

УДК 01:001.8

С.Г. БУБЛИК, кандидат технічних наук,
заступник завідувача відділу,
ДУ «Інститут досліджень науково-технічного потенціалу
та історії науки ім. Г.М. Доброва НАН України»,
бульвар Тараса Шевченка, 60, Київ, 01032, Україна,
<https://orcid.org/0000-0002-8463-9981>,
e-mail: boublyk@gmail.com

ТЕМАТИЧНЕ ЯДРО ПУБЛІКАЦІЙ ДОСЛІДНИКА: НАУКОМЕТРИЧНИЙ ПІДХІД

Бібліометричні профілі дослідників формуються з різних за тематичною спрямованістю наукових публікацій. Найбільш впливові із публікацій утворюють тематичне ядро бібліометричного профілю дослідника. Актуальним завданням наукометрії повинен стати не лише моніторинг публікаційної активності певного дослідника чи визначення його h-індексу, а й дослідження тематичного ядра його публікацій. Мета статті — провести наукометричне оцінювання тематичного ядра публікацій дослідника. Для досягнення поставленої мети проведено модифікацію методів лінгвостатистичного аналізу тематичної концентрації слів у тексті та h-індексу для обґрунтування визначення тематичного ядра публікацій. Для формування профілю дослідника (румунського фізика-лінгвіста Іоана-Йовітца Попеску) використовувалися дані бібліометричної бази Google Scholar. Їх використання пояснюється простою доступу до бібліометричних профілів; можливістю охоплення матеріалів конференцій та неангломовних публікацій, а також проведення ретроспективного аналізу динаміки посилань на публікації та h-індексу.

Поява публікацій з різних напрямів досліджень у тематичному ядрі поточного бібліометричного профілю дослідника є не випадковою подією, а наслідком певних закономірних процесів формування тематичних ядер за окремими тематичними напрямками (дисциплінарними, галузевими чи міждисциплінарними). Встановлення переліку всіх основних тематичних публікацій розширює можливості для об'єк-

© БУБЛИК С.Г.,
2019

¹ Стаття написана за матеріалами доповіді на Міжнародній конференції з наукометрії та бібліометрії: Бублик С.Г., Велентейчик Т.М. Наукометричне оцінювання тематичного ядра публікацій дослідника. Міжнародна конференція з наукометрії та бібліометрії (Київ, 4 червня 2019 р.). URL: <https://drive.google.com/file/d/1LDyhk-inNNcA1CD7TfoPcOEaVQ49a4KU/view?usp=sharing>.

тивного оцінювання творчого різноманіття дослідницької діяльності ученого. Показано, що тематичне ядро бібліометричного профілю дослідника презентує всі основні тематичні напрями його досліджень. Використання тематичного ядра публікацій як методичного інструменту збільшує можливості бібліометричного та наукометричного аналізу публікаційної діяльності дослідника. Отримані результати свідчать, що більш важливим в оцінюванні ефективності дослідницької діяльності може стати визначення переліку публікацій, які складають тематичне ядро бібліографічного профілю дослідника, ніж його h-індексу.

Ключові слова: наукометрія, h-індекс, тематичне ядро публікацій, бібліометричний профіль дослідника.

Вступ. Основним лейтмотивом життя кожного дослідника є процес здобуття нових знань. Результати дослідницької діяльності в залежності від схильностей дослідника та обраної царини наукових досліджень трансформуються у різновиди наукових продуктів — наукові публікації (статті, монографії, наукові реферати, тези доповідей та інше), патенти, промислові зразки. Тематичне розмаїття наукових публікацій визначається колом наукових інтересів кожного дослідника. Сфера наукових інтересів дослідника протягом його наукової кар'єри має природну тенденцію до зростання: через збільшення предметного поля у межах певної наукової галузі або через зміну первинної наукової спеціалізації. Отже, з часом бібліометричні портрети (профілі) дослідників складатимуться із різних за тематичною спрямованістю наукових публікацій.

В одній частині наукових публікацій дослідники розвивають вже існуючі тематичні сфери власних досліджень, удосконалюючи дослідницькі методики або отримуючи нові результати на їхній базі. Але інша їх частина може стати проривною в певному науковому напрямі чи навіть галузі. Публікації як першої, так і останньої групи формують унікальне тематичне поле результатів досліджень, а найбільш впливові з них — тематичне ядро.

Інтуїтивно можна зрозуміти, що вміст тематичного ядра публікацій дослідника повинен віддзеркалювати найбільш ефективні (з точки зору впливу на обрані наукові напрями) його публікації. Фактично ж тематичне ядро публікацій дослідника є відображенням успішності та ефективності отриманих ним результатів досліджень за певними тематичними напрямками (галузями, дисциплінами), тобто його унікальним науковим «автопортретом».

На думку наукознавців, навіть «створення і дослідження персональних бібліографій наукових праць вчених, у тому числі розробка методології та методичного інструментарію таких досліджень, мають актуальне теоретичне і прикладне значення» [1, с. 106]. Дослідження індивідуальних тематичних ядер публікацій дослідників може надати можливості отримання якісної оцінки їх наукової успішності з точки зору різноманіття їх наукових напрямів.

Нагальна необхідність визначення тематичного ядра публікацій як одного з наукометричних критеріїв ефективності дослідника полягає у тому, що за сприятливих умов діяльності кожен дослідник прагне більше розвиватися в обраних тематичних напрямках, розкрити свої творчі таланти. За несприятливих умов діяльності, які домінують в українській науці протягом останніх десятиліть, нерідкими є випадки наукових публікацій, коли «автори <...> не намагаються довести новизну своїх пропозицій у порівнянні із наявними у публікаціях інших вчених» [2, с. 8].

Дослідження тематичного ядра публікацій може надати додаткову можливість для розв'язання однієї «з центральних проблем в оцінці науково-дослідної діяльності та діяльності на індивідуальному рівні», якою є «взаємозв'язок продуктивності публікацій та якості наукових публікацій...» [3, с. 55]. Відзначимо, що показниками якості публікацій вважаються **популярність та престиж публікації** (виділено автором) [4], тобто ті ознаки, які властиві їх тематичному ядру.

Актуальним завданням наукометрії повинен стати не лише моніторинг публікаційної активності певного дослідника чи визначення його h-індексу (індексу Гірша), а й дослідження тематичного ядра публікацій.

Аналіз останніх досліджень і публікацій. Завдання щодо визначення тематичного ядра публікацій будь-якого із дослідників не поставало раніше через обмежені можливості кількісних методів аналізу та оцінювання праці науковця, оскільки ані загальна кількість публікацій, ані загальна кількість цитувань (посилань на них) не можуть служити єдиним критерієм ефективності наукової діяльності [5, 6]. Так, «на цитування роботи впливають тематична спрямованість наукового дослідження <...> вибір журналу для розміщення публікації» [7, с. 84]. Окрім того, практика цитування в одній галузі досліджень відрізняється від практики в іншій та навіть між галузями досліджень у межах дисципліни [8].

Відповідно, використання будь-якого з цих методів (за кількістю публікацій чи посилань) надаватиме лише кумулятивну оцінку структури публікацій дослідника за різними тематичними напрямками, без можливості виділення тематичного ядра.

Найбільш апробованим методом наукометричного аналізу публікацій, який інтегрує показники продуктивності та популярності публікацій дослідника, є h-індекс [9]. Хоча використання методу h-індексу як основного наукометричного інструменту й викликає багато суперечок у науковій спільноті через наявні обмеження щодо визначення актуального поля публікацій дослідників, насамперед щодо їх наукової важливості [10–13].

Дійсно, ядро Гірша (h-core) складається із найбільш популярних публікацій з кількістю посилань, не меншою за індекс Гірша, та ігнорує фактичну кількість публікацій, які мають меншу кількість посилань на них. Втім перевагами використання h-індексу над іншими наукометричними методами є [14]:

- поєднання як продуктивності, так й впливу публікацій в єдиному вимірі, який є легкодоступним і дуже інтуїтивно зрозумілим;
- його легко обчислювати: для цього потрібні лише дані про кількість посилань, взяті з WoS, Scopus або Google Scholar;
- застосовність на різних рівнях — дослідник, журнал чи відділ;
- об'єктивність і порівнянність на рівні дисциплін, коли між показниками немає значної різниці;
- надійність — на нього не впливають малопримітні (найменш популярні) публікації, оскільки він ігнорує їх через можливі проблеми з обчисленням; це має особливо важливе значення в разі використання Google Scholar.

Прогностична властивість методу h -індексу, яка полягає у визначенні важливих за майбутнім впливом публікацій через ретроспективний аналіз кількості посилань [15], надає підстави стверджувати про можливість визначення тематичного ядра поза межами «ядра Гірша». Така можливість неявно була доведена й у роботі Сіднея Реднера. На репрезентативній вибірці, яка включала 255 вчених-теоретиків у галузі конденсованої речовини та статистичної фізики з Північної Америки та Європи, він визначив, що розмір подвійного ядра Гірша співвідноситься із квадратним коренем загальної кількості посилань на усі публікації дослідника майже як один до одного [16]. Тим самим він припустив, що подвійне ядро Гірша повинно забезпечувати таку ж інтегровану міру ширини та глибини запису цитування особи, як і сам h -індекс.

У 2007 році румунським фізиком-лінгвістом Іоаном-Йовітцем Попеску спільно з лінгвістами Карлом-Гайнцем Бестом та Габріелем Альтманном було розроблено концепцію тематичної концентрації слів у тексті на підставі використання точки h (аналога h -індексу) у рангово-частотній закономірності Ципфа [17]. Надалі ця концепція була розвинена до методу аналізу тематичної концентрації значущих (ключових) слів у тексті з подальшою успішною апробацією [18—21]. Зокрема, у роботі чеських лінгвістів здійснено модифікацію вимірювання тематичної концентрації ключових слів через впровадження поняття «вторинна тематична концентрація», що збільшує можливості для аналізу тематичних характеристик тексту [20]. Згідно з цим підходом, імовірність розташування тематичних слів у рангово-частотній закономірності Ципфа з рангом до $2h$ є значно вищою, ніж у ядрі Гірша.

Закономірності розподілу слів у тексті та ранговий розподіл статей за кількістю посилань на них підпорядковуються спорідненим за математичною сутністю степеневим законам, що можна розглядати як вияв стійкого закону розподілу ймовірностей. Подібність статистичних рангово-частотних розподілів слів у тексті (закон Ципфа) та публікаційних посилань (h -розподіл) є виявами одних і тих самих закономірностей (Ципф, Бредфорд) [14].

Виходячи зі спільності статистичних закономірностей рангово-частотних розподілів як слів у тексті, так і впливовості публікацій за частотою по-

силань на них, пропонується використати методику визначення тематичної концентрації значущих (ключових) слів для визначення тематичного ядра публікацій дослідника.

Метою дослідження є наукометричне оцінювання тематичного ядра публікацій дослідника. Основними завданнями є методологічне визначення тематичного ядра публікацій та його апробація на бібліометричному профілі дослідника.

Виклад основного матеріалу дослідження і отриманих результатів. Тематика досліджень багатьох учених змінюється з плином часу внаслідок проведення міждисциплінарних досліджень або через зміну наукової дисципліни/галузі. В результаті бібліометричний профіль дослідника складатиметься із публікацій за різними тематичними напрямками. Наукометричне визначення впливових публікацій за різними тематичними напрямками досліджень вченого пропонується здійснювати за допомогою методу h -індексу, модифікованого для завдань квантитативної лінгвістики (лінгвостатистики).

Згідно з методом тематичної концентрації [20], публікації, які входять до ядра Гірша (рангом від 1 до h), вважатимемо **основними тематичними** публікаціями, з рангом від $h + 1$ до $2h$ — **вторинними тематичними** публікаціями. Отже, усі публікації, які мають ранг від 1 до $2h$ (статистично дорівнює подвійному ядру Гірша), утворюють тематичне ядро публікацій дослідника.

Для статистичної перевірки такого припущення проведемо наукометричне дослідження бібліометричного профілю румунського фізика-лінгвіста Іоана Йовітца Попеску (Ioan-Iovitz Popescu) [22], який з 2006 року почав займатися лінгвістикою (змінивши не лише дисциплінарний, а й галузевий напрям досліджень) і став відомим ученим у сфері квантитативної лінгвістики [23].

Використання бібліометричної бази Google Scholar як бази даних для формування профілю дослідника обумовлюється як простотою доступу до бібліометричних профілів, охопленням матеріалів конференцій і неангломовних публікацій, так і можливістю проведення ретроспективного аналізу динаміки посилань та індексу Гірша [24, 25].

За базовий бібліометричний профіль дослідника обираємо останній завершений рік (на момент написання статті — 2018). Для зручності статистичного оброблення даних позначаємо усі цитовані публікації аббревіатурою POP, з нумерацією відповідно до рангу 2018 року (фільтр Google Scholar, посилання): від публікації з найбільшим посиланням, *The quantum effects of electromagnetic fluxes* (POP001), до публікації з найменшим посиланням, *On basic physical laws* (POP222).

Тематична диференціація статистичного масиву публікацій здійснювалася з використанням наукометричних методів — сленгового (для відбору за контекстом) та бібліометричного (за виданнями), для уточнення сленгового. Результати сленгового аналізу публікацій за лінгвістичним тематичним

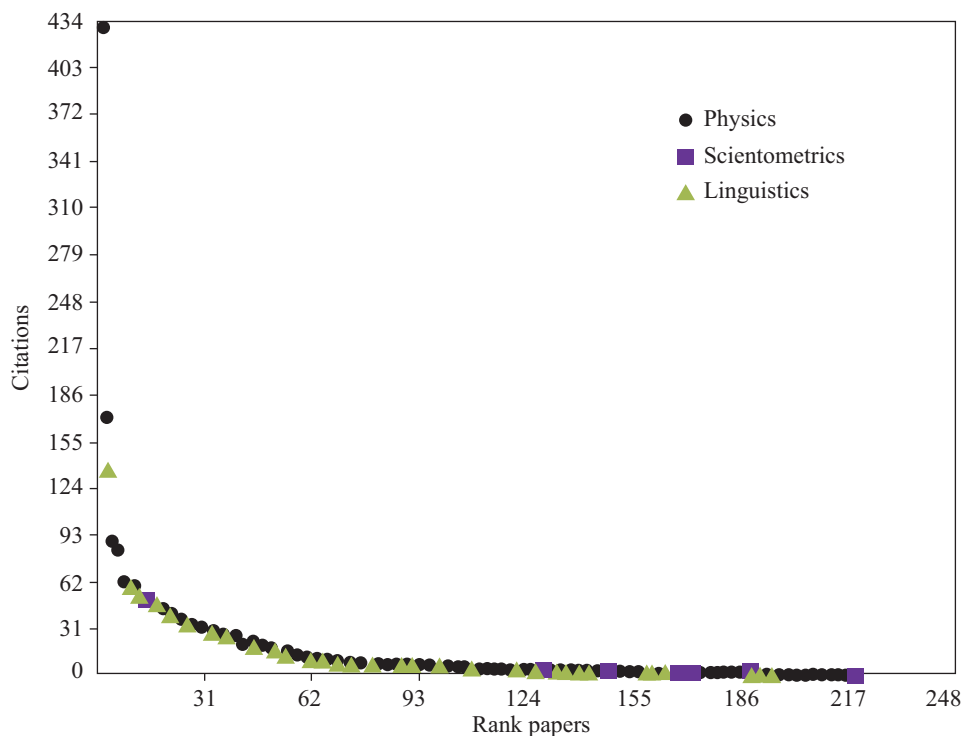


Рис. 1. Бібліометричний профіль Попеску за 2018 рік ($h = 31$), за тематичними напрямками публікацій

Джерело: розраховано та побудовано автором за [22].

напрямом у бібліометричному профілі Google Scholar окремо звірялися із актуальним переліком лінгвістичних публікацій Попеску [26].

Сленговий аналіз публікацій із бібліометричного профілю Попеску [22] дозволив виділити три основні напрями його наукових досліджень: з фізики (Physics), лінгвістики (Linguistics), а також наукометрії (Scientometrics). Для спрощення дослідження вважатимемо усі публікації за будь-яким тематичним напрямом такими, що розкривають одну тематику (галузева тематичність).

Із первинного переліку з 222 публікацій (2018 рік) переважна частина (154) стосувалася фізичного напрямку, 58 — лінгвістики, 10 — наукометрії. Для наочності рангового розташування публікацій за різними тематичними напрямками наведемо бібліометричний профіль Попеску за 2018 рік (рис. 1).

Для зручності аналізу графічних даних бібліографічного профілю горизонтальна вісь координат на рис. 1 градуйована як кратна до індексу Гірша. Аналіз графічних даних свідчить, що до ядра Гірша (основні тематичні публікації) належала 31 публікація, з яких одна — з наукометричного напрямку, шість — лінгвістичного та 24 — з фізичного. До вторинних тематичних пуб-

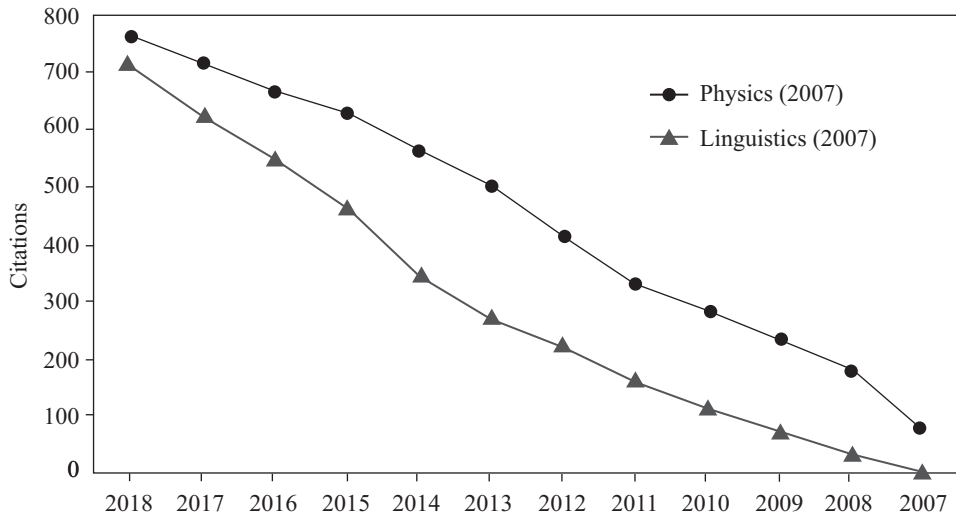


Рис. 2. Кумулятивна кількість посилань на публікації за окремими тематичними напрямками, з 2007 року

Джерело: розраховано та побудовано автором за [22].

лікацій належали 32 публікації: одна — з наукометричного напрямку, дев'ять — лінгвістичного та 22 — фізичного.

Збільшення кількості публікацій вторинних тематичних напрямків пов'язане з тим, що не всі публікації з однаковою частотою посилань за рангом входять до математично визначеної нижньої межі тематичного ядра. Оскільки факт виникнення нових посилань є певною мірою випадковою подією, то досі не знайдено точного механізму включення/вилучення тієї чи іншої публікації, кількість посилань на яку дорівнює h -індексу (або подвійному його значенню), але за рангом перевищує його. Тому у ході дослідження було прийнято, що будь-яке **перевищення кількості публікацій**, які за кількістю посилань дорівнюють h або $2h$, але за рангом виходять за межі ядра Гірша (подвійного ядра Гірша), є некритичним щодо **визначення їх належності** до основних чи вторинних тематичних напрямків.

Для з'ясування закономірностей формування тематичного ядра публікацій дослідника здійснимо ретроспективні дослідження публікацій та посилань на них за окремими тематичними напрямками і бібліометричного профілю в цілому.

Початковим відліком для ретроспективного аналізу публікацій Попеску вважатимемо 2006 рік. У цей рік ним була оприлюднена перша публікація з лінгвістики, а також остання з наукометрії. Здається, що ці події були взаємопов'язані, оскільки саме Попеску увів точку h в лінгвістиці [27], вірячи у величезні можливості використання h -індексу в наукометрії [19]. Відтоді Попеску захопився прикладним застосуванням наукометричних законів, зокрема закону Ципфа, у галузі мовознавства, та увійшов до міжна-

родної школи лінгвістів, яка включає вчених Німеччини, Польщі, Чехії, Словаччини, України та інших країн. Водночас напрям досліджень з фізики втрачає для нього актуальність: остання його публікація з питань квантового генератора когерентного гама-опромінювання (*gamma ray lasers*) датується 2009 роком [23]. Протягом 2007—2017 років Попеску опублікував 65 наукових праць, на які були здійснені посилання до кінця 2018 року. Із них лише 9 (1 — одноосібно) стосувалися наукових проблем у галузі фізики, решта (56) були присвячені результатам досліджень у галузі квантитативної лінгвістики [22].

Окремим підтвердженням того, що лінгвістичний напрям досліджень з 2006 року стає для Попеску більш пріоритетним, ніж фізичний, є результати аналізу динаміки кількості посилань на його публікації за різними тематичними напрямками з 2007 року (рис. 2). Так, протягом 2007—2018 рр. його лінгвістичні публікації мали 718 посилань проти 765 посилань з фізики, але за динамікою накопичення вони вже переважали фізику.

Розглянемо еволюцію формування тематичного ядра публікацій бібліометричного профілю дослідника з 2006 року. Для більшої ефективності аналізу великого масиву даних зробимо вибірку за певними роками: 2006 (рік першої публікації з лінгвістики), 2008 (двохрічний проміжок часу можливої появи посилання на публікації 2006 року), 2011 (п'ятирічний проміжок), 2015 (десятирічний проміжок), а також 2018 (останній повний рік формування бібліометричного профілю). Впровадження таких часових інтервалів пов'язане із існуванням популярних журнальних метрик, коли двохрічний проміжок часу відповідає умовам обчислення коефіцієнта впливовості академічного видання (Thomson Reuters) [28], п'ятирічний — умовам чинника

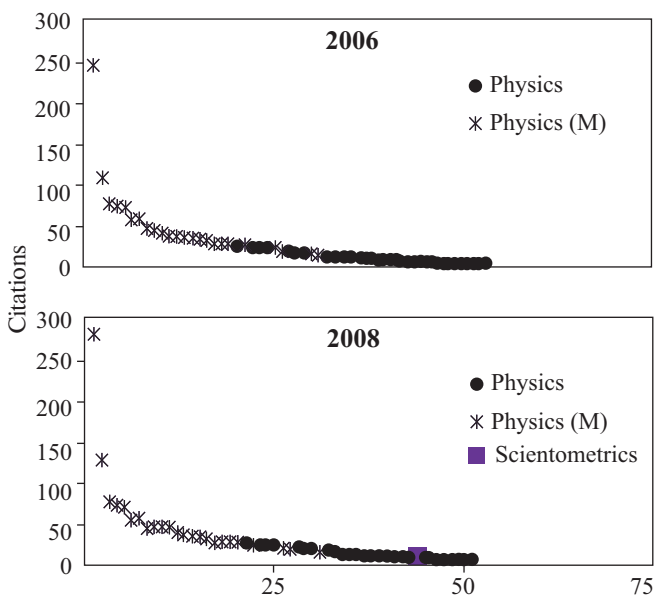


Рис. 3. Тематичне ядро публікацій Попеску за 2006 та 2008 роки. З-поміж публікацій з фізики (Physics) виокремлено основні (main) тематичні публікації тематичного ядра 2018 року, Physics (M)

Джерело: розраховано та побудовано автором.

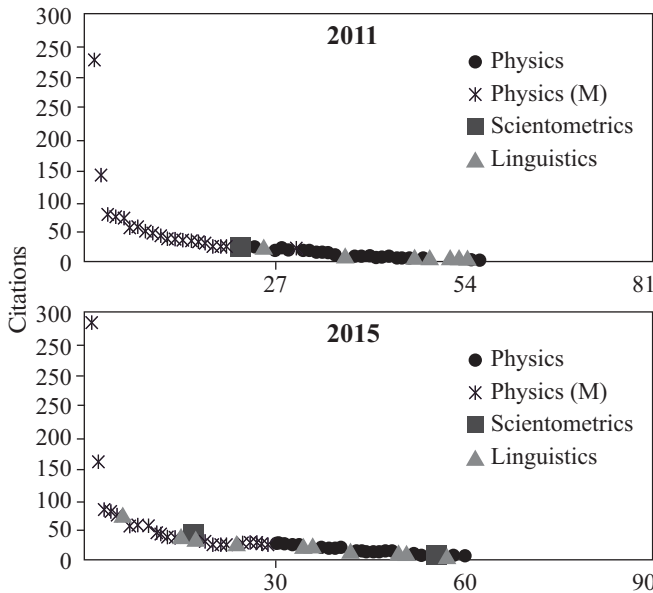


Рис. 4. Тематичне ядро публікацій Попеску за 2011 та 2015 роки. З-поміж публікацій з фізики (Physics) виокремлено основні (main) тематичні публікації тематичного ядра 2018 року, Physics (M)
Джерело: розраховано та побудовано автором.

впливу JCR (Journal Citation Reports) [29]. Вибір двох- і п'ятирічного часового інтервалу для аналізу даних є довільним.

Аналіз графічних даних (рис. 3) свідчить, що у 2006 році тематичне ядро складалося лише з публікацій з фізики, а частина майбутніх основних публікацій перебувала у вторинній тематичній області. У 2008 році до тематичного ядра як вторинної тематичності увійшла публікація з наукометрії від 2003 року.

У 2011 році до тематичного ядра додалася низка публікацій з лінгвістики: одна від 2009 року як основної тематичності, а також шість як вторинної тематичності, видання 2006—2010 років (рис. 4). Одна публікація з наукометрії від 2003 року увійшла до первинної області тематичного ядра, друга від 1997 року — до вторинної області. У 2015 році визначилися усі основні публікації з фізики (24) підсумком 2018 року, останні з яких були опубліковані у 2002 році. Кількість публікацій з лінгвістики, які стали основними у бібліометричному профілі, зростає до чотирьох, 2006, 2009 (2) та 2010 років видання. Збільшилася до 10 загальна кількість публікацій з лінгвістики, які увійшли до тематичного ядра внаслідок появи нових публікацій, 2008, 2009 та 2011 років видання.

Отже, аналіз еволюції формування тематичного ядра публікацій бібліометричного профілю дослідника протягом 2008—2015 років підтвердив прийняті умови щодо визначення перспективності публікацій (як для визначення імпаکت-фактора журналу), а саме:

- у 2008 році до вторинної області увійшла публікація 2003 року (6 років);
- у 2011 році до первинної області увійшла публікація 2009 року (2 роки), до вторинної — дві публікації 2010 року (1 рік!), по одній — 2009 (2 роки), 2008 (3 роки), 2007 (4 роки), 2006 (5 років);

• у 2015 році до первинної області увійшли публікації 2010 року (5 років), до вторинної — 2011 року (4 роки).

Здається, що поява тих чи інших публікацій з різних напрямів досліджень у тематичному ядрі поточного бібліометричного профілю дослідника є не *випадковою подією*, а наслідком певних закономірних процесів формування тематичних ядер за окремими тематичними напрямами (дисциплінарними, галузевими чи міждисциплінарними). На підтвердження цього припущення проведемо аналіз еволюції бібліометричного профіля дослідника за лінгвістичним та наукометричним напрямами його досліджень (табл. 1).

У щойно сформованому у 2008 році ядрі Гірша за лінгвістичним тематичним напрямом вже виділилися впливові тематичні публікації, але вони не увійшли до поточного тематичного ядра бібліометричного профілю дослідника через незначну кількість посилань на них. Причин тому могло бути декілька, основними з яких є брак часу для популяризації публікацій та незначний розмір ядра Гірша ($h_T = 3$). Натомість усі публікації з лінгвістики, які склали відповідне ядро Гірша у 2011, 2015 та 2018 роках, містилися у поточному тематичному ядрі бібліометричного профілю дослідника відповідних років.

Приклад еволюції ядра Гірша за наукометричним тематичним напрямом (табл. 2) є цікавим і з точки зору визначення ефективності дослідниць-

Таблиця 1. Еволюція ядра Гірша за лінгвістичним тематичним напрямом

Rank	$h_T = 13$ (2018)		$h_T = 10$ (2015)		$h_T = 7$ (2011)		$h_T = 3$ (2008)	
	P	C	P	C	Papers	C	P	C
1	POP003	134	POP003	78	POP003	27	POP003	4
2	POP011	55	POP012	43	POP012	15	POP021	4
3	POP012	52	POP011	39	POP017	12	POP017	3
4	POP017	47	POP017	32	POP026	10	<i>POP063</i>	3
5	POP021	40	POP021	27	POP011	9		
6	POP026	35	POP026	26	POP021	9		
7	<i>POP033</i>	29	<i>POP039</i>	20	<i>POP039</i>	9		
8	<i>POP039</i>	26	<i>POP045</i>	15				
9	<i>POP045</i>	21	<i>POP033</i>	14				
10	<i>POP051</i>	17	<i>POP056</i>	11				
11	<i>POP052</i>	16						
12	<i>POP054</i>	15						
13	<i>POP056</i>	14						

P (papers) — публікації; C (citations) — посилання; h_T — тематичний h-індекс.

Примітка: У 2006 році $h_T = 0$. Публікації, які у 2018 році входять до ядра Гірша, позначені напівжирним шрифтом; публікації, які входять до вторинної тематичної області, позначені курсивом

Джерело: сформовано автором.

Таблиця 2. Еволюція ядра Гірша за наукометричним тематичним напрямом

Rank	$h_T = 3$ (2018)		$h_T = 3$ (2015)		$h_T = 3$ (2011)		$h_T = 3$ (2008)		$h_T = 3$ (2006)	
	P	C	P	C	P	C	P	C	P	C
1	POP014	50	POP014	43	POP014	29	POP014	10	POP014	5
2	<i>POP055</i>	15	<i>POP055</i>	11	<i>POP055</i>	10	<i>POP055</i>	6	<i>POP055</i>	5
3	POP129	4	POP129	4	POP129	4	POP129	4	POP129	4

P (papers) — публікації; C (citations) — посилання; h_T — тематичний h-індекс.

Примітка: Публікації, які у 2018 році входять до ядра Гірша, позначені напівжирним шрифтом; публікації, які входять до вторинної тематичної області, позначені курсивом.

Джерело: сформовано автором.

кої діяльності за абсолютним показником h-індексу, тобто за якого значення h-індексу можна вважати сформованим тематичне ядро публікацій дослідника. Навіть за незначного показника h-індексу ($h_T = 3$) тематичного напрямку неосновних досліджень дві найпопулярніші публікації входять до тематичного ядра бібліографічного профілю дослідника. Цей факт свідчить про більшу важливість тематичного ядра для визначення ефективності дослідницької діяльності, ніж показника h-індексу. Водночас встановлення переліку всіх основних тематичних публікацій надає можливості для об'єктивного оцінювання творчого різноманіття наукової діяльності дослідника та його талантів.

Отже, підсумком 2018 року тематичне ядро бібліометричного профілю Попеску складалося із 63 публікацій. Ядро Гірша бібліометричного профілю дослідника формувалося з публікацій, які є основними за тематичними напрямками його досліджень, — з фізики, лінгвістики та наукометрії. Вторинна область бібліометричного профілю складалася переважно з основних публікацій з тих самих напрямів досліджень (21) та частини вторинних (12) за окремими напрямками з фізики й лінгвістики. Оскільки розділення на окремі тематичні напрями робилося за галузевою ознакою (фізика, лінгвістика), вторинні у галузевому профілі публікації можуть виявитися як основні в дисциплінарному напрямі (напрямах). Так, відомо, що тематика публікацій з фізики складалася із трьох напрямів (Gas discharge and plasma physics, Laser spectroscopy, Gamma ray lasers), в яких проводив свої дослідження Попеску як фізик [23].

Окремо відзначимо, що за збереження нинішніх темпів нарощування посилань на публікації з лінгвістики та природного згасання посилань на публікації з фізики (рис. 2) у 2019 та наступних роках відбуватиметься поступове зрушення основного тематичного напрямку від фізики до лінгвістики (рис. 1, 4).

Висновки і перспективи подальших досліджень. Входження публікацій лінгвістичного чи наукометричного напрямів до поточного тематичного ядра бібліометричного профілю дослідника у 2008, 2011, 2015 та 2018 роках

були невинними подіями, оскільки такі публікації формували ядро Гірша за локальними напрямками. Припинення активних досліджень у наукометричному напрямі (останні результати були оприлюднені у 2006 році одночасно з кардинальною зміною основного дисциплінарного напрямку з фізики на лінгвістику) унеможливило представлення у тематичному ядрі бібліометричного профілю всіх основних публікацій з наукометрії (2 із 3).

Вдалих початок нових досліджень з квантитативної лінгвістики, які здійснювалися у міжнародному консорціумі лінгвістів, дозволив швидко проявитися новому, основному напрямку наукової діяльності Попеску, лінгвістиці. Внаслідок цього сталося входження публікацій із лінгвістики до первинної та вторинної областей тематичного ядра, переважно у двох- та п'ятирічний період часу.

Отже, тематичне ядро публікацій бібліометричного профілю дослідника є засобом, що дозволяє оцінити увесь спектр його основних тематичних публікацій, а також виділити найбільш перспективні з них. Використання тематичного ядра публікацій як методичного інструменту збільшує можливості бібліометричного та наукометричного аналізу публікаційної діяльності дослідника. Здається, що більш важливим в оцінюванні ефективності дослідницької діяльності може стати визначення переліку публікацій, які складають тематичне ядро бібліографічного профілю дослідника, а не його h-індексу. Також підтвердилася доцільність застосування двох- та п'ятирічних часових інтервалів для оцінювання перспективності тематичних публікацій.

В подальшому планується визначити мінімальний розмір тематичного ядра публікацій, що може стати критерієм оцінки перспективності дослідницької діяльності науковця та можливостей реалізації його творчих талантів.

СПИСОК ЛІТЕРАТУРИ

1. Рибачук В.П., Грачов О.О., Кухтенко Т.О., Віденіна Н.Г. До питання щодо визначення загальних та особливих бібліометричних характеристик наукової діяльності вчених. *Наука та наукознавство*. 2005. № 4. Додаток. С. 105—112.
2. Порев С.М. Між наукою і грою. Політика досліджень, університети і підприємницькі екосистеми: моногр. К.: «Малотиражка», 2018. 179 с.
3. Rybachuk V., Quist G. Classification of individual age-phase dynamics profiles of researchers' scientific activity. *Science and Science of science*. 2013. No 3. P. 55—62.
4. Мриглюд О.І. Від фізики до наукометрії: задачі аналізу складних систем (за матеріалами наукового повідомлення на засіданні Президії НАН України 29 березня 2017 року). *Вісн. НАН України*. 2017. № 7. С. 54—63. DOI: 10.15407/visn2017.07.054.
5. Налімов В.В., Мульченко З.М. Наукометрия. Изучение науки как информационного процесса. М.: Наука, 1969. 192 с.
6. Боргоякова К. С. Методика определения цитируемости научных публикаций исследователя. *Библиосфера*. 2017. № 1. С. 103—109. DOI: 10.20913/1815-3186-2017-1-103—109.
7. Чайковський Ю.Б., Сілкина Ю.В., Потоцька О.Ю. Наукометричні бази та їх кількісні показники (Частина II. Фактори, що впливають на кількісні показники наукометричних баз). *Вісн. НАН України*. 2013. № 9. С. 84—92.

8. Anauati V., Galiani S., Gálvez R.H. (2016). Quantifying the Life Cycle of Scholarly Articles Across Fields Of Economic Research. *Economic Inquiry, Western Economic Association International*, 2016. No 54(2). P. 1339—1355. April. DOI: 10.1111/ecin.12292.
9. Hirsch. J. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. U.S.A.* 2005. No 102. P. 16569—16572.
10. Sidiropoulos A., Katsaros D., Manolopoulos Y. Generalized Hirsch h-index for disclosing latent facts in citation networks. *Scientometrics*. 2007. No 72(2). P. 253—280. DOI: 10.1007/s11192-007-1722-z.
11. Гуськов А.Е. Российская наукометрия: обзор исследований. *Библиосфера*. 2015. № 3. С. 75—86.
12. Порев С.М., Сандига І.В. Шлях науки університету: моногр. Київ: Хімджест, 2016. 112 с.
13. Lande D.V., Andrushchenko V.B., Balagura I.V. Wiki-index of authors popularity. *arXiv preprint arXiv*. 2017. 1702.04614.
14. Mingers J., Leydesdorff L. A review of theory and practice in scientometrics. *European Journal of Operational Research*. 2015. No 246(1). P. 1—19.
15. Hirsch J.E. Does the h index have predictive power? *Proceedings of the National Academy of Sciences of the United States of America*. 2007. No 104(49). P. 19193—19198.
16. Redner S.J. On the meaning of the h-index (Letter). *Stat. Mech. Theory Exp*. 2010. 3. L03005. 1-6. DOI: 10.1088/1742-5468/2010/03/L03005.
17. Popescu I.-I., Best K.-H., Altmann G. On the dynamics of word classes in text. *Glottometrics*. 2007. No 14. P. 58—71.
18. Popescu I.-I., Altmann G. Thematic concentration in texts. *Issues in Quantitative Linguistics*. 2011. Vol. 2. E. Kelih, V. Levickij, Y. Matskulyak (Eds.). (pp. 110—116). Lüdenscheid: RAM-Verlag.
19. Čech R., Popescu I.-I., Altmann G. Methods of analysis of a thematic concentration of the text. *Czech and Slovak Linguistic Review*. 2013. No 3. P. 4—21.
20. Čech R., Garabik R., Altmann G. Testing the thematic concentration of text. *Journal of Quantitative Linguistics*. 2015. No 22(3). P. 215—232.
21. Ruina Chen, Haitao Liu. Thematic Concentration as a Discriminating Feature of Text Types. *Journal of Quantitative Linguistics*. 2017. No 25(1). P. 1—24. DOI: 10.1080/09296174.2017.1339441.
22. Popescu, Ioan-Iovitz. Google Scholar Citations (03.07.2019). Retrieved from https://scholar.google.com/citations?hl=ru&user=iqhVvzEAAAAJ&view_op=list_works.
23. Ioan-Iovitz Popescu. 80 years of life. Retrieved from <http://www.iipopescu.com/>.
24. Alonso S., Cabrerizo F.J., Herrera-Viedma E., Herrera F. h-index: A Review Focused in its Variants, Computation and Standardization for Different Scientific Fields. *Journal of Informetrics*. 2009. No 3. P. 273—289. DOI: 10.1016/j.joi.2009.04.001.
25. Назаровець С.А. Аналіз цитувань в епоху 2.0. Сучасні проблеми діяльності бібліотеки в умовах інформаційного суспільства: матер. III наук.-практ. конф. (Львів, 29 вер. 2011 р.). Львів, 2011. С. 202—208.
26. Ioan-Iovitz Popescu. Linguistic publications updated January 2019. Retrieved from <http://www.ram-verlag.eu/journals-e-journals/glottometrics/http://www.cechradek.cz/publ.html#Articles>.
27. Popescu I.-I., Altmann G. Some aspects of word frequencies. *Glottometrics*. 2006. No 13. P. 23—46.
28. Garfield E. Journal impact factor: a brief review. *Can. Med. Assoc. J.* 1999. No 161. P. 979—980.
29. Jacso P. Five-year impact factor data in the Journal Citation Reports. *Online Information Review*. 2009 (June). No 33(3). P. 603—614.

Одержано 02.09.2019

REFERENCES

1. Rybachuk, V.P, Grachov, O.A., Kukhtenko, T.A, Videnina, N.G. (2005). Defining the general and specific bibliometric characteristics of research activities of scientists. *Science and Science of Science*, 4 (Addendum), 105–112 [In Ukrainian].
2. Porev, S.M. (2018). *Between Science & Game. Research policy, universities and entrepreneurial ecosystems*. Kyiv: Malotirazhka, 179 [in Ukrainian].
3. Rybachuk, V. & Quist, G. (2013). Classification of individual age-phase dynamics profiles of researchers' scientific activity. *Science and Science of science*, 3, 55–62.
4. Mryglod, O.I. (2017). From physics to scientometrics: analysis of complex systems (by the materials of scientific report at the meeting of the Presidium of NAS of Ukraine March 29, 2017). *Bulletin of the NAS of Ukraine*, 7, 54–63 [in Ukrainian]. DOI: 10.15407/vsn2017.07.054.
5. Nalimov, V.V., Mulchenko, Z.M. (1969). *Scientometrics. Studying science as an information process*. Moscow: Nauka, 192 [in Russian].
6. Borgoyakova, K.S. (2017). Methods to determine scientific publications citation of a researcher. *Bibliosphere*, 1, 103–109 [in Russian]. DOI: 10.20913/1815-3186-2017-1-103-109.
7. Chaikovskiy, Yu.B., Silkina, Yu.V., Pototskaya, O.Yu. (2013). Scientometric Databases and their Quantitative Indices. Part II. Factors influencing the quantitative indices of scientometric databases, *Bulletin of the NAS of Ukraine*, 9, 84–92 [in Ukrainian].
8. Victoria Anauati & Sebastian Galiani & Ramiro H. Gálvez (2016). Quantifying the Life Cycle of Scholarly Articles Across Fields Of Economic Research. *Economic Inquiry, Western Economic Association International*, 54(2), 1339–1355. April. DOI: 10.1111/ecin.12292.
9. Hirsch, J. (2005). An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. U.S.A.*, 102, 16569–16572.
10. Sidiropoulos, A., Katsaros, D., & Manolopoulos, Y. (2007). Generalized Hirsch h-index for disclosing latent facts in citation networks. *Scientometrics*, 72(2), 253–280. DOI: 10.1007/s11192-007-1722-z.
11. Guskov, A.Ye. (2015). Russian Scientometrics: A review of Researches. *Bibliosphere*, 3, 75–86 [in Russian].
12. Porev, S. & Sandya, I. (2015). *The Way of University Science*. Kyiv: Himgest, 112 [in Ukrainian].
13. Lande, D.V., Andrushchenko, V.B., Balagura, I.V. (2017). Wiki-index of authors popularity. *arXiv preprint arXiv:1702.04614*.
14. Mingers, J. & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research*, 246(1), 1–19.
15. Hirsch, J.E. (2007). Does the h index have predictive power? Proceedings of *the National Academy of Sciences of the United States of America*, 104(49), 19193–19198.
16. Redner, S. J. (2010). On the meaning of the h-index (Letter). *Stat. Mech. Theory Exp.*, 3. L03005. 1–6. DOI: 10.1088/1742-5468/2010/03/L03005.
17. Popescu, I.-I., Best, K.-H., Altmann, G. (2007). On the dynamics of word classes in text. *Glottometrics*, 14, 58–71.
18. Popescu, I.-I. & Altmann, G. (2011). Thematic concentration in texts. *Issues in Quantitative Linguistics* (Vol. 2). E. Kélih, V. Levickij & Y. Matskulyak (Eds.). (pp. 110–116). Lüdenscheid: RAM-Verlag.
19. Čech, R., Popescu, I.-I., & Altmann, G. (2013). Methods of analysis of a thematic concentration of the text. *Czech and Slovak Linguistic Review*, 1, 4–21.
20. Čech, R., Garabik, R., Altmann, G. (2015). Testing the thematic concentration of text. *Journal of Quantitative Linguistics*, 22(3), 215–232.
21. Ruina Chen & Haitao Liu (2017). Thematic Concentration as a Discriminating Feature of Text Types. *Journal of Quantitative Linguistics*, 25(1), 1–24. DOI: 10.1080/09296174.2017.1339441.

22. Popescu, Ioan-Iovitz. Google Scholar Citations (03.07.2019). Retrieved from https://scholar.google.com/citations?hl=ru&user=iqhVvzEAAAAJ&view_op=list_works.
23. Ioan-Iovitz Popescu. 80 years of life. Retrieved from <http://www.iipopescu.com/>.
24. Alonso, S., Cabrerizo, F.J., Herrera-Viedma, E., Herrera, F. (2009). h-index: A Review Focused in its Variants, Computation and Standardization for Different Scientific Fields. *Journal of Informetrics*, 3:4, 273—289. DOI: 10.1016/j.joi.2009.04.001.
25. Nazarovets, S.A. (2011). Citation analysis in the era 2.0. Proceedings from: *The Third Scientific and Practical Conference "Modern problems of library activity in the conditions of information society"*. (pp. 202—208). Lviv [in Ukrainian].
26. Ioan-Iovitz Popescu. Linguistic publications updated January 2019. Retrieved from <http://www.ram-verlag.eu/journals-e-journals/glottometrics/http://www.cechradek.cz/publ.html#Articles>.
27. Popescu, Ioan-Iovitz, Altmann, Gabriel (2006). Some aspects of word frequencies. *Glottometrics*, 13, 23—46.
28. Garfield, E. (1999). Journal impact factor: a brief review. *Can. Med. Assoc. J.*, 161, 979—980.
29. Jacso, P. (2009). Five-year impact factor data in the Journal Citation Reports. *Online Information Review*, 33(3), 603—614.

Received 02.09.2019

С.Г. Бублик, кандидат технических наук,
заместитель начальника отдела,
ГУ «Институт исследований научно-технического потенциала
и истории науки им. Г.М. Доброва НАН Украины»,
бульвар Тараса Шевченко, 60, Киев, 01032, Украина,
<https://orcid.org/0000-0002-8463-9981>,
e-mail: boublyk@gmail.com

ТЕМАТИЧЕСКОЕ ЯДРО ПУБЛИКАЦИЙ ИССЛЕДОВАТЕЛЯ: НАУКОМЕТРИЧЕСКИЙ ПОХОД

Библиометрические профили исследователей формируются из разных по тематической направленности научных публикаций. Наиболее влиятельные из публикаций образуют тематическое ядро библиометрического профиля исследователя. Актуальным заданием наукометрии должен стать не только мониторинг публикационной активности определенного исследователя или определение его h-индекса, но и исследование тематического ядра его публикаций. Цель статьи — провести наукометрическое оценивание тематического ядра публикаций исследователя. Для достижения поставленной цели проведена модификация методов лингвостатистического анализа тематической концентрации слов в тексте и h-индекса для обоснования определения тематического ядра публикаций. Для формирования профиля исследователя (румынского физика-лингвиста Иоана-Йовитца Попеску) использовались данные библиометрической базы Google Scholar. Их использование объясняется простотой доступа к библиометрическим профилям; возможностью охвата материалов конференций и неанглоязычных публикаций, а также проведения ретроспективного анализа динамики ссылок на публикации и h-индекса.

Появление публикаций по разным направлениям исследований в тематическом ядре текущего библиометрического профиля исследователя является не случайным событием, а следствием определенных закономерных процессов формирования тематических ядер по отдельным тематическим направлениям (дисциплинарным, отраслевым или междисциплинарным). Установление перечня всех основных тематических публикаций расширяет возможности для объективного оценивания творческого многообразия исследовательской деятельности ученого и его талантов. Показано, что тематическое ядро библиометрического профиля исследователя представляет все основные

тематические направления его исследований. Использование тематического ядра публикаций как методического инструмента увеличивает возможности библиометрического и наукометрического анализа публикационной деятельности исследователя. Полученные результаты свидетельствуют, что более важным в оценивании эффективности исследовательской деятельности может стать определение перечня публикаций, которые образуют тематическое ядро библиографического профиля исследователя, чем его h-индекса.

Ключевые слова: наукометрия, h-индекс, тематическое ядро публикаций, библиометрический профиль исследователя.

C. G. Boublyk, PhD (Engineering), deputy head of department, Dobrov Institute for Scientific and Technological Potential and Science History Studies of the NAS of Ukraine, 60, Taras Shevchenko boulevard, Kyiv, 01032, Ukraine, <https://orcid.org/0000-0002-8463-9981>, e-mail: boublyk@gmail.com

THE THEMATIC CORE OF THE RESEARCHER'S PUBLICATIONS: THE SCIENTOMETRIC APPROACH

Bibliometric profiles of researchers are formed from scientific publications in various thematic fields. The most influential publications create the thematic core of the researcher's bibliometric profile. Apart from monitoring of the publication activity of a given researcher or computing his/her h-index, an important task of scientometrics has to involve a study of the thematic core of his/her publications. The article's objective is to perform scientometric assessment of the thematic core of the researcher's publications. This objective is achieved by modifying the methods for linguistic statistical analysis of the thematic concentration of words in the text and h-index, to substantiate determining the thematic core of publications. Data from the bibliometric database Google Scholar are used to form the researcher's profile (the case of Ioan-Iovitz Popescu, a Romanian physicist-linguist). Their use is explained by easy access to bibliometric profiles; by the possibility of including conference proceedings and non-English publications and making a retrospective analysis of the dynamics of references to publications and h-index.

The occurrence of publications in various research fields in the thematic core of the current bibliometric profile of the researcher is not an incidence, but a consequence of regular processes involved in the formation of thematic cores in individual thematic fields (disciplinary or interdisciplinary ones). Making a list of all the main thematic publications broadens the possibilities for sound assessment of the diversity of the creative activities of a researcher and his/her talents. It is shown that the thematic core of the researcher's bibliometric profile represents all the thematic fields of his/her research. Use of the thematic core of publications as a methodical tool increases the capacities for bibliometric and scientometric analysis of the publication activity of a researcher. Results of the study demonstrate that compiling a list of publications making the thematic core of the researcher's bibliographic profile can be more important in the research performance assessment, than computing his/her h-index.

Keywords: scientometrics, h-index, thematic core of publications, bibliometric profile of researcher.