

Marzena Błaszczczyńska  
marzena@man.poznan.pl  
Karolina Bohdanowicz  
karolina@man.poznan.pl  
Natalia Jeszke  
njeszke@man.poznan.pl  
Tomasz Parkoła  
tparkola@man.poznan.pl  
Poznańskie Centrum Superkomputerowo-Sieciowe

## Prostota, jakość i widoczność jako kluczowe czynniki budowy systemu SINUS

**Streszczenie:** Uniwersytety oraz placówki naukowe w Polsce wykorzystują biblioteki cyfrowe, repozytoria cyfrowe, bazy bibliograficzne oraz systemy informacji naukowej do promowania swoich osiągnięć badawczych. Jednym z trendów rozwojowych w tym zakresie jest próba konsolidacji rozwiązań i uruchomienia zintegrowanego środowiska promocji zasobów naukowych. W tym kontekście w roku 2015 Poznańskie Centrum Superkomputerowo-Sieciowe rozpoczęło współpracę z Biblioteką Politechniki Poznańskiej. Głównym celem tej współpracy była optymalizacja procesów zarządzania dorobkiem naukowych Politechniki Poznańskiej. W efekcie docelowe rozwiązanie miało zastąpić istniejące systemy działające już w bibliotece, zapewnić kompleksową informację dla użytkowników końcowych, zautomatyzować pracę związaną z gromadzeniem i udostępnianiem zasobów oraz dostarczyć mechanizmy elastycznego eksportu danych. W ramach przeprowadzonych prac wykonana została dogłębna analiza istniejących rozwiązań w zakresie oprogramowania. W rezultacie zdecydowano się na opracowanie nowego systemu dostosowanego do potrzeb Biblioteki Politechniki Poznańskiej. Warto podkreślić, że prace projektowe i rozwojowe polegały na dostarczeniu systemu prostego w obsłudze, z wbudowanymi funkcjami poprawiającymi/zapewniającymi jakość dostarczanych danych oraz systemu dobrze widocznego w wyszukiwarkach internetowych. Po czterech latach pracy nad systemem SINUS uważamy, że szczegółowa analiza potrzeb, elastyczne rozwiązania oraz prostota, jakość i widoczność prac naukowych są najważniejszymi czynnikami sukcesu przy budowaniu kompleksowych rozwiązań w zakresie promocji i dostępu do dorobku naukowego pracowników uczelni.

**Słowa kluczowe:** otwarte repozytoria, biblioteki cyfrowe, systemy informacji naukowej, Current Research Information System (CRIS)

SINUS jest systemem informacji naukowej, który ma również właściwości otwartego repozytorium, biblioteki cyfrowej oraz narzędzia do oceny dorobku naukowego pracowników uczelni. Prostota, jakość i widoczność są kluczowymi cechami tego systemu.

Wiele uczelni i instytutów korzysta z bibliotek cyfrowych lub repozytoriów w celu wymiany i promocji informacji na temat osiągnięć naukowych uczelni. Otwartość placówek badawczych stwarza możliwości ponownego wykorzystania informacji w Polsce i za granicą. Ponadto wiele instytucji dzieli się swoimi osiągnięciami naukowymi w bibliograficznych bazach danych lub poprzez systemy typu CRIS. Cechą takich systemów jest gromadzenie metadanych, niekiedy z linkami do treści. Bibliograficzne bazy danych wykorzystywane są w większości przypadków do gromadzenia, przetwarzania i prezentowania danych statystycznych według potrzeb własnych instytucji lub Ministerstwa Nauki i Szkolnictwa Wyższego. Wymienione systemy działają niezależnie, co prowadzi do powielania informacji i utrudnia użytkownikowi końcowemu dostęp do pełnej informacji o publikacji. Ponadto aktualizacja informacji w kilku systemach znacznie zwiększa nakład pracy redaktorów. Praca z wieloma narzędziami jest zazwyczaj niejednorodna, przez co często metadane nie są

spójne, tym samym jakość informacji na poziomie całej uczelni jest niska. Powyższe systemy nie pozwalają również na wprowadzenie informacji o nietypowych zasobach, takich jak projekty architektoniczne, dokumentacja techniczna czy oprogramowanie.

W 2015 r. Poznańskie Centrum Superkomputerowo-Sieciowe rozpoczęło współpracę z Biblioteką Politechniki Poznańskiej w celu opracowania i wdrożenia narzędzia optymalizującego proces zarządzania osiągnięciami naukowymi Politechniki Poznańskiej. Optymalizacja ta odbywa się zarówno na poziomie metadanych, jak i treści cyfrowych. Zgodnie z założeniami nowy system miał zastąpić cztery inne systemy, które były jednocześnie wykorzystywane na uczelni w podobnym celu, by zapewnić kompleksową informację dla użytkowników końcowych, zautomatyzować część pracy osób odpowiedzialnych za zbieranie i udostępnianie zasobów oraz wspierać połączenie z systemami zewnętrznymi (w szczególności w związku z krajowym systemem Polskiej Bibliografii Naukowej, który zobowiązuje polskie uczelnie do raportowania osiągnięć naukowych).

Pierwszym etapem prac była analiza istniejących krajowych i zagranicznych systemów informatycznych, które wykazywały cechy oraz funkcjonalności pomocne w budowie narzędzia dla Politechniki Poznańskiej. W przypadku systemów repozytorium zweryfikowano DSpace i E-Print. Okazało się jednak, że ich modele danych były zbyt ubogie do uruchomienia systemu informacji naukowej. W przypadku systemów podobnych do CRIS zweryfikowano rozwiązanie VIVO. Ten semantyczny system spełnia wymagania dotyczące instytucjonalnego repozytorium oraz samego systemu CRIS. Ponadto ma graficzny interfejs dostępny przez przeglądarkę internetową. Jednak podczas końcowej weryfikacji wewnętrznej skupionej na użyteczności, wydajności i funkcjonalności okazało się, że system nie spełnił oczekiwań uczelni. Po kilku miesiącach testów stało się oczywiste, że wydajność systemu znacząco spada wraz z wprowadzeniem tysięcy zasobów opatrzonych obszernymi opisami. Ponadto system nie był wyposażony w potrzebne funkcje detekcji duplikatów oraz tworzenia złożonych opisów bibliograficznych. Wyeliminowanie zaobserwowanych problemów w kontekście systemu VIVO uznano za zbyt czasochłonne w porównaniu z implementacją pełnego rozwiązania od nowa. W związku z tym pod koniec 2015 r. podjęto decyzję o zaprojektowaniu indywidualnego rozwiązania, które spełni wszystkie wymagania uczelni (repozytorium i funkcje CRIS) oraz zapewni ergonomię wprowadzania danych.

Drugi etap prac dotyczył zaprojektowania od podstaw systemu o nazwie SINUS, którego funkcjonalność koncentruje się na trzech głównych cechach: prostocie, jakości i widoczności. Priorytetami związanymi z prostotą są przejrzystość interfejsu portalu internetowego, szeroka gama opcji wyszukiwania oraz wygodny sposób przeglądania dostępnych zasobów. Użytkownicy nowoczesnych stron internetowych oczekują dobrze zaprojektowanej architektury informacji oraz przyjaznych, angażujących, ale bezinwazyjnych interakcji. Dlatego też w poszukiwaniu inspiracji i dobrych praktyk w zakresie użyteczności interfejsów przeanalizowano nowoczesne rozwiązania stron internetowych spoza dziedziny nauki.

Prostota to cecha wspierająca nawigację, podkreślająca hierarchię poszczególnych treści i ułatwiająca przeglądanie. Miarą sukcesu stron internetowych jest widoczność ich treści w wyszukiwarkach, dlatego w systemie SINUS zastosowano najlepsze praktyki optymalizacji dla wyszukiwarek internetowych (SEO), poprawiając widoczność strony w Google Search i innych. Hierarchia stron internetowych została podzielona na odrębne sekcje (ludzie, organizacje, publikacje, czasopisma, konferencje), a każdy element otrzymał swoją własną stronę. Dla każdej strony stworzono unikalne tytuły do wyświetlania krótkich, ale

opisowych informacji (np. prefiks strony z nazwą osoby/organizacji/konferencji lub tytułem publikacji/czasopisma). Tam, gdzie ma to zastosowanie i jest uzasadnione, użyto adresów URL ze słowami, które są istotne dla zawartości i struktury strony, a nie tylko identyfikatorów, które są trudne do odczytania/zrozumienia przez ludzi. Aby uczynić stronę systemu SINUS bardziej widoczną, stworzono również plik mapy strony. Plik mapy strony zawiera adresy URL do strony głównej, profile wszystkich publikacji oraz: czasopisma, w których wydano publikacje, konferencje, na których zaprezentowano publikacje, wszystkich autorów publikacji i wszystkie jednostki organizacyjne Politechniki Poznańskiej. Plik ten jest codziennie aktualizowany. W środowisku produkcyjnym wykorzystano Google Search Console do monitorowania obecności strony w wynikach wyszukiwania, optymalizacji strony pod kątem wyszukiwarek i rozwiązywania problemów z indeksowaniem.

Ostatnim etapem było usprawnienie i ułatwienie procesu wprowadzania danych. Utworzono formularze podzielone na etapy i wprowadzono wiele usprawnień w zakresie ergonomii ich wypełniania (np. listy ze słownikami, wykrywanie błędów, autouzupełnianie struktur danych). Zastosowano również mechanizm wykrywania duplikatów w ramach publikacji, profili pracowników i konferencji. SINUS ma dwa rodzaje wykrywania duplikatów. Pierwszy opiera się na mechanizmie nadawania unikalnych numerów, a drugi na zewnętrznych identyfikatorach, takich jak: ORCID, ResearcherID, Google Scholar, PBN ID (identyfikator osoby w Polskiej Bibliografii Naukowej) oraz DOI dla publikacji. W systemie SINUS autorzy mogą nadsyłać swoje publikacje we własnym zakresie (samoarchiwizacja), a dane te są weryfikowane i aktualizowane przez pracowników Biblioteki Politechniki. Opisane funkcje wprowadzają większą efektywność pracy bibliotekarzy, szybszy proces zbierania i udostępniania publikacji, ale przede wszystkim – wyższą jakość danych.

SINUS jest również wyposażony w funkcje statystyczne. Może być wykorzystywany do oceny osiągnięć naukowców oraz, w szerszej perspektywie, całych instytucji prowadzących działalność naukowo-badawczą. Ocena jakości wyników badań i charakterystyki instytucji opiera się na wytycznych Ministerstwa Nauki i Szkolnictwa Wyższego. Celem wytycznych jest wyłonienie najlepszej kadry naukowej i dydaktycznej, a także najlepszych instytutów i wydziałów oraz jeszcze szerzej – wiodących dyscyplin naukowych. W dłuższej perspektywie czasowej oznacza to lepsze finansowanie jednostek i wsparcie dla konkretnych obszarów badań. Władze uczelni mają szybki wgląd w prace badawcze całej jednostki naukowej poprzez generowanie danych do pliku XLS. Formularze Excel pozwalają na łatwe przetwarzanie danych statystycznych, co może być wykonane przez wykwalifikowanego pracownika. W ten sposób SINUS umożliwia również delegowanie zadań związanych z oceną osiągnięć instytutów na pracowników administracji i bibliotekarzy. Ocena osiągnięć opiera się na parametrach bibliometrycznych oraz kryteriach określonych w ustawie Ministra Nauki i Szkolnictwa Wyższego. SINUS przetwarza takie parametry, jak Impact Factor czy informację o indeksowaniu publikacji w Web of Science. Oprócz wspomnianych wskaźników zewnętrznych ministerstwo wprowadziło własne kryteria oceny osiągnięć w postaci listy oceniania czasopism, zasad punktacji monografii, konferencji naukowych oraz patentów. SINUS pozwala na wdrożenie wszystkich rodzajów wskaźników i kryteriów oceny osiągnięć naukowców.

Dalszy rozwój systemu SINUS będzie koncentrował się na potrzebach użytkowników końcowych. Ponieważ regulacje prawne dotyczące parametryzacji nie ograniczają się jedynie do publikacji naukowych, ale obejmują również prace typu: patenty, formularze wynalazków, prezentacje na konferencjach krajowych i międzynarodowych, system SINUS poszerzy swoje funkcjonalności w tym zakresie. Dodanie nowych funkcji, zwłaszcza w kontek-

ście ochrony danych naukowych, poszerzy spektrum możliwości i zaowocuje kompleksowym rozwiązaniem łączącym funkcje repozytorium cyfrowego, systemu CRIS oraz systemu DAM.

Doświadczenia zebrane w ramach prowadzonych prac rozwojowych nad systemem SINUS jednoznacznie wskazują, że:

- Należy testować istniejące rozwiązania i sprawdzać je pod kątem wymagań konkretnych placówek naukowych. Otwarte oprogramowanie nie oznacza, że jest najlepszym rozwiązaniem w danej dziedzinie, podobnie popularność konkretnego systemu w kraju.
- Cyfrowy ekosystem ogólnoswiatowej sieci wymusza zastosowanie prostych rozwiązań oraz dobrej jakości danych z odpowiednim indeksowaniem w wyszukiwarkach. System powinien angażować odbiorcę, prowadząc go do konkretnych zasobów, należy zatem wprowadzić dobre praktyki procesu User Centered Design, np. testować rozwiązania na docelowych grupach użytkowników.
- Techniczne podstawy systemu powinny być zbudowane w taki sposób, aby zmiany, które są nieuniknione, mogły być wprowadzane jak najsprawniej, np. poprzez zastosowanie elastycznych ram technicznych.

#### Bibliografia:

1. SMITH, M., BARTON, M., BASS, M., BRANSCHOFKY, M., McCLELLAN, G., STUVE, D., TANSLEY, R., WALKER, J. DSpace: An Open Source Dynamic Digital Repository. *D-Lib Magazine* [online]. 2003, Vol. 9, nr 1. [Dostęp 22.07.2019]. ISSN 1082-9873. Dostępny w: <http://www.dlib.org/dlib/january03/smith/01smith.html>.
2. TANSLEY, R., BASS, M., STUVE, D., BRANSCHOFKY, M., CHUDNOV, D., McCLELLAN, G., SMITH, M. The DSpace institutional digital repository system: current functionality. W: *2003 Joint Conference on Digital Libraries. Proceedings* [online]. IEEE, 20 June 2003. [Dostęp 22.07.2019]. ISBN 0-7695-1939-3. Dostępny w: <https://ieeexplore.ieee.org/document/1204846>.
3. BEAZLEY, M. Eprints Institutional Repository Software: A Review. *Partnership: the Canadian Journal of Library and Information Practice and Research* [online]. 2010, Vol. 5, nr 2. ISSN 1911-9593. Dostępny w: <https://journal.lib.uoguelph.ca/index.php/perj/article/view/1234/1874>.
4. ROSA, C.A., CRAVEIRO, O., DOMINGUES, P. Open Source Software for Digital Preservation Repositories: a Survey. W: *International Journal of Computer Science & Engineering Survey (IJCSES)* [online]. 2017, Vol. 8, nr 3. [Dostęp 22.07.2019]. ISSN 0976-2760. Dostępny w: <https://arxiv.org/ftp/arxiv/papers/1707/1707.06336.pdf>.