##### Príloha č. 8 - Metodika a spôsob rozdelenia účelových finančných prostriedkov pre excelentné výskumné súčasti verejných vysokých škôl v roku 2023

##### 1. Cieľom tohto rozdelenia účelových finančných prostriedkov je podporiť excelentné výskumné pracoviská verejných vysokých škôl. Táto metodika identifikuje podľa indikátorov, uvedených nižšie, štvrtinu najlepších súčastí verejných vysokých škôl za každú oblasť vedy (M1: prírodné vedy, M2: technické vedy, M3: lekárske vedy, M4: poľnohospodárske, lesnícke a veterinárske vedy, M5: spoločenské vedy, M6a: humanitné vedy, M6b: umenie. Presné zaradenie vedných oblastí sa nachádza v prílohe A, hárok E3).

##### 2. Hlavné indikátory, ktoré sa použijú na identifikáciu excelentných súčastí:

##### a) excelentná publikačná činnosť podľa bodu 4, celková váha 60%.

##### b) excelentný výkon pri získavaní výskumných grantov podľa bodu 5, celková váha 40%.

##### c) počet tvorivých zamestnancov na ustanovený týždenný pracovný čas (100% úväzok), zaradených do vedných oblastí vysokými školami k 31. 12. 2021. Tieto údaje sa nachádzajú v prílohe A, hárok E2.

##### 3. Časovým rámcom údajov pre indikátory z bodu 2, v súlade s metodikou rozpisu dotácie, sú roky 2020 a 2021.

##### 4. Excelentná publikačná činnosť sa určuje nasledovne:

##### a) V oblastiach M1: prírodné vedy, M2: technické vedy, M3: lekárske vedy, M4: poľnohospodárske, lesnícke a veterinárske vedy sa zohľadňujú váhou 100% indexované publikácie (WoS, Scopus), teda kategórie ADC, ADD, ADM, ADN. Zoznam sa nachádza v prílohe B.

##### b) V oblasti M5: spoločenské vedy sa zohľadňujú váhou 75% indexované publikácie (WoS, Scopus), teda kategórie ADC, ADD, ADM, ADN. Tieto výstupy sa bonifikujú v súlade s metodikou rozpisu (prvý kvartil JCR je zohľadnený váhou 6, druhý kvartil váhou 4, tretí kvartil váhou 1 a štvrtý kvartil váhou 0,5). Zoznam sa nachádza v prílohe B. Váhou 25% sa zohľadňujú monografie (AAA, AAB). Tieto výstupy sa bonifikujú tak, že vydavateľstvá zaradené v databáze Nordic List (https://kanalregister.hkdir.no/publiseringskanaler/Forside.action?request\_locale=en) v kategórii 2 (prestížne vedecké vydavateľstvo) sú bonifikované váhou 6, vydavateľstvá zaradené v databáze Nordic List v kategórii 1 (vedecké vydavateľstvo) váhou 4, vydavateľstvá zaradené v databáze CVTI (zoznam zahraničných vydavateľstiev: http://cms.crepc.sk/Data/Sites/1/pdf/zoznam-vydavatelstva/zoznam-vydavatelstva-01-2020.pdf) sú zohľadnené váhou 1 a všetky ostatné monografie sú zohľadnené váhou 0,5). Zoznam aj so zaradením sa nachádza v prílohe C.

##### c) V oblasti M6a: humanitné vedy sa zohľadňujú váhou 60% indexované publikácie (WoS, Scopus), teda kategórie ADC, ADD, ADM, ADN. Tieto výstupy sa bonifikujú v súlade s metodikou rozpisu (prvý kvartil JCR je zohľadnený váhou 6, druhý kvartil váhou 4, tretí kvartil váhou 1 a štvrtý kvartil váhou 0,5). Zoznam sa nachádza v prílohe B. Váhou 40% sa zohľadňujú monografie (AAA, AAB). Tieto výstupy sa bonifikujú tak, že vydavateľstvá zaradené v databáze Nordic List (https://kanalregister.hkdir.no/publiseringskanaler/Forside.action?request\_locale=en) v kategórii 2 (prestížne vedecké vydavateľstvo) sú bonifikované váhou 6, vydavateľstvá zaradené v databáze Nordic List v kategórii 1 (vedecké vydavateľstvo) váhou 4, vydavateľstvá zaradené v databáze CVTI (zoznam zahraničných vydavateľstiev: http://cms.crepc.sk/Data/Sites/1/pdf/zoznam-vydavatelstva/zoznam-vydavatelstva-01-2020.pdf) sú zohľadnené váhou 1 a všetky ostatné monografie sú zohľadnené váhou 0,5). Zoznam aj so zaradením sa nachádza v prílohe C.

##### d) v oblasti M6b umenie sa zohľadňujú váhou 100% umelecké výstupy v kategóriách ZZZ, ZZY, ZYZ, ZYY podľa váh platných pre rozpis dotácie 2022. Údaje sa nachádzajú v rozpise dotácie na rok 2022, hárok E4g\_M6\_umenie. Osobitne sa zohľadňujú oblasti performatívne umenie a vizuálne umenie.

##### 5. Excelentný výkon pri získavaní výskumných grantov sa určuje nasledovne: pre všetky oblasti okrem M6b (umenie) sa zohľadňuje objem grantových prostriedkov, získaných súťažným spôsobom v kategórii zahraničné výskumné granty. Zoznam sa nachádza v prílohe D. Zahraničné výskumné granty sa priraďujú iba tým pracoviskám, v ktorých vysoké školy uviedli príslušných zamestnancov v danej oblasti.

##### 6. Pri určovaní excelentného výkonu sa použijú štatistické modely (body 10 až 12), ktoré zohľadňujú reálnu empirickú distribúciu údajov (body za publikačnú činnosť, počet € za zahraničné výskumné granty), aby sa zabránilo skresleniam, vyplývajúcim z použitia pre dané distribúcie nevhodných bodových odhadov (priemery, mediány), a aby sa zabránilo aj skresľujúcej redukcii variability, ktorá vzniká použitím normalizačných prepočtov (percentá, podiely). Podrobnosti a kódy na výpočet sú rozpísané v bode 11.

##### 7. Výkony za všetky publikácie podľa bodu 4 sa vynásobia 2, aby sa zachoval diskrétny (bodový) charakter publikačných výstupov s vylúčením zlomkov, pretože publikácie sú reálne v diskrétnych počtoch (counts), teda váhy publikačných výkonov sa budú zohľadňovať nie na škále 0,5 – 1 – 4 – 6, ale na škále 1 – 2 – 8 – 12.

##### 8. Hodnotí sa každá súčasť verejnej vysokej školy, ktorá má v danej oblasti (M1: prírodné vedy, M2: technické vedy, M3: lekárske vedy, M4: poľnohospodárske, lesnícke a veterinárske vedy, M5: spoločenské vedy, M6a: humanitné vedy, M6b: umenie) zaradených aspoň 5 zamestnancov.

##### 9. Pri hodnotení sa zohľadňuje počet zamestnancov danej súčasti, a to prostredníctvom logaritmickej škály a offset (offset umožňuje prepočítať výsledky modelu na pomer na jedného zamestnanca).

##### 10. Pri výpočte za každú oblasť sa použijú regresné modely, v ktorých sú závislými premennými výkony (počet bodov za publikácie, objem grantov v €), a nezávislou premennou počet zamestnancov (logaritmická škála a offset). Oprávnenosť týchto modelov vychádza z veľmi jednoduchej premisy, že čím má súčasť verejnej vysokej školy vyšší počet zamestnancov, tým vyšší vedecký výkon by mala vytvárať.

##### 11. V prípade publikačných výstupov, presne v súlade s príslušnou štatistickou literatúrou, mali najlepšiu zhodu s údajmi modely, ktoré použili Sichelovu (generalizovanú inverznú Gaussovsko-Poissonovú) distribúciu, vyvinutú práve pre bibliometrické dáta (Sichel, H. S. (1992): Anatomy of the generalized inverse Gaussian-poisson distribution with special applications to bibliometric studies, Information Processing & Management, 28(1), 5-17). Ako uvádza Sichel, bibliometrické dáta sú veľmi často „šikmé“ (skewed) a majú „ťažké chvosty“ (heavy tails) vďaka prítomnosti extrémnych hodnôt (extrémne hodnoty sú také údaje, ktoré sa výrazne líšia od ostatných údajov v danej premennej). Ak by sa tieto vlastnosti nezohľadnili, bežné výpočty (priemer a pomer na zamestnanca) by boli veľmi skreslené: niekoľko pracovísk s veľmi vysokým výkonom by ich vychýlilo, takže pracoviská s menším výkonom by boli veľmi znevýhodnené. Ako je jasne vidieť v bode 17 Dodatok, všetky distribúcie sú veľmi výrazne šikmé s jasnou prítomnosťou extrémnych hodnôt.

Sichelova (alebo generalizovaná inverzná Gaussova-Poissonova distrubúcia) má 3 parametre (μ, σ, ν) a jej pravdepodobnostná funkcia je nasledovná:



##### pričom premenná y nadobúda kladné diskrétne hodnoty y = 0, 1, 2.... a μ ˃ 0, σ ˃ 0, -∞ < ν < ∞ a zároveň platí



K λ (t) je modifikovanou Besselovou funkciou tretieho druhu.

Na výpočet je potrebné použiť R knižnicu „gamlss“ (generalizované aditívne modely), príslušný kód je nasledovný:

Na načítanie dát sa odporúča použiť prílohu „zdroj\_data23.xlsx“, v ktorej sa nachádzajú všetky relevantné premenné (počty zamestnancov, publikačné výkony, objemy grantov). Na výpočet regresných modelov sa odporúča štatistický program R, na načítanie dát je možné použiť napríklad knižnicu „readxl“:

install.packages("readxl")

a privolať ju príkazom:

library(readxl)

Za predpokladu, že príloha nazvaná „zdroj\_data23.xlsx“ je uložená v priečinku s názvom R, dátový súbor sa načíta nasledovne:

dat = read\_excel('C:\\R\\zdroj\_data23.xlsx')

Premenná „zamPRIR“ predstavuje počet zamestnancov v oblasti prírodných vied, premenná „pubPRIR“ predstavuje publikačné výkony daných pracovísk. Tieto premenné extrahujeme z dátového súboru pomocou príkazov:

zamPRIR = as.vector(dat$zamPRIR)[1:24]

pubPRIR = as.vector(dat$pubPRIR)[1:24]

Pri vypočítaní iných modelov pre iné vedné oblasti je potrebné zmeniť názvy, a aj rešpektovať príslušný počet pracovísk (údaj v hranatých zátvorkách [1:24] zodpovedá počtu pracovísk v oblasti prírodných vied).

Po nainštalovaní knižnice príkazom:

install.packages("gamlss")

a jej privolaním príkazom

library(gamlss)

je výpočet príslušného regresného modelu nasledovný:

m1 = gamlss(pubPRIR~offset(log(zamPRIR)),family=SICHEL,method=RS(200))

Hodnoty, predikované modelom (popísané v bode 15 a uvedené v dátových prílohách) získame príkazom:

fitted(m1)

Vypočítanú presnú hodnotu priemerného pomeru publikácií vidíme pomocou príkazu:

summary(m1)

kde sa nachádza ako „Mu cofficients, Estimate“. Ak túto hodnotu prevedieme z logaritmickej škála príkazom exp(x), dostaneme pomer publikačných bodov na jedného zamestnanca.

##### V prípade grantov, rovnako v súlade so štatistickou literatúrou, mali najlepšiu zhodu s údajmi modely, ktoré použili Tweedie distribúciu, zohľadňujúcu mimoriadnu šikmosť a veľký počet núl (Kurz, F. C. (2017): Tweedie distributions for fitting semicontinuous health care utilization cost data, BMC Medical Research Methodology, 17:171). Opäť autor presvedčivo ukázal, že kontinuálne premenné, ktoré sú mimoriadne šikmé, obsahujú extrémne hodnoty a excesívny počet núl (v bode 17 dodatok je zo zobrazení úplne zrejmé, že to platí aj pre objemy grantov), sa najlepšie modeluje pomocou Tweedie distrinúcie, ktorej pravdepodobnostná funkcia je nasledovná:

 pričom jej parametre λ, α a β majú k distribučným parametrom μ, ϕ a *p* nasledujúci vzťah:



Na výpočet sa použila knižnica „glmmTMB“, jej inštalácia a privolanie:

install.packages("glmmTMB")

a privolá sa príkazom:

library(glmmTMB)

Za predpokladu, že si pripravíme premennú s grantovými výkonmi:

graPRIR = as.vector(dat$graPRIR)[1:24]

príslušný model sa vypočíta nasledovne:

m2 = glmmTMB(graPRIR~offset(log(zamPRIR)),family=tweedie)

Vypočítanú presnú hodnotu priemerného pomeru grantového objemu vidíme pomocou príkazu:

summary(m2)

kde sa nachádza ako „Conditional model, Estimate“. Ak túto hodnotu prevedieme z logaritmickej škála príkazom exp(x), dostaneme pomer grantového objemu na jedného zamestnanca.

##### 12. Za každý regresný model sa vypočítali predikované (fitted) hodnoty, teda hodnoty za každú súčasť, ktoré by mala táto súčasť dosiahnuť, pokiaľ by mala v danej oblasti vzhľadom na svoj počet zamestnancov a výsledky ostatných súčastí v danej oblasti priemerné výkony. Následne sa vypočíta pre každú súčasť rozdiel medzi pozorovanými (reálnymi) výkonmi a týmito predikovanými hodnotami (technicky povedané, ide o reziduály): kladné hodnoty znamenajú nadpriemerné výkony danej súčasti v danej oblasti, záporné výkony znamenajú podpriemerné výkony. Následne sa tieto hodnoty štandardizujú do z-skóre, aby bolo možné ich porovnávať, a vypočíta sa finálne z-skóre podľa váhovaní v súlade s bodmi 2 a 4. Všetky tieto údaje sa za jednotlivé oblasti nachádzajú v prílohe A, hárky E4a-E4g. Presný popis všetkých údajov sa nachádza v bode 15.

##### 13. Za každú oblasť sa určí štvrtina najlepších súčastí, ktoré dostanú alokácie podľa bodu 14.

##### 14. Pri rozpise alokácie sa postupuje nasledovne:

##### a) z celkovej alokácie za excelentný výskum (18 658 146,30 €) sa vypočíta 60% podiel za publikácie (11 194 887,78 €), a z tohto podielu sa vyčlení v súlade s metodikou rozpisu dotácie na rok 2022 10% na umeleckú tvorivú činnosť (1 119 488,78 €). Tieto zdroje sa rozpíšu v oblasti M6b umenie osobitne, a to oddelene pre performatívne umenie a oddelene pre vizuálne umenie.

##### b) zvyšná alokácia (17 538 657,52 €) sa rozdelí do jednotlivých oblastí podľa reálnych excelentných výkonov: za každú oblasť okrem M6b umenie (teda M1: prírodné vedy, M2: technické vedy, M3: lekárske vedy, M4: poľnohospodárske, lesnícke a veterinárske vedy, M5: spoločenské vedy, M6a: humanitné vedy) sa zohľadnia celkové výkony za excelentné publikácie a medzinárodné výskumné granty. Presný prepočet alokáci na oblasti sa nachádza v prílohe A, hárok E1 (vrátane rozdelenia alokácií v oblasti M6b Umenie pre performatívne umenie a vizálne umenie).

##### c) pri alokácii na štvrtinu excelentných súčastí sa vezme do úvahy počet zamestnancov a výška vedeckého výkonu tak, že sa preváhuje excelentný výkon spôsobom, že súčasti s vyšším výkonom dostanú na jedného zamestnanca vyššiu alokáciu než súčasti s nižším výkonom.

##### 15. Podrobné vysvetlenie údajov z prílohy A, hárky E4a-E4g:

##### a) stĺpec „zam2021“ uvádza počet zamestnancov na ustanovený týždenný pracovný čas (100% úväzok) k 31. 12. 2021. Počet zamestnancov menší než 5 (nehodnotené súčasti) je zvýraznený červenou.

##### b) stĺpec „final23“ uvádza z-skóre celkového excelentného výkonu, váhované podľa bodu 2, z údajov podľa písm k) a m). Štvrtina najlepších súčastí v danej oblasti je zvýraznená zelenou.

##### c) stĺpec „suma23“ uvádza alokovanú sumu podľa písmen d) – g).

##### d) stĺpec/údaj „výkon\_exc23“ predstavuje súčet excelentného vedeckého výkonu za štvrtinu najlepších súčastí v danej oblasti (zvýraznený zelenou).

##### e) stĺpec „podiel\_exc23“ uvádza podiel daného excelentného pracoviska na štvrtine excelentného výkonu.

##### f) stĺpec „váha\_zam23“ uvádza váhy, použité pri prepočte jednotkovej alokácie na excelentného zamestnanca.

##### g) stĺpec/údaj „suma/zam23“ uvádza jednotkovú sumu na prepočet pre jedného excelentného zamestnanca.

##### h) stĺpec „index23“ uvádza súčet bodov za roky 2020 a 2021 za indexované publikácie (vynásobené podľa bodu 7), stĺpec „mon23“ uvádza v prípade oblastí M5 (spoločenské vedy) a M6a (humanitné vedy) navyše aj súčet bodov za monografie (vynásobené podľa bodu 7).

##### i) stĺpec „fit\_index23“ uvádza hodnoty, predikované modelom podľa bodu 12 za indexované publikácie, stĺpec „fit\_mon23“ uvádza v prípade oblastí M5 (spoločenské vedy) a M6a (humanitné vedy) hodnoty, predikované modelom podľa bodu 12 za monografie.

##### j) stĺpec „res\_index23“ uvádza rozdiel medzi stĺpcami „vyk\_index23“ a „fit\_index23“ (neštandardizované reziduály), stĺpec „res\_mon23“ uvádza v prípade oblastí M5 (spoločenské vedy) a M6a (humanitné vedy) rozdiel medzi stĺpcami „vyk\_mon23“ a „fit\_mon23“ (neštandardizované reziduály).

##### k) stĺpec „z\_pub\_index23“ uvádza z-skóre zo stĺpca „res\_index23“, v prípade oblastí M5 (spoločenské vedy) a M6a (humanitné vedy) je v stĺpci „z\_pub\_mon“ uvedené z-skóre zo stĺpca „res\_mon23“, stĺpec „z\_pub23“ v prípade oblastí M5 (spoločenské vedy) a M6a (humanitné vedy) uvádza z-skóre za všetky publikácie (štandardizované reziduály).

##### l) stĺpec „granty23“ uvádza objem získaných zahraničných výskumných grantov, stĺpec „fit\_granty23“ uvádza hodnoty, predikované modelom podľa bodu 12 a stĺpec „res\_granty23“ uvádza rozdiel medzi „granty23“ a „fit\_granty23“ (neštandardizované reziduály).

##### m) stĺpec „z\_granty23“ uvádza z-skóre zo stĺpca „res-granty23“ (štandardizované reziduály).

16. V prípade korekcií vo vstupných údajoch sa môžu výsledky ešte zmeniť.

17. Dodatok: Zobrazenia distribúcií údajov

 Publikácie: prírodné vedy Publikácie: technické vedy



 Publikácie: lekárske vedy Publikácie: poľnohospodárske vedy

 

 Publikácie: spoločenské vedy Publikácie: humanitné vedy

 

 Granty: prírodné vedy Granty: technické vedy

 

 Granty: lekárske vedy Granty: poľnohospodárske vedy

 

 Granty: spoločenské vedy Granty: humanitné vedy

 