuje.

Povaha infrastruktury, její zaměření a portfolio služeb, které poskytuje v oblasti humanitních a společenských věd, vyžadují významný podíl expertní činnosti kvalifikovaných pracovníků. Převážnou část rozpočtu CNC proto tvoří **osobní náklady**, které zahrnují dvě složky: především jsou to mzdy interních zaměstnanců hostitelské instituce VVI CNC zaměstnaných formou HPP (v roce 2023 předpokládáme přibližně 15 FTE, do dalších let počet FTE mírně klesne) a v menší míře pak odměny externích spolupracovníků na akvizici, transkripci a anotaci jazykových dat zaměstnaných na DPČ a DPP.

Následující souhrn ukazuje rozdělení celkového počtu úvazků interních zaměstnanců do kategorií předpokládané pro rok 2023 (uváděná výše úvazku je po zaokrouhlení), včetně jejich mzdových hladin (uváděna je měsíční hrubá mzda přepočtená na celý úvazek včetně osobního ohodnocení, bez odměn):

- projektový management, koordinace a administrativa: 1 FTE, ████████ Kč;
- vědečtí a výzkumní pracovníci: 8 FTE, ████████ Kč;
- techničtí pracovníci (včetně programování a vývoje software): 6 FTE, ████████ Kč.

Mzdy všech interních pracovníků CNC jsou stanoveny v souladu se mzdovým předpisem Univerzity Karlovy jako hostitelské instituce. Mzdové rozpětí v každé kategorii je dáno především rozdílnou pracovní zkušeností juniorních a seniorních pracovníků, odlišností pracovní náplně a dosaženou akademickou hodností.

Provozní náklady se skládají ze tří hlavních částí: cestovních nákladů, služeb a režie.

Cestovní náklady budou použity především na cesty na relevantní národní a mezinárodní konference v oboru působnosti VI, na nichž budou pracovníci CNC prezentovat výsledky infrastrukturního výzkumu nebo představovat a propagovat služby, data či aplikace VI, dále na jednání o zapojení do mezinárodních projektů infrastrukturní povahy a na cesty jinak spojené s infrastrukturou CNC.

Služby budou použity na nákup hardware a software, opravy zařízení, externí programátorské práce, výdaje spojené s akvizicí a revizí textů, nákup odborné literatury a kancelářských potřeb.

Výše režie byla stanovena v souladu s interní směrnicí FF UK a v jednotlivých letech se pohybuje mezi 10–11,5 % celkového rozpočtu CNC.

Členské poplatky nejsou v rámci účelové podpory plánovány.

Investice nejsou vzhledem ke komplementárnímu způsobu financování výzkumných infrastruktur součástí účelové podpory, jejich financování se předpokládá v rámci projektu podaného do výzvy OP JAK.

CNC

PŘÍLOHA II – DETAILNÍ ROZPOČET PROJEKTU A UZNANÉ NÁKLADY PROJEKTU (V TIS. KČ)

	2023		2024		2025		2026		Celkem	
	Uznané náklady	Dotace MŠMT	Uznané náklady	Dotace MŠMT	Uznané náklady	Dotace MŠMT	Uznané náklady	Dotace MŠMT	Uznané náklady	Dotace MŠMT
Osobní náklady	13 599	13 599	11 482	11 482	10 998	10 998	10 806	10 806	46 885	46 885
Investice	0	0	0	0	0	0	0	0	0	0
Členské poplatky	0	0	0	0	0	0	0	0	0	0
Provozní náklady	2 006	2 006	1 857	1 857	1 857	1 857	1 857	1 857	7 577	7 577
Celkem	15 605	15 605	13 339	13 339	12 855	12 855	12 663	12 663	54 462	54 462