

Nový rozmer v poskytovaní obsahových služieb pre vedu a výskum – platforma DIMENSIONS



Začiatkom roku 2018 bola do života uvedená nová databáza Dimensions. Bola vyvinutá pre potreby akademickej komunity a v spolupráci s ňou, s cieľom, podať z jedného miesta dôležité informácie o výskume v celej komplexnosti a súvislostiach, počínajúc jeho plánovaním v podobe grantov, cez výstupy výskumu publikované vo vedeckej literatúre, pokračujúc klinickými štúdiami a končiac patentmi s potenciálnym uplatnením v hospodárskej praxi. V porovnaní s klasickými databázami, napr. Web of Science, Scopus, Google Scholar, teda okrem publikácií a citácií, vrátane altmetrických skóre, obsahuje granty, klinické štúdie a patenty. Údaje pochádzajú z diverzifikovaných voľne prístupných, ale aj licencovaných zdrojov, ktoré sú navzájom prepojené hypertextovými odkazmi, na základe ktorých je možné rýchlo získať prehľad o predmete záujmu. Využitie sú v nich pokročilé informačno-komunikačné technológie – strojové učenie, spracovanie prirodzeného jazyka a umelá inteligencia a indexácia úplných textov. Existujú tri verzie databázy, z ktorých bibliografická je voľne prístupná na <https://app.dimensions.ai/>.

V tomto príspevku sú uvedené základné fakty o databázovej platforme Dimensions, o jej špecifických prvkoch, funkcionalitách a obsahu v porovnaní s inými bibliografickými databázami, o výhodách, ktoré prináša pre výskumníkov a výskumné inštitúcie a pozíciu, akú v nej zaujíma Slovensko. Informuje o zisteniach, ktoré priniesli nedávne vedecké publikácie zamerané na štúdium tejto novej a inovátorskej databázy. Článok je zároveň pozvaním pre členov akademickej obce, aby sa s touto databázou oboznámili a využívali jej mimoriadny potenciál a schopnosti.

Tento nový rozmer prináša technologická firma Digital Science, do portfólia ktorej patrí šesť spoločností: Readcube, Altmetric, Figshare, Symplectic, DS Consultancy a ÜberResearch. Inovátorské úsilie formulované v ich spoločnom projekte, ktorý slúži potrebám vedeckej komunity v kľúčových bodoch výskumného cyklu, vyústilo v rámci spolupráce s viac ako stovkou vedúcich výskumných a vedeckých inštitúcií a jednotlivcov na celom svete do vzniku novej bibliografickej a citačnej platformy Dimensions. Jej tvorcovia ju charakterizujú ako moderný, inovatívny, vzájomne prepojený vedomostný systém, ktorého cieľom je prelomiť bariéry pri vyhľadávaní a v inováciách tým, že umožňuje vyhľadať a rýchlejšie sprístupniť najvýznamnejšie informácie, analyzovať výstupy výskumu z hľadiska akademického. ale aj širšieho kontextu a získať prehľad o súčasnom dianí vo vede pri plánovaní výskumných činností do budúcnosti (1). Je určená pre výskumníkov, akademické a vládne inštitúcie a agentúry financujúce vedu a výskum na jednej strane a na druhej slúži vydavateľom vedeckej a odbornej literatúry.

1. Dimensions v porovnaní s inými databázami

Podľa propagačných materiálov (1) platformu Dimensions nie je možné priamo porovnávať s databázami, akými je PubMed, Scopus alebo Web of Science alebo voľne prístupným indexom Google Scholar. To, čo ponúka, sa vyznačuje odlišným štýlom, charakterom a prístupom. Svojím inovačným konceptom sa snaží vymykať z rámca toho, čo ponúkajú „klasické-tradičné“ bibliografické databázy. Prístup je založený na identifikovaní a zlepšení tých častí vyššie uvedených databázových systémov, ktoré sú najlepšie prepracované alebo najdôležitejšie z hľadiska práce používateľa tak, aby ju skvalitnili a uľahčili.

1.1 Publikácie, citácie/altmetrické skóre, patenty, klinické štúdie, granty, pokročilé IKT

Hoci ju nie je možné priamo porovnávať so skôr uvedenými databázovými platformami, pretože je vybudovaná na odlišných princípoch a vyznačuje sa odlišnými črtami, poskytuje solídny citačný graf podobne ako Scopus alebo Web of Science. Je však zrejme, že citácie pochádzajú z rôznorodejších zdrojov, neobmedzujú sa len na vedeckú literatúru publikovanú vo vedeckých časopisoch a knihách, ale obsahujú aj dokumenty kategórie sivej literatúry, priemyselného vlastníctva publikovaného formou patentov a údaje pochádzajúce z klinických štúdií. Tieto rôznorodé druhy publikácií slúžia aj ako zdroj informácií pri vyhľadávaní podľa zadanej témy podobným prístupom, aký je uplatňovaný v Google Scholars – využívaná je indexácia úplných textov publikácií. Pokročilé syntaktické vyhľadávanie, ktoré umožňuje, aby vyhľadávanie bolo špecifické, sa neobmedzuje len na biomedicínsky výskum, ako v prípade PubMed. Jednou z vrstiev v množine poskytovaných informácií sú aj údaje o grantoch, ktorými sa otvára výskumný cyklus, ako prvá prezentácia výskumného zámeru, ktorá má presvedčiť financujúci subjekt o dôležitosti tohto zámeru. Údaje o grantoch majú mimoriadne špecifické postavenie, pretože sú pohľadom do budúcnosti – na základe zamýšľaného výskumu, ktorý sa ešte len bude realizovať v nasledujúcich rokoch, umožňujú analyzovať trendy, vývoj a smerovanie danej vednej oblasti. Výstupom je mnohovrstvová informácia, ktorá so zadanou témou najužšie súvisí a poskytuje na tému hlbší a komplexnejší pohľad v rôznych súvislostiach a vzťahoch. Platí to aj pre plánovanie výskumu a kanály, prostredníctvom ktorých budú výsledky výskumu komunikované (1).

1.2 Zdroje údajov

1.2.1 Časopisy

Metadátovej základ databázy je vybudovaný na širokom spektre dátových zdrojov, voľne prístupných a pod licenciou, z ktorých vynikajú PubMed, PubMed central, Arxiv a najmä Crossref. Prispieva doň viac ako 50 000 vydavateľov a je otvorený návrhom akademickej komunity alebo vydavateľov na zaradenie ďalších. Implementované sú časopisy z nórskeho registra vedeckých časopisov, seriálov a vydavateľov (Norwegian Register for Scientific Journals, Series and Publishers), beta verzie Web of Science, časopisy zo zoznamu ERA zostaveného Austrálskou radou pre výskum (Australian Research Council/ARC) v spolupráci s Radou pre národné zdravie a výskum v medicíne (National Health and Medical Research Council/NHMRC) pre potreby národného hodnotenia excelentného výskumu ERA, aktualizovaného v roku 2018 (ERA 2015, resp. ERA 2018), zo zoznamu voľne prístupných vysoko kvalitných časopisov zo všetkých vedných oblastí v DOAJ, ktorých úplné znenia textov sú dostupné v bezplatnej verzii Dimensions. Napokon je to PubMed, ktorý je vyhľadávacím nástrojom v abstraktoch a odkazoch na citovanú literatúru v publikáciách v oblasti vied o živej prírode a biomedicíny – prevažne z MEDLINE. V Dimensions tento vyhľadávací nástroj filtruje len publikácie označené identifikátorom PubMed (PMID). Najväčšími prispievateľmi sú vydavateľstvá vedeckej literatúry (v abecednom poradí) *Cambridge University Press, DeGruyter, Elsevier, IEEE, Oxford University Press, Royal Society of Chemistry Sage Publications, Springer Nature, Taylor & Francis, Wiley a Wolters Kluwer*, no zastúpenie majú aj stredné a malé vydavateľstvá. Pre vydavateľov je vytvorená samoobslužná platforma, cez ktorú môžu svoj obsah integrovať do Dimensions priamo bez vynaloženia väčšieho úsilia. Publikácie s úplným znením textu, ktoré tvoria viac ako 50 % obsahu databázy Dimensions, sú indexované, čo znamená, že je prehľadávaný podstatne väčší počet publikácií a sprístupnené sú tie publikácie, ktoré sú tematicky najzhodnejšie s požadovaným vyhľadávacím príkazom – kľúčovým slovom. Publikácie sú v Dimensions rozčlenené na 5 typov: článok, konferenčný príspevok, kapitola, preprint a monografia. Bezplatná verzia Dimensions sprístupňuje všetky publikácie, a tie sú zaradené do vedných oblastí podľa klasifikačnej schémy FoF (157 vedných oblastí).

1.2.2 Klinické štúdie

Podľa definície Štátneho ústavu na kontrolu liečiv klinické skúšanie lieku je každý výskum na človeku, ktorým sa určujú alebo potvrdzujú klinické účinky, farmakologické účinky alebo iné farmakodynamické účinky, ktorým sa preukazuje akýkoľvek nežiaduci účinok a ktorým sa zisťuje absorpcia, distribúcia, metabolizmus a vylučovanie jedného alebo viacerých skúšaných liekov určených na humánne použitie, s cieľom zistiť ich neškodnosť a účinnosť; klinickým skúšaním sa hodnotí aj biologická dostupnosť a biologická rovnocennosť (bioekvivalencia) skúšaného lieku (2). Údaje o klinických skúškach pochádzajú z voľne prístupných registrov v USA ClinicalTrials.gov, európskeho registra (EU-CTR <https://www.clinicaltrialsregister.eu/>), japonského UMIN Clinical Trials Registry (UMIN-CTR <http://www.umin.ac.jp/ctr/>), registra ISRCTN, ktorý je časťou Springer Nature (<https://www.isrctn.com/>), registra austrálskych a novozélandských klinických skúšok (ANZCTR <http://www.anzctr.org.au/Default.aspx>), čínskeho registra (CHICTR <http://www.chictr.org.cn/abouten.aspx>), holandského (NTR <http://www.trialregister.nl/trialreg/index.asp>) a nemeckého registra klinických skúšok (GTRS https://www.drks.de/drks_web/setLocale_EN.do). Odkazmi sú prepojené na výskumníkov, výskumné inštitúcie, krajiny, financujúce subjekty. Sú kategorizované do vedných oblastí podľa klasifikačnej schémy FoF (157 vedných oblastí) a prístupné len v platenej verzii Dimension Plus.

1.2.3 Patenty

Údaje o patentoch ako právne chránenom priemyselnom vlastníctve a potenciálne komerciovateľných výstupoch vedeckého výskumu uplatniteľných v praxi poskytuje spoločnosť IFI Claims <https://www.ificlaims.com/>, ktorá patrí do portfólia firiem Digital Science. Dimensions obsahuje údaje zo svetovej databázy (WIPO), nadnárodných (EPO) a národných databáz úradov priemyselného vlastníctva: Patentový a známkový úrad v USA (USPTO), Nemecký patentový a známkový úrad (DPMA), Kanadský úrad pre duševné vlastníctvo (CIPO), Duševné vlastníctvo India (IPI), Úrad pre duševné vlastníctvo vo VB (IPO UK), Francúzsky národný úrad pre priemyselné vlastníctvo (INPI), Odbor pre duševné vlastníctvo v Hong Kongu (IPD). Patenty sú odkazmi prepojené navzájom a sú prepojené na publikácie, inštitúcie, financujúce subjekty a granty, pričom je uplatnená kategorizácia do vedných oblastí podľa klasifikačnej schémy FoF (157 vedných oblastí). Patenty sú prístupné len v platenej verzii Dimension Plus.

1.2.4 Granty

Agregovanie veľkej databázy grantov spoločnosťou ÜberResearch, ktorá patrí do portfólia firiem Digital Science, v spolupráci so subjektmi financujúcimi výskum začalo ešte v roku 2013. Tá sa následne stala súčasťou platformy Dimensions. Granty neodrzakujú všetok financovaný výskum, pretože sú poskytované rôznymi formami (súťažne-predkladaním projektov alebo jednorazovou všeobecnou podporou), preto sa niektoré ani neobjavia v databáze podporených grantov. Poskytujú teda čiastočný pohľad na vývoj trendov, pohyby vo vedných disciplínach, úmysly a zámery vo výskume v nasledujúcich rokoch, akési nazretie do budúcnosti. Analýza tohto prostredia umožňuje rozhodovacej sfére (vedná politika, stratégie, plánovanie) včasne zasiahnuť, formulovať stratégie, nielen späťne identifikovať nesprávne rozhodnutia a konštatovať fakty. Spoločnosť ÜberResearch agreguje databázu grantov s financovaním 1,2 biliónov US dolárov. Granty sú odkazmi prepojené na publikácie, inštitúcie, financujúce subjekty a publikácie s uplatnením kategorizácie podľa klasifikačnej schémy FoR (157 vedných oblastí). Granty sú prístupné len v platenej verzii Dimension Plus.

1.3 Klasifikačná schéma vedných odborov FoR na mikroúrovni

Všetky súčasti databázovej platformy Dimensions – výskumníci, výskum, (časopisecké a knižné publikácie, patenty, klinické štúdie a projekty), výskumné organizácie, citácie ohlasy na publikácie vyjadrené altmetrickými ukazovateľmi, vydavatelia, subjekty financujúce výskum a podporené granty, sú navzájom prepojené konzistentným spôsobom. Jeden z prvkov, ktorý zabezpečuje konzistentnosť je jednotná klasifikácia do vedných oblastí. Všetky publikované výstupy výskumu, či už sú to publikácie, patenty, klinické štúdie a projekty, sú obsahom jednotne zaradené do vedných oblastí podľa systému kategórií vedných a vzdelávacích odborov zavedeného v Austrálii a na Novom Zélande (*Australian and New Zealand Standard Research Classification* – ANZSRC) pod označením FoR (Field of Research) (3). Zaradovanie do kategórií vedných oblastí sa nedeje podľa vedeckého zamerania, klasifikácie časopisov (obsahové zameranie článkov nie vždy zodpovedá zaradeniu časopisu z hľadiska vednej disciplíny; pri takomto triedení je tiež problematická a nedostatočná obsahová špecifikácia publikácií v časopisoch označených ako multidisciplinárne), ako je to v ostatných databázach, ale na základe analýzy textu jednotlivých publikácií, teda na mikroúrovni, uplatnením najnovších informačno-komunikačných technológií pri práci s obsahom (strojové učenie ML, spracovanie prirodzeného jazyka NLP a umelá inteligencia). Práve to umožňuje súčasne získať informácie o danej téme nielen z článkov a kníh, ale aj z ostatných typov výstupov výskumu (patent, klinická štúdia, grant) a zaručuje ich konzistentnosť. Klasifikačná schéma FoR má 3 hierarchické stupne, z ktorých je v DIMENSIONS použitý 2. stupeň obsahujúci 157 skupín vedných oblastí vychádzajúcich z 22 sekcií – hlavných kategórií vedných odborov. Oba stupne (viac a menej granulovaný) systému kategórií je možné zvoliť aj v analytickom nástroji InCites spol. Clarivate Analytics, ktorý spracúva údaje z databázy Web of Science Core Collection. To umožňuje vzájomné porovnanie oboch databázových nástrojov hlavne pri ich použití na hodnotenie výstupov výskumu, ale aj kvality samotných nástrojov. Z hľadiska financujúcich subjektov je v Dimensions zavedený aj členenie podľa iných kategórií, napr. klasifikačná schéma pre výskum, ochorenia a kategorizáciu chorôb Research, Condition, and Disease Categorisation (RCDC) používaná Národnými ústavmi zdravia (NIH) v USA a klasifikácia výskumu v zdravotníctve Health Research Classification System (HRCS) používaného takmer všetkými subjektmi financujúcimi biomedicínsky výskum vo Veľkej Británii. Klasifikácia HRCS má dva kódovníky – kódy pre výskumné činnosti Research Activity Codes (RAC) a zdravotnícke kategórie (HC).

1.4. Voľný prístup

1.4.1 Dimensions. Mimoriadne priťažlivá a veľmi aktuálna je myšlienka voľného prístupu k takmer 97 miliónom metadáto- vým záznamom o publikáciách a k citáciám prevyšujúcim počet 1 miliardy pre výskumníkov na vedecké nekomerčné účely cez aplikáciu Dimensions (obrázok 1) <https://app.dimensions.ai/> a prostredníctvom rozhrania pre programovanie aplikácií (API); surovým údajom a metrikám cez voľne prístupné Dimensions Metrics API a Dimensions Badges. Pomocou nich si môžu jednotlivci alebo inštitúcie vložiť na svoje webové stránky alebo aplikácie symbol Dimensions s aktuálnym počtom citácií (obrázok 2) a altmetrickým skóre (obrázok 3) zároveň s prepojením na príslušné zdroje.

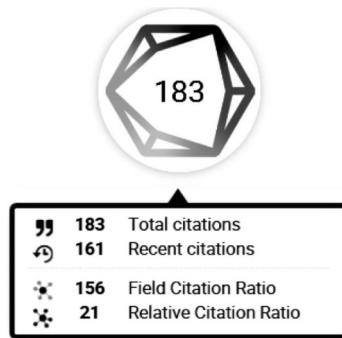
The screenshot shows the Dimensions database interface. At the top, there is a search bar with the query 'e.g. plastic AND instrument'. Below the search bar, there are several sections: 'FILTERS' on the left, 'PUBLICATIONS' in the center, and 'ANALYTICAL VIEWS' on the right. The 'PUBLICATIONS' section shows a list of results with columns for 'Title, Author(s), Bibliographic reference - About the metrics', 'Effect of end group of amorphous perfluoro-polymer electrets on electron trapping', and 'Seonwoo Kim, Kuniko Suzuki, Ai Sugie, Hiroyuki Yoshida, Masafumi Yoshida, Yuji Suz...'. The 'ANALYTICAL VIEWS' section shows a list of 'FIELDS OF RESEARCH' with counts, such as '1103 Clinical Sciences' (5,058,191) and '0601 Biochemistry and Cell Bio...' (2,870,261). There is also an 'OVERVIEW' section and a 'RESEARCHERS' section with a list of names and counts, such as 'Roman Gladyshevskii' (7,887) and 'K Cenzual' (7,706).

Obrázok 1. Uživatelské rozhranie voľne prístupnej verzie databázy Dimensions

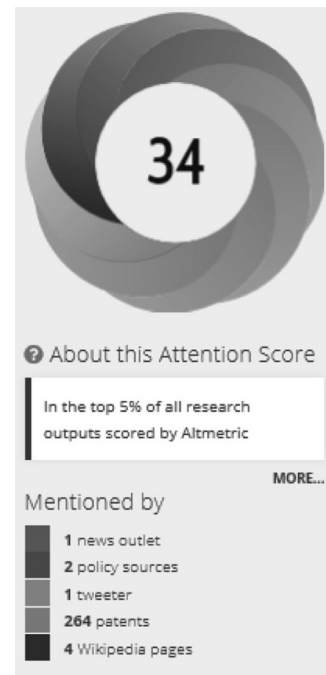
Okrem tejto voľne prístupnej verzie existujú dve verzie s prístupmi spoplatnenými za, podľa vyjadrenia jej tvorcov, prijateľnú cenu: Dimension Plus a Dimension Analytics.

1.4.2 Dimensions Plus (obrázok 4) je určená inštitúciám a ich výskumníkom. Poskytuje prístup k všetkým údajom v databáze, t. j. okrem publikácií a citácií ako vo voľne prístupnej Dimensions, je to prístup k viac ako 36 miliónom patentov, 427 tisíc klinických štúdií, 4,1 miliónom grantov a 361 miliónom politických dokumentov spolu s prepojeniami. K dispozícii je aj viac funkcionalít v porovnaní s voľne prístupnou verziou; k nim patria, napr. filter pre inštitúciu, financujúci subjekt a filter pre krajinu, ktoré nie sú prítomné vo voľne dostupnej verzii. Tiež obsahuje filtre špecifické pre patenty – majiteľ, krajina udelenia patentu, vedná oblasť, klasifikácia podľa medzinárodného patentového triedenia (IPC), klasifikácia podľa kooperatívneho patentového triedenia (CPC), právny stav, krajina udelenia patentu, rok publikovania, stav prihlášky a rok udelenia patentu, filtre špecifické pre klinické štúdie – rok začatia a rok ukončenia klinickej štúdie, rok v ktorom klinická štúdia prebieha, sponzor/inštitúcia spolupracujúca na klinickej štúdiu, financujúci subjekt, ktorý spolupracuje na klinickej štúdiu, fáza klinickej štúdie, typ ochorenia, vedné oblasti, register, z ktorého pochádzajú údaje o klinickej štúdiu a napokon sú to filtre pre granty – číslo grantu, riešiteľská organizácia a jej krajina, riešiteľ, financujúci subjekt, dátum a rok začatia a dátum a rok ukončenia projektu, suma v eurách. V tejto verzii je možné získať prístup k API.

1.4.3 Dimensions Analytics slúži pre potreby subjektov financujúcich vedu a vydavateľov. Má zabudované nástroje na pokročilé analytické spracovanie dát, napr. porovnanie výkonov inštitúcií alebo financujúcich subjektov a na tvorbu výstupných správ. Tiež je možné pomocou týchto nástrojov spracovávať vlastné údaje.



Obrázok 2. Symbol Dimensions pre počet citácií s príslušnými metrikami



Obrázok 3. Symbol Dimensions pre altmetrické skóre

The screenshot shows the Dimensions Plus interface. On the left, there are filter categories like PUBLICATION YEAR, RESEARCHER, FUNDER, etc. The main area displays search results for 'e.g. plastic AND instrument', showing publication titles, authors, and Altmetric scores. On the right, there is an 'ANALYTICAL VIEWS' section with a 'FIELDS OF RESEARCH' list and an 'OVERVIEW' line graph showing 'Publications' from 2009 to 2018.

Obrázok 4. Uživatelské rozhranie spoplatenej verzie databázy Dimensions Plus

1.5 Metriky

Dimensions zavádza nové metriky odvodené od počtu citácií. Ten vyjadruje, koľkokrát bola určitá publikácia uvedená v zozname citovanej literatúry v publikáciách – v databáze Dimensions. Tvorcovia Dimensions presadzujú a podporujú myšlienku, aby sa samotní užívatelia zapojili do procesu vývinu a implementovania nových metrických podľa vlastných potrieb.

1.5.1 Relatívny pomer citácií (RCR)

Tento ukazovateľ slúži ako príklad aktívnej účasti používateľov na tvorbe databázy Dimensions. Bol vyvinutý v Národnom ústave zdravia (NIH) v Bethesde v Marylande (4) a následne implementovaný do Dimensions. Je to metrika na úrovni článku nezávislá od vednej disciplíny, ktorá udáva, ako je citačne úspešný určitý článok v porovnaní s ostatnými článkami v rovnakej vednej oblasti. Vedná oblasť je určená na základe vednej oblasti, do ktorej sú zaradené články citujúce hodnotený článok. Kvantifikuje vplyv tohto článku na základe novátorského využitia ko-citačnej siete na normalizáciu citácií (podľa oblasti výskumu FoR), ktoré tento článok získa. Priemerný RCR pre všetky články sa rovná 1. Hodnota nad 1 značí, že citačne je hodnotený článok nad priemerom a hodnota pod 1 znamená, že citačne hodnotený článok získal menej citácií ako je priemer pre danú vednú oblasť. RCR nie je uvedené v prípade, keď článok nemá žiadnu citáciu a v prípade, keď doba od jeho publikovania je kratšia ako 2 roky.

1.5.2 Pomer citácií vo vednej oblasti (FCR)

Ďalším novým ukazovateľom je (Field Citation Ratio FCR), ktorý je tiež na úrovni článku. Udáva, ako je hodnotený článok relatívne citačne úspešný v porovnaní s článkami približne rovnakého veku zaradenými v rovnakej vednej oblasti. Referenčná hodnota je 1. Nad 1 poukazuje na to, že hodnotený článok má v porovnaní s priemerným počtom citácií článkov z rovnakej vednej oblasti a definovaného roku publikovania a veku, viac citácií a vice versa. Priemerná hodnota sa počíta ako geometrický priemer, ktorý znižuje vplyv odľahlých hodnôt, t. j. extrémne vysokých počtov citácií na niektoré články (5). Článok bez citácie má hodnotu FCR rovnú 0. Rovnako, ako RCR ani FCR nie je uvedený, keď doba od publikovania článku je kratšia ako 2 roky.

1.5.3 Vysoký počet citácií (HCI)

Ukazovateľ vysokého počtu citácií (Highly Cited Indicator, HCI) je odvodený of FCR. Porovnávaná je hodnota FCR hodnoteného článku s hodnotami FCR článkov v rovnakej vednej oblasti a veku ako hodnotený článok. Za publikácie s vysokým počtom citácií sú označené tie, ktoré sú na základe počtu citácií zaradené do horných 10 %, 5 % a 1%. Nakoľko vychádza z ukazovateľa FCR, je normalizovaným ukazovateľom podľa vednej oblasti, veku článku a roku publikovania. Články, ktorých vek je pod 2 roky nemusia mať uvedený indikátor vysokého počtu citácií. Ak je tento ukazovateľ už raz priradený, zostáva.

1.5.4 Nedávne citácie

Je to počet citácií, ktoré daný článok získal za obdobie posledných 2 rokov. Zaznamenáva sa pre publikácie staršie ako dva roky a indikuje, či v nedávnej dobe bola publikácia citovaná.

1.5.5 Priemerná citovanosť (Average Citation Rate, ACR)

Je nenormalizovanou metriku časopisu a vednej oblasti. Umožňuje porovnávať citovanosť časopisov na základe počtu citácií na články, ktoré sú v príslušnom časopise publikované a citovanosť vedných oblastí na základe počtu citácií na články, ktoré sú zaradené do príslušných vedných oblastí. Počíta sa ako pomer počtu citácií skupiny článkov publikovaných v určitom roku oproti počtu citácií, ktoré získali články publikované v predchádzajúcom roku. Uvádza sa pre obdobia 2, 3 a 5 rokov, čo dovoľuje jednoduché a relevantné porovnanie časopisov.

1.5.6 Altmetrické skóre (Altmetric Attention Score)

Je vážený počet, ktorý číselne vyjadruje pozornosť, akú články vyvolali na internete – v masmédiách, politických dokumentoch, blogoch, vo Wikipédii, voľne prístupných expertných posudkoch, v akademických sylabách a v sociálnych médiách vrátane Reddit, Twitter and Facebook. Ku skóre prispieva každý zo zdrojov a presné údaje je možné nájsť na *visiting the Altmetric Support page*. Príklad zobrazenia citačných metrik v databáze Dimensions je uvedený na obrázku 5.

2. Práca v databáze Dimensions

Vyhľadávanie prebieha štandardne ako v iných databázach, použitím kľúčového slova, s možnosťou ďalej spresňovať a cieľit' vyhľadávanie použitím filtrov na vymedzenie časového úseku, vednej oblasti, zdrojového dokumentu a autora. Na analytickom paneli sa zobrazujú vyhodnotené vyhľadane údaje a trendy. Pre každú publikáciu je k dispozícii informácia o počte citácií, bibliometrické skóre špecifické pre Dimensions a pozornosť, akou publikácia upútala v elektronických sociálnych sieťach, vyjadrenou alternatívnymi metrickými ukazovateľmi ako komentármi a zmienkami v blogoch, na Twitter-i, Facebook-u, vo Wikipédii, v Mendeley-i a v CiteULike. Úplné znenia článkov s voľným prístupom sú dostupné ako pdf dokumenty, ktoré je možné prevziať a uložiť do vlastného počítača alebo do knižnice (Readcube). Na stránke publikácie sú uvedené odkazy na použité literatúru, citujúce publikácie a hypertextové odkazy na granty a financujúcu inštitúciu, klinické štúdie a patenty, na ktoré môže byť publikácia naviazaná. Pri inštitucionálnom predplatnom je možné súčasne vstúpiť aj do týchto informačných zdrojov, a tak získať úplnejší prehľad o danej tematike. Detailne sú základné operácie v databáze Dimensions opísané v práci (6).

3. Detské choroby

Myšlienka vzájomne prepojiť z jedného miesta rozličné zdroje obsahujúce obrovské množstvo informácií o rôznych aspektoch výskumu a s nimi súvisiacich údajov je vizionárska, odvážna, ambiciózná a je spojená s rizikami. Jedno z nich predstavuje kvalita údajov, od ktorej sa odvíja využitie databázy (i na hodnotiace účely) a zároveň aj jej spoľahlivosť a dôveryhodnosť. Hoci je Dimensions vybudovaná na pokročilých a najnovších IK technológiách, indexácii textov, spracovaní prirodzeného jazyka, strojvom učením, umelej inteligencii, automatizácii procesov, bez prirodzenej inteligencie človeka a jeho zásahov sa neazobíde. Tvorcovia databázy si tento fakt veľmi dobre uvedomujú a pomerne silným hlasom pozývajú užívateľov na spoluprácu na jej vylepšení a skvalitnení dát. Nemožno nesúhlasiť aj s kritickými závermi, ku ktorým dospeli viacerí autori, ktorí túto databázu

skúmali pomerne krátko po ohlásení jej vstupu na trh v januári 2018 (7). Orduña-Malea a Delgado-López-Cózar (6), ktorí sa detailne zaoberali voľne prístupnou verziou Dimensions, vidia jej omedzenia v absencii filtrov pre krajinu a inštitúciu, v nemožnosti pri vyhľadávaní súčasne použiť viaceré filtre, nutnosti zadávať ich jednotlivo, vo výskyte duplicitných mien a v rôznych variantoch a citlivosti na diakritiku, v nekonzistentnosti pri indexácii článkov v časopisoch súvisiacou s automatickým spracovaním textu (ako článok boli nájdené, napr. zoznam redakčnej rady, posudzovateľov). Pri práci s databázou autor-ka tohto príspevku zaznamenala kuriózný prípad, keď do zoznamu publikácií istého autora boli nesprávne zaradené publikácie autorky s rovnakým menom ako dotyčný autor, avšak s krstným ženským menom a priezviskom s prechýľovacou príponou-ová (ako príklad uvediem fiktívne meno Daniel Žiak/Daniela Žiaková). Za závažnejšie, autori vyššie uvedenej publikácie, považujú zaradenie článkov a najmä monografií do klasifikačnej schémy FoR. V prípade monografií pravdepodobne len na základe abstraktov. Klasifikačná schéma FoR sa dobre uplatňuje pre časopisy v anglickom jazyku, nedá sa však aplikovať na časopisy v inom (v ich konkrétnom prípade, španielskom) jazyku. Určitá časť starších článkov nemá priradenú kategóriu FoR vôbec. Kriticky hodnotia chyby pri automatickej klasifikácii článkov, kde zaznamenali výrazné disproporcie, čo následne ovplyvňuje hodnoty metrických ukazovateľov, predovšetkým FCR a RCR. K podobnému zisteniu prišiel aj Bornmann (8), vývojový partner Dimensions, pri štúdiu zaradenia 191 svojich vlastných článkov do kategórií FoR. Väčšinu z nich pokladá za nesprávne zatriedené, no na druhej strane pripúšťa, že bol testovaný len malý súbor dát a v špecifickej oblasti. Tak, ako predošli autori zdôrazňuje význam správnej kategorizácie pre spoľahlivosť metrických na ňu naviazaných a považuje za potrebné preskúmať zatriedenie oveľa väčšieho počtu článkov a z rôznych vedných oblastí. Za silné stránky Dimensions možno považovať transparentnosť, s akými sú hneď ako prvé na užívateľskom rozhraní zobrazené dôležité údaje, aktualizované štatistiky a dáta. Pozitívne je vnímaná aj skutočnosť, že vývojári databázy berú do úvahy skúsenosti užívateľov s prácou v databáze a implementujú ich. Automatická aktivácia všetkých filtrov a údajov na analytickom paneli pri každom vyhľadávaní dovoľuje užívateľovi okamžite získať prehľad a rýchlo si o dátach vytvorí úsudok. Filtrácia pre otvorený prístup prebieha na úrovni článku, nie časopisu, čo umožňuje vyhľadanie aj takých článkov v režime voľného prístupu, ktoré sú publikované v hybridných časopisoch (v nich sú ako voľne prístupné a publikované len niektoré články). Ďalšie kladné zistenie je možné nájsť v porovnávej štúdií (9), ktorá vyšla len tri mesiace po oficiálnom predstavení platformy Dimensions. Snažila sa odpovedať na otázku, či Dimensions predstavuje pre Scopus a Web of Science konkurenciu. Zamerala sa na platformu Dimensions z aspektu článkov, ktoré obsahuje a možnosti využitia tejto novej databázovej platformy na hodnotenie publikačných výstupov vedy a ohlasov na ne. Na tento účel autor vybral články v oblasti potravinového výskumu z obdobia 2008 – 2018 a z náhodne vybranej vzorky 10 000 článkov (len typu article s DOI) v Scopus-e z roku 2012, pretože, ako tvrdí, odvolávajú sa na práce viacerých autorov: „Scopus konzistentne preukazuje najvyšší počet akademických časopisov, a teda predstavuje najlepšiu prax, čo sa týka pokrytia zdrojov vedeckých informácií“. Autor zistil, že medzi počtom citácií v Scopus-e a v Dimensions existuje silná korelácia (0,96 v úzkej vednej oblasti v roku 2012) s podobným počtom citácií a porovnateľnou hodnotou priemerného počtu citácií. Takmer všetky články v databáze Scopus boli prítomné aj v platforme Dimensions (97 % v roku 2012). Svoj výskum autor uzatvára konštatovaním, že platforma Dimensions sa ukazuje, ako možná alternatíva databáz Scopus a Web of Science na účely všeobecnej citačnej analýzy a využitia citačných údajov na niektoré typy hodnotenia výstupov výskumu. Autor víta vznik novej databázy, pretože: „Existencia alternatív prináša subjektom hodnotiacim vedy a výskum tri hlavné výhody: voľne prístupné alternatívy môžu znížiť náklady na hodnotenie a umožniť neformálne sebahodnotenie mnohým výskumníkom, ktorí by inak neplatili za sprístupnenie údajov; vzhľadom na to, že žiadny z citačných indexov nie je dokonalý, existencia alternatív umožňuje vzájomnú krížovú kontrolu údajov; určitý citačný index môže mať v porovnaní s inými prednosťami, napr. širšie zastúpenie vedeckej literatúry alebo iné prednosti, ktoré sú dôležité z hľadiska potrieb a úloh, ktorú má jednotlivé hodnotenie splniť.“ A to predstavuje ďalšie z pozitív databázy Dimensions.

4. Dimensions a údaje za Slovensko

Začiatkom júna 2018 boli v Dimensions Plus prístupné údaje o 83 099 publikáciách s autorstvom v Slovenskej republike, z ktorých bolo takmer 22 % voľne prístupných. Tieto údaje sú porovnateľné s údajmi zistenými v databáze Web of Science Core Collection. Najviac boli zastúpené publikácie z vedných oblastí podľa klasifikačnej schémy FoR a časopisy, ako je uvedené v tabuľke 1. Z publikácií prevládali články, ktorých bolo 69 624 (84 %), po nich nasledovali konferenčné príspevky a kapitoly so zastúpením po asi 8 % a nepatrné percento tvorili preprinty a monografie. Dimensions Plus obsahovala 890 patentov, 62 klinických štúdií, ktoré začali od rokov 2000 až 2018 (49 pochádzalo z databázy ClinicalTrials.gov v USA) a 11 102 grantov. Z nich, napríklad SAV mala uvedených 2 820 grantov s počtom viac ako 200 v oblastiach materiálového výskumu, biochémie a bunkovej biológie, celkovú hodnotu grantov vyčíslených na takmer 380 mil. eur a priemernú hodnotu 136 tisíc eur pre granty začínajúce a prebiehajúce v rokoch 2008 – 2018. Podľa údajov v databáze Dimensions SAV bolo priradených 55 patentov a účasť na 10 klinických štúdiách. Je potrebné dodať, že databáza je neustále dopĺňaná o nové údaje a súčasný stav bude pravdepodobne výrazne vyšší.

Pozvanie namiesto záveru

Bolo by veľkou škodou, ak by slovenská akademická komunita neprijala pomocnú ruku, ktorú jej spoločnosť Digital Science podáva vo forme voľného prístupu do bibliografickej a citačnej databázy Dimensions.

Publication metrics

[About](#)

Dimensions Badge



61 Total citations
50 Recent citations

47 Field Citation Ratio
4.99 Relative Citation Ratio

Altmetric



News (13)
Blogs (14)
Twitter (485)
Peer reviews (1)
Facebook (7)
Google+ (3)
F1000 (1)
Mendeley (276)
CiteULike (5)

Research Categories

Fields of Research

1503 Business and Management

External sources

- Full text at publisher site
- Abstract at PubMed
- Full text at PMC

Obrázok 5. Súhrnná informácia o metrikách publikácie v databáze Dimensions

Vedná oblasť podľa klasifikačnej schémy FoR	počet publikácií	Názov časopisu	počet článkov
Materiálové inžinierstvo	6324	Bratislava Medical Journal	1487
Fyzikálna chémia (vrátane štrukt.)	5592	Cheminform	795
Klinické vedy	5641	Proceedia Engineering	746
Biochémia a bunková biológia	4766	Czechoslovak Journal of Physics	698
Umelá inteligencia a spracovanie obrazu	3116	Folia Microbiologica	680
Ostatné fyzikálne vedy	2916	Acta Physica Polonica A	625
Genetika	2728	Čechoslovakja Filozofija	562
Verejné zdravie a starostlivosť	2442	General Physiology and Biophysics	534
Kardiorespiračné lekárstvo a hematológia	2094	Physics Letters B	521
Čistá matematika	2014	Neoplasma	494

Tabuľka 1. Zastúpenie vedných odborov podľa klasifikačnej schémy FoR a časopisov v databáze Dimensions pre Slovensko

Literatúra:

- (1) Bode Christian, Herzog Christian, Hook Daniel, McGrath Robert. 2018. Dimensions Report. A Guide to the Dimensions Data Approach. Digital Science. 2018, 24 s.
- (2) Definície najdôležitejších pojmov, používaných pri klinickom skúšaní liekov. [cit. 31-8-2018]. Dostupné na: https://www.sukl.sk/buxus/docs/Klinicke_skusanie_liekov/Definicie_najdolezitejsich_pojmov_KS_1_8_2016.pdf
- (3) 1 2 9 7 . 0 Australian and New Zealand Standard Research Classification (ANZSRC). 2 0 0 8. [cit. 31-8-2018]. Dostupné na: [http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/2A3A6DB3F4180D03CA25741A000E25F3/\\$File/12970_2008.pdf](http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/2A3A6DB3F4180D03CA25741A000E25F3/$File/12970_2008.pdf)
- (4) Hutchins Ian B., Yuan Xin, Anderson James M., Santangelo George M. 2016. [online]. Relative Citation Ratio (RCR): A New Metric That Uses Citation Rates to Measure Influence at the Article Level. In PLoS Biology, vol.1, No. 9, e1002541, 2016. [cit. 31-8-2018]. Dostupné na: <https://doi.org/10.1371/journal.pbio.1002541>
- (5) Thelwall Mike. 2016. [online]. The precision of the arithmetic mean, geometric mean and percentiles for citation data: An experimental simulation modelling approach journal of Informetrics, vol. 10, No. 1, Feb. 2016, s. 110-123. [cit. 31-8-2018]. Dostupné na: <https://doi.org/10.1016/j.joi.2015.12.001>
- (6) Orduña-Malea, E., Delgado-López-Cózar, E. 2018. [online]. Dimensions: re-discovering the ecosystem of scientific information. El Profesional de la Información, vol. 27, No. 2, s. 420-431. [cit. 31-8-2018]. Dostupné na: <https://doi.org/10.3145/epi.2018.mar.21>
- (7) Schonfeld, Roger C. 2018. [online]. A New Citation Database Launches Today: Digital Science's Dimensions. Scholarly Kitchen. Jan 15, 2018. [cit. 31-8-2018]. Dostupné na: <https://scholarlykitchen.sspnet.org/2018/01/15/new-citation-database-dimensions/>
- (8) Bornmann, Lutz. 2018. [online]. Field classification of publications in Dimensions: a first case study testing its reliability and validity. Scientometrics, vol. 11, No. 1, s. 637-640. [cit. 31-8-2018]. Dostupné na: DOI: 10.1007/s11192-018-2855-y
- (9) Thelwall Mike. 2018. [online]. Dimensions: A competitor to Scopus and Web of Science? Journal of Informetrics, vol. 12, No. 2, 2018, s. 430-435. [cit. 31-8-2018]. Dostupné na: <https://doi.org/10.1016/j.joi.2018.03.006>

PodĎakovanie

Článok bol pripravený v rámci implementácie národného projektu CVTI SR Informačný systém výskumu a vývoja/prístupy do databáz pre potreby výskumných inštitúcií (NISPEZ IV), kód ITMS Projektu: 313011I407, spolufinancovaného z Európskeho fondu regionálneho rozvoja (EFRR/ERDF) v rámci Operačného programu Výskum a inovácie. Obrazový materiál je publikovaný s povolením spol. Digital Science.

Ing. Adriana Shearman, CSC.

(Centrum vedecko-technických informácií SR) ■