



Jarosław Protasiewicz<sup>1</sup>

Laboratorium Inteligentnych Systemów Informatycznych  
Ośrodek Przetwarzania Informacji – Instytut Badawczy

## System wspomaganie wyboru recenzentów



Dr inż. Jarosław Protasiewicz - Kierownik Laboratorium Inteligentnych Systemów Informatycznych (<http://lis.opi.org.pl/>) oraz adiunkt w Ośrodku Przetwarzania Informacji Instytut Badawczy (<http://www.opi.org.pl/>).

Profil LinkedIn <http://www.linkedin.com/pub/jaroslaw-protasiewicz/2a/49a/267>.

Profil na WIT <http://info.wsisiz.edu.pl/~jprotas>.

**Streszczenie:** Referat nie jest oryginalną pracą, lecz wyborem najciekawszych fragmentów z dwutomowej publikacji „Procedury recenzowania i doboru recenzentów”. Przedstawiony został szczegółowo problem heurystyk zniekształceń poznawczych i związanych z nimi błędów, które mogą wystąpić w procesie recenzowania. Stanowiły one podwaliny do przeprowadzonego w roku 2011 badania ankietowego wśród recenzentów i wnioskodawców, którzy brali udział w procesie przyznawania grantów za pomocą systemu OSF. W artykule zamieszczono podsumowanie wyników ankiety. Wyciągnięte wnioski oraz badania różnych systemów recenzowania skłoniły do zaproponowania polskiego systemu wspomagającego dobór recenzentów. W pracy zamieszczono skrócony opis koncepcji i funkcjonalności systemu informatycznego, który jest dostępny publicznie (<http://sssr.opi.org.pl/><sup>2</sup>).

**Słowa kluczowe:** dobór recenzentów, systemy informatyczne, uczenie maszynowe

**Abstract:** The paper is not an original work, but the selection of the most interesting parts of the two-volume publication "Procedures for review and selection of reviewers". We have presented in details the problem of heuristics cognitive distortions and errors related to them that may occur in the process of reviewing. They provided the foundation for a questionnaire survey among reviewers and applicants who had took part in the grants system by using the OSF. The survey was conducted in 2011. The article summarizes the survey results. The conclusions reached and our research of the various reviewing systems allowed us to propose the support system for selection of reviewers in Polish scope. The paper contains a brief description of the concept and functionality of the information system, which is available on-line (<http://sssr.opi.org.pl/>).

**Keywords:** selection of reviewers, computer systems, machine learning

### Prezentacja

## Wprowadzenie

(Jarosław Protasiewicz)

Projekty naukowe i rozwojowe w Polsce finansowane są głównie ze środków budżetowych, unijnych i innych funduszy celowych. Gremia przyznające pieniądze napotykają dwa zasadnicze problemy. Pierwszym jest określenie zasadności merytorycznej finansowania danych projektów, natomiast drugim — wybranie

<sup>1</sup> Autor występuje w roli redaktora, który dokonał jedynie wyboru fragmentów tekstu z dwutomowej publikacji: PROTASIEWICZ, J. (red.) *Procedury recenzowania i doboru recenzentów*. Warszawa: Ośrodek Przetwarzania Informacji — Instytut Badawczy, 2012. Pełny tekst publikacji jest dostępny bez ograniczeń pod adresem: <http://www.opi.org.pl/Wydawnictwa/newsId/28.html>.

<sup>2</sup> Wszystkie odesłania do stron internetowych przedstawiają wersję aktualną w dn. 16.05.2013 r.



przedsięwzięć najlepszych spośród wielu, często równie dobrych, w sytuacji ograniczenia środków pieniężnych. Podstawowym narzędziem stosowanym do oceny merytorycznej projektów są recenzje sporządzane przez przedstawicieli świata nauki oraz ekspertów z sektora komercyjnego. Recenzja często w sposób decydujący warunkuje sfinansowanie danego projektu. Wybór najlepszych projektów leży w interesie państwa, ponieważ umiejętnie wydane środki mogą znacząco wpłynąć na poprawę konkurencyjności gospodarki i rozwój ekonomiczny kraju, a nawet na poczucie praworządności i sprawiedliwości społecznej. Stąd tak istotna jest niezależność osób opiniujących, ich odpowiednie dopasowanie do problematyki projektu oraz brak konfliktu interesów. W ten sposób dochodzimy do kwestii odpowiedniego doboru recenzentów i ekspertów oraz recenzowania projektów badawczych, rozwojowych i artykułów naukowych.

Istotność tego problemu skłoniła zespół Laboratorium Inteligentnych Systemów Informatycznych<sup>3</sup> OPI do podjęcia się opracowania systemu wspomagającego dobór recenzentów<sup>4</sup>. Przeprowadzono badanie ankietowe, którego celem było zebranie opinii na temat procesu oceniania projektów i prac naukowych w Polsce. Ankiety skierowano do osób recenzujących wnioski grantowe składane w naszym kraju oraz do badaczy starających się o granty. Na większość pytań należało odpowiedzieć precyzyjnie, natomiast ostatnie pytanie miało charakter otwarty. Wyniki badania zostały szczegółowo omówione w tomie pierwszym opracowania *Procedury recenzowania i doboru recenzentów*.

Biorąc pod uwagę krytykę systemów grantowych w Polsce oraz dysponując analizą metod recenzowania, zaproponowano wstępną koncepcję systemu wspomagania wyboru recenzentów. Koncepcja miała za zadanie wyznaczyć kierunki dalszych poszukiwań optymalnego rozwiązania. Na jej podstawie opracowano projekt docelowego systemu wspomagającego dobór recenzentów. Dołożono również starań, aby nie powielać błędów charakterystycznych dla innych systemów. System zawiera pięć modułów merytorycznych: zbierania danych, klasyfikacji, identyfikacji osób, ekstrakcji słów kluczowych oraz rankingowania. Moduły te, to właściwe algorytmy odpowiedzialne za przetwarzanie danych. Użytkownik nie ma do nich bezpośredniego dostępu, dopiero wynik ich działania jest prezentowany przez interfejs użytkownika.

### Heurystyki zniekształceń poznawczych

(Agata Kopacz, Anna Maria Wieczorek, Jarosław Protasiewicz, Sławomir Dadas)

Jądrem recenzji naukowej wydaje się być dobór ekspertów, którzy mają ocenić manuskrypt. Efekt końcowy procesu to często pytanie o to, w jaki sposób dobierani

<sup>3</sup> LIS Laboratory of Intelligent Information Systems [on-line]. [Dostęp 16.05.2023]. Dostępny w World Wide Web: <http://lis.opi.org.pl>.

<sup>4</sup> Subprojekt „System wspomagania wyboru recenzentów”, zadanie 5.1 projektu systemowego „Wsparcie systemu zarządzania badaniami naukowymi oraz ich wynikami”, w ramach priorytetu I, działanie 1.1., poddziałanie 1.1.3. Programu Operacyjnego Innowacyjna Gospodarka 2007–2013.



są recenzenci<sup>5</sup>. Wyniki badań sugerują, że kluczowym elementem rozstrzygającym o rzetelności, ważności i sprawiedliwości procesu jest właśnie właściwy dobór oceniającego. Jest to pierwszy moment, kiedy mogą wystąpić nieprawidłowości decydujące o losach artykułu lub wniosku o sfinansowanie prac B+R. W przeglądach weryfikujących ideę recenzowania pomija się ryzyko popełnienia błędów poznawczych, które towarzyszą aspektom formalnym i równie znacząco mogą zaburzać właściwe podejmowanie decyzji.

Przemyślenia własne skłaniają do refleksji, że wiele problemów dotyczy recenzenta w kontekście wystawianych przez niego opinii, a więc kolejnego etapu procesu; stąd tak ważne jest dokonanie właściwego i optymalnego wyboru eksperta. Pewne wątpliwe kwestie mogą stanowić potencjalne źródło zagrożenia, zarówno na etapie doboru recenzenta, jak i podczas wystawiania przez niego opinii oraz podejmowania ostatecznej decyzji przez redaktora lub — w przypadku grantów aplikacyjnych — osobę przyznającą finansowanie. Tabela 1 pokazuje występujące skrzywienia.

Fakt istnienia ograniczeń w ocenie oraz heurystyk i wynikających z nich błędów poznawczych można wykorzystać podczas tworzenia systemu informatycznego wspierającego dobór recenzentów. Wiedza na ich temat pozwala wyeliminować je z procesu lub zminimalizować ich wpływ.

Tab. 1. Skrzywienia według miejsca w procesie recenzowania

| Rodzaj skrzywienia  | Wybór recenzenta | Wystawianie recenzji | Decyzja końcowa |
|---|------------------|----------------------|-----------------|
| Konflikt interesów  | X                | X                    | X               |
| Manipulowanie wrażeniem   |                  | X                    |                 |
| Dyskryminowanie nowych i kontrowersyjnych idei oraz zasada podczepienia |                  | X                    |                 |
| Partykularyzm tematyczny i orientacja ideologiczna                      |                  | X                    | X               |
| Efekt ekspozycji i <i>becoming famous overnight</i> *                   |                  | X                    |                 |
| Efekt skupienia   |                  | X                    |                 |
| Tendencja do konfirmacji  |                  | X                    |                 |
| Sugerowanie się afiliacją i statusem oraz efekt halo                    | X                | X                    | X               |
| Efekt płci  |                  | X                    |                 |
| Efekt sugerowania recenzentów   |                  | X                    |                 |
| Syndrom grupowego myślenia i efekt kongruencji**                        |                  | X                    | X               |
| Efekt pierwszeństwa   |                  | X                    | X               |
| Efekt pewności wstecznej  |                  | X                    |                 |
| Efekt kontrastu i przeddecyzyjna stronniczość                           |                  | X                    | X               |
| Efekt <i>file drawer</i> ***  |                  | X                    |                 |

\* Efekt dotyczy ograniczonej możliwości unikania nieświadomych wpływów przeszłości.

\*\*Efekt dotyczy bezpośredniego testowania założeń, przy ignorowaniu możliwości testowania pośredniego.

\*\*\*Efekt polega na rekomendowaniu badań, które zakończyły się pozytywnymi rezultatami, potwierdzającymi postawione hipotezy, a odrzucaniu tych z negatywnymi, niespełnionymi przewidywaniami.

Źródło: Opracowanie własne autorów.

<sup>5</sup> HEMLIN, S. Peer review agreement or peer review disagreement: Which is better? *Journal of Psychology of Science and Technology* 2009, Vol. 2, nr 1, s. 5–12.



### Badania opinii recenzentów i wnioskodawców

(Agata Kopacz, Marek Młodożeniec, Jarosław Protasiewicz)

Dla zaledwie 47% naukowców publikacja w recenzowanym czasopiśmie świadczy o wysokiej jakości manuskryptu<sup>6</sup>. Podważana jest rzetelność, trafność, uniwersalizm czy obiektywizm recenzji. [...] Informacje uzyskane z zagranicznego piśmiennictwa<sup>7, 8, 9</sup> były inspiracją do zweryfikowania, jak recenzowanie przebiega w naszym kraju i jak postrzegają je rodzimi naukowcy. Sporządzono ankietę, której przedmiotem było zebranie opinii o procesie oceniania. Skierowano ją do osób opiniujących wnioski składane w Polsce oraz starających się o granty w Polsce. Badanie miało na celu wyodrębnienie wad i zalet obowiązującej procedury, tak aby w przyszłości możliwe było wypracowanie rozwiązań podnoszących jakość recenzowania i eliminujących aktualne problemy z nim związane.

Ankietę rozesłano do około 40 tys. osób związanych z nauką w Polsce. Udział w badaniu zdecydowało się wziąć ponad 8 tys. z nich, co stanowiło zwrot na poziomie 20%. W pytaniu otwartym ankiety wypowiedziało się ponad 2,5 tys. respondentów, czyli 32% badanych.

W pierwszym etapie dokonano analizy ilościowej rozkładu osób badanych. Wzięto pod uwagę takie kryteria, jak: płeć, stopień i tytuł naukowy, dziedzinę i instytucję respondenta, jego rolę w procesie recenzenckim, program grantowy, w jakim ma doświadczenie oraz liczbę złożonych i zatwierdzonych wniosków oraz wykonanych recenzji. Okazało się, że większość ankietowanych to mężczyźni (64%) i osoby ze stopniem doktora (43%). Większość (45%) miała również praktykę jedynie w aplikowaniu o granty. Przeważająca część osób (57%) złożyła od jednego do trzech wniosków. Największa grupa (31%) otrzymała grant jeden raz, a co piąta (20%) nie dostała żadnego grantu. Wśród dziedzin nauki reprezentowanych przez badanych rozkład był równomierny, żadna dziedzina nie zdominowała pozostałych. Większość respondentów zadeklarowała współpracę z jednostkami państwowych uczelni (78%). Aż 95% uczestników miało doświadczenie w programach grantowych Ministerstwa Nauki i Szkolnictwa Wyższego.

Kolejnym etapem była analiza ilościowa danych poglądów na temat procedury recenzowania. Zdecydowana większość ankietowanych (96%) uznała ją za potrzebną, jednak zaledwie co drugi (52%) uważał, że ocenianie przebiega w sposób obiektywny, a 45% wyraziło opinię, że proces jest uczciwy. Rozkład odpowiedzi na oba ostatnie pytania zależał od roli wypowiadającego się. I tak, 76% recenzentów

<sup>6</sup> MACNAB, N., THOMAS, G. Quality in research and the significance of community assessment and peer review: education's idiosyncrasy. *International Journal of Research & Method in Education* 2007, Vol. 30, nr 3, s. 339–352.

<sup>7</sup> MOED, H.F. The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. *Science and Public Policy* 2007, Vol. 34, nr 8, s. 575–583.

<sup>8</sup> MAYO, N.E. i in. Peering at peer review revealed high degree of chance associated with funding of grant applications. *Journal of Clinical Epidemiology* 2006, Vol. 59, nr 8, s. 842–848.

<sup>9</sup> WENNERÅS, C., WOLD, A. Nepotism and sexism in peer review. *Nature* 1997, nr 387 (6631), s. 341–343.



oceniło proces jako obiektywny, a 61% jako uczciwy. Wśród wnioskodawców za obiektywizmem opowiedziało się zaledwie 35% badanych, a za uczciwością — 36%. Pod względem oceny stopnia dopasowania recenzentów zaledwie co drugi ankiетowany zgodził się z opinią, że recenzenci są dobrze dobrani do wniosków, które opiniują. Ponownie, wśród recenzentów uważało tak 74% osób, a wśród wnioskodawców — 36%. Podobna różnica poglądów zaistniała w pytaniu o posiadaną przez ekspertów wiedzę: 83% recenzentów i 45% aplikujących uznało ją za wystarczającą. Także o rzetelności recenzji przekonanych było 74% ekspertów i 37% wnioskodawców. Poproszeni o wybranie modelu recenzowania, badani w większości (60%) poparli model *double blind*, opowiadając się za całkowitą anonimowością. Najmniejszym powodzeniem cieszył się system *single blind* (anonimowość zachowuje jedynie recenzent). Wśród powodów zaburzających właściwy przebieg recenzowania najczęściej wymieniano: sugerowanie się afiliacją i statusem wnioskodawcy (58%), interes własny (58%), znajomość z recenzentem (57%), rywalizację (55%), ograniczenia w wiedzy recenzenta (54%) oraz przeświadczenie o własnej nieomyślności (44%).

Ostatni etap polegał na przeprowadzeniu analizy jakościowej pytania otwartego, w którym uczestnicy badania mogli swobodnie wypowiedzieć się na temat procesu recenzowania w Polsce. W wyniku analizy zdefiniowano pięć głównych kategorii problemów: „recenzowanie jako całość”, „ocena”, „jakość pracy”, „formalizm” i „anonimowość”. Większość uwag dotyczyła całości procesu (ponad 30%), oceny (niecałe 30%) i jakości pracy (ponad 20%). Zarówno anonimowość, jak i formalizm pojawiły się w około 10% wypowiedzi. Wśród bezpośrednio poruszanych przez badanych tematów największe wątpliwości budził sposób wyboru recenzentów (ponad 15% komentarzy) oraz stosowane przez recenzentów kryteria oceny (14%). O anonimowości, problemie biurokracji oraz sposobie agregacji ocen wspomniano w 10% przypadków. Blisko 8% osób dostrzegło zagadnienie jakości wystawianych recenzji, a 7% — kwestię uczciwości. Poniżej 5% wypowiedzi dotyczyło możliwości prowadzenia dialogu podczas recenzowania, kontroli recenzentów, wiedzy ekspertów, możliwości odwoływania się, rozbieżności wystawianych ocen, braku wytycznych do oceny, subiektywizmu oceny oraz wąskiej skali ocen.

Zdecydowana większość wypowiedzi — ponad 90% — miała charakter negatywny. Zaledwie 1,5% opinii zawierało pochlebne uwagi na temat procesu recenzowania. Taki rozkład wyników nie napawa optymizmem. Należy jednak mieć na uwadze fakt, że połowa respondentów nie otrzymała do tej pory grantu lub miała doświadczenie z jednym grantem, co mogło przełożyć się na wysoki poziom niezadowolenia wyrażonego w przeprowadzonej ankiecie oraz brak wglądu w rzeczywisty przebieg procesu recenzowania. Rozbieżności w wypowiedziach dwóch stron procesu, tj. recenzentów i wnioskodawców, sugerują, by do przedstawionych wyników podchodzić z pewną rezerwą. Wyniki są bowiem w pewnym stopniu obarczone błędem braku obiektywizmu osób wypowiadających się, jak i ludzką skłonnością do nadmiernego przekonania o swojej nieomyślności i słuszności własnych osądów.



### System informatyczny wspomagający dobór recenzentów

(Jan Artysiewicz, Sławomir Dadas, Małgorzata Gałęzewska, Marek Kozłowski, Agata Kopacz, Anna Maria Plewa, Jarosław Protasiewicz, Tomasz Stanisławek)

Badania wybranych systemów<sup>10,11,12,13</sup> wspomagających proces recenzowania wykazały, że nie ma możliwości zaadaptowania istniejących implementacji do potrzeb polskiego środowiska naukowego. W tej sytuacji rozsądnym krokiem wydaje się opracowanie i wdrożenie własnych rozwiązań.

Ośrodek Przetwarzania Informacji — Instytut Badawczy ma pewne doświadczenia w obsłudze procesu recenzowania<sup>14</sup>. Skuteczność dotychczasowych narzędzi została zweryfikowana poprzez ankietę rozesłaną do wnioskodawców i recenzentów w systemach informatycznych obsługiwanych przez OPI. Najważniejsza konkluzja brzmi: niezbędna jest praca nad systemem wspomagania wyboru recenzentów oraz polepszeniem organizacji całego procesu.

#### Koncepcja systemu

Przedstawiony projekt nowego systemu zakłada, że ranking recenzentów będzie budowany na podstawie aktualnego dorobku naukowego poszczególnych osób. Zbadana zostanie zgodność słów kluczowych pobranych z wniosku oraz słów kluczowych związanych z potencjalnym oceniającym, przy czym słowa kluczowe naukowca będą wydobyte automatycznie z jego dorobku.

W pierwszym etapie pracy nad systemem wspomagania wyboru recenzentów stworzony zostanie dynamiczny i adaptacyjny model gromadzenia, analizy i oceny ekspertów dla poszczególnych artykułów lub wniosków badawczych. Następnie model ten posłuży do utworzenia inteligentnej bazy gromadzącej informacje o potencjalnych recenzentach; na podstawie tych danych system będzie przedstawiał ranking oceniających dla zadanego problemu. Zapytanie do bazy będzie miało postać: „Wskaż recenzentów do oceny projektu, do oceny artykułu etc.” System odpowiadający na takie pytanie będzie musiał zbierać informacje o pracownikach naukowych, tworzyć wiedzę na ich temat oraz formować ranking recenzentów<sup>15</sup>, na co składają się poniżej opisane procesy biznesowe.

<sup>10</sup> Elsevier Editorial System [on-line]. [Dostęp 08.08.2012]. Dostępny w World Wide Web: <http://www.elsevier.com/wps/find/authors.authors/EES-author>.

<sup>11</sup> ScholarOne Manuscripts. The online manuscript submission and peer review process. W: ScholarOne [on-line]. [Dostęp 08.08.2012]. Dostępny w World Wide Web: [http://scholarone.com/media/manuscripts\\_fs.pdf](http://scholarone.com/media/manuscripts_fs.pdf).

<sup>12</sup> RODRIGUEZ, M.A., BOLLEN, J., VAN DE SOMPEL, H. The convergence of digital-libraries and the peer-review process. *Journal of Information Science* 2006, Vol. 32, nr 2, s. 149–159.

<sup>13</sup> FLACH, P.A. i in. Novel tools to streamline the conference review process: Experiences from SIGKDD'09. *ACM SIGKDD Explorations Newsletter* 2009, Vol. 11, nr 2, s. 63–67.

<sup>14</sup> Zob. System OSF (<https://osf.opi.org.pl>); System PNRF (<https://pnrf.opi.org.pl>); System PSRP (Polish-Swiss Research Programme, po polsku: PSPB — Polsko-Szwajcarki Program Badawczy; <https://psrp-system.opi.org.pl>); System LSI-PO IG (<https://poig-wnioski.opi.org.pl>).

<sup>15</sup> Raport „Projekt systemu” opracowany w ramach projektu „System wspomagania wyboru recenzentów”. Warszawa: OPI 2011. Niepubl.



#### a) Zbieranie informacji o pracownikach naukowych

Zaangażowane będą roboty internetowe i lokalne, czyli programy gromadzące — w sieci lub lokalnych bazach danych — informacje o poszczególnych osobach. Dla każdego pracownika naukowego znajdującego się w Bazie Wiedzy o Nauce Polskiej (BWNP) wywoływany jest cyklicznie robot internetowy (jedna instancja robota równa się jednemu procesowi serwera aplikacji). Robot lokalny przeszukuje bazy OPI, natomiast robot internetowy — zasoby zewnętrzne. W pierwszej fazie projektu informacje zbierane są z baz artykułów, a następnie z sieci WWW.

Zebrana wiedza o pracowniku naukowym jest zapisywana w bazie danych. Wykorzystane zostaną następujące źródła:

- BWNP — dane o pracownikach naukowych,
- OSF — dane o recenzowaniu i klasyfikacji wniosków,
- WWW — przeszukiwane przy wykorzystaniu istniejących wyszukiwarek internetowych lub własnych mechanizmów,
- bazy artykułów — dostępne w Internecie.

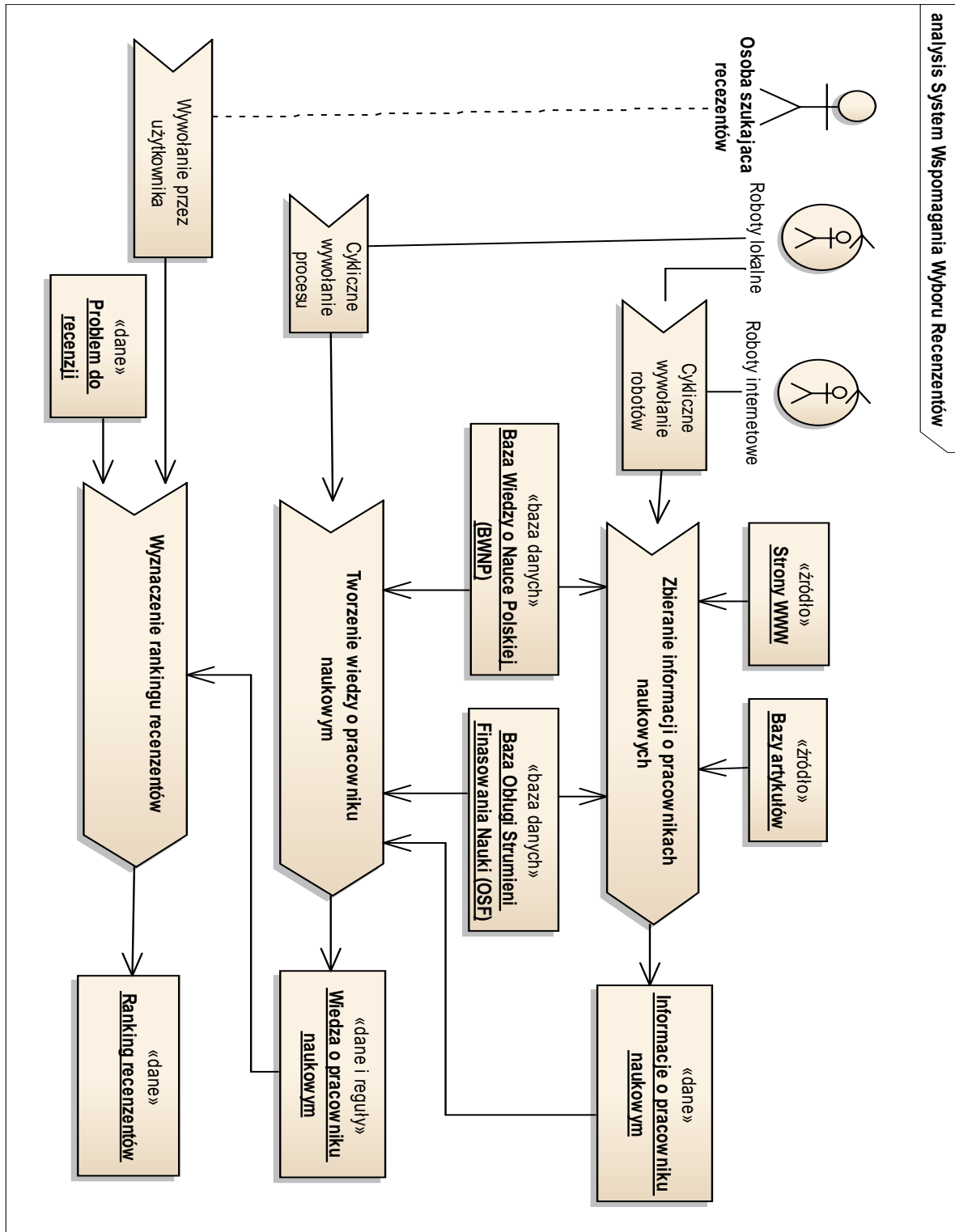
Ostatecznie pozyskane zostaną informacje o pracowniku naukowym, czyli lista jego artykułów naukowych zawierająca: tytuł artykułu, nazwę czasopisma, listę autorów, pozycję w kolejności autorów, streszczenie, słowa kluczowe, treść artykułu (opcjonalnie) oraz różne klasyfikacje dziedzin/dyscyplin pracownika naukowego.

#### b) Tworzenie bazy wiedzy o pracowniku naukowym

Robot lokalny, czyli proces działający na serwerze aplikacji, będzie odpowiedzialny za przetworzenie surowych informacji o naukowcu w wiedzę o nim. Źródłem danych są tekstowe informacje o danej osobie oraz BWNP i OSF. Przekształcenie informacji w wiedzę polega na wyodrębnieniu zbioru słów unikatowych dla każdego naukowca. Słowa będą wyodrębniane z artykułów naukowych (tytułów, słów kluczowych, abstraktów). Podjęta będzie też próba zbudowania zestawu konceptów oraz relacji między słowami (ontologii) dla pracownika naukowego.

#### c) Wyznaczenie rankingu recenzentów

Użytkownik systemu może przesłać zapytanie do bazy wiedzy w dwojaki sposób: podać słowa kluczowe lub wskazać dokument, dla którego poszukuje recenzentów. W drugim przypadku słowa kluczowe z dokumentu wyodrębni odpowiedni algorytm. W efekcie pozyskana zostanie wiedza o problemie do recenzji, czyli zbiór słów kluczowych, które mogą być także powiązane regułami.



Rys. 1. Koncepcja systemu wspomagania wyboru recenzentów

Źródło: Opracowanie własne autorów.

Do obliczenia rankingów recenzentów służy algorytm rankingowy, czyli program implementujący algorytmy dopasowania profili pracowników naukowych do





dokumentu przeznaczony do recenzji. Ranking może być utworzony poprzez znalezienie najlepszego dopasowania recenzowanego problemu (słów kluczowych z dokumentu) do wiedzy o pracownikach naukowych (słowa kluczowymi z profili uczonych), łącznie z wykorzystaniem ontologii. Należy zastanowić się, jak uwzględnić w zestawieniu dane o recenzjach wniosków z różnych systemów grantowych. Ranking sygnalizuje również potencjalne konflikty interesów. Koncepcja systemu przedstawiona jest na rysunku 1.

#### Założenia i ograniczenia

System automatycznie gromadzi wiedzę o potencjalnych recenzentach. Problem wielojęzyczności dokonań naukowców (głównie publikacji) jest bardzo trudny do rozwiązania, ponieważ wymaga tłumaczenia słów kluczowych na wiele języków. Przyjęto zatem założenie, że w tej fazie projektu system będzie brał pod uwagę publikacje napisane tylko w językach angielskim i polskim. Dzięki udostępnieniu serwisu WWW dowolny użytkownik będzie miał dostęp do systemu z publicznej sieci, przez przeglądarkę internetową. Usługa sieciowa (*web service*) pozwoli innym systemom informatycznym kierować pytania do bazy wiedzy o recenzentach. Interfejs użytkownika ma dwie wersje językowe: angielską i polską. Rdzeniem, na którym zbudowany został system, jest Baza Wiedzy o Nauce Polskiej<sup>16</sup> — istniejące profile naukowców dotyczą tylko osób znajdujących się w tej bazie.

System został wykonany i wdrożony z wykorzystaniem następujących technologii:

- język programowania Java i technologia J2EE,
- serwer aplikacji Boss,
- baza danych Oracle,
- system operacyjny Linux.

#### Warstwy i składowe systemu

W związku z przyjętymi ograniczeniami zaprojektowano system składający się z następujących elementów (rys. 2):

- 1) baza danych,
- 2) moduły merytoryczne:
  - zbierania danych,
  - klasyfikacji,
  - identyfikacji osób,
  - ekstrakcji słów kluczowych,
  - rankingowania,
- 3) interfejs użytkownika.

Pięć modułów merytorycznych odpowiada za przetwarzanie danych. Bezpośredni użytkownicy ich nie dostrzegają, poprzez interfejs użytkownika widoczne są jedynie efekty ich działań.

<sup>16</sup> Baza wiedzy. W: *Nauka Polska* [on-line]. [Dostęp 08.08.2012]. Dostępny w World Wide Web: <http://nauka-polska.pl>.



#### Funkcjonalności systemu

Interfejs użytkownika jest dostępny przez przeglądarkę WWW. Do wybranych funkcjonalności będzie można dotrzeć w ramach usługi sieciowej (*web service*); skorzystają z niej także inne systemy informatyczne. Interfejs ma angielską i polską wersję językową. Menu użytkownika zawiera kilka funkcjonalności.

#### a) Ranking recenzentów działający w trzech trybach:

- Kryterialnym — recenzenci zostają dobrani na podstawie wpisywanych przez użytkownika słów kluczowych. Interfejs umożliwia wprowadzenie takich informacji, jak: słowa kluczowe w języku polskim i angielskim, dziedziny naukowe, wnioskodawcy, wykonawcy.
- Dokumentalnym — tworzy ranking na podstawie dokumentu, który można dodać z poziomu interfejsu. Na podstawie wpisanego tekstu system wyodrębnia słowa kluczowe, a następnie bierze je pod uwagę przy tworzeniu rankingu recenzentów w metodzie kosinusowej.
- Przetwarzanie wsadowe — działa w oparciu o zagregowane pliki wniosków, dla których zwraca sugerowane listy recenzentów w postaci pliku xls/xml. Pozwala na wykonanie doboru recenzentów dla grupy wniosków.

Finalny ranking recenzentów jest wypadkową rankingów pozyskanych z metody kosinusowej i pełnotekstowej. Ranking kosinusowy działa w oparciu o mierzenie kąta dopasowania słowa kluczowego profilu naukowca do zapytania. Pełnotekstowy ranking działa w oparciu o bibliotekę Apache Lucene. Stopień wpływu metod na finalny wynik można zmieniać, wprowadzając wagi suwakami.

#### b) Analiza dokumentu

Umożliwia wyodrębnienie słów kluczowych z wprowadzonego tekstu lub z dołączonego pliku. Po wykonaniu analizy system zwraca nam statystyki tekstu (język tekstu, język modelu ekstrakcji słów kluczowych, liczba słów, liczba znaków etc.), a dodatkowo wyróżnia kolorem granatowym słowa kluczowe. W wyniku analizy powstaje również lista wyodrębnionych słów wraz z ich formą oryginalną, formą zlematyzowaną oraz prawdopodobieństwem bycia słowem kluczowym. Kolejnym wynikiem analizy dokumentu dostępnym w zakładce „chmura fraz kluczowych” jest graficzna reprezentacja listy słów kluczowych, gdzie rozmiar czcionki zależy od istotności danego słowa.

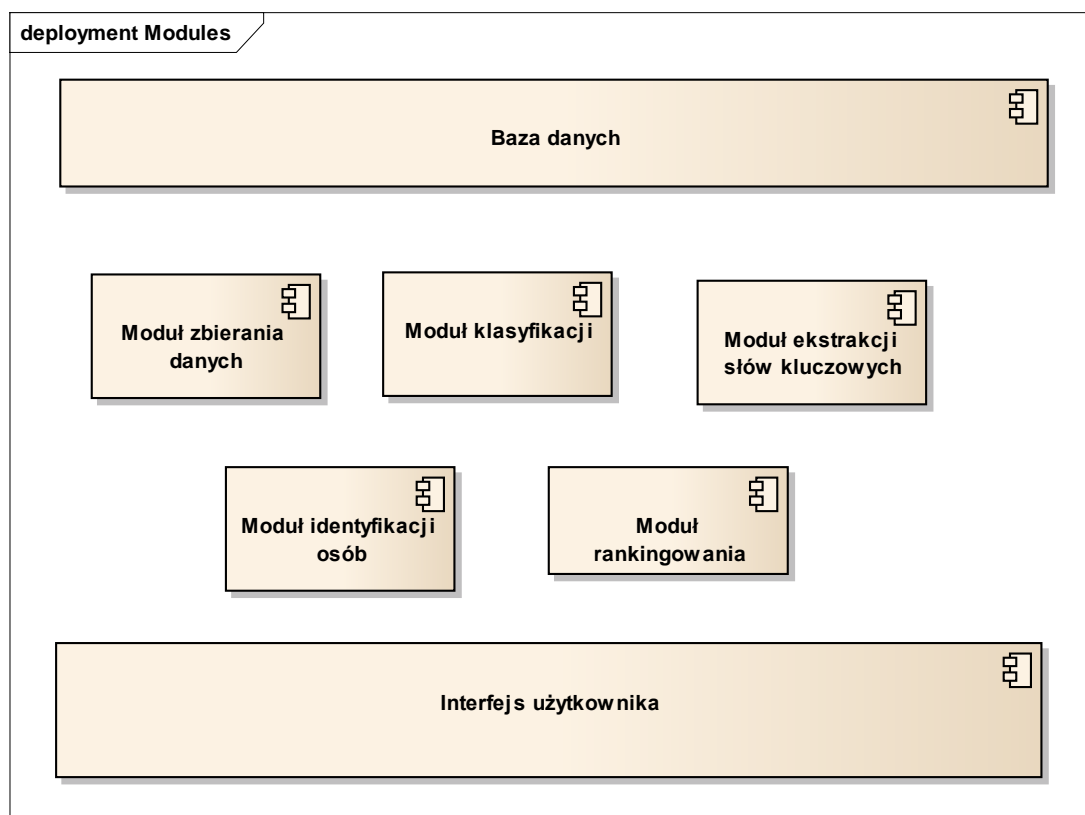
#### c) Bazy danych — system umożliwia przeglądanie pięciu zbiorów danych, są to:

- Dziedziny naukowe — podzielone na dwie sekcje:
  - pierwsza przedstawia reprezentację siedmiu klasyfikacji nauki w postaci grafu lub tabeli; umożliwia to wybór dziedziny naukowej należącej do dowolnej z tych klasyfikacji,
  - druga sekcja zawiera szczegóły dziedzin (pojawia się po kliknięciu na któryś z węzłów sekcji pierwszej).
- Ludzie — system umożliwia odnajdywanie informacji o ludziach nauki poprzez ich imiona, nazwiska, stopnie lub tytuły naukowe oraz afiliacje. Dodatkowo w wyszukiwaniu zaawansowanym można ustalić słowo kluczowe opisujące osobę oraz przypisaną do niej dziedzinę.

- Słowa kluczowe — formularz umożliwia wyszukiwanie słów kluczowych poprzez wprowadzenie nazwy, określenie języka oraz dziedziny, do której dana fraza została przypisana.
- Publikacje — proste wyszukiwanie publikacji możliwe jest poprzez podanie dowolnej frazy z tytułu, nazwiska autora oraz roku (lub przedziału lat) wydania. Dodatkowo zaawansowana opcja rozszerza zakres kryteriów wyszukiwania poprzez podanie źródła, typu publikacji oraz dziedziny, do której należy.
- Źródła — w systemie rozróżniamy trzy typy źródeł: książki, czasopisma i konferencje. Ich wyszukiwanie możliwe jest poprzez wybranie jednego z dostępnych typów, określenie dziedziny, do której należy, oraz wprowadzenie dowolnej liczby znaków z tytułu źródła.

#### d) Administracja

Narzędzia administracyjne zawierają takie operacje, jak: logowanie do systemu, zmiana hasła, zmiana danych osobowych użytkownika. System znajduje się w Internecie, ale dostęp do wrażliwych danych mają wyłącznie autoryzowani odbiorcy. Ograniczonym dostępem zostały również objęte narzędzia administracyjne oraz funkcja generowania rankingów recenzentów.



Rys. 2. Elementy systemu wspomagania wyboru recenzentów

Źródło: Opracowanie własne autorów.

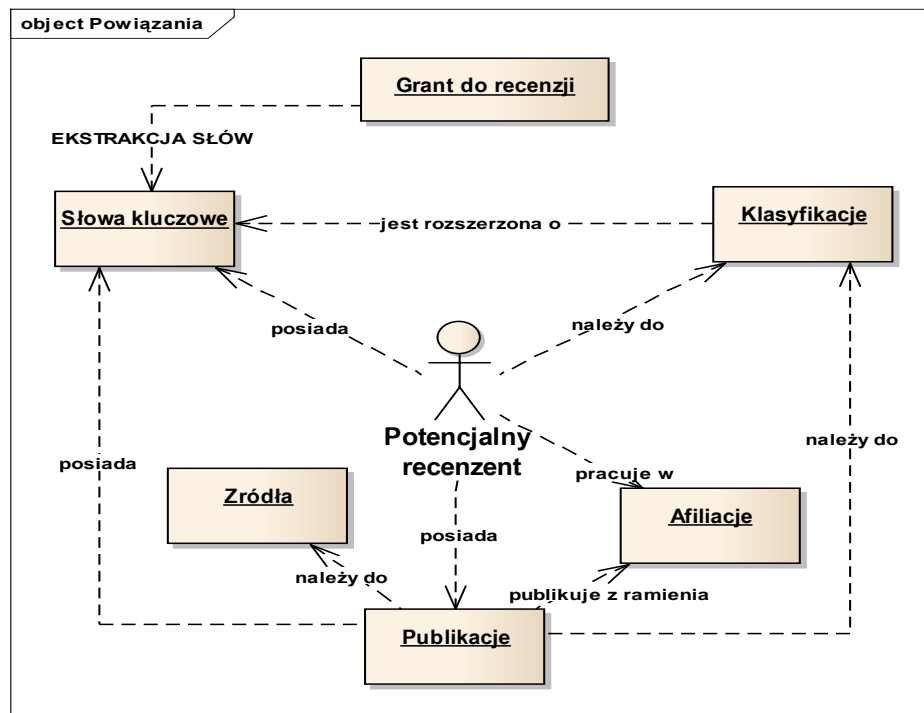


## Podsumowanie

(Jarosław Protasiewicz)

W wyniku działania algorytmów zawartych w zaprojektowanych modułach powstała nie tylko baza danych o potencjalnych recenzentach, ale także system zawierający relacje pomiędzy tymi danymi. Na rysunku 3 pokazano związki między potencjalnym recenzentem, słowami kluczowymi, publikacjami, afiliacjami, klasyfikacjami, źródłami. Potencjalny recenzent przynależy do odpowiednich dziedzin i dyscyplin, które są reprezentowane przez model siedmiu powiązanych ze sobą taksonomii nauki. Klasyfikacje nauki są rozszerzane przez słowa kluczowe. Oczywiście pracownik naukowy posiada także wektor ważonych słów kluczowych tworzący jego profil, jest autorem publikacji oraz jest związany z różnymi instytucjami. Obiekty publikacji posiadają źródła oraz — podobnie jak osoby — należą do odpowiednich klasyfikacji nauki i posiadają afiliacje.

Z całą pewnością nadal rozwijane będą algorytmy systemu i testowane różne nowe, obiecujące metody. W najbliższym czasie planowane jest uruchomienie *roboty internetowej* analizującego strony internetowe pracowników naukowych. Prawdopodobnie mechanizmy ekstrakcji słów kluczowych będą wymagały ciągłego doskonalenia, a także budowy bardziej wyrafinowanych modeli, być może oddzielnych dla każdej dziedziny nauki i języka. Ponadto w algorytmie grupowania hierarchicznego arbitralnie przyjęto wartości parametrów podobieństwa. Planowane jest zastosowanie technik optymalizacyjnych do wyznaczania tych parametrów i ponowne testy algorytmu. Ostatecznie metodę generowania rankingu potencjalnych recenzentów należy traktować jako pierwszą propozycję, która będzie weryfikowana i modyfikowana już podczas używania systemu przez końcowych użytkowników. Autorzy nie zakładają, że przedstawiony systemem jest dziełem kompletnym i zamkniętym. Jest to raczej przyczynek do dalszej pracy, której efekty być może pozwolą lepiej dobrać recenzentów.



Rys. 3. Związki pomiędzy danymi  
Źródło: Opracowanie własne autorów.

Projekt systemu został implementowany przez zespół laboratorium inteligentnych systemów informatycznych OPI (<http://lis.opi.org.pl>). Po zakończeniu prac programistycznych przyjęte założenia poddano weryfikacji na drodze doświadczeń. System dostępny jest publicznie (<http://sssr.opi.org.pl>). Przygotowano także wersję systemu dedykowaną dla Narodowego Centrum Badań i Rozwoju, która została wdrożona do użycia.

### Bibliografia:

1. Baza wiedzy. W: *Nauka Polska* [on-line]. [Dostęp 08.08.2012]. Dostępny w World Wide Web: <http://nauka-polska.pl>.
2. *Elsevier Editorial System* [on-line]. [Dostęp 08.08.2012]. Dostępny w World Wide Web: <http://www.elsevier.com/wps/find/authors.authors/EES-author>.
3. FLACH, P.A. i in. Novel tools to streamline the conference review process: Experiences from SIGKDD'09. *ACM SIGKDD Explorations Newsletter* 2009, Vol. 11, nr 2, s. 63–67.
4. HEMLIN, S. Peer review agreement or peer review disagreement: Which is better? *Journal of Psychology of Science and Technology* 2009, Vol. 2, nr 1, s. 5–12.
5. *LIS Laboratory of Intelligent Information Systems* [on-line]. [Dostęp 16.05.2023]. Dostępny w World Wide Web: <http://lis.opi.org.pl>.
6. MACNAB, N., THOMAS, G. Quality in research and the significance of community assessment and peer review: education's idiosyncrasy. *International Journal of Research & Method in Education* 2007, Vol. 30, nr 3, s. 339–352.



7. MAYO, N.E. i in. Peering at peer review revealed high degree of chance associated with funding of grant applications. *Journal of Clinical Epidemiology* 2006, Vol. 59, nr 8, s. 842–848.
8. MOED, H.F. The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. *Science and Public Policy* 2007, Vol. 34, nr 8, s. 575–583.
9. *Obsługa Strumieni Finansowania* [on-line]. [Dostęp 16.05.2023]. Dostępny w World Wide Web: <http://osf.opi.org.pl>.
10. PROTASIEWICZ, J. (red.) *Procedury recenzowania i doboru recenzentów* [on-line]. Tom. 1. Warszawa: Ośrodek Przetwarzania Informacji — Instytut Badawczy, 2012 [Dostęp 16.05.2013]. Dostępny w World Wide Web: <http://www.opi.org.pl/Wydawnictwa/newsId/28/newsId/28.html>.
11. PROTASIEWICZ, J. (red.) *Procedury recenzowania i doboru recenzentów* [on-line]. Tom 2. Warszawa: Ośrodek Przetwarzania Informacji — Instytut Badawczy, 2012 [Dostęp 16.05.2013]. Dostępny w World Wide Web: <http://www.opi.org.pl/Wydawnictwa/newsId/28/newsId/28.html>.
12. *Raport „Projekt systemu” opracowany w ramach projektu „System wspomaganie wyboru recenzentów”*. Warszawa: Ośrodek Przetwarzania Informacji, 2011. Niepubl.
13. RODRIGUEZ, M.A., BOLLEN, J., VAN DE SOMPEL, H. The convergence of digital-libraries and the peer-review process. *Journal of Information Science* 2006, Vol. 32, nr 2, s. 149–159.
14. ScholarOne Manuscripts. The online manuscript submission and peer review process. W: *ScholarOne* [on-line]. [Dostęp 08.08.2012]. Dostępny w World Wide Web: [http://scholarone.com/media/manuscripts\\_fs.pdf](http://scholarone.com/media/manuscripts_fs.pdf).
15. Subprojekt „System wspomaganie wyboru recenzentów”, zadanie 5.1 projektu systemowego „Wsparcie systemu zarządzania badaniami naukowymi oraz ich wynikami”, w ramach priorytetu I, działanie 1.1., poddziałanie 1.1.3. Programu Operacyjnego Innowacyjna Gospodarka 2007–2013.
16. *The OSF system* [on-line]. [Dostęp 16.05.2023]. Dostępny w World Wide Web: <http://osf.opi.org.pl>
17. WENNERÅS, C., WOLD, A. Nepotism and sexism in peer review. *Nature* 1997, nr 387 (6631), s. 341–343.

---

Protasiewicz, J. System wspomaganie wyboru recenzentów. W: Bibliograficzne bazy danych i ich rola w rozwoju nauki. II Konferencja naukowa Konsorcjum BazTech, Poznań, 17-19 kwietnia 2013 [on-line]. Stowarzyszenie EBIB, 2013 [Dostęp: 30.08.2013]. Materiały konferencyjne EBIB, nr 24, Dostępny w World Wide Web: [http://open.ebib.pl/ojs/index.php/Mat\\_konf/article/view/44](http://open.ebib.pl/ojs/index.php/Mat_konf/article/view/44). ISBN 978-83-63458-06-5.