

Maria Przystek-Samokowa*
Instytut Informacji Naukowej i Studiów Bibliologicznych
Uniwersytet Warszawski



**„BIBLIOGRAFIA JĘZYKOZNAWSTWA SŁAWISTYCZNEGO”
I BIBLIOGRAFICZNA BAZA DANYCH ŚWIATOWEGO JĘZYKOZNAWSTWA
SŁAWISTYCZNEGO „iSYBISLAW”
JAKO PRZYKŁAD DZIAŁAŃ INTERDYSCYPLINARNYCH**

[“THE BIBLIOGRAPHY OF SLAVONIC LINGUISTICS” AND THE BIBLIOGRAPHIC
DATABASE OF THE WORLD SLAVONIC LINGUISTICS “iSYBISLAW” AS THE
EXAMPLE OF INTERDISCIPLINARY PROJECTS]

Abstrakt: W artykule przedstawiono przykład interdyscyplinarnego działania językoznawców sławistów i pracowników informacji naukowej przy tworzeniu i udostępnianiu *Bibliografii językoznawstwa sławistycznego* oraz bibliograficznej bazy danych światowego językoznawstwa sławistycznego iSYBISLAW, a także wybrane zagadnienia interdyscyplinarnego projektu badawczego, dotyczącego analizy zasobu kartoteki słów kluczowych zaprezentowanego systemu.

„BIBLIOGRAFIA JĘZYKOZNAWSTWA SŁAWISTYCZNEGO” – INFORMACJA NAUKOWA – INFORMACJA SŁAWISTYCZNA – iSYBISLAW – JĘZYK SŁÓW KLUCZOWYCH

Abstract: The article describes interdisciplinary activities of the experts in Slavonic linguistics and information professionals involved in creation of and granting access to *Bibliografia językoznawstwa sławistycznego* [*Bibliography of Slavonic Linguistics*] iSYBISLAW – a bibliographic database of the world Slavonic linguistics. The author discusses selected issues of an interdisciplinary research project concerning the analysis of the keywords' authority file the system is based upon.

“BIBLIOGRAPHY OF SLAVONIC LINGUISTICS” – INFORMATION SCIENCE – iSYBISLAW – KEYWORDS LANGUAGE – SLAVONIC INFORMATION

* Dr MARIA PRZASTEK-SAMOKOWA, starszy wykładowca w Instytucie Informacji Naukowej i Studiów Bibliologicznych Uniwersytetu Warszawskiego; absolwentka filologii słowiańskiej, specjalność filologia bułgarska (UW); doktor nauk humanistycznych w zakresie językoznawstwa sławistycznego (Uniwersytet Jagielloński). Dwie najważniejsze publikacje: (2008). *Audyty informacyjny – dwa modele*. [W:] K. Materska, E. Chuchro, B. Sosińska-Kalata red. (2008). *Organizowanie środowiska informacji i wiedzy*. Warszawa: SBP, s. 181–184 [współaut.: J. Koszewska]; (2008) *Narzędzia lingwistyczne – pomoc czy przeszkoda w organizowaniu i wyszukiwaniu informacji?*, dok. elektr. CD [W:] *Problemy hierarchizacji i kategoryzacji informacji o kulturze regionu*. Gdańsk: Wojewódzka i Miejska Biblioteka Publiczna w Gdańsku. E-mail: m.przastek-samok@uw.edu.pl

[MARIA PRZASTEK-SAMOKOWA, PhD in Slavonic Linguistics (Jagiellonian University); Senior Lecturer in the Institute of Information Science and Book Studies, Warsaw University; MA in Slavonic (Bulgarian) Philology (Warsaw University). Two the most important publications: (2008) *Audyty informacyjny – dwa modele* [*Information Audit – Two Models*]. [In:] K. Materska, E. Chuchro, B. Sosińska-Kalata eds. (2008). *Organizowanie środowiska informacji i wiedzy* [*Organization of the Information and Knowledge Environment*]. Warszawa: SBP, s. 181–184 [co-author: J. Koszewska]; (2008) *Narzędzia lingwistyczne – pomoc czy przeszkoda w organizowaniu i wyszukiwaniu informacji?* [*Linguistic Tools – Help or Handicap in Information Organizing and Retrieval?*], e-doc. CD [In:] *Problemy hierarchizacji i kategoryzacji informacji o kulturze regionu* [*Problems of Prioritization and Categorization of the Information Related to Regional Culture*]. Gdańsk: Wojewódzka i Miejska Biblioteka Publiczna w Gdańsku. E-mail: m.przastek-samok@uw.edu.pl]

*

* *

TROCHĘ HISTORII...

Bibliografia językoznawstwa slawistycznego jest wspólnym przedsięwzięciem naukowym Instytutu Sławi-
styki PAN i Towarzystwa Naukowego Warszawskiego. Prace rozpoczęto w połowie lat 90. XX w. *Bibliografia*
jako seria wydawnicza stała się kontynuacją *Przeglądu bibliograficznego językoznawstwa slawistycznego*, dru-
kowanego przez 85 lat w „Roczniku Sławistycznym”, a następnie jako jego odrębnie oprawiona część. W poło-
wie lat 90. podjęto decyzję o odłączeniu *Bibliografii* od „Rocznika...”. Zaważyły na niej co najmniej dwa argu-
menty – specyfika prac bibliograficznych nie pozwalała na rytmiczne, zgodne z ukazywaniem się kolejnego
„Rocznika...” publikowanie *Przeglądu...*, oraz przykład innych bibliografii dziedzinowych, przygotowywanych
przez zespoły specjalistów i stanowiących wydawnictwa samodzielne [Rudnik-Karwatowa 1995, s. 17].

W 1993 r. w Instytucie Sławi-
styki PAN powołano Centrum Sławistycznej Informacji Naukowej, którego
głównym celem, przy wsparciu organizacyjnym Towarzystwa Naukowego Warszawskiego, było przekształcenie
tradycyjnej bibliografii w nowoczesny system informacyjno-wyszukiwawczy. Zaczęto tworzenie dokumentacyj-
nej bazy danych, obsługiwanej przez autonomiczny system SYBISLAW (SYstem BIBliografii SLAWistycznej),
a pochodną tej bazy była tradycyjna postać bibliografii [Teraźniejszość i przyszłość bibliografii..., dok. elektr.].

Pierwszy tom *Bibliografii językoznawstwa slawistycznego* za rok 1992 ukazał się w roku 1995. W sumie,
w latach 1995–2003 opublikowano 6 tomów *Bibliografii...* za lata 1992–1997. Merytorycznie przygotowywał ją
międzynarodowy zespół slawistów językoznawców pod kierunkiem dr Zofii Rudnik-Karwatowej, zaś nad jej
aspektem formalnym czuwała początkowo dr Alina Grabowska (wówczas Biblioteka Narodowa), a następnie
(i po dziś dzień – także od strony informatycznej) mgr Zenon Mikos (Biblioteka Sejmowa). Autorka tego tekstu
była członkiem zespołu w latach 1994–2003 oraz 2007–2009.

SYBISLAW I PIERWSZA BAZA DANYCH

Stworzenie dokumentacyjnej bazy danych i wydawanie kolejnych tomów bibliografii umożliwiło użyt-
kownikom, głównie slawistom, pozyskiwanie informacji niezbędnej w ich pracy. W Centrum Sławistycznej In-
formacji Naukowej pracowano jednak stale nad unowocześnianiem systemu informacyjno-wyszukiwawczego.
Jak już wówczas twierdzili autorzy, SYBISLAW nie pretendował do miana docelowego oprogramowania. Z za-
łożenia miał umożliwić tworzenie i rozbudowę dokumentacyjnej bazy danych w oparciu o istniejący język in-
formacyjno-wyszukiwawczy układu klasyfikacyjnego, a także dzięki specjalnemu podprogramowi wydawni-
czemu miał pomóc w publikacji bibliografii. SYBISLAW pozwalał na wyszukiwanie informacji we wszystkich
polach opisu formalnego, według dowolnie wybranych cech formalnych, takich jak: tytuł, odpowiedzialność,
język dokumentu. Możliwe było także wyszukiwanie informacji według klas występujących w schemacie klasy-
fikacyjnym oraz według słów kluczowych [Teraźniejszość i przyszłość bibliografii..., dok. elektr.]. Ponieważ
pochodną bazy danych stanowiła postać książkowa bibliografii, jednym z ważniejszych zadań było zaprojekto-

wanie odpowiedniej procedury porządkowania i wyszukiwania dokumentów i danych o tych dokumentach, zapisanych w różnych językach i alfabetach [Rudnik-Karwatowa 1996, s. 159–162].

Należy też zaznaczyć, iż omawiana tu dokumentacyjna baza danych w owym okresie była dostępna jedynie lokalnie. Zawierała dane o dokumentach w ich oryginalnych językach. Oznacza to, że opisy bibliograficzne odwzorowują alfabet dokumentów źródłowych, a także dokumentów wtórnych, jakimi są przekłady, we wszystkich językach słowiańskich i innych językach używanych w Europie, a także poza jej granicami [Teraźniejszość i przyszłość bibliografii..., dok. elektr.]. Kolejnym zadaniem, którego podjął się zespół CSIN, było więc zaprojektowanie odpowiedniej procedury porządkowania i wyszukiwania dokumentów i danych o tych dokumentach, zapisanych w różnych językach i alfabetach [Rudnik-Karwatowa 1996].

Już wstępna koncepcja systemu informacji dokumentacyjnej językoznawstwa slawistycznego zakładała, iż dotychczasowy system SYBISLAW – i generowana dzięki niemu książkowa postać bibliografii – powinien być sukcesywnie modernizowany i etapowo rozwijany. Autorzy projektu uważali, że stworzenie nowoczesnego systemu nie powinno polegać na prostym przekształceniu istniejącego ówczesnie, ponieważ miał on zbyt wiele wad [Bojar, Rudnik-Karwatowa 1998]. Uznano, że zasadniczej zmiany wymaga język informacyjno-wyszukiwawczy jako podstawowy składnik systemu.

W SYBISLAW-ie używany był język klasyfikacji przedmiotowej. Mimo przeprowadzonej w systemie skomputeryzowanej modyfikacji, polegającej na wprowadzeniu do klasyfikacji tak istotnych dla językoznawstwa tematów, jak np. semantyka i pragmatyka lingwistyczna, klasyfikacja okazała się nadal zbyt płytka i nierelevantna w stosunku do obecnego stanu wiedzy. Inną wadą tej klasyfikacji było jednokrotne przypisanie dokumentu – do wybranej klasy tematycznej – ze względu na temat główny. Nie można było uniknąć arbitralnych rozstrzygnięć w tej kwestii nawet poprzez wskazanie tematów pobocznych w odsyłaczach znajdujących się na końcu każdego działu tematycznego drukowanej postaci *Bibliografii językoznawstwa slawistycznego* [Teraźniejszość i przyszłość bibliografii..., dok. elektr.].

Autorzy projektu zdecydowali, że docelowym językiem systemu będzie język deskryptorowy o strukturze fasetowej, cechujący się postkoordynacją, polegającą na tym, że każde wyrażenie wchodzące w skład charakterystyki wyszukiwawczej dokumentu może samodzielnie pełnić funkcję wyszukiwawczą, umożliwiając bezpośredni dostęp do zbioru informacyjnego. Dzięki rozbudowanej paradygmatyce, wykorzystującej relacje hierarchiczne i asocjacyjne, język deskryptorowy zapewni wyszukiwanie na relewantnym poziomie głębokości, eliminując szum informacyjny [Teraźniejszość i przyszłość bibliografii..., dok. elektr.].

Ponadto, ze względu na wygodę użytkowników, którzy samodzielnie będą prowadzić wyszukiwanie w trybie online, a także w celu ułatwienia indeksatorom opracowywania informacji dla systemu, ustalono, że będzie to język o notacji i strukturze naturalnej. Język deskryptorowy będzie prezentowany przez tezaurus wielojęzyczny z dostępem we wszystkich językach słowiańskich, a także – ze względu na konieczność udostępniania informacji w Internecie – w języku angielskim. W pierwszej kolejności będzie tworzony JIW oparty na języku polskim. Wynika to stąd, że system informacyjny językoznawstwa slawistycznego jest inicjatywą polskiego środowiska slawistycznego i jest instytucjonalnie związany z Instytutem Slawistyki PAN. Następnie będą włączane do struktury tworzonego języka ekwiwalenty w języku rosyjskim, jako najszerszej stosowanym w środowisku slawistycznym, i kolejno – w innych językach słowiańskich. W efekcie użytkownicy będą mogli dotrzeć do każdej informacji według dowolnego wyrażenia (w dowolnym języku i alfabecie), w dowolnej strefie opisu, dzięki

ustanowieniu klas ekwiwalencji zarówno w obrębie języka opisu formalnego, jak i w języku opisu rzeczowego, w którym również mogą występować te same wyrażenia zapisane w różnych językach i/lub w różnych alfabetach [Teraźniejszość i przyszłość bibliografii..., dok. elektr.]. Zadanie docelowe realizowane będzie etapami.

JĘZYK SŁÓW KLUCZOWYCH JĘZYKOZNAWSTWA SŁAWISTYCZNEGO I JEGO SŁOWNIK

W 1999 r. opublikowano – w postaci tradycyjnej – *Słownik słów kluczowych językoznawstwa sławistycznego* [Rudnik-Karwatowa, Karpińska 1999] będący słownikiem dziedzinowym o notacji paranaturalnej, zaczerpniętej z terminologii w języku polskim. Następnie w roku 2006 – jako dokument elektroniczny na płycie CD oraz w wersji demo na stronie CSIN-u [Słownik..., dok. elektr.] – ukazała się wersja poprawiona i uzupełniona. *Słownik* przedstawia słownictwo języka słów kluczowych, jego organizację, a także związki między jego wyrażeniami i wyrażeniami języka naturalnego i stanowi podstawę do opracowania przyszłego teaurusu dla polskiej terminologii dokumentów językoznawstwa sławistycznego.

Zasady tworzenia *Słownika* zostały szczegółowo przedstawione we wstępie do obu jego wersji. Za autorkami słownika przedstawiam najistotniejsze:

- zakres terminologii językoznawczej obejmuje przede wszystkim językoznawstwo sławistyczne, a także istotne dla sławistyki terminy z teorii języka oraz językoznawstwa ogólnego;
- zastosowana metoda indukcyjno-dedukcyjna zapewnia aktualność słownictwa, z uwzględnieniem innowacji terminologicznych oraz odzwierciedlenie pełnego zasobu tematycznego językoznawstwa sławistycznego;
- do słownika przejęto jedynie te terminy, które odzwierciedlają tematy reprezentatywnej liczby obecnie publikowanych dokumentów, które są obecnie używane (a spośród dawniejszych – te, na które powołują się współcześni autorzy);
- terminy bardziej szczegółowe, reprezentujące tematy pojedynczych dokumentów zostały pominięte.

Istnienie kontrolowanej listy słów kluczowych w postaci *Słownika słów kluczowych* na obecnym etapie prac nad nowoczesnym systemem informacyjnym językoznawstwa sławistycznego umożliwi indeksatorom pełne i relewantne przedstawianie treści dokumentów, a użytkownikom – wyższą efektywność wyszukiwania informacji zawartych w bazie danych bibliograficznych, liczącej obecnie ponad 40 000 rekordów [Rudnik-Karwatowa, Karpińska 1999, s. 6–9; Teraźniejszość i przyszłość bibliografii..., dok. elektr.].

iSYBISLAW JAKO BIBLIOGRAFICZNA BAZA DANYCH ONLINE

Od roku 2008 autonomiczny system SYBISLAW jest zastępowany nowoczesnym systemem informacyjnym iSYBISLAW. Jest to pionierskie przedsięwzięcie w dziedzinie językoznawstwa sławistycznego w postaci bibliograficznej bazy danych, zawierającej charakterystyki formalne i rzeczowe indeksowanych dokumentów w ich oryginalnych językach (słowiańskich). Oznacza to, że opisy bibliograficzne odwzorowują alfabet dokumentów źródłowych, a także dokumentów wtórnych, jakimi są przekłady, we wszystkich językach słowiańskich i innych językach używanych w Europie, a także poza jej granicami. Na bazę składają się opisy zarówno opra-

cowañ monograficznych, jak te¿ artykułów opublikowanych w czasopismach i pracach zbiorowych. Jedn z kartotek haseł wzorcowych – obok kartotek: osb, serii, klasyfikacji i jzyków – jest kartoteka sww kluczowych. Aktualnie do tworzenia charakterystyk i instrukcji wyszukiwawczych wykorzystuje si opisany wy¿ej istniejcy i stale rozbudowywany jzyk sww kluczowych. Jego sownik odgrywa i odgrywa trudn do przecenienia rol w tworzeniu nowoczesnego systemu informacyjno-wyszukiwawczego iSYBISLAW.

Jak zapowiadaj uczestnicy prac nad przedstawianym przeze mnie systemem – sownik sww kluczowych bdzie podstaw do skonstruowania optymalnego dla systemu slawistycznej informacji dokumentacyjnej jzyka deskryptorowego o strukturze fasetowej. Udostpniony bdzie jako tezaurus wielojzyczny wszystkich jzyków sowiaskich i jzyka angielskiego – ze wzgldu na umo¿liwienie korzystania ze w Internecie. Ponadto, ze wzgldu na wygod u¿ytkownikw, ktrzy samodzielnie bd prowadzi wyszukiwanie w trybie online, a tak¿e w celu ulatwienia indeksatorom opracowywania informacji dla systemu, bdzie to jzyk o notacji i strukturze naturalnej. Jzyk deskryptorowy bdzie prezentowany przez tezaurus wielojzyczny z dostpem we wszystkich jzykach sowiaskich, a tak¿e – ze wzgldu na konieczno udostpniania informacji w Internecie – w jzyku angielskim [Rudnik-Karwatowa, Mikos, Bojar 2007, s. 22–23, 27].

W pierwszej kolejnoci bdzie tworzony JIW oparty na jzyku polskim. Wynika to std, ¿e system informacyjny jzykoznawstwa slawistycznego jest inicjatyw polskiego rodowiska slawistycznego i jest instytucjonalnie zwizany z Instytutem Slawistyki PAN. Nastpnie do struktury tworzonego jzyka bd wlczane ekwiwalenty w jzyku rosyjskim, jako najszerszej stosowanym w rodowisku slawistycznym, i kolejno – w innych jzykach sowiaskich. W efekcie u¿ytkownicy bd mogli dotrze do ka¿dej informacji wedug dowolnego wyra¿enia (w dowolnym jzyku i alfabecie), w dowolnej strefie opisu, dziki ustanowieniu klas ekwiwalencji zarwno w obrbie jzyka opisu formalnego, jak i w jzyku opisu rzeczowego, w ktrym rwnie¿ mog wystpowa te same wyra¿enia zapisane w rwnych jzykach i/lub w rwnych alfabetach [Tera¿niejszo i przysz bibliografii..., dok. elektr.].

PROJEKT WLASNY BADANIA INTERDISCYPLINARNEGO

Moja kilkuletnia – jako jzykoznawcy slawisty – wsppraca z zespoem przygotowujcym tradycyjn post *Bibliografii jzykoznawstwa slawistycznego* oraz opracowywanie informacji, a nastpnie wprowadzanie ich do nowoczesnego systemu informacyjnego iSYBISLAW, rosnca fascynacja sowami kluczowymi, jak i samym jzykiem sww kluczowych, d¿enie do poznania zasad budowy sownikw sww kluczowych dla wybranych dziedzin, stay si inspiracj do podjcia problematyki indeksowania kontrolowanego i niekontrolowanego na przykladzie kartoteki wzorcowej sww kluczowych bibliograficznej bazy danych iSYBISLAW.

W swoich badaniach chciaabym si skupi – midzy innymi – nad takimi kwestiami, jak:

- problem normalizacji terminologii jzykoznawczej,
- problem identyfikacji wyra¿e terminologicznych,
- problem terminw preferowanych, jak te¿ przestrzeganie kryteriw ich wyboru (frekwencja, poprawno terminologiczna, aktualno, rodzimo, zwi¿ i wyrazisto strukturalna),
- problem synonimii i bliskonaczoci terminw,
- analiza form sww kluczowych,

- analiza nazw osobowych,
- analiza tworzenia otwartych list identyfikatorów oraz stopień szczegółowości i szerokości leksyki tworzącej analizowany słownik,
- wymagania stawiane rozwojowi słownictwa kontrolowanego przez dyscyplinę, jaką jest językoznawstwo sławistyczne,
- rozwój specjalistycznego języka słów kluczowych na tle tworzonych obecnie ontologii
- oraz na tle rozwoju języków słowiańskich w ogóle.

WYKORZYSTANE ŹRÓDŁA I OPRACOWANIA

- Bojar, B.; Z. Rudnik-Karwatowa (1998). Koncepcja nowoczesnego systemu informacji sławistycznej. [W:] J. Siatkowski red. (1998). *Z polskich studiów sławistycznych, Seria 9: Językoznawstwo: prace na XII Międzynarodowy Kongres Sławistów w Krakowie 1998*. Warszawa: Wydaw. Energeia, s. 41–47.
- CSIN, dok. elektr. <http://www.ispan.waw.pl/cnis/BJS.htm> [dostęp: 25.03.2010].
- Rudnik-Karwatowa, Z. et al. (2007). Nowoczesny system informacji sławistycznej: zadania, dotychczasowe wyniki i perspektywy. *Zagadnienia Informacji Naukowej* nr 2, s. 19–40.
- Rudnik-Karwatowa, Z. (1996). Problemy przetwarzania tekstów stosujących różne alfabety: automatyczne porządkowanie i wyszukiwanie informacji w bibliograficznej bazie danych. [W:] Z. Vetulani et al. red. (1996). *Język i technologia*. Warszawa: Akademska Oficyna Wydawnicza PLJ, s. 159–162.
- Rudnik-Karwatowa, Z. (2002). Słowa kluczowe – elementarne jednostki leksykalne języka informacyjno-wyszukiwawczego. [W:] A. Pstyga; K. Szcześniak red. (2002). *Słowo z perspektywy językoznawcy i tłumacza*. Gdańsk: Wydaw. Uniwersytetu Gdańskiego, s. 13–18.
- Rudnik-Karwatowa, Z.; H. Karpińska (1999). *Słownik słów kluczowych językoznawstwa sławistycznego*. Warszawa: Sławistyczny Ośrodek Wydawniczy (SOW), 109 s.
- Rudnik-Karwatowa, Z. (1995). Wstęp. [W:] Z. Rudnik-Karwatowa red. (1995). *Bibliografia językoznawstwa sławistycznego za rok 1992*. Warszawa: Sławistyczny Ośrodek Wydawniczy (SOW), s. 17.
- Teraźniejszość i przyszłość bibliografii językoznawstwa sławistycznego, dok. elektr. <http://www.ispan.waw.pl/cnis/BJS.htm> [odczyt: 25.03.2010].