



OŚRODEK
PRZETWARZANIA
INFORMACJI
PAŃSTWOWY INSTYTUT BADAWCZY

POL-on

System Informacji o Nauce
i Szkolnictwie Wyższym

TOM I



redaktor naukowy
Marek Michajłowicz

**Systemy informatyczne
wspierające naukę i szkolnictwo wyższe**

**POL-on:
System Informacji o Nauce
i Szkolnictwie Wyższym
TOM I**

Redakcja naukowa:
Marek Michajłowicz



**OŚRODEK
PRZETWARZANIA
INFORMACJI**
PAŃSTWOWY INSTYTUT BADAWCZY

Systemy informatyczne wspierające naukę i szkolnictwo wyższe
POL-on: System Informacji o Nauce i Szkolnictwie Wyższym
TOM I

Redakcja naukowa:

Marek Michajłowicz

Redakcja techniczna:

Krystyna Fetera, Robert Siewiorek

Autorzy:

Artur Idzi: rozdział 3, 4
dr Jakub Kierzkowski: rozdział 3
Iwona Kucharska: rozdział 2
Marek Michajłowicz: rozdział 1, 2, 3, 4
Małgorzata Paszkowska: rozdział 2
Emil Podwysocki: rozdział 4
dr Piotr Rodzik: rozdział 1, 2
Andrzej Sadłowski: rozdział 3, 4
Antoni Sobkowicz: rozdział 4

Recenzent:

dr hab. Zenon A. Sosnowski, prof. Politechniki Białostockiej

Wydawca:

Ośrodek Przetwarzania Informacji – Państwowy Instytut Badawczy
al. Niepodległości 188b
00-608 Warszawa
e-mail: opi@opi.org.pl
www.opi.org.pl



© Copyright by Ośrodek Przetwarzania Informacji – Państwowy Instytut Badawczy
Warszawa 2021
Wszelkie prawa zastrzeżone

ISBN 978-83-63060-19-0 (całość)
ISBN 978-83-63060-20-6 (tom pierwszy)

Projekt okładki i opracowanie graficzne:

Karina Maszewska

Skład i łamanie:

dr Jakub Kierzkowski

Druk:

Drukarnia DSS Szczepan Szymański
ul. Wolska 108, 05-119 Wola Aleksandra

Szanowni Państwo,

nauka i szkolnictwo wyższe to niezwykle ważne obszary, których rozwój wpływa na funkcjonowanie całego społeczeństwa. Kapitał intelektualny naszego państwa jest wysoki, a w Ministerstwie Edukacji i Nauki (MEiN) stale podejmujemy działania, które wspierają polskich badaczy. Dbamy, aby transfer wiedzy pomiędzy różnymi jednostkami naukowymi i akademickimi przebiegał efektywnie. Inicjujemy także wiele przedsięwzięć, ułatwiających współpracę pomiędzy środowiskiem polskiej nauki a światem biznesu. Dzięki temu nasza gospodarka jest innowacyjna i stale wzrasta jej konkurencyjność na arenie międzynarodowej.



Do efektywnego transferu wiedzy oraz kreowania polityki naukowej i edukacyjnej w wymiarze krajowym i międzynarodowym potrzebne są wysokiej jakości źródła informacji. Wymusza to zastosowanie kompleksowych systemów informatycznych, które będą na bieżąco gromadzić dane i przekształcać je w użyteczne treści. Na szczęście w Polsce mamy jedno z najlepszych rozwiązań technologicznych, które pełni wymienione funkcje. Na zlecenie Ministerstwa Edukacji i Nauki, Ośrodek Przetwarzania Informacji – Państwowy Instytut Badawczy (OPI PIB) opracował system, który otrzymał nagrodę Eunis (European University Information Systems). Europejskie jury doceniło stałe udoskonalanie zakresu funkcjonalności i zwiększanie elastyczności całej architektury polskiego rozwiązania. Monografia, którą trzymają Państwo w ręku, poświęcona jest właśnie temu polskiemu, innowacyjnemu systemowi IT.

Zintegrowany System Informacji o Szkolnictwie Wyższym i Nauce POL-on to ponad 40 modułów, które stale są rozwijane przez ekspertów z OPI PIB. To innowacyjne narzędzie dostarcza obiektywnych informacji o nauce i szkolnictwie wyższym – studentom, kandydatom na studia, pracownikom nauki i przedsiębiorcom. W ramach POL-onu funkcjonuje m.in. ogólnopolskie repozytorium pisemnych prac dyplomowych, w którym przechowuje się obecnie powyżej 10 TB danych. Gromadzone są w nim pisemne prace dyplomowe i udostępniane Jednolitemu Systemowi Antyplagiatowemu (JSA). Częścią systemu jest też aplikacja SEDN, która służy do ewaluacji jakości nauki i szkolnictwa wyższego. W ramach POL-onu działa także Polska Bibliografia Naukowa. To innowacyjne narzędzie, opracowane i rozwijane przez ekspertów z OPI PIB, jest największym działającym systemem publicznym w Polsce, uwzględniając zakres zbieranych danych.

Pierwsze moduły POL-onu zostały uruchomione już w 2011 r. Świątujemy więc w tym roku jubileusz dziesięciolecia systemu. Oczywiście, dzisiejszy POL-on jest już zupełnie inny. Na przestrzeni lat znacznie się on rozwinął. Dzięki działaniom MEiN ewoluował system szkolnictwa wyższego i nauki, a wraz z nim również POL-on. Niniejsza publikacja w znakomity sposób przedstawia genezę systemu i jego główne cechy. Monografia zawiera także analizę porównawczą, która uwzględnia wiele analogicznych rozwiązań IT funkcjonujących w innych państwach. W szczegółowy sposób autorzy publikacji

przedstawiają także implementację nowych funkcjonalności, które wprowadzone zostały wraz z rozwojem polskiego prawodawstwa z zakresu nauki i szkolnictwa wyższego.

Warto także podkreślić, że POL-on dostosowany jest do oczekiwań kadry naukowej i akademickiej. Eksperti OPI PIB stale współpracują z przedstawicielami nauki z całej Polski. Prowadzą aktywny dialog i słuchają potrzeb osób, na co dzień odpowiedzialnych za obsługę POL-onu w polskich uczelniach i instytutach. Jak wskazuje tytuł niniejszej publikacji, system został zaprojektowany, aby wspierać polską naukę i szkolnictwo wyższe. Właśnie taką funkcję POL-on efektywnie realizuje.

Zachęcam Państwa do lektury monografii „Systemy informatyczne wspierające naukę i szkolnictwo wyższe. POL-on: System Informacji o Nauce i Szkolnictwie Wyższym”.

Wojciech Murdzek

Sekretarz Stanu,
Pełnomocnik Rządu do spraw reformy
funkcjonowania instytutów badawczych

Ministerstwo Edukacji i Nauki

SPIS TREŚCI

Wstęp	7
Rozdział 1. Sieć informacyjna	9
<i>Marek Michajłowicz, dr Piotr Rodzik</i>	
1.1. Geneza systemu	9
1.2. Rejestry danych	13
1.3. Istotne fakty na temat systemu	15
1.4. Statystyki dotyczące użycia systemu oraz zbieranych w nim danych	18
1.5. Wpływ systemu na otoczenie	22
Rozdział 2. Koncepcja systemu	27
<i>Iwona Kucharska, Marek Michajłowicz, Małgorzata Paszkowska, dr Piotr Rodzik</i>	
2.1. Analiza porównawcza systemów podobnej kategorii	27
2.1.1. Systemy typu CRIS	29
2.1.2. Systemy ewaluacji dorobku naukowego	33
2.1.3. Wybrane centralne systemy informacji o dorobku studentów	34
2.2. Obsługiwane procesy biznesowe	37
2.3. Kluczowe cechy i wymagania systemu	46
2.3.1. Model pojęciowy systemu	46
2.3.2. Bezpieczeństwo i ochrona danych osobowych	50
2.3.3. Rozwiązywanie konfliktów danych	51
Słownik pojęć	55
Spis ilustracji	57
Spis tabel	58
Bibliografia	59

WSTĘP

Zintegrowany System Informacji o Nauce i Szkolnictwie Wyższym POL-on stanowi unikalne w skali światowej rozwiązanie w zakresie systemów agregujących kluczowe informacje i procesy dotyczące szeroko rozumianej nauki i szkolnictwa wyższego w obrębie jednego państwa. Jego unikalność nie polega jednak na założeniu, że jest to jedyny w swoim rodzaju tego typu zbiór danych i systemów z tego obszaru, będących do dyspozycji instytucji państwowych¹. Można znaleźć wiele podobnych systemów. Praktycznie każde państwo posiada zbliżone zbiory do realizacji swoich podstawowych zadań związanych z kontrolą jakości i finansowania sfery nauki i szkolnictwa wyższego. Specyfika systemu POL-on polega na unifikacji danych do jednego specyficznego i wysoce ustrukturyzowanego modelu, którego podstawę stanowią akty prawne obligatoryjnie narzucone wszystkim instytucjom kształtującym szkolnictwo wyższe i naukę. Konsekwencją takiego podejścia jest bardzo wysoki stopień unifikacji i centralizacji zbiorów. Tego typu podejście w połączeniu z bardzo szerokim spektrum pokrycia danych oraz relatywnie krótkim czasem wytworzenia², spowodowały, że został on doceniony przez środowisko międzynarodowe i organizacje takie jak np. EUNIS³, która zrzesza instytucje naukowe zaangażowane w tworzenie rozwiązań IT dla sektora nauki i szkolnictwa wyższego w Europie i na świecie. Wyrazem tego było przyznanie systemowi POL-on nagrody EUNIS Elite Award w roku 2018⁴.

Niniejsza publikacja ma na celu przedstawienie głównych cech systemu i jego zastosowań w formie analizy porównawczej uwzględniającej inne tego typu systemy na świecie oraz opisanie doświadczeń i ich wyników w obszarze wdrażania tak rozległego przedsięwzięcia. Szczególna uwaga poświęcona tu została zwłaszcza technicznym aspektom w sferze praktyki wdrażania systemów informatycznych. Budowa systemu stanowiła wyzwanie, ponieważ trzeba było znaleźć rozwiązania wielu różnych problemów. Nie wszystkie z nich udało się rozstrzygnąć w sposób optymalny, niektóre wciąż wymagają nowego podejścia, dla większości – z punktu widzenia skuteczności projektu – znaleziono rozwiązania o charakterze unikatowym. Każde z tych doświadczeń stanowi ciekawy materiał do dalszych analiz i wymiany doświadczeń na gruncie naukowym.



¹ w tym przypadku od roku 2008 rolę taką pełniło Ministerstwo Nauki i Szkolnictwa Wyższego Rzeczypospolitej Polskiej, a od października roku 2020 przejęło ją Ministerstwo Edukacji i Nauki

² pierwsze moduły systemu zostały uruchomione w październiku 2011 roku, pół roku po rozpoczęciu projektu
³ www.eunis.org/about-us (dostęp: 18.06.2021)

⁴ www.eunis.org/blog/2018/06/12/2018-eunis-elite-award-winners (dostęp: 18.06.2021)

Publikacja podzielona została na następujące rozdziały:

- **Rozdział pierwszy** – zawiera syntetyczny opis genezy i struktury systemu. Stanowi on punkt wyjścia do bardziej szczegółowego przedstawienia kolejnych zagadnień;
- **Rozdział drugi** – skupia się na analizie porównawczej systemu na tle innych tego typu rozwiązań oraz definiuje problemy, dla których opis praktycznych rozwiązań został przedstawiony w kolejnych rozdziałach;
- **Rozdział trzeci** – koncentruje się na procesie tworzenia systemu jako opisie praktycznych doświadczeń związanych z metodykami zarządzania projektem oraz procesem wytwarzania oprogramowania;
- **Rozdział czwarty** – przedstawia techniczne aspekty i opisy doświadczeń na gruncie architektury całego systemu oraz jego poszczególnych komponentów.

Publikacja ma na celu nie tylko przedstawienie kluczowych faktów dotyczących systemu, ale przede wszystkim opis doświadczeń i zastosowanych rozwiązań, które mogą stanowić wartość dla innych osób uczestniczących w realizacji podobnych przedsięwzięć lub szukających inspiracji do ich pogłębiania na niwie naukowej (co stanowi główną intencję i nagrodę dla jej autorów).

Pojęcie „system POL-on” będzie używane do określenia całości rozwiązań informatycznych wchodzących w skład Zintegrowanego Systemu Informacji o Nauce i Szkolnictwie Wyższym w praktyce stanowiącego cały ekosystem systemów i usług szczegółowo wymienionych w rozdziale 1.

Nie wszystkie z podsystemów będą omawiane w sposób szczegółowy. Niektóre z nich zastępują na oddzielne opracowania. W praktyce najistotniejszym obszarem zainteresowania autorów, zwłaszcza w części dotyczącej opisów eksperymentów był główny system POL-on, stanowiący centralny zbiór rejestrów, Ogólnopolskie Repozytorium Prac Dyplomowych oraz Polska Bibliografia Naukowa.

Szczególny akcent w pracy poświęcony został opisowi doświadczeń związanych z migracją systemu do nowej wersji.

ROZDZIAŁ 1

SIEĆ INFORMACYJNA

Marek Michajłowicz
dr Piotr Rodzik

Poniższy rozdział ma na celu przedstawienie jak największej liczby faktów dotyczących systemu POL-on, by dać czytelnikowi jak najlepszą wiedzę na temat jego skali i znaczenia w opisywanej dziedzinie. Stanowi to też formę wprowadzenia w bardziej dogłębne zagadnienia opisujące doświadczenia z pracy nad systemem i przeprowadzone eksperymenty badawcze, przedstawione w kolejnych rozdziałach.

Celem opisu nie jest wyłącznie prezentacja danych i aspektów prawnych, ale też próba przedstawienia i określenia ich na gruncie społecznego wpływu systemu na wiele dziedzin, w których można zaobserwować jego oddziaływanie.

1.1. Geneza systemu

System POL-on powstał w latach 2010-2014 jako efekt projektu „Stworzenie systemu informacji o szkolnictwie wyższym”, współfinansowanego ze środków Unii Europejskiej w ramach Programu Operacyjnego Kapitał Ludzki (priorytet IV, działanie 4, poddziałanie 4.1.3). Liderem projektu było Ministerstwo Nauki i Szkolnictwa Wyższego (MNiSW). Projekt był realizowany przez trzech wykonawców: Ośrodek Przetwarzania Informacji – Państwowy Instytut Badawczy (OPI PIB), Index Copernicus International (ICI) oraz Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego (ICM UW). Nazwa systemu stanowi nawiązanie do pierwiastka chemicznego odkrytego przez sławną polską uczoną Marię Skłodowską-Curie⁵. Jednocześnie jest to także kombinacja części słowa „POL” (w domyśle: Polska) oraz „on” (ang. włączyć).

POL-on jest naturalnym następcą Bazy Wiedzy o Nauce Polskiej (BWNP), która powstała w OPI w 1998 roku i była utrzymywana jako kontynuacja bazy Nauka Polska. Baza Nauka Polska istniała od 1990 roku, a od roku 1999 była również bezpłatnie dostępna w sieci internet. W roku 2018 baza BWNP przestała istnieć, a jej zasoby zostały



⁵ rok 2011, czyli ten w którym powstał system, był również rokiem poświęconym wspomnianej uczonej

przejęte przez POL-on. W ten sposób w POL-onie można znaleźć dane, które były wprowadzone przed powołaniem tego systemu do życia⁶, przy czym dane pochodzące z BWNP traktowane są w POL-on jako dane archiwalne.

Ostatecznie, Ośrodek Przetwarzania Informacji PIB był autorem modułów⁷: Jednostki; Rejestr Uczelni Niepublicznych; Pracownicy Naukowi; Kierunki studiów; Studenci, Absolwenci, Doktoranci; Pomoc materialna; Sprawozdania finansowe; Plan rzeczowo-finansowy; Dotacje dla szkolnictwa wyższego; Dotacje statutowe; Nieruchomości i infrastruktura; Laboratoria i aparatura; Biblioteki; Inwestycje; Patenty i dokonania; Projekty naukowe; Działalność upowszechniająca naukę. Index Copernicus stworzyło podsystemy: System Ewaluacji Czasopism Naukowych i Technologicznych oraz System Ewaluacji Dorobku Naukowego (SEDN), a Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego było autorem podsystemów: Polska Bibliografia Naukowa (PBN), Polski Index Cytowań (POL-index) i Ogólnopolskie Repozytorium Prac Dyplomowych (ORPD).

Celem powstałego systemu było wspomaganie pracy Ministerstwa Nauki i Szkolnictwa Wyższego i innych ministerstw, a także Głównego Urzędu Statystycznego (GUS) oraz Centralnej Komisji do Spraw Stopni i Tytułów (CK) poprzez gromadzenie informacji (zarówno na poziomie jednostkowym, jak informacji zbiorczych) dotyczących ogólnie rozumianej polskiej nauki. Początkowo system obejmował głównie dane dotyczące szkolnictwa wyższego⁸ i był oparty na nowelizacji ustawy Prawo o szkolnictwie wyższym, opublikowanej w dniu 18 marca 2011 roku ([35]), oraz na aktach wykonawczych powstałych w oparciu o tę ustawę – rozporządzeniach: w sprawie tytułów zawodowych nadawanych absolwentom studiów, warunków wydawania oraz niezbędnych elementów dyplomów ukończenia studiów i świadectw ukończenia studiów podyplomowych oraz wzoru suplementu do dyplomu ([24]), w sprawie danych zamieszczanych w ogólnopolskim wykazie studentów ([25]), w sprawie centralnego wykazu nauczycieli akademickich i pracowników naukowych ([26]) i innych.

W latach 2012 i 2013 POL-on został rozbudowany o moduły służące raportowaniu przez jednostki naukowe informacji o własnym dorobku naukowym oraz posiadanym przez nie potencjale badawczym. Wykonano moduły: „Nieruchomości i infrastruktura”, „Laboratoria i aparatura”, „Biblioteki”, „Inwestycje”, „Patenty i dokonania”, „Projekty naukowe” i „Działania upowszechniające naukę”. Wykonano też moduły związane z finansowaniem jednostek naukowych: „Sprawozdania finansowe”, „Plan rzeczowo-finansowy”, „Dotacje dla szkolnictwa wyższego” oraz „Dotacje statutowe”. Reszta systemu była sukcesywnie przebudowywana wraz ze zmianami obowiązujących przepisów prawa (np. [28]) oraz decyzjami pracowników ministerstw interpretującymi zapisy znajdujące się w aktach prawnych.



⁶ przez ustawę Prawo o szkolnictwie wyższym ([35])

⁷ stanowiących główną część systemu POL-on

⁸ w pierwszej fazie rozwoju system zawierał: rejestr uczelni niepublicznych, wykaz jednostek szkolnictwa wyższego i jednostek naukowych, wykaz prowadzonych przez szkoły wyższe kierunków studiów oraz wykaz studentów, a także centralny wykaz nauczycieli akademickich i pracowników naukowych

POL-on wspomagał wybory do Centralnej Komisji do Spraw Stopni i Tytułów w 2012 roku – pierwsze w Polsce wybory całkowicie przeprowadzone drogą elektroniczną⁹.

System wspierał również procesy oceny jednostek naukowych przeprowadzone w 2013 roku (i na tych samych zasadach w latach późniejszych), będąc wówczas kontenerem do przekazywania danych pochodzących z innego programu¹⁰.

W 2014 roku zakończył się, współfinansowany z funduszy UE, projekt realizacji systemu, a kierownictwo projektu objął Marek Michajłowicz. W tym samym roku system POL-on został w pełni oddany do użytku i utrzymania.

W 2015 i 2016 roku POL-on został znacznie przebudowany. Rozszerzono zakres obsługiwanych przez system danych. Wiązało się to z publikacją w styczniu 2015 roku znowelizowanej ustawy o finansowaniu nauki ([36]) oraz – w czerwcu 2015 roku – rozporządzenia w sprawie Systemu Informacji o Nauce ([29]). W tych latach nastąpiło też przejście systemów PBN, POL-index, ORPPD¹¹, SEDN do utrzymania i rozwoju przez Ośrodek Przetwarzania Informacji-Państwowy Instytut Badawczy.

Ośrodek Przetwarzania Informacji powstał w roku 1990. Jego szczególna rola w stworzeniu Bazy Wiedzy o Nauce Polskiej (BWNP) oraz informatyzacja bardzo złożonego procesu obsługi strumieni finansowania nauki (OSF), budowy Systemu Wspomagania Wyboru Recenzentów (SWWR) opartego o złożone algorytmy rekomendacji recenzentów (poprzez *text mining*) oraz wsparcie procesów ewaluacji uczyniły go jedną z wiodących instytucji odpowiedzialnych za wdrażanie rozwiązań informatycznych dla obszaru nauki w Polsce. W ramach docenienia szczególnego wkładu Ośrodka Przetwarzania Informacji w pierwszej fazie tworzenia systemu POL-on i jego rosnącej roli jako instytucji informatyzującej obszar nauki i szkolnictwa wyższego, minister podjął decyzję o bezpośredniej legitymizacji jego roli w dalszym utrzymaniu i rozbudowie tego systemu – rola instytutu naukowego, jako głównego operatora technicznego systemu, została wpisana do znowelizowanej ustawy (Art. 4c ust.5 [36]). Ponadto Ośrodek posiadał właściwie rozwiniętą infrastrukturę informatyczną i kadrę pozwalającą na utrzymywanie, przebudowę i rozszerzanie zakresu systemu na oczekiwanym, odpowiednio wysokim, poziomie.

POL-on wspomagał również wybory do Centralnej Komisji do Spraw Stopni i Tytułów przeprowadzone w roku 2016, kiedy to procedury obsługujące proces wyborczy zostały w znaczący sposób zmodernizowane – więcej na temat przeprowadzenia wyborów w 2016 roku można znaleźć w wystąpieniu prezentowanym na kongresie EUNIS (European University Information Systems Organization) w 2019 roku i następnie opublikowanym w European Journal of Higher Education IT ([18]).

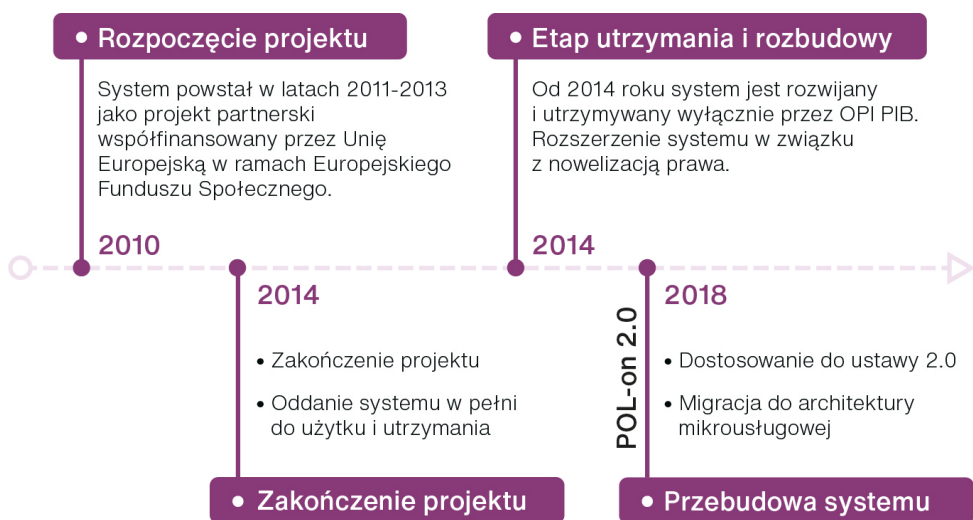


⁹ szczegóły prawne można znaleźć w [27]

¹⁰ specjalny program dedykowany ewaluacji jednostek naukowych

¹¹ Ogólnopolskie Repozytorium Prac Dyplomowych stworzone przez ICM z czasem zostało przekształcone w Ogólnopolskie Repozytorium Pisemnych Prac Dyplomowych (ORPPD)

W 2017 roku (i na tych samych zasadach w latach późniejszych) przeprowadzono ewaluację dorobku naukowego jednostek naukowych, raportując tylko te dane, które zostały wprowadzone do systemu POL-on¹². System został również dostosowany do potrzeb statystyki publicznej (integracja z GUS) w ramach projektu „Integracja Systemu Informacji o Szkolnictwie Wyższym (POL-on) z Głównym Urzędem Statystycznym”, finansowanego ze środków Programu Operacyjnego Kapitał Ludzki. Od 2017 roku POL-on stanowi jedyne oficjalne źródło danych dotyczących sprawozdawczości uczelni na potrzeby GUS oraz oficjalne źródło danych do algorytmu podziału dotacji dla szkolnictwa wyższego. Finalnie¹³ system POL-on obejmował prawie czterdzieści modułów poświęconych raportowaniu i przetwarzaniu danych dotyczących nauki i szkolnictwa wyższego w Polsce. Zawsze był w pełni oparty na przepisach obowiązującego prawa. Na każdym etapie tworzenia systemu użytkownicy – którymi byli zarówno pracownicy szkół wyższych i innych jednostek naukowych, jak i pracownicy ministerstw – byli szeroko wspierani przez: systemy dedykowanej pomocy, zespoły Helpdesk¹⁴, odpowiadające na pytania użytkowników i reagujące na nieprawidłowości w działaniu systemu, systemy szkoleń oraz filmy instruktażowe do wybranych modułów systemu, zamieszczane na stronach pomocy.



Ilustracja 1.1. Kluczowe fazy tworzenia i rozbudowy systemu

W 2018 roku system POL-on został wyróżniony nagrodą EUNIS Elite Award for Excellence w kategorii systemów informatycznych w szkolnictwie wyższym w Europie.



¹² podjęto taką decyzję, aby uniknąć rozbieżności pomiędzy danymi raportowanymi w ankiecie jednostki i w systemie POL-on

¹³ w pierwszej wersji

¹⁴ również podstawowy zespół zorganizowany w OPI PIB

Również w roku 2018 została uchwalona nowa ustawa Prawo o szkolnictwie wyższym i nauce ([37]) określająca przepisy przejściowe ([38]), pojawiły się też towarzyszące tym ustawom rozporządzenia ([31]) i ([32]), co spowodowało konieczność wprowadzenia daleko idących modyfikacji systemu POL-on. W efekcie podjęto decyzję o napisaniu od nowa wszystkich modułów systemu w innej, bardziej nowoczesnej, technologii (wersja 2).

Wszystkie moduły systemu są obecnie wykonywane, utrzymywane i modyfikowane przez OPI PIB, co gwarantuje ich jednolitość i pełną dostępność zarówno od strony kodu, jak dokumentacji technicznej. Wiele funkcji systemu w zakresie udostępnienia zawartych w nim informacji oraz ich integracji zostanie rozszerzonych w związku z realizacją projektu Zintegrowanego Systemu Usług dla Nauki – etap II¹⁵.

1.2. Rejestry danych

Zmiany wprowadzone w aktach prawnych przełożyły się na zmiany zawartości systemu informatycznego, który ma przetwarzać dane dotyczące szkolnictwa wyższego i nauki w Polsce. Obecnie system ten (zgodnie z Art. 342. ust. 3 Prawa o szkolnictwie wyższym i nauce [37]) ma obejmować:

- wykaz nauczycieli akademickich i innych osób prowadzących zajęcia,
- wykaz osób prowadzących działalność naukową oraz wykaz osób biorących udział w jej prowadzeniu,
- wykaz instytucji systemu szkolnictwa wyższego i nauki wraz z wykazem kierunków kształcenia zarówno na studiach wyższych jak i doktoranckich,
- informacje o prowadzonych (lub współprowadzonych) przez instytucję szkółach doktorskich,
- wykaz studentów,
- repozytorium pisemnych prac dyplomowych dla studiów pierwszego, drugiego stopnia oraz studiów jednolitych,
- wykaz osób ubiegających się o stopień doktora,
- bazę dokumentów w postępowaniach awansowych,
- bazę osób upoważnionych do podpisywania dokumentów w instytucjach naukowych,
- bazę dokumentów planistyczno-sprawozdawczych instytucji naukowych w Polsce.

Szczegółowy wykaz danych, które mają znaleźć się w systemie, określa ustawa [37] i wydane na jej podstawie Rozporządzenie w sprawie danych przetwarzanych w Zintegrowanym Systemie Informacji o Szkolnictwie Wyższym i Nauce POL-on [32].

W efekcie, w systemie mają być gromadzone następujące dane:

- dla pracownika naukowego i osoby prowadzącej zajęcia w szkole wyższej oprócz danych osobowych i danych o zatrudnieniu powinny się znaleźć informacje o nu-



¹⁵ radon.nauka.gov.pl (dostęp: 18.06.2021)

merze elektronicznego identyfikatora naukowca ORCID^{*}, posiadanych tytułach zawodowych, stopniach i tytułach naukowych, prowadzonych przez daną osobę zajęciach zarówno na studiach wyższych jak i w szkole doktorskiej, dyscyplinach naukowych, w których prowadzi ona badania naukowe, zaliczeniu w zatrudniającej instytucji do osób prowadzących badania naukowe, zdobytym poza uczelnią doświadczeniu (w przypadku innych osób prowadzących zajęcia), pełnionych w wyższych uczelniach funkcjach kierowniczych, orzeczonych względem danego pracownika karach dyscyplinarnych oraz o dorobku naukowym danego pracownika (autorstwie lub współautorstwie patentów, praw ochronnych i wyłącznych praw do ochrony roślin, autorstwie lub współautorstwie artykułów naukowych, monografii i rozdziałów w monografiach, prowadzeniu redakcji naukowych monografii oraz o autorstwie lub współautorstwie osiągnięć artystycznych);

- dla instytucji naukowych oprócz danych identyfikacyjnych podmiotu i danych adresowych (wraz z historią przekształceń instytucji) oraz informacji o osobach kierujących podmiotem w wykazie powinny znaleźć się informacje o prowadzeniu działalności poza siedzibą, posiadanych kategoriach naukowych, pozwoleniach na utworzenie studiów, prowadzonych studiach, opłatach pobieranych od studentów, prowadzonym kształceniu specjalistycznym, współczynnikach kosztowności, prowadzonych szkołach doktorskich, organach nadających stopnie naukowe i stopnie w zakresie sztuki, aparaturze naukowo-badawczej i infrastrukturze informatycznej (o wartości przekraczającej 500.000 zł), inwestycjach, nakładach na badania naukowe i prace rozwojowe, źródłach pochodzenia środków i wynikach finansowych oraz przychodach z tytułu komercjalizacji wyników działalności naukowej lub know-how związanego z tymi wynikami, a także informacje o tym czy dany podmiot prowadzi badania niejawne i informacje o prowadzonych przez pracowników podmiotu badaniach naukowych, wpływie działalności naukowej instytucji na funkcjonowanie społeczeństwa i gospodarki oraz realizowanych przez nią projektach obejmujących badania naukowe, prace rozwojowe lub upowszechnianie nauki, finansowanych ze środków, o których mowa w art. 365 pkt 7, 9, 11 i 12 ustawy, oraz przez instytucje zagraniczne lub organizacje międzynarodowe;
- dla studentów, oprócz danych osobowych i poziomu zaawansowania realizowanych kierunków studiów, system gromadzi informacje o rodzaju i okresie, na które przyznano świadczenia socjalne oraz stypendia naukowe ministra, a także szczegółowe dane dotyczące ukończenia studiów i przyznanego tytułu zawodowego;
- dla osób ubiegających się o stopień doktora oprócz danych osobowych i danych o zaawansowaniu na kolejnych kierunkach studiów doktorskich (o ile są one realizowane) powinny się znaleźć informacje o numerze ORCID oraz o dorobku naukowym danej osoby (autorstwie lub współautorstwie patentów, praw ochronnych i wyłącznych praw do ochrony roślin, autorstwie lub współautorstwie artykułów naukowych, monografii i rozdziałów w monografiach, prowadzeniu redakcji naukowych monografii oraz o autorstwie lub współautorstwie osiągnięć artystycznych), a także szczegółowe dane dotyczące uzyskania stopnia naukowego doktora, jak również, informacje o zatrudnieniu danej osoby;



• Termin wyjaśniony w słowniku

- dla repozytorium pisemnych prac dyplomowych wprowadza się metadane samej pracy i jej autora, informacje o promotorze (promotorach) pracy i jej recenzentach oraz o instytucji nadającej tytuł zawodowy oraz realizowanym przez studenta kierunku, a także informacje o egzaminie dyplomowym, którym zakończone było procedowanie danej pracy;
- dla dokumentów w postępowaniu awansowym podaje się dane osoby ubiegającej się o stopień lub tytuł naukowy oraz opinie recenzentów, a w przypadku osób ubiegających się o stopień doktora również treść rozprawy doktorskiej wraz z jej streszczeniem i datą złożenia;
- dla bazy osób upoważnionych do podpisywania dokumentów w instytucjach naukowych podaje się informacje o danej osobie (dane osobowe), cyfrowe odwzorowanie wzoru podpisu i parafy tej osoby oraz funkcje pełnione przez nią w instytucji oraz informacje o podpisywanych dokumentach: dyplomach ukończenia studiów, dyplomach doktorskich i dyplomach habilitacyjnych i odpisach tych dokumentów oraz zaświadczeniach o ukończeniu studiów i studiów podyplomowych;
- dla bazy dokumentów planistyczno-sprawozdawczych instytucji naukowych w Polsce podaje się: plany rzeczowo-finansowe (uczelni publicznych), roczne plany finansowe Centrum Łukasiewicz i instytutów Sieci Łukasiewicz, sprawozdania z wykonania planów rzeczowo-finansowych uczelni publicznych, sprawozdania i raporty z wykorzystania środków finansowych, o których mowa w art. 365 oraz roczne sprawozdania finansowe uczelni publicznych zbadane przez firmę audytorską.

1.3. Istotne fakty na temat systemu

Formą wprowadzenia w opisywaną problematykę systemu będzie przedstawienie kluczowych informacji, definiujących stosowany w dalszej części pracy aparat pojęciowy w zakresie jego wewnętrznej struktury oraz środowiska, z którym kooperuje. Ponieważ przedmiotem opisu jest system informatyczny, w praktyce wymienione zostaną wraz z ilustracjami wszystkie systemy znajdujące się w ramach ekosystemu¹⁶. Dodatkowo przedstawione zostaną statystyki dotyczące użycia systemu oraz jego zawartości informacyjnej, wraz z krótkim podsumowaniem.

W skład Zintegrowanego Systemu Informacji o Nauce i Szkolnictwie Wyższym POL-on wchodzi:

- POL-on (System centralny zawierający zbiór głównych rejestrów)¹⁷ – główny system wokół, z którym komunikują się pozostałe aplikacje;
- Polska Bibliografia Naukowa (PBN)¹⁸ – system zawierający dane o dorobku publikacyjnym polskich naukowców;



¹⁶ jako współpracę uznajemy wymianę danych

¹⁷ polon.nauka.gov.pl (dostęp: 18.06.2021)

¹⁸ pbn.nauka.gov.pl (dostęp: 18.06.2021)

- System Oceny Osiągnięć Naukowych (SEDN)¹⁹ – system obsługujący proces ewaluacji jednostek naukowych ministra nauki i szkolnictwa wyższego, wykorzystujący dane z systemu POL-on oraz PBN;
- Polski System Cytowania (POL-Index) – system zbierający informacje o artykułach opublikowanych w czasopiśmie naukowych do roku 2015, mający na celu wyliczenie tzw. Polskiego Współczynnika Wpływu²⁰. System został wyłączony z użytku w roku 2020,
- Ogólnopolskie Repozytorium Rozpraw (ORPPD)²¹ – repozytorium, w którym składowane są pisemne prace dyplomowe polskich absolwentów od roku 2009,
- Ekonomiczne Losy Absolwentów (ELA)²² – portal prezentujący wyniki realizowanych co roku badań losów studentów, absolwentów oraz doktorantów, dotyczących ich sytuacji na rynku pracy. Badanie bazuje na anonimizowanych danych z systemu POL-on oraz Zakładu Ubezpieczeń Społecznych. Wyniki prezentowane są w formie szeregu interaktywnych raportów oraz infografik;
- Elektroniczny System Wyborczy do Centralnej Komisji ds. Stopni i Tytułów Naukowych²³ – system obsługujący elektroniczny proces wyborczy wśród naukowców polskich do kluczowego gremium odpowiedzialnego za nadawanie i pilnowanie jakości polskich stopni i tytułów naukowych;
- Nauka Polska²⁴ – to najstarsza baza danych od 1999 roku udostępniana bezpłatnie w sieci internet. Zwiera dane o nauce polskiej w zakresie instytucji, osób, prac naukowych, publikacji oraz konferencji. Od roku 2018 jest częścią systemu POL-on w wymiarze archiwalnym;
- Portal RADON²⁵ – system raportowy RAD-on (Raporty, Analizy, Dane) – rozbudowana platforma analiz oraz udostępniania danych publicznych do dalszego wykorzystania (open data);

Warto już na wstępie zwrócić uwagę na dualizm pojęć w odniesieniu do POL-on rozumianego jako ekosystem wymienionych wyżej aplikacji oraz jego głównej aplikacji jako jedynej korzystającej z adresu internetowego, nawiązującego w sposób bezpośredni do tej nazwy. W potocznym podejściu bardzo często używa się rozróżnienia na tzw. „duży POL-on”, rozumiany jako cały ekosystem współpracujących ze sobą aplikacji i usług, oraz „mały POL-on”, czyli główną aplikację zawierającą największy zbiór rejestrów. Podejście to jest pochodną ewolucji systemu w czasie i rozbudowy go o kolejne autonomiczne elementy oraz brakiem jednoznacznych rozgraniczeń dla poszczególnych części ekosystemu na gruncie ustawowym²⁶.

Strukturę całego ekosystemu obrazuje Ilustracja 1.2.



¹⁹ sedn.nauka.gov.pl (dostęp: 18.06.2021)

²⁰ pci-www.vls.icm.edu.pl/polindex/info/polski-wspolczynnik-wplywu (dostęp: 18.06.2021)

²¹ polon.nauka.gov.pl/orpd (dostęp: 18.06.2021)

²² ela.nauka.gov.pl (dostęp: 18.06.2021)

²³ obecnie Rada Doskonałości Naukowej

²⁴ nauka-polska.pl (dostęp: 18.06.2021)

²⁵ radon.nauka.gov.pl (dostęp: 18.06.2021)

²⁶ systemy takie jak PBN, SEDN, RADON nie są zdefiniowane na gruncie prawnym, lecz stanowią raczej efekt decyzji projektowych



- Studia.gov.pl²⁷ – wyszukiwarka kierunków i danych dla kandydatów na studia;
- Inwentorium²⁸ – system rekomendujący informacje z obszaru nauki oraz biznesu. Baza wykorzystuje częściowo informacje z POL-on oraz Nauki Polskiej w obszarze danych o projektach, przedsiębiorstwach innowacyjnych, instytucjach naukowych oraz ekspertach i konferencjach;
- JSA²⁹ – jednolity system antyplagiatowy, wykorzystywany na mocy prawa przez wszystkie podmioty szkolnictwa wyższego. System wykorzystuje zbiory

²⁹ isa.opi.org.pl (dostęp: 18.06.2021)

ogólnopolskiego repozytorium pisemnych prac dyplomowych, wybrane konta użytkowników oraz strukturę jednostek i kierunków studiów z POL-on;

- OSF³⁰ – system do obsługi strumieni finansowania, wymieniający z systemem POL-on³¹ informacje na temat instytucji nauki i szkolnictwa wyższego, uruchomionych projektów naukowych oraz publikacji;
- ORCID³² – system rozwijany w ramach projektu ujednolicania danych o dorobku naukowców. Zostanie on bardziej szczegółowo omówiony w rozdziale drugim, przy okazji analizy porównawczej systemów przetwarzających dane o dorobku publikacyjnym;
- Empatia³³ – rządowy portal do rejestracji wniosków o świadczenia. W zakresie świadczeń socjalnych, w których brany jest pod uwagę status studenta. Informacje wymieniane są z systemem POL-on;
- REGON³⁴ – prowadzona przez Główny Urząd Statystyczny wyszukiwarka podmiotów gospodarki narodowej, z którą komunikuje się rejestr instytucji nauki i szkolnictwa wyższego systemu POL-on;
- TERYT³⁵ – prowadzony przez Główny Urząd Statystyczny słownik podziału terytorialnego stosowany na terenie Rzeczypospolitej Polskiej;
- Login.gov.pl³⁶ – utrzymywane przez rząd RP rozwiązanie pojedynczego logowania (ang. *Single Sign On*) do systemów domeny państwowej. Jest to komponent uwierzytelniający architektury informatycznej państwa, z którym na moment wydania publikacji zintegrowany jest system RAD-on w zakresie usługi danych obywatela, prezentującej szczegółowe wyniki na temat danej osoby, przetwarzanych w systemach źródłowych, w tym głównie POL-on;

1.4. Statystyki dotyczące użycia systemu oraz zbieranych w nim danych

W celu przedstawienia skali użycia systemu oraz rozmiaru i zakresu przechowywanych w nim danych warto odwołać się do niektórych statystyk.

Ważnym wskaźnikiem określającym stopień użycia systemu oraz jego skalę są dane dotyczące liczby kont unikalnych użytkowników wraz ze wskazaniem typów jednostek, w których one występują.



³⁰ osf.opi.org.pl (dostęp: 18.06.2021)

³¹ w dużym stopniu poprzez bazę nauki polskiej

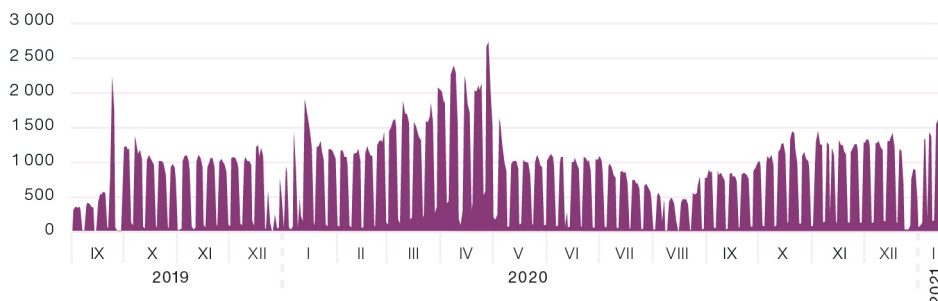
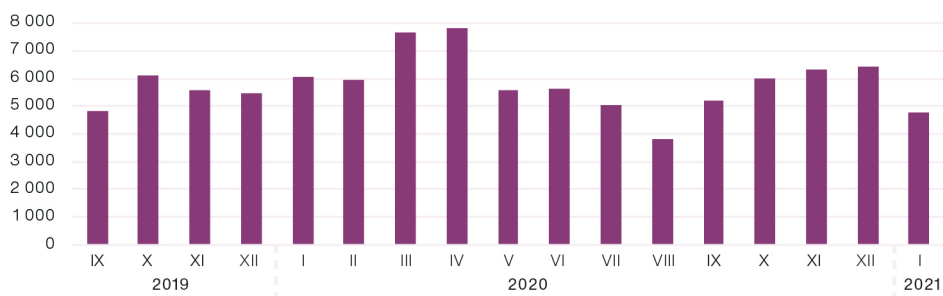
³² orcid.org (dostęp: 18.06.2021)

³³ empatia.mpips.gov.pl (dostęp: 18.06.2021)

³⁴ wyszukiwarkaregon.stat.gov.pl/appBIR/index.aspx

³⁵ eteryt.stat.gov.pl/eTeryt

³⁶ login.gov.pl

23 509liczba
aktywnych użytkowników**781**średnia dzienna liczba
aktywnych użytkowników**Dzienna liczba aktywnych użytkowników****Liczba aktywnych użytkowników w danym miesiącu****Ilustracja 1.3.** Statystyka aktywnych użytkowników systemu POL-on

Szczególnie istotne jest przytoczenie statystyk dotyczących danych zawartych w systemie, żeby zobrazować jego skalę i zakres.

Z punktu widzenia instytucjonalnego system stanowi istotne źródło do zbierania i wymiany danych dla prawie ośmiuset podmiotów nauki i szkolnictwa wyższego w Polsce.

Tabela 1.1. Statystyki POL-on według stanu na 30.09.2020 – wykaz instytucji

Typ instytucji	Liczba
Uczelnia niepubliczna	376
Instytut naukowy	268
Uczelnia publiczna	134
Uczelnia kościelna	16
Instytucja nadzorująca lub monitorująca	14

W obszarze danych o zatrudnieniu i dorobku pracowników naukowych i dydaktycznych z dziedziny nauki i szkolnictwa wyższego zakres danych w dużej części pokrywa funkcjonalności systemów kadrowych dla wspomnianych wyżej instytucji.

Tabela 1.2. Statystyki POL-on według stanu na 30.09.2020 – wykaz pracowników

Rodzaj pracownika	Liczebność
Nauczyciel akademicki	139 453
Inna osoba prowadząca zajęcia	48 866
Pracownik badawczo-techniczny lub inżynieryjno-techniczny	23 045
Pracownik naukowy	16 308
Pracownik naukowo-techniczny lub inżynieryjno-techniczny	13 744

Jednym z najcenniejszych zasobów systemu jest na bieżąco aktualizowany rejestr danych o studentach i absolwentach w Polsce. Zbierane oficjalnie od roku 2011 dane urosły już do rozmiaru prawie 10 milionów unikalnych rekordów. Dzięki integracji danych system po raz pierwszy umożliwił jednoznaczne ustalenie liczby osób faktycznie studiujących oraz oddzielił to od pojęcia „podjęcia studiów”, co do tej pory było utożsamiane.

Tabela 1.3. Statystyki POL-on według stanu na 30.09.2020 – wykaz studentów i doktorantów

Typ danych	Liczebność
Liczba studentów	3 870 796
Liczba studiowań	6 826 246
Liczba absolwentów	1 958 743
Liczba doktorantów i uczestników szkół doktorskich	94 740

Od 2014 roku system został rozszerzony o dane o dorobku naukowym. Zakres ten podlegał największym zmianom z uwagi na strategiczne decyzje w obszarze ewaluacji jednostek naukowych i związane z tym zmiany w obowiązującym prawie. Zakres zbieranych danych nie ma ciągłości w tym obszarze, ponieważ determinuje go przeprowadzana co cztery lata ewaluacja jednostek naukowych oraz przyjęte kryteria.

Tabela 1.4. Statystyki POL-on według stanu na 30.09.2020 – osiągnięcia naukowe

Typ danych	Liczebność
Publikacje	1 360 175
Projekty naukowe	81 302
Osiągnięcia artystyczne	124 473
Zgłoszenia patentowe	22 841
Patent otrzymany	14 771
Prawo ochronne prawo z rejestracji	2 790
Wykorzystane prawa autorskie	2 195
Wyłączne prawo do odmiany roślin	151

System stanowi też rozległe źródło danych dokumentowych oraz formularzy. Obsługuje on większość obowiązków sprawozdawczych podmiotów nauki i szkolnictwa wyższego.

Tabela 1.5. Statystyki POL-on według stanu na 30.09.2020 – baza dokumentów planistyczno-sprawozdawczych

Typ dokumentu	Liczebność
Sprawozdanie z planu rzeczowo finansowego	2 932
Plany rzeczowo finansowe	2 931
Sprawozdania finansowe	1 289
Formularz GUS S10	2 553
Formularz GUS S11	2 318
Formularz GUS S12	2 531
Sprawozdanie mobilności GUS	547
Ankiety rekrutacyjne EN-1	7 519
Sprawozdanie z działalności uczelni	56 634
Ankieta jednostki	2 061

Warto przywołać przy tej okazji również niektóre liczby prezentujące rozmiar wspomnianych wyżej formularzy, ponieważ obrazują one skalę wpływu systemu w zakresie cyfryzacji niektórych procesów w podmiotach szkolnictwa wyższego. Odwołując się do liczby stron niektórych z nich można określić jak wiele pracy wymagało ich przygotowanie przed wprowadzeniem systemu informatycznego do ich obsługi. Szczególnie znamienna jest tutaj statystyka formularza S-10 Sprawozdanie o studiach wyższych składanego do Głównego Urzędu Statystycznego. Przed jego wdrożeniem w systemie POL-on wykonywane były symulacje bazujące na danych z roku 2014, w których wypełniony formularz sprawozdania dla dużych uczelni, takich jak np. Uniwersytet Warszawski lub Uniwersytet Jagielloński, przekraczał ponad 160 stron. Jeszcze większy rozmiar

osiągał, składany co roku, formularz sprawozdania z działalności uczelni, który przed przeniesieniem do systemu w roku 2011, liczył nawet około 500 stron³⁷.

1.5. Wpływ systemu na otoczenie

Polskie uczelnie i jednostki naukowe są zobowiązane przez prawo do przekazywania różnych informacji i przechowywania ich w systemie centralnym. Informacje dotyczą działalności dydaktycznej i naukowej tych instytucji, a mianowicie studentów, pracowników naukowych i nauczycieli akademickich, stopni i tytułów naukowych, wydziałów i studiów, prac dyplomowych, projektów, patentów, publikacji, konferencji, nagród, laboratoriów i ich wyposażenia, nieruchomości, inwestycji itp.³⁸. Dane mogą być przenoszone do systemu za pomocą narzędzi masowego importu plików XML oraz RESTful API³⁹ lub wprowadzane ręcznie poprzez interfejsy webowe. Niezależnie od metody dostarczania danych, wszystkie rekordy są walidowane zgodnie z zasadami wynikającymi z prawa i najlepszych praktyk. Zasady te są kluczowym elementem architektury systemu, ponieważ chronią i weryfikują w czasie rzeczywistym zgodność niezależnych polityk uczelni i instytutów badawczych z obowiązującym prawem, zwłaszcza w zakresie organizacji studiów i zatrudniania pracowników naukowych.

POL-on działa w heterogenicznym środowisku informacyjnym, co utrudnia integrację w jednym centrum danych. Ogólną ideę środowiska demonstruje grafika przedstawiona na ilustracji wcześniej. Podstawowy system obejmuje około czterdziestu modułów przetwarzających dane i realizujących zadania biznesowe. Najważniejsze moduły są wyliczone na środku ilustracji. Na górze widzimy uczelnie i jednostki naukowe, które najczęściej przekazują swoje dane do POL-onu. Zwykle posiadają one unikalne systemy informacyjne w różnych technologiach i architekturach, takich jak ERP (ang. *enterprise resource planning – ERP*)³⁹, systemy antyplagiatowe, repozytoria, systemy do zbierania informacji o dorobku naukowym i tym podobne. Integrują się one z systemem POL-on poprzez interfejsy omówione powyżej. Dodatkowo POL-on jest ściśle sprzężony z sześcioma współpracującymi ze sobą systemami tworzącymi cały ekosystem informacyjny nauki i szkolnictwa wyższego w Polsce.

Forma wdrażania systemu ma na gruncie strategicznym swoją specyfikę wyróżniającą go na tle podobnych tego typu przedsięwzięć w innych krajach. Dla porównania można podać model zastosowany w Niemczech lub Austrii, gdzie centralne organy zdecydowały się na zakup rozwiązań typu ERP i CRM (ang. *Customer Relationship Management*)⁴⁰ opartych o technologię konkretnego dostawcy (SAP)⁴¹ i sfinansowały wdrożenie ich we



³⁷ wynikało to w dużej mierze z podawania w nim informacji o charakterze opisowym

³⁸ szczegółowy zakres danych omówiony został w podrozdziale „Rejestry danych”

³⁹ kategoria systemów służących do efektywnego zarządzania zasobami przedsiębiorstwa

⁴⁰ system informatyczny, który automatyzuje i wspomaga procesy związane z relacjami klientów z organizacją

⁴¹ SAP SE wcześniej SAP AG, niem. *Systeme, Anwendungen und Produkte in der Datenverarbeitung*, ang. *Systems Applications and Products in Data Processing* – niemieckie przedsiębiorstwo informatyczne, będące dostawcą oprogramowania biznesowego ERP, na postawie sap.com (dostęp: 18.06.2021)

- kontrolnej – kluczową funkcją systemu jest kontrola zjawisk zachodzących w obszarze nauki i szkolnictwa wyższego, zwłaszcza wykrywanie nieprawidłowości i nadużyć.

W procesie pozyskiwania danych do centralnych repozytoriów kluczową kwestią są zagadnienia związane z kontrolą jakości danych. Najczęstsze problemy z ich przekazywaniem wynikają z następujących przyczyn [1]:

- uczelnia nie ma jednego zintegrowanego systemu informatycznego ze swoją centralną bazą danych.
- uczelnia ma zintegrowany system, ale ten system nie wspiera w wystarczającym stopniu:
 1. kontroli jakości gromadzonych danych;
 2. weryfikacji danych pod kątem zgodności ciągłości historycznej;
 3. archiwizowania ważnych danych historycznych z blokowaniem zmian;
 4. analizy spójności danych pod kątem eksportu do POL-onu;
 5. eksportu do POL-onu;
 6. uczelnia nie wdrożyła efektywnych procedur w zakresie:
 - a. poprawy jakości danych podlegających sprawozdawczości;
 - b. realizacji procesu eksportu, w tym podziału zadań między jednostki;
 - c. nadzoru i egzekwowania wykonania zadań.

Przyjęto rozwiązania, mające na celu dostosowanie wewnętrzne podmiotów zasilających system centralny do wymogów związanych z ich aktualizacją i kontrolą jakości stanowiły istotny element determinujący ich wewnętrzną organizację. W wielu przypadkach okazało się, że stosowane formy kształcenia lub inne informacje znajdujące się w systemach źródłowych były niezgodne z obowiązującą wykładnią prawną.

W wymiarze strategicznym powstanie systemu przyczyniło się do wielu korzyści, takich jak:

- ograniczenie lub wyeliminowanie zjawiska nadużyć w obszarze stypendialnym – bez centralnej bazy i odpowiednich procesów monitorujących byłoby to niewykonalne;
- ograniczenie zjawiska plagiatyzmu – z pewnością żaden kraj nie może pochwalić się całkowitym sukcesem w tym obszarze, ale wprowadzenie jednolitej bazy plików prac dyplomowych (ORPPD) oraz kooperującego systemu, jakim jest Jednolity System Antyplagiatowy (JSA) wyposażyło uczelnie i instytucje nadzorujące w skonsolidowany i kompleksowy system kontroli tego zjawiska, oparty o nowoczesne technologie;
- uwiarygodnienie statystyk związanych z edukacją i nauką. Szczególnie widocznym przykładem były dane raportowane przez uczelnie w zakresie liczby studentów do roku 2015, czyli momentu gdy system POL-on zaczął przejmować obowiązki raportowe z tego obszaru;
- standaryzacja metod ewaluacji naukowej oraz podziału środków – system stanowi praktyczną implementację zunifikowanych norm i pojęć wynikających z prawa, co wraz z praktycznym stosowaniem jednolitych metodologii obliczeń, stanowi element ewaluacji i podziału środków publicznych poprzez zdefiniowane algorytmy;

- ograniczenie biurokracji związanej z przygotowaniem i wielokrotną dystrybucją dokumentów w formie papierowej na potrzeby organów administracji centralnej – szczególnym przykładem skuteczności systemu jest całkowite wyeliminowanie formularza Sprawozdania z Działalności Uczelni (tzw. sprawozdania rektora), którego rozmiar dla dużych uczelni jeszcze kilka lat temu obejmował około 1000 stron szczegółowych danych i statystyk zebranych z uczelni z całego roku akademickiego;
- minimalizacja opóźnień w procesie analizy zjawisk zachodzących w nauce i szkolnictwie wyższym – w poprzednich procesach istotny był w zasadzie tylko stan na koniec roku kalendarzowego, według którego wyliczane były dane na potrzeby kluczowych procesów związanych ze statystyką publiczną oraz algorytmami podziału subwencji;
- minimalizacja kosztów – system jest w całości finansowany przez Ministerstwo Nauki i Szkolnictwa Wyższego⁴³. Uczelnie, instytucje nauki i organy nadzorujące nie ponoszą żadnych dodatkowych kosztów związanych z jego funkcjonowaniem. W przypadku wielu podmiotów było to jedyne narzędzie informatyczne wspomagające również procesy kształcenia i raportowania dorobku⁴⁴;
- ujednolicenie metodologii zbierania danych oraz wyliczania wskaźników do wszelkiego typu rankingów. Jedną z głównych motywacji stojących u podstaw wdrożenia systemu było wsparcie m.in. tego procesu zwłaszcza w kontekście zwiększenia „widoczności” (ang. *visibility*) polskich uczelni⁴⁵ w rankingach międzynarodowych. Obecnie dane z POL-on są głównym źródłem informacji w kluczowych rankingach organizowanych w Polsce⁴⁶;
- zapewnienie możliwości kontroli i weryfikacji danych dotyczących poszczególnych osób. System zawiera szczegółowy zbiór danych osobowych dotyczących przede wszystkim studentów, kadry naukowo-dydaktycznej oraz doktorantów. Informacje te dają możliwość weryfikacji wielu procesów, związanych m.in. z migracjami międzynarodowymi oraz weryfikacją statusów, które mają wpływ na bezpieczeństwo państwa. Informacja w wymiarze osobowym daje też możliwość dokładnego określenia rzeczywistej liczby osób (studentów) w stosunku do liczby zdarzeń (studiowań). W przeszłości, przy braku integracji zbiorów statystyki te były utożsamiane, co prowadziło do różnego typu błędnych wniosków oraz związanych z nimi decyzji⁴⁷;
- wprowadzenie rozbudowanej historii danych – kluczowym wymaganiem w stosunku do systemu jest jego możliwość uzyskania informacji na temat zmian historycznych (rozliczalność) oraz możliwość ich przekrojowych analiz na przestrzeni czasu. W wielu systemach kooperujących z POL-on aspekt ten nie był tak mocno wykorzystywany. Wiele procesów skupiało się raczej na bieżącej rejestracji zjawisk, a ich wersjonowanie odbywało się w dużej mierze w oparciu o tzw. log techniczny składnicy danych. Wdrożenie systemu umożliwiło, dzięki unifikacji modelu pojęciowego, przekrojowe analizy w czasie oraz weryfikację wszelkich zjawisk występujących w ujęciu historycznym;



⁴³ Obecnie Ministerstwo Edukacji i Nauki

⁴⁴ w szczególności niedużych uczelni niepublicznych

⁴⁵ intvp.edu.pl (dostęp: 18.06.2021)

⁴⁶ takich jak m.in. ranking fundacji Perspektywy, Związek Liderów Sektora Usług Biznesowych, Rzeczpospolita

⁴⁷ najczęstszym zjawiskiem było zawyżanie liczby ogólnej studentów w Polsce

- stworzenie szeregu usług i narzędzi informatycznych dla kandydatów na studia⁴⁸, naukowców, przedsiębiorców⁴⁹ oraz obywateli chcących pozyskać obiektywne informacje na temat nauki i szkolnictwa wyższego w formie danych oraz analiz⁵⁰.



⁴⁸ studia.gov.pl (dostęp: 18.06.2021)

⁴⁹ system rekomendujący inventorum.opi.org.pl (dostęp: 18.06.2021)

⁵⁰ [portal radon.nauka.gov.pl](https://portal.radon.nauka.gov.pl)

ROZDZIAŁ 2

KONCEPCJA SYSTEMU

Iwona Kucharska
Marek Michajłowicz
Małgorzata Paszkowska
dr Piotr Rodzik

Poniższy rozdział skupia się na analizie systemów informatycznych z szeroko rozumianej domeny nauki i szkolnictwa wyższego, mogących stanowić punkt odniesienia dla systemu POL-on. Analiza ta bazuje na założeniu, że porównaniu powinny podlegać systemy, mające podobną skalę, czyli zbierające dane przynajmniej z obszaru danego państwa oraz formę realizacji jako przedsięwzięcia nadzorowanego i wspieranego przez organy centralne. Tylko takie porównanie wydaje się, zdaniem autorów, uzasadnione. Stanowi ono punkt wyjścia do przedstawienia kluczowych procesów obsługiwanych przy pomocy opisywanego rozwiązania oraz jego kluczowych cech determinujących określone podejście do niektórych kwestii technicznych oraz architektury.

2.1. Analiza porównawcza systemów podobnej kategorii

Nie ma jednego modelu budowy systemów centralnych o dorobku naukowym i dydaktycznym w skali poszczególnych państw. W praktyce nie wszystkie organy krajowe dysponują lub chcą dysponować takimi zbiorami. Instytucje nauki posiadają z reguły bardzo dużą autonomię w zakresie formy prowadzenia działań naukowych i dydaktycznych, co często utrudnia lub wręcz uniemożliwia zbyt daleko idącą centralizację i unifikację. Jest ona jednak niezbędna do prowadzenia skoordynowanych działań na wyższym poziomie. Element różnicujący stanowi w istotniej mierze jej skala i forma realizacji. W pewnym uproszczeniu możemy zidentyfikować modele oparte na rozwiązaniach rozproszonych sieci (CRIS[•], repozytoria), dla których tworzone mogą być repozytoria oferujące infrastrukturę i interfejsy komunikacyjne dla współpracujących instytucji naukowych. Przykładem takiego podejścia są chociażby Cyfrowe archiwa i repozytoria akademickie DABAR w Chorwacji [6] oraz systemy do obsługi bibliotek (*Library Management System* (LMS)), korzystające z protokołu OAI-PMH[•], opracowanego standardu dedykowanego do zbierania danych do repozytoriów [19]. W Norwegii w formule podobnej do stosowanej w warunkach polskich, czyli poprzez scedowanie budowy rozwiązań informatycznych przez norweskie ministerstwo nauki do podległego Uniwersyteckiego Centrum Technologii Informacyjnych w Oslo dysponującego odpowiednim zapleczem

oraz know-how, utworzone zostało centralne repozytorium typu CRIS do zastosowania krajowego o nazwie CRISTIN⁵¹. Kolejną architekturą systemu informatycznego dla nauki jest chmura, w której infrastruktura współpracujących instytucji jest zintegrowana z platformą programową. Przykładem takiego podejścia jest Sciebo – chmura kampusowa dla Nadrenii Północnej-Westfalii w Niemczech [40] lub platforma interoperacyjności Higher Education Institutions Closer wprowadzona w Portugalii [21].

Po drugiej stronie spektrum stosowanych rozwiązań znajdują się scentralizowane systemy takie jak POL-on [14] lub Norweski Wspólny System Studencki⁵².

Musimy zauważyć, że interesariusze systemu to nie tylko decydenci, ale także kierownicy badań, administratorzy, rady badawcze, organizacje zajmujące się transferem technologii, media i ogół społeczeństwa. Wszyscy potrzebują łatwego i otwartego dostępu do szerokich informacji naukowych, które mogą być gromadzone w systemie centralnym [4]. Taki system może również być odpowiedzią na rosnące wymagania sprawozdawcze dla naukowców, którzy często muszą wielokrotnie wprowadzać te same dane.

Dokonując analizy porównawczej trzeba jasno zdefiniować płaszczyzny, na których może ona występować. Jedną z takich płaszczyzn stanowią systemy stworzone do celów raportowania i ewaluacji dorobku naukowego. Ich kształt wynika w głównej mierze z formy takiej ewaluacji. W przypadku systemów ewaluacji opartych na zasadzie oceny parametrycznej (ocena ilościowa) kluczowym elementem jest strukturalizacja i jakość danych. Wymagania są w tym przypadku zdecydowanie wyższe niż w przypadku systemów opartych na ocenie eksperckiej, gdzie dominującym kryterium jest czynnik analizy jakościowej.

W kontekście systemów centralnych zbierających dane o studentach i procesach dydaktycznych kluczowym czynnikiem jest ustalenie, czy w ogóle taki system – z punktu widzenia konkretnego państwa – jest potrzebny do prowadzenia przez nie skoordynowanej polityki opartej na faktach (ang. *fact based policy*). Dane można pozyskiwać w różny sposób, najczęściej wystarczające jest zbieranie raportów, sprawozdań i zagregowanych informacji bezpośrednio od uczelni. Dużą rolę odgrywa również potencjał ludnościowy danego państwa oraz współczynnik scholaryzacji⁵³. W krajach o stosunkowo niewielkim potencjale ludnościowym, gdzie liczba uniwersytetów jest niewielka, potrzeba tworzenia dedykowanych rejestrów na poziomie centralnym wydaje się niewielka.

Poniższe podrozdziały są próbą analizy wybranych rozwiązań w zależności od typu i celu pozyskiwania danych.



⁵¹ cristin.no (dostęp: 18.06.2021)

⁵² fellesstudentsystem.no (dostęp: 18.06.2021)

⁵³ jest to pojęcie stosowane w statystyce publicznej rozumiane jako relacja liczby osób uczących się (stan na początku roku szkolnego) na danym poziomie kształcenia (niezależnie od wieku) do liczby ludności (stan w dniu 31 XII) w grupie wieku określonej jako odpowiadająca temu poziomowi nauczania

2.1.1. Systemy typu CRIS

W XXI wieku zaczęliśmy w sposób wyjątkowy doceniać związek pomiędzy danymi a informacją i wiedzą. Postępująca informatyzacja, po raz pierwszy w historii, pozwala nam na przeistaczanie opisów rzeczywistości (danych) w wielokontekstowe informacje a informacje w cenną wiedzę. Wiedza zaś daje nam możliwość bardziej świadomego kreowania teraźniejszości na rzecz lepszej przyszłości. Na intensywności przybrały inicjatywy mające na celu gromadzenie danych, informacji i wiedzy o tym aspekcie życia społecznego, który jest jego motorem napędowym – o nauce. Dodatkowo w XXI wieku unaocznione zostało nam, że masowość danych wzmacnia ich potencjał przetwarzania się w rzetelną wiedzę. Masowość potrafi zakryć aberracje i ukazać nam uśredniony obraz stanu rzeczy. Inicjatywy takie jak popularyzacja systemów typu CRIS, opartych na wspólnym języku – CERIF⁵⁴ – wspierają dążenie do gromadzenia danych w sposób masowy i zunifikowany.

Analiza przyczyn powstawania oraz przegląd deklarowanych celów działania systemów typu CRIS w Europie, wyłania obraz wspólnych problemów jakie mają one rozwiązać i tożsamyh pytań na jakie mają odpowiadać. Ujęcie to wydaje się być dobrym punktem wyjścia do zestawiania ze sobą różnych systemów typu CRIS. Pomimo, iż w Europie funkcjonuje wiele systemów typu CRIS operujących na poziomie lokalnym (uczelnianym, regionalnym) zestawienie obejmie systemy typu CRIS operujące na poziomach krajowych. Takie bowiem systemy stanowią analogię do systemu POL-on będącego przedmiotem niniejszej monografii.

Jednym z przeważających wyzwań, do sprostania którym zostają powołane systemy typu CRIS, jest stanie się efektywnym i wydajnym narzędziem pozwalającym na dostarczanie odpowiedzi na pytania zadawane przez jego fundatorów. Odpowiedzi te kryją się często w zdenormalizowanych danych rozproszonych po wielu instytucjach. Zgromadzenie tych danych na potrzeby fundatorów z poziomu krajowego wymaga uwspólnienia języka na poziomie krajowym. Zgromadzenie danych na potrzeby fundatorów z poziomu ogólnoeuropejskiego wymaga uwspólnienia języka na poziomie międzynarodowym.

W celu umożliwienia realizacji tego zadania powstał CERIF. Jest to ustandaryzowany model danych, którego celem jest stworzenie uniwersalnych formatów zapisów informacji o badaniach naukowych i ich rezultatach, naukowcach czy instytucjach naukowych. Model ten jest rekomendowany krajom członkowskim przez Unię Europejską w celu osiągnięcia maksymalnej interoperacyjności danych co ma wpłynąć pozytywnie na konkurencyjność Europejskich badań naukowych [9]. Pieczę nad modelem Unia powierzyła organizacji euroCRIS⁵⁵ zrzeszającej europejskich dostawców i operatorów systemów typu CRIS. CERIF obejmuje szeroki zakres wiedzy o świecie nauki, obsługując takie encje domeny jak: osoba, organizacja, projekt, publikacja, patent, nagrody czy sprzęt badawczy [9]. Warto zwrócić uwagę na fakt, że o ile ustandaryzowany model danych



⁵⁴ Common European Research Information Format (CERIF)

⁵⁵ eurocris.org (dostęp: 18.06.2021)

w jak najszerszym zakresie daje niezaprzeczone korzyści dla interoperacyjności, nie zawsze możliwa jest jego pełna implementacja. Przykładowo aplikacja PBN, obsługująca w ramach Zintegrowanego Systemu Informacji o Szkolnictwie Wyższym i Nauce informacje o publikacjach naukowych, nie implementuje formatu CERIF z powodu bardzo silnych uwarunkowań prawnych regulujących zakres gromadzonych danych.

Adaptacja opracowywanego centralnie (na poziomie europejskim) modelu danych stanowi wyzwanie dla twórców systemów. W 2014 roku do grona systemów kompatybilnych z modelem CERIF zaliczano jedynie 7 systemów europejskich [17]. Już wtedy jednym z nich był FRIS – Flanders Research Information Space⁵⁶ – regionalny portal gromadzący informacje o flandryjskich naukowcach i badaniach naukowych. Flandryjski rząd chciał dzięki temu stworzyć swoim naukowcom możliwość zwiększenia widoczności ich pracy. Przyjęcie międzynarodowego modelu danych sprawia, że są one łatwe do zrozumienia przez szerokie grono odbiorców. FRIS udostępnia swoje dane w sposób otwarty za pomocą między innymi CERIF SOAP API⁵⁷.

W związku z korzyściami z uwspólnionego języka, kolejne systemy dołączają do grona implementacji modelu CERIF. W 2019 roku w zaawansowanej fazie mapowania swojego modelu na CERIF [34] był system NARCIS – National Academic Research and Collaborations Information System⁵⁸. NARCIS to narodowy, holenderski system gromadzący informacje o holenderskich naukowcach, publikacjach i projektach naukowych oraz zbiorach danych. Pomimo, że posiada on swój natywny model danych, jego twórcy zdecydowali się na podjęcie wysiłku w celu ujednoliczenia swojej domeny z ogólnoeuropejskim modelem. Dzięki uwspólnieniu języka możliwe jest zapewnienie wysokiej interoperacyjności danych i udzielenie odpowiedzi każdemu, kto potrafi w tym języku zadać pytanie.

Nauka i badania oparte na ogromnych zbiorach danych mogą się rozwijać dzięki zwiększaniu się ich dostępności. Otwarty dostęp do publikacji naukowych i naukowych zbiorów danych jest kluczowym elementem zarówno w aspekcie użyteczności jak i wpływu na aktywność naukową [8]. Otwarty model komunikacji naukowej wspiera postęp naukowy. Jednym z zadań stawianych przed systemami typu CRIS jest bycie platformą zapewniającą otwarty dostęp⁵⁹ do zbiorów danych i publikacji naukowych oraz monitoring polityk związanych z tym zagadnieniem.

Twórcy wspomnianego już systemu FRIS uważają za istotne położenie nacisku na otwieranie dostępu do publikacji naukowych i zestawów danych. Motorem napędowym takiego podejścia był pilotaż programu Komisji Europejskiej Open Data z 2017 roku. Inicjatywa jest też wspierana przez flandryjski rząd⁶⁰.



⁵⁶ researchportal.be (dostęp: 18.06.2021)

⁵⁷ fris4.researchportal.be/ws (dostęp: 18.06.2021)

⁵⁸ narcis.nl (dostęp: 18.06.2021)

⁵⁹ otwarty dostęp, OD, ang. *open access*, OA

⁶⁰ researchportal.be/en/open-science (dostęp: 18.06.2021)

Monitoring sytuacji związanej z otwartym dostępem wśród norweskich naukowców jest jednym z zadań CRISTIN – Current Research Information System in Norway⁶¹. CRISTIN jest krajowym systemem norweskim, który pozwala na dostęp dzięki instytucjonalnym, lokalnym danym uwierzytelniającym. Charakteryzuje się tym, że pozwala wielu użytkownikom na współpracę nad jednym rekordem przy jednoczesnym wyznaczaniu „superużytkowników” z każdej instytucji odpowiedzialnych za poprawność i kompletność danych. System jest wykorzystywany do monitorowania stanu otwartości publikacji norweskich naukowców [11] po tym jak w 2017 roku norweski rząd ustanowił narodowe wytyczne dla otwartego dostępu do publikacji naukowych.

Najpowszechniejszym zadaniem realizowanym przez europejskie systemy informacyjne gromadzące informacje o nauce jest zapewnianie dostępności do informacji o stanie nauki na potrzeby różnych interesariuszy w tym opinii publicznej. Modelowym przykładem systemu w znacznej mierze dedykowanemu temu celowi jest polski system RAD-on⁶². System RAD-on jest agregatem informacji pochodzących ze Zintegrowanego Systemu Informacji o Nauce i Szkolnictwie Wyższym i współpracuje z takimi jego elementami jak POL-on, PBN czy ORPPD. Głównym zadaniem systemu RAD-on jest transformacja danych zbieranych z różnych systemów domenowych w ujednoliconą w danym kontekście biznesowym informację. System oferuje każdemu otwarte raporty, predefiniowane analizy oraz surowe dane.

Bardzo rozbudowaną wyszukiwarką dostępną w ramach portalu (a więc zapewniającą dostęp do informacji użytkownikom niekorzystającym np. z *open api*) może poszczycić się system ETIS – Estonian Research Information System⁶³. Estońskie rozwiązanie koncentruje się na informacjach o instytucjach naukowych, naukowcach, ich projektach czy publikacjach naukowych. Jednocześnie system służy do obsługi procesu aplikowania o granty i do raportowania postępów w grantowych projektach naukowych. Wyszukiwarka projektów naukowych w portalu, ustanowionym przez estońskie ministerstwo odpowiedzialne za edukację i naukę, oferuje przeszukiwanie bazy projektów dzięki aż osiemnastu kryteriom wyszukiwania, które można łączyć. Po wyszukaniu interesującego nas projektu, możemy zapoznać się z informacjami o jego szczegółach – w tym budżecie przyznany na projekt ze środków publicznych.

Rozliczalność środków publicznych i ich transparentna dystrybucja w świecie nauki to idee przyświecające powstaniu szwedzkiego systemu SweCRIS⁶⁴. System umożliwia każdemu obywatelowi Szwecji (oraz każdej osobie posiadającej dostęp do sieci internet) dostęp do informacji o tym, w jaki sposób fundusze na naukę są dzielone wśród szwedzkich odbiorców. Dwanaście szwedzkich agencji (zarówno publicznych jak i prywatnych), dystrybuujących granty zasilają system danymi. Gromadzone są informacje o adresatach i przeznaczeniu środków. System zunifikował sposób raportowania danych przez



⁶¹ cristin.no (dostęp: 18.06.2021)

⁶² radon.nauka.gov.pl (dostęp: 18.06.2021)

⁶³ etis.ee (dostęp: 18.06.2021)

⁶⁴ vr.se/swecris (dostęp: 18.06.2021)

Realizacja potrzeb wielu instytucji jednocześnie to pożądaný model działania dla systemu typu CRIS. Doskonałym przykładem realizacji takiego modelu jest system POL-on będący przedmiotem analizy w niniejszej monografii. Inne europejskie systemy również dążą do tego aby wysiłek z ich tworzenia przekładał się na korzyści w eliminacji zbędnej i powtarzalnej biurokracji. W 2019 roku do grona instytucji członkowskich systemu CRISTIN dołączyła *Statistisk sentralbyrå* [11] – norweskie biuro statystyczne będące głównym dostawcą raportów i analiz statystycznych o Norwegii⁶⁵.

Pomimo uniwersalnych przyczyn powstawania, systemy typu CRIS w Europie różnią się od siebie zakresem swoich funkcjonalności oraz powiązań z systemami wewnętrznymi. Podczas gdy 80 procent systemów typu CRIS, które odpowiedziały na kwestionariusz EuroCRIS z 2016 roku, wspiera obsługę danych dotyczących publikacji naukowych, jedynie połowa z nich gromadzi pełne teksty tych publikacji. Jedynie 30 procent systemów ma za zadanie wesprzeć swoimi funkcjonalnościami ocenę badaczy. 68 procent zbadanych systemów typu CRIS jest zintegrowanych z systemami kadrowymi, 63 procent z repozytoriami poszczególnych instytucji, ale już tylko 8 procent z nich jest powiązanych z systemami bibliotecznymi [20].

Age Group	Percentage
18-24	15%
25-34	14%
35-44	13%
45-54	12%
55-64	11%
65-74	10%
75-84	9%
85+	2%

⁶⁶ nauka-polska.pl (dostęp: 18.06.2021)

2.1.2. Systemy ewaluacji dorobku naukowego

Polityka naukowa była prowadzona „od zawsze”, ale w drugiej połowie XX wieku wprowadzono systemy uzależnienia środków otrzymywanych przez instytucję na badania naukowe od jej doskonałości naukowej. Stąd tylko krok dzielił poszczególne państwa od wprowadzenia krajowych systemów oceny doskonałości naukowej dla poszczególnych instytucji – czyli systemów ewaluacji działalności naukowej. Było to konieczne w obliczu znacznego ograniczenia środków, które dany kraj mógł przeznaczyć na badania naukowe.

Pierwszym z systemów ewaluacji był brytyjski system *Research Assessment Exercise* z 1986 r. System ten początkowo opierał się wyłącznie na zasadzie oceny eksperckiej. Liczba ocenianych jednostek oraz parametrów oceny znacznie zmieniła się w trakcie kolejnych ocen przeprowadzonych w latach 1986, 1989, 1992, 1996, 2001 i 2008. Następną wersją brytyjskiego systemu oceny był *Research Excellence Framework* wprowadzony w roku 2014. System ten posłużył do przeprowadzenia oceny w 2014 roku. Funkcjonuje on w nieco zmienionej formule do dzisiaj i po raz kolejny będzie wykorzystany w ewaluacji w 2021 roku.

Mniej więcej w tym samym czasie (od początku lat 80 ubiegłego wieku) rozwijał się holenderski system oceny działalności jednostek naukowych. Finalnie opracowano *Standard Evaluation Protocol*, który ma posłużyć do oceny działalności naukowej w roku 2021⁶⁷ instytucji naukowych działających w Holandii⁶⁸.

Po systemie brytyjskim pojawiły się kolejne systemy ewaluacji działalności jednostek naukowych: we Francji, w Hiszpanii (*Sexenios de investigación*), Portugalii (*Research Unit Evaluation*), Australii (*Excellence in Research for Australia*), Nowej Zelandii (*Performance-based Research Funding* – patrz np. [13]), Polsce i innych krajach. Wszystkie z wymienionych tu państw nadal prowadzą ewaluację działalności naukowej własnych instytucji.

W ewaluacji używa się zwykle trzech metod mierzących efektywność naukową danej jednostki. Są to: metody oparte na ocenie eksperckiej (*peer review-based models*), metody oparte na liczbie publikacji (*publication count-based models*) i metody oparte na liczbie cytowań (*citation-based models*). Dostępne również (i najczęściej stosowane) są metody mieszane, przy czym zwykle występuje w nich przewaga metod ilościowych. Więcej informacji związanych z zastosowaniem tych rozwiązań można znaleźć w rozdziale trzecim monografii [5]. Metody mieszane bazują na ocenie ilościowej części osiągnięć naukowych instytucji (np. publikacji) oraz na ocenie innych osiągnięć przez ekspertów (np. wpływ efektów prac badawczych na rozwój danej dziedziny wiedzy). Największy problem tkwi w doborze i zastosowaniu kryterium ilościowego,



⁶⁷ za lata 2015-2021

⁶⁸ knav.nl/nl/actueel/publicaties/standard-evaluation-protocol-2015-2021 (dostęp: 18.06.2021)

ponieważ w zależności od wybranego kryterium można otrzymać inne rezultaty, które ściśle zależą od dziedziny nauki, do której to kryterium zostało zastosowane (patrz np. [12]).

Na początku lat 90. ubiegłego wieku ewaluację rozpoczęto również w Polsce. Oceny dokonywano *ex post*. Procedura ewaluacji początkowo była bardzo niedoskonała⁶⁹, coraz bardziej złożona i wymagała określenia dla każdej ocenianej jednostki naukowej coraz większej liczby wskaźników. W pierwszej fazie ewaluacja bazowała na corocznie składanych przez jednostki Ankietach Jednostki Naukowej, które były grupowane⁷⁰ przy okazji każdej przeprowadzanej⁷¹ ewaluacji. Później ministerstwo zrezygnowało z konieczności corocznego składania przez jednostki Ankiety Jednostki Naukowej. Wprowadzono najpierw konieczność podawania przez jednostki wymaganych danych dopiero przed samą ewaluacją, za cały okres oceny, a ostatnio oparto przeprowadzaną ewaluację wyłącznie na danych zawartych w systemie POL-on⁷². Ocena do tej pory⁷³ została przeprowadzona w latach 1999, 2003, 2006, 2010, 2013 i 2017. Oddzielnie były oceniane szkoły wyższe, instytuty PAN, instytuty badawcze i inne zależnie od liczby pracowników deklarujących prowadzenie badań naukowych – jednostki naukowe ewaluowane były w zdefiniowanych przez Ministra Nauki grupach wspólnej oceny. Od 2003 roku ewaluacja była prowadzona przy udziale wspomagającego ją oprogramowania komputerowego, które w ostatnich latach przekształciło się w system do ewaluacji jednostek naukowych. Na rok 2021 planowana jest kolejna ewaluacja na zmienionych zasadach – oceniane będą całe instytucje w poszczególnych dyscyplinach naukowych, w których realizowane są badania naukowe w danej instytucji. Szczegółowe informacje na ten temat można znaleźć w [37] i [31].

2.1.3. Wybrane centralne systemy informacji o dorobku studentów

Rozwiązania bazujące na centralizacji zbiorów danych o studentach wspierane i finansowane przez jednostki centralne nie są dominującym podejściem w większości państw. Decyzje co do metod informatyzacji pozostawia się raczej samym uczelniom. Stąd mnogość różnego typu rozwiązań i dostawców je oferujących. Występuje osobna kategoria systemów informacji o studentach SIS (ang. *Student Information Systems*⁷⁴). Systemy informacyjne dla studentów zapewniają możliwość rejestracji studentów na kursy, dokumentowanie ocen, wyników sprawdzianów uczniów i innych ocen; tworzenie grafików studenckich; śledzenie obecności uczniów; i zarządzanie wieloma innymi potrzebami danych studentów w szkole wyższej.



⁶⁹ metoda oceny eksperckiej

⁷⁰ ponownie składane i uzupełniane

⁷¹ zwykle co cztery lata

⁷² dane te instytucje wprowadzały systematycznie

⁷³ na zmienianych zasadach i przy różnej liczbie wykorzystywanych parametrów

⁷⁴ gartner.com/reviews/market/k-12-student-information-systems (dostęp: 18.06.2021)

W Polsce najpopularniejszym tego typu systemem jest Uczelniany System Obsługi Studiów (USOS), rozwijany od roku 2000 na polskich uczelniach. W skali całej Polski system utrzymywany jest w ramach projektu USOS realizowanego przez Międzyuniwersyteckie Centrum Informatyzacji – konsorcjum powołane w celu utrzymywania i rozwoju systemów informatycznych wspierających zarządzanie szkołami wyższymi. Z systemu korzystają uczelnie będące członkami MUCI (udziałowcy konsorcjum) lub uczelnie będące tzw. Członkami Stowarzyszonymi MUCI w ramach projektu USOS [7]. W każdej uczelni system ten wdrażany jest *de facto* jako oddzielna instancja. Z racji swojej popularności systemy te stanowią główne źródło danych o studentach do rejestru prowadzonego w systemie POL-on, który pełni funkcję kontrolną i agregującą w skali ogólnopolskiej.

Innym rozwiązaniem, zbliżonym bardziej do systemu POL-on, bo wdrażanym i utrzymywanym w skali całego państwa, jest szwedzki centralny system informacji o dorobku studentów LADOK⁷⁵. Wszystkie szwedzkie uniwersytety korzystają z LADOK lub LADOK NOVAU, należącego do konsorcjum [3] składającego się z 37 podmiotów⁷⁶. LADOK to informatyczny system dla studentów i system dokumentacji dla uniwersytetu lub kolegium uniwersyteckiego. Skupia się na administrowaniu studiami. System jest wdrażany lokalnie i zarządzany przez instytucje. Akronim „Ladok” stanowi skrót od szwedzkiego *Lokalt adb-baserat dokumentationssystem* („lokalny system dokumentacji”). Jest to system rejestracji studentów i dokumentacji ocen, którego nie należy mylić z systemem rekrutacji i różnymi systemami zarządzania uczeniem się na uczelniach. Z Ladoka korzysta obecnie około 20.000 nauczycieli i administratorów, a około 234.000 uczniów ma dostęp do systemu. System zawiera również dane i rekordy 3,2 miliona studentów⁷⁷.

Kolejnym rozwiązaniem wspierającym zarządzanie danymi o studentach w wymiarze państwowym jest norweski Wspólny System Studencki (*Common Student System, Felles Studentsystem, FS*)⁷⁸. Jest to jedyny system do zarządzania studiami opracowany dla uniwersytetów, kolegiów naukowych i krajowych kolegiów uniwersyteckich w Norwegii, rozwijany przez Unit – Norweską Dyрекcję ds. ICT i Połączonych Usług w Szkolnictwie Wyższym i Badaniach. System jest krytycznym narzędziem dla obszaru szkolnictwa wyższego w Norwegii obejmującym ponad 14 modułów funkcjonalnych, pokrywających wszystkie procesy związane z administracją toku studiów⁷⁹. Z systemem współpracuje też szereg usług, do których należy hurtownia danych STAR, gromadząca raporty z danych istotnych do podejmowania decyzji zarządczych [16]. Dodatkowo z FS zintegrowany jest Krajowy Rejestr Wykluczeń (RUST – System informacji o wykluczeniach studentów)⁸⁰, który został stworzony przez norweskie Ministerstwo Szkolnictwa Wyższego w celu zapewnienia krajowych procedur informowania o wykluczeniu i zapobieganiu przyjęciu na podstawie

⁷⁵ ladok.se (dostęp: 18.06.2021)

⁷⁶ student.ladok.se (dostęp: 18.06.2021)

⁷⁷ dane na podstawie ladok.se/ladok/detta-ar-ladok z dnia 4.10.2020

⁷⁸ fellesstudentsystem.no (dostęp: 18.06.2021)

⁷⁹ fellesstudentsystem.no/english/applications/fs-client (dostęp: 18.06.2021)

⁸⁰ fellesstudentsystem.no/applikasjoner/rust (dostęp: 18.06.2021)

fałszywych dokumentów. Utworzenie rejestru stanowi skutek działań prewencyjny w odniesieniu do akademickich wykroczeń i działa na rzecz zwiększenia wiarygodności sektora szkolnictwa wyższego w Norwegii. Dodatkowe korzyści płynące z rozwiązania krajowego to zapewnienie dobrej jakości danych i wprowadzenie wspólnej procedury dla procesów administracyjnych w sektorze szkolnictwa wyższego. Formalne utworzenie rejestru było skomplikowane z uwagi na kwestie prawne i regulacyjne, które należało przeanalizować i rozwiązać. Kluczowa dla powodzenia projektu była współpraca pomiędzy Ministerstwem Szkolnictwa Wyższego, Norweskim Urzędem Ochrony Danych i zarządzanie projektami w Uniwersyteckim Centrum Technologii Informacyjnych w Oslo⁸¹.

Uniwersytety i kolegia muszą zarejestrować wszelkie sankcje nałożone na ich kandydatów/studentów w swoim systemie zarządzania studiami (np. FS). W kilku przypadkach informacje o osobie i sankcji są przekazywane do RUST. Za pośrednictwem RUST inne uniwersytety i kolegia otrzymują informacje o sankcjach w bezpieczny sposób tylko wtedy, gdy dana osoba jest kandydatem lub studentem w danej uczelni.

Powody zarejestrowania studentów lub kandydatów w RUST to:

- fałszywe dyplomy
- niepokojące/kłopotliwe zachowanie, rażące naruszenie poufności, rażąca nieprzyzwoitość w praktyce lub nauczaniu
- oszustwo, usiłowanie oszustwa, współudział w oszustwie

Po upływie okresu sankcji dane osobowe i informacje o decyzji są automatycznie usuwane z rejestru.

W systemie POL-on podobne rozwiązanie występuje wyłącznie w zakresie danych o pracownikach naukowych i nakładanych na nich karach dyscyplinarnych i wyrokach sądowych.

Dokonując analizy porównawczej wymienionych wyżej systemów trzeba zwrócić uwagę na ich fundamentalną różnicę względem modelu zastosowanego w Polsce, opartego na dualizmie i synergii systemów do obsługi toku studiów na uczelniach (wymiar lokalny) i centralnego rejestru, do którego przekazywana jest część danych w celach realizacji polityk państwowych (wymiar centralny). Tego typu model nie znalazł adekwatnego odpowiednika na gruncie analizy porównawczej autorów niniejszej pracy. Jest to szczególnie widoczne w obszarze danych związanych z procesami dydaktycznymi (studenci, doktoranci, absolwenci), gdzie dominują modele zorientowane wyłącznie na wymiar lokalny. W analizie porównawczej trzeba zwrócić też uwagę na rozmiar i zakres danych. System POL-on jest pod tym względem wyjątkowy, ponieważ unifikuje do jednego zbioru dosyć rozbudowany zakres informacyjny na temat ponad jednego miliona studentów



⁸¹ usit.uio.no/english (dostęp: 18.06.2021)

w skali roku. Nie realizuje jednak zbyt wielu procesów uczelnianych, a koncentruje się głównie na analizie zaraportowanych danych⁸².

2.2. Obsługiwane procesy biznesowe

W rozdziale pierwszym w ramach szczegółowego opisu systemu, przedstawionych zostało wiele faktów dotyczących jego struktury oraz zakresu informacyjnego. W tym momencie istotne jest również opisanie w sposób bardziej szczegółowy procesów realizowanych za jego pomocą.

W systemie POL-on gromadzone są dane instytucji systemu szkolnictwa wyższego i nauki, w tym uczelni, federacji podmiotów systemu szkolnictwa wyższego i nauki, instytutów naukowych PAN, instytutów badawczych, międzynarodowych instytutów naukowych, Centrum Łukasiewicz, instytutów działających w ramach Sieci Badawczej Łukasiewicz oraz innych podmiotów prowadzących głównie działalność naukową w sposób samodzielny i ciągły. Wymienione instytucje rejestrują w systemie odpowiednio dane nauczycieli akademickich, pracowników naukowych, studentów oraz osób ubiegających się o stopień/ tytuł naukowy. Dodatkowo zasilają repozytorium pisemnych prac dyplomowych oraz bazę dokumentów w postępowaniach awansowych, a także składają raporty i sprawozdania w bazie dokumentów planistyczno-sprawozdawczych, (zakres ten został szczegółowo przedstawiony w rozdziale 1).

Zbiór danych gromadzonych w systemie jest ściśle określony aktami prawnymi. Wykorzystywany jest przez Ministerstwo Edukacji i Nauki w celu nadzoru nad systemem szkolnictwa wyższego i nauki oraz realizacją zadań przez NAWA, NCBiR oraz NCN. W szczególności do wykonywania zadań związanych z ustalaniem i realizacją polityki naukowej państwa, przeprowadzaniem ewaluacji jakości kształcenia, szkół doktorskich i jakości działalności naukowej oraz prowadzeniem postępowań w sprawie nadania stopnia doktora, stopnia doktora habilitowanego i tytułu profesora, a także ustalaniem wysokości subwencji i dotacji.

Wśród wielu procesów biznesowych obsługiwanych przez system POL-on kilka jest szczególnie wartych uwagi. Pierwszym z nich jest kontrola nadużyć w zakresie niezgodności zarejestrowanych w systemie danych z prawem, których przykładami są:

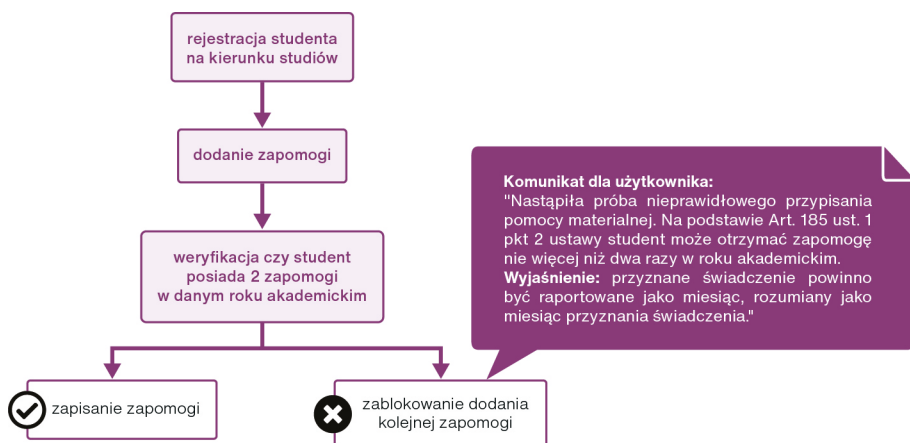
- pobieranie więcej niż dwa razy w danym roku akademickim pomocy materialnej w postaci zapomogi;
- pobieranie pomocy materialnej w kilku podmiotach jednocześnie;
- kształcenie się w więcej niż jednej szkole doktorskiej jednocześnie;
- zatrudnienie pracownika więcej niż raz w danym momencie na podstawowym miejscu pracy;



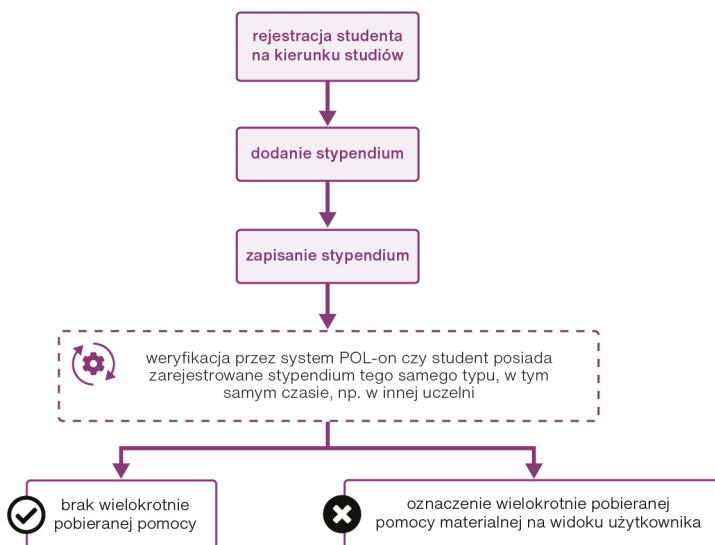
⁸² szczegółowe statystyki w tym zakresie zostały przedstawione w pierwszym rozdziale. W przypadku systemów stanowiących porównanie liczba studentów nie przekracza 300-400 tys. rocznie

- reprezentacja więcej niż dwóch dyscyplin naukowych przez pracownika w danym momencie;
- rejestracja kluczowych danych naukowych pracownika po terminie.

System wykrywa sytuację błędną i blokuje możliwość wprowadzenia takich danych (Ilustracja 2.1) lub umożliwia ich wprowadzenie, ale oznacza dane jako wymagające dodatkowej weryfikacji czy wyjaśnienia (Ilustracja 2.2).



Ilustracja 2.1. Weryfikacja ilości zarejestrowanej pomocy materialnej typu „zpomoga”



Ilustracja 2.2. Weryfikacja ilości zarejestrowanej pomocy materialnej typu „stypendium”

Kolejnym ważnym procesem jest podział dotacji publicznej, który jest ustalany na podstawie algorytmów określonych odrębnymi rozporządzeniami. Ich wartość jest wyliczana stosownie do danych wprowadzonych do systemu POL-on przez instytucje. Dodatkowo na podstawie danych z systemu ewaluacji dokonywana jest ocena instytucji w zakresie jakości kształcenia, szkół doktorskich czy jakości działalności naukowej.

Ocena jakości kształcenia polega na cyklicznej ocenie dokonywanej z inicjatywy PKA⁸³ lub na wniosek uczelni bądź ministra. W trakcie przeprowadzania oceny programowej brane są pod uwagę dane z zakresu programów studiów i standardów kształcenia, kadry dydaktycznej i naukowej, infrastruktury wykorzystywanej do realizacji programu studiów, współpracy z otoczeniem społeczno-gospodarczym, umiędzynarodowienia oraz wsparcia studentów w procesie uczenia się.

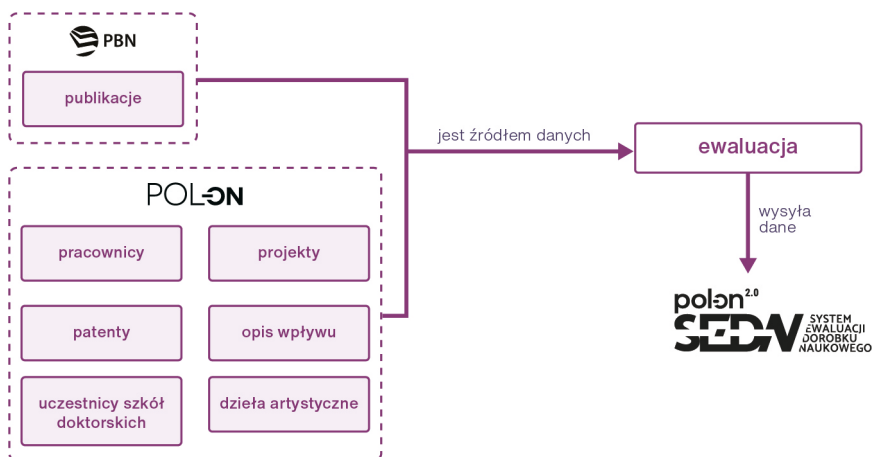
Ewaluacja szkół doktorskich to nowy proces, realizowany przez KEN⁸⁴ nie rzadziej niż raz na 6 lat lub w innym terminie na wniosek ministra. Pierwsze szkoły doktorskie rozpoczęły kształcenie z dniem 1 października 2019 roku, zatem pierwsza ocena odbędzie się w IV kwartale 2024 roku, ponieważ zgodnie z zapisami ustawy Prawo o szkolnictwie wyższym i nauce z dnia 20 lipca 2018 r. pierwsza ocena działalności szkoły ma mieć miejsce nie wcześniej niż po upływie co najmniej 5 lat od dnia rozpoczęcia kształcenia w tej szkole.

Ocena jakości działalności naukowej dotyczy uczelni akademickich, instytutów Polskiej Akademii Nauk, instytutów międzynarodowych, uczelni zawodowych, instytutów badawczych oraz podmiotów prowadzących głównie działalność naukową w sposób samodzielny i ciągły, posiadających siedzibę na terytorium Rzeczypospolitej Polskiej i obejmuje osiągnięcia wszystkich pracowników prowadzących działalność naukową. Ocena w tym zakresie przeprowadzana jest na wniosek, a w pozostałych przypadkach co 4 lata. Przy czym, pierwsza ocena, ze względu na pandemię, zostanie przeprowadzona dopiero w roku 2022. Pod uwagę brany jest poziom naukowy lub artystyczny prowadzonej działalności, efekty finansowe badań naukowych i prac rozwojowych, a także wpływ działalności naukowej na funkcjonowanie społeczeństwa i gospodarki (Ilustracja 2.3).



⁸³ Polska Komisja Akredytacyjna jest niezależną instytucją działającą na rzecz zapewniania i doskonalenia jakości kształcenia

⁸⁴ Komisja Edukacji Narodowej



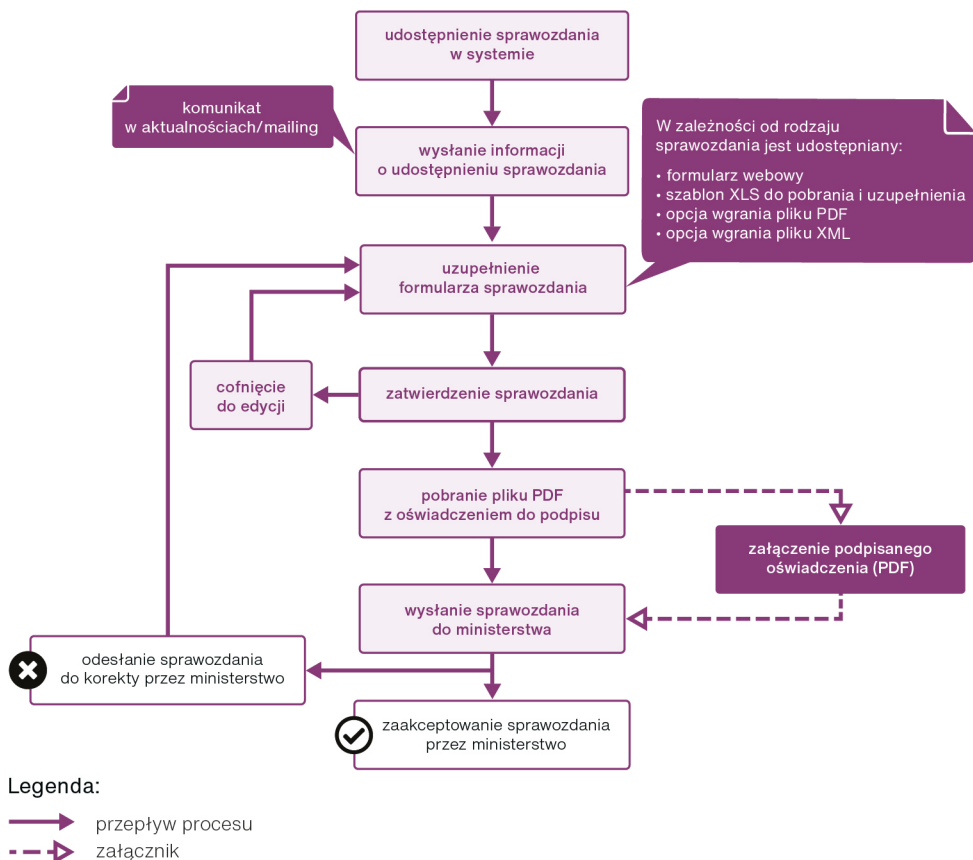
Ilustracja 2.3. Ewaluacja podmiotów w zakresie jakości kształcenia, szkół doktorskich i działalności naukowej

Proces sprawozdawczości finansowej jest realizowany w ramach modułu „Baza dokumentów planistyczno-sprawozdawczych”. Umożliwia składanie, agregowanie i docelowo archiwizowanie niżej wymienionych dokumentów:

- plany rzeczowo-finansowe uczelni publicznych;
- roczne plany finansowe Centrum Łukasiewicz i instytutów Sieci Badawczej Łukasiewicz;
- sprawozdania z wykonania planów rzeczowo-finansowych uczelni publicznych;
- sprawozdania i raporty z wykorzystania środków finansowych, np.: sprawozdanie z wykorzystania dotacji na zadania związane z zapewnieniem osobom niepełnosprawnym warunków do pełnego udziału w procesie przyjmowania na studia, do szkół doktorskich, kształcenia na studiach i w szkołach doktorskich lub prowadzenia działalności naukowej, sprawozdanie z wykorzystania środków funduszu stypendialnego czy raporty roczne i końcowe z wykorzystania środków finansowych na realizację inwestycji związanych z kształceniem oraz działalnością naukową;
- roczne sprawozdania finansowe uczelni publicznych;
- roczne sprawozdania finansowe Centrum Łukasiewicz i instytutów Sieci Badawczej Łukasiewicz.

Większość sprawozdań i raportów składana jest raz w roku, z wyjątkiem raportów końcowych, które składane są celem rozliczenia się ze środków, które instytucja uzyskała z ministerstwa. Dane sprawozdawcze są wprowadzane do systemu na podstawie wygenerowanych i udostępnionych szablonów, opracowanych przez Ministerstwo Edukacji i Nauki. W przypadku sprawozdania finansowego dane wprowadzane są w formacie XML zgodnym z wytycznymi Ministerstwa Finansów, określonymi w opublikowanym XSD*. Po zaimportowaniu do systemu dane są zatwierdzane przez pracownika podmiotu oraz generowane jest oświadczenie o zgodności danych. Następnie, już podpisane przez rektora uczelni lub dyrektora instytutu (poza systemem), jest załączane

w momencie wysyłania sprawozdania do ministerstwa. Pracownik resortu odpowiedzialny za sprawozdawczość weryfikuje przesłane dane i akceptuje lub odsyła je do korekty, jeśli zachodzi taka potrzeba. Po dokonaniu korekty przez pracownika uczelni formularz sprawozdania jest odsyłany do ministerstwa tą samą ścieżką (Ilustracja 2.4).



Ilustracja 2.4. Proces sprawozdawczości w systemie POL-on

Wszelkie sprawozdania, wgrywane do systemu POL-on, po roku od ich złożenia są archiwizowane, a następnie przechowywane w systemie przez 10 lat.

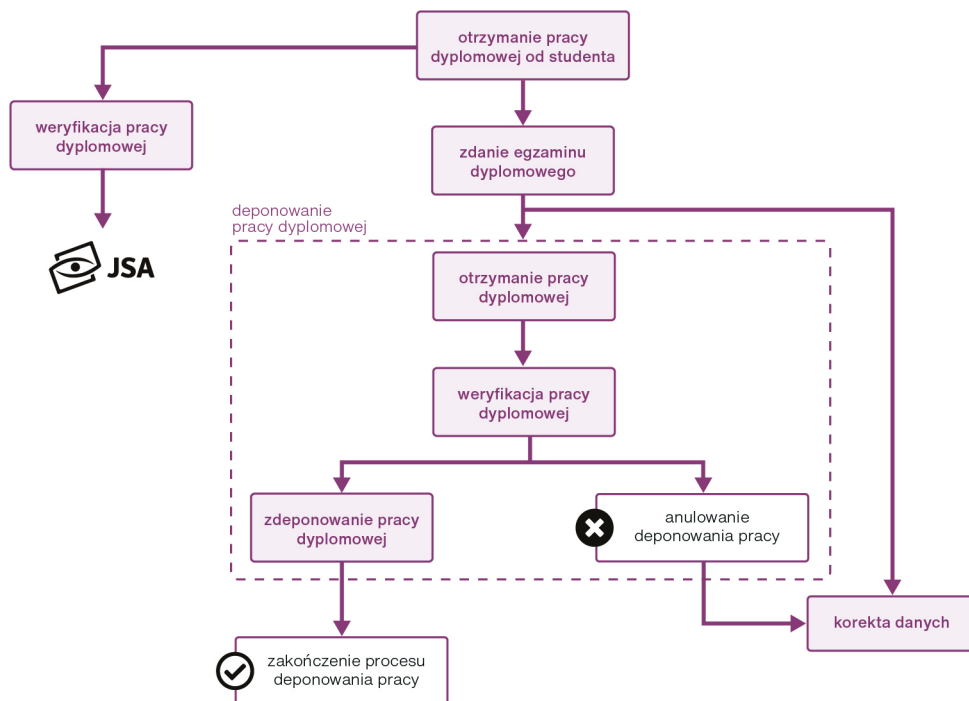
Sprawozdawczość statystyczna za pośrednictwem systemu POL-on zdecydowanie upraszcza dotychczasowe statystyki przekazywane przez uczelnie w formie papierowych formularzy do GUS, z którymi jeszcze do niedawna borykało się szkolnictwo wyższe. Obecnie na podstawie danych wprowadzonych do systemu POL-on w odpowiednich modułach wyliczane są zagregowane wartości do sprawozdań S-10, S-11 i S-12. Sprawozdania są już zatem wstępnie wypełnione, co z pewnością odciąża użytkowników i wymaga jedynie weryfikacji i ewentualnej poprawy danych w systemie przez insty-

```
graph TD; U[Uczelnia] --> POLON; subgraph POLON [POLON]; D[dane zagregowane]; W[wyliczenia systemu]; end; POLON --> GUS[GUS]; GUS --> S10[S-10];
```

Ogólnopolskie Repozytorium Pisemnych Prac Dyplomowych (ORPPD) to zbiór obronionych prac dyplomowych, który stanowi jedno z wielu źródeł referencyjnych dla systemu JSA⁸⁵. Przed obroną każda praca dyplomowa jest badana z wykorzystaniem Jednolitego Systemu Antyplagiatowego następnie, jeśli nie wykryto plagiatu i doszło do obrony pracy, wprowadzana jest do ORPPD. Proces deponowania prac dyplomowych w systemie przedstawia Ilustracja 2.6.

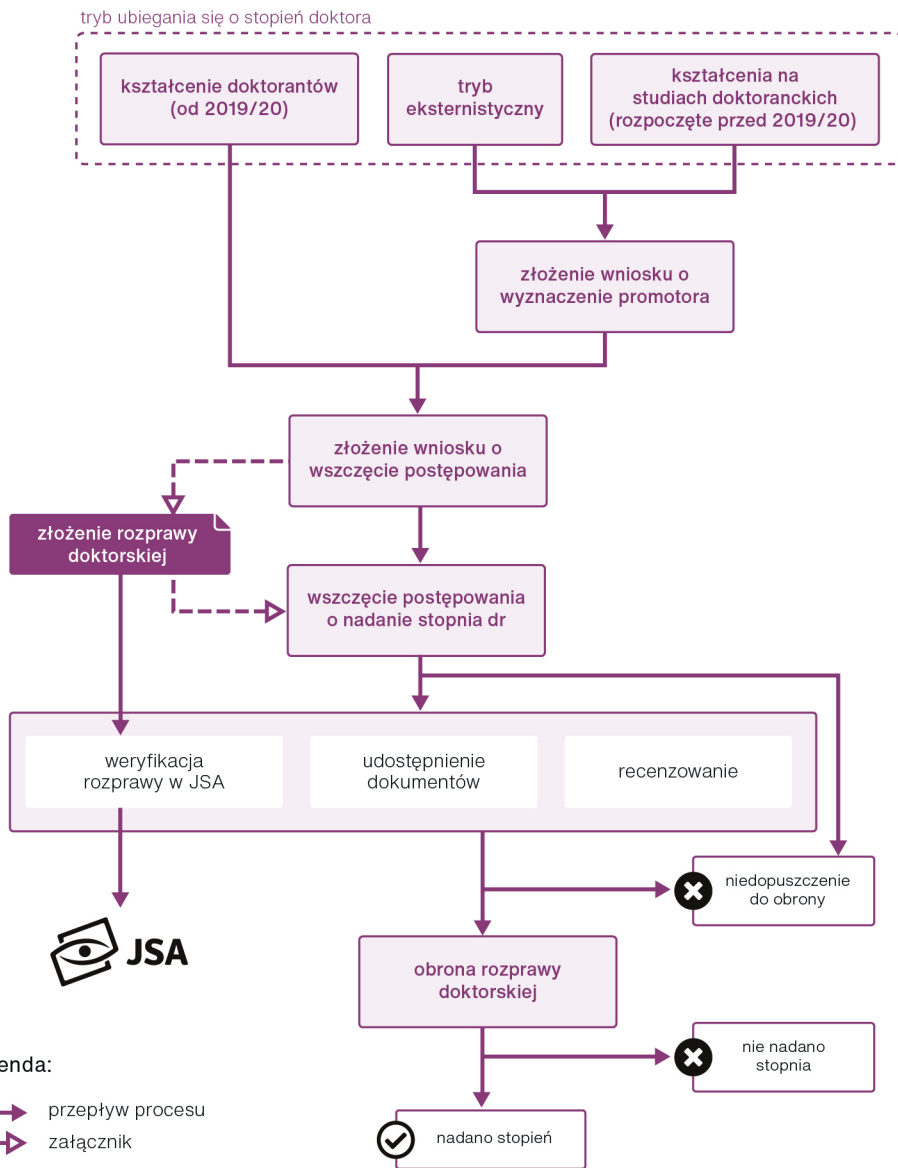
Ogólnopolskie Repozytorium Pisemnych Prac Dyplomowych (ORPPD) to zbiór obronionych prac dyplomowych, który stanowi jedno z wielu źródeł referencyjnych dla systemu JSA⁸⁵. Przed obroną każda praca dyplomowa jest badana z wykorzystaniem Jednolitego Systemu Antyplagiatowego następnie, jeśli nie wykryto plagiatu i doszło do obrony pracy, wprowadzana jest do ORPPD. Proces deponowania prac dyplomowych w systemie przedstawia Ilustracja 2.6.





Ilustracja 2.6. Deponowanie prac dyplomowych w ORPPD 2.0

Po obronie pracy magisterskiej możliwa jest kontynuacja procesu awansowego polegającego na ubieganiu się o stopień naukowy. W systemie POL-on gromadzone są dane zarówno dotyczące kształcenia doktorantów, jak i późniejszego nadania stopnia, ponieważ proces awansowy to nie tylko decyzja o nadaniu stopnia, to również szereg informacji dotyczących kolejnych etapów kariery naukowej osób ubiegających się o stopień lub tytuł, wraz z dokumentami, które towarzyszą temu procesowi. Baza postępowań awansowych zawiera zbiór informacji umożliwiających weryfikację, czy i kiedy dana osoba uzyskała stopień lub tytuł oraz w jakiej dyscyplinie (Ilustracja 2.7). Natomiast treści rozpraw zamieszczane w bazie postępowań awansowych stanowią składową zbioru referencyjnego wykorzystywanego przez JSA do badania prac.



Ilustracja 2.7. Proces ubiegania się o stopień doktora jako przykład procesu awansowego, którego dane gromadzone są w systemie POL-on

W bazie znajdują się również informacje o ewentualnym pozbawieniu stopnia/tytułu, jeśli zaistniała taka sytuacja.

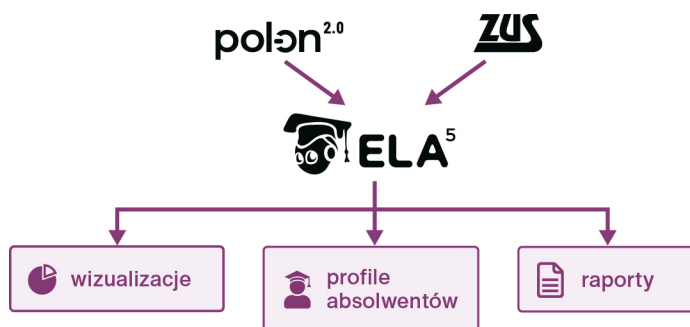
System POL-on wspiera również kontrolę dokumentów (NAWA⁸⁶) poprzez uwierzytelnianie przez dyrektora NAWA dokumentów wydawanych w związku z przebiegiem lub ukończeniem studiów, takich jak:

- dyplom ukończenia studiów oraz suplement do dyplomów;
- odpis wyżej wymienionych dokumentów, w tym odpis w języku obcym;
- duplikat wyżej wymienionych dokumentów;
- zaświadczenie o ukończeniu studiów, przeznaczone do obrotu prawnego z zagranicą.

Proces bazuje na porównaniu danych o podpisach i funkcjach osób upoważnionych do podpisywania, wzorach uwierzytelnianych dokumentów oraz wzorach pieczęci składowanych centralnie w systemie POL-on z rzeczywiście okazanym dokumentem. Dodatkowo dyrektor NAWA może podejmować decyzję o kształceniu cudzoziemców, a system POL-on umożliwia monitorowanie przebiegu tego kształcenia.

Udostępnianie publicznie danych gromadzonych w systemie POL-on to proces, który ma na celu realizację polityki otwartych danych w administracji publicznej.

Celem ogólnopolskiego systemu monitorowania Ekonomicznych Losów Absolwentów szkół wyższych (ELA) jest gromadzenie danych w oparciu o współpracę systemu POL-on oraz Zakładu Ubezpieczeń Społecznych. Następnie w oparciu o zebrane dane dostarczane są wiarygodne informacje o sytuacji absolwentów polskich uczelni na rynku pracy, w postaci wizualizacji czy raportów (Ilustracja 2.8).



Ilustracja 2.8. Proces zbierania danych wykorzystywanych w systemie ELA

Aby wspierać procesy decyzyjne kandydatów na studia udostępniono portal studia.gov.pl⁸⁷ zawierający wyszukiwarkę „Wybierz Studia”, w której można znaleźć sprawdzone i rzetelne informacje o kierunkach studiów pochodzące z czterech źródeł:



⁸⁶ Narodowa Agencja Wymiany Akademickiej

⁸⁷ studia.gov.pl/wybiez-studia/pages/search/wizard (dostęp: 18.06.2021)

- stale aktualizowanych danych z oficjalnej bazy o szkolnictwie wyższym – POL-on;
- ocen kierunków studiów dokonywanych przez Polską Komisję Akredytacyjną;
- ocen działalności naukowej dokonywanych przez Komitet Ewaluacji Jednostek Naukowych;
- informacji przesłanych przez uczelnie.

Uczelnie, które otrzymały status Krajowych Naukowych Ośrodków Wiodących – najlepszych polskich jednostek naukowych w naukach ścisłych, medycznych, biologicznych, rolniczych i naukach o Ziemi – są zaznaczone logo „KNOW”. Status ten nadawany jest przez ministra właściwego do spraw szkolnictwa wyższego i nauki na okres 5 lat. Zebrane dane można przeszukiwać na wiele sposobów. Wystarczy zaznaczyć jedną lub kilka kategorii dostępnych na stronie głównej, tj. nazwę uczelni, lokalizację, kierunki studiów, przedmioty maturalne (oficjalną listę wymaganych przedmiotów można znaleźć na stronie rekrutacji uczelni), studia dzienne (stacjonarne) i zaoczne (niestacjonarne), rodzaj uczelni.

2.3. Kluczowe cechy i wymagania systemu

W poprzednich rozdziałach przedstawione zostały kluczowe fakty dotyczące systemu oraz dokonano jego analizy porównawczej do innych tego typu rozwiązań na gruncie międzynarodowym. W celu pogłębienia tego opisu trzeba przywołać jego kilka istotnych cech oraz wymagań związanych z jego funkcjonowaniem. Poniżej przedstawione zostały wybrane aspekty, które zdeterminowały obecny kształt systemu.

2.3.1. Model pojęciowy systemu

W systemach informacyjnych o tak rozległej skali, gdzie dane pozyskiwane są wyłącznie ze źródeł zewnętrznych i stanowią tak naprawdę skumulowany efekt informacji pozyskiwanych z procesów odbywających się na uczelniach i w szeroko rozumianym świecie naukowym, forma opisu danych stanowi kluczowe zagadnienie.

W podejściu do architektury systemów informatycznych wyróżnić można wiele wzorców, w większości bazują one jednak na wyodrębnieniu poszczególnych warstw systemu z określonym zakresem odpowiedzialności [22]:

- warstwa prezentacji (interfejs użytkownika);
- warstwa logiki aplikacyjnej;
- warstwa dostępu do danych;
- warstwa zapisu danych.

Architektura informacyjna (warstwa dostępu do danych) w przypadku systemu takiego jak POL-on, którego celem jest prawie wyłącznie pozyskiwanie danych, stanowi kluczowy element. Można powiedzieć, że POL-on jest systemem w mniejszym stopniu nastawionym na aspekt behawioralny (czynności, procesy, usługi), a bardziej

na deskryptywny, raportowy. Nie stanowi on też jedynego punktu oddziaływania na środowisko, które opisuje. W praktyce ten często uwarunkowany określonymi celami strategicznymi punkt patrzenia na dane stanowi największą trudność w integracji z zasilającymi go systemami podmiotów i instytucji. Dla przykładu sposób przyznawania grantów stanowi jeden z najbardziej kompleksowych i rozbudowanych procesów biznesowych, które realizowane są w oparciu o inne systemy⁸⁸. W przypadku systemu POL-on zakres danych o projektach naukowych stanowił zawsze finalne podsumowanie o dosyć ograniczonym zakresie informacyjnym, uwarunkowanym głównie ewaluacją jednostek naukowych, która również ewaluowała w kolejnych perspektywach [41].

Kolejnym przykładem może być podejście do kwestii danych związanych z procesami ewaluacji nauki, które skupiały się na bardzo specyficznych pojęciach związanych również z procesem oceny parametrycznej jednostek naukowych, przez co przez dłuższy czas były niekompatybilne z modelami i ontologiami opisującymi w sposób zobiektywizowany dorobek publikacyjny w ujęciu światowym⁸⁹. Dopiero od roku 2018 w związku z nowelizacją prawa i oparciu zasad nowej ewaluacji o integrację z ORCID nastąpiło większe zbliżenie modelu publikacji PBN do ogólnie przyjętych standardów.

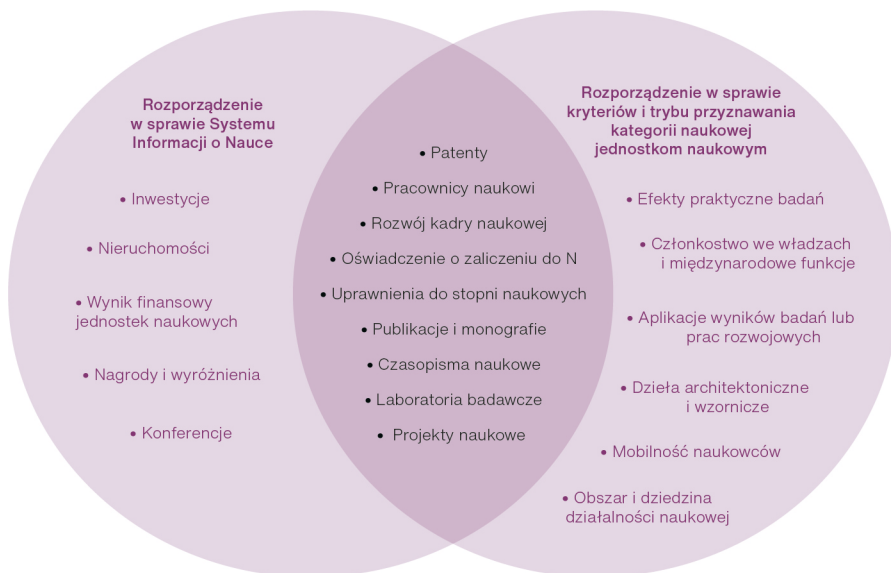
Poniższy uproszczony diagram pokazuje w jaki sposób dane zbierane w systemie od roku 2014 na podstawie Rozporządzenia w sprawie Systemu Informacji o Nauce⁹⁰ przełożyły się na kryteria procesu ewaluacji. Jak widać zakresy te różniły się od siebie w dosyć istotny sposób, co przełożyło się na konieczność rozbudowy i zmiany definicji wielu encji i atrybutów systemu w momencie wykorzystania go do wsparcia procesu ewaluacji jednostek naukowych w roku 2017.



⁸⁸ na tej niwie największym w Polsce systemem jest rozwijany przez Ośrodek Przetwarzania Informacji i specjalizowane agencje system ZSUN/OSF, osf.opi.org.pl (dostęp: 18.06.2021)

⁸⁹ w roku 2017 analiza przeprowadzona w obszarze zgodności dotychczasowego modelu danych PBN z danymi opisującymi publikacje w systemie ORCID wskazywała, że ponad połowa atrybutów opisujących publikacje nie znajdowała swojego odpowiednika (dotyczyło to na przykład w dużej mierze pól charakteryzujących sposób wydania, takich jak np. liczba arkuszy wydawniczych, sposób wydania, a także cech związanych z nagrodami i wyróżnieniami oraz informacji o powiązanych konferencjach i otwartym dostępie

⁹⁰ to była jedna z kilku definicji systemu POL-on przed jego ujednoliceniem na gruncie prawnym



Ilustracja 2.9. Prezentacja różnic w zakresach danych na poziomie aktów prawnych implementowanych w systemie

Można powiedzieć, że podejście do modelu danych w systemach POL-on, PBN polegało bardziej na klasycznym procesie modelowania struktur bazodanowych, aniżeli próbie stworzenia lub odwołania się do istniejących ontologii [33]. W związku z bardzo krótkim czasem powstania poszczególnych modułów⁹¹ ich zakres pojęciowy był zdeterminowany w dużej mierze technologią (relacyjna baza danych SQL) oraz klasycznymi artefaktami procesu wytwórczego (diagramy ERD[•] lub konceptualne modele klas UML[•]). Ponieważ były one tworzone w oparciu o bardzo nieusystematyzowaną dziedzinę i w procesie zmieniających się regulacji prawnych, ich stopień przełożenia na popularne w tym czasie modele typu CERIF był niewielki.

Sytuacja ta zaczęła ulegać zmianie w związku z budową systemu RAD-on, który zawierał w sobie komponent analityczny (hurtownia danych oraz narzędzia analityki biznesowej) oraz umożliwienie otwartego dostępu do danych (*OPEN DATA*). W procesie tym uwzględnione zostały elementy stosowania metadanych dla procesów ETL oraz budowy sieci semantycznych (RDF⁹²).



⁹¹ w pierwszej fazie tworzenia systemu średni czas wytworzenia pojedynczego modułu oscylował w okolicach 3-4 miesięcy

⁹² *Resource Description Framework* – język/metoda pozwalająca na opisywanie zasobów sieci Web, ze składnią opartą na XML, opracowana przez W3C. Służy przedstawieniu (nie zaś wyświetlaniu) wiedzy zawartej w internecie, w sposób zrozumiały dla komputerów (łatwo przetwarzany przez programy komputerowe). Może być wykorzystany w celu tworzenia semantycznego internetu

Zanim to jednak nastąpiło trzonem modelu semantycznego systemu, który przez dłuższy czas oddziaływał na proces integracji ze środowiskiem zewnętrznym był opisany za pomocą XSD schemat plików XML przekazywanych przez podmioty zasilające system. Schemat oprócz technicznej warstwy semantycznej służącej do transferu danych, zawierał warstwę domeny biznesowej wywodzącej się z pojęć aktów prawnych, która była bardzo często różna od stosowanej w systemach uczelnianych i w dominujących ontologiach z tej dziedziny.

Spośród wielu przypadków różnic w interpretacji i definicji określonych pojęć wymienić można:

- pojęcie „korekty” i „biznesowej zmiany danych” – w pierwszym przypadku mamy do czynienia z poprawą, będącą wynikiem typowej pomyłki; druga sytuacja stanowi jasno zdefiniowany w czasie moment, gdy dany obiekt zmienia jeden ze swych atrybutów (np. zmiana nazwiska w wyniku zmiany stanu cywilnego, zmiana stanowiska);
- szczególne trudności wynikały z ujednolicenia pojęć związanych z organizacją nauczania. W 2011 roku w wyniku dostosowania systemu szkolnictwa wyższego do założeń tzw. reformy bolońskiej [10], zniknęło pojęcie definicji kierunków studiów. Od tego momentu każdy z kierunków mógł posługiwać się dowolną formą nazewnictwa, co znacząco utrudniło i skomplikowało analizy odwołujące się do tego atrybutu⁹³;
- pojęcie semestru studenta i jego liczebności w roku akademickim w systemach typu USOS jest wynikiem skomplikowanych operacji wyliczania danych na podstawie konkretnych dat (w praktyce stosowanym pojęciem był tzw. etap studiów);
- niejasności związane z określeniem różnicy pomiędzy językiem filologii obcej (będącym przedmiotem studiów) a językiem, w którym prowadzone są studia;
- szczególnym problemem w pierwszym etapie powstawania systemu było pojęcie specjalizacji, która na gruncie prawa mogła dotyczyć wyłącznie kierunków filologicznych. W praktyce jednak w wewnętrznych systemach to pojęcie było stosowane bardzo szeroko;
- definicja podstawowego miejsca pracy pracowników naukowych i dydaktycznych przez wiele lat była definiowana inaczej w kontekście uczelni oraz instytutów naukowych;
- definicja przekształcenia uczelni, filii uczelni lub instytutu naukowego: część użytkowników systemu jako przekształcenie rozumie akcję, które są de facto likwidacją podmiotu albo zmianą jego nazwy albo zmianą organu nadzorującego albo zmianą typu. W systemie starano się natomiast stosować tę definicję wyłącznie do połączenia kilku instytucji, włączenia jednej z nich do innej albo do podziału jednej z nich na wiele innych, gdyż zmiany te były konieczne do uchwycenia w czasie;
- dane dotyczące liczby godzin w wykazie pracowników oraz liczba godzin zajęć określonych w programie studiów realizowanych w danym roku akademickim podawane w systemie POL-on stanowiły istotny problem interpretacyjny dla wielu uczelni.



⁹³ dla przykładu kierunek o nazwie „informatyka” zyskał wiele reprezentacji np. „informatyka w biznesie”, „informatyka stosowana” etc.

Przykładów tego typu różnic interpretacyjnych było dużo więcej. Warto zwrócić jednak uwagę, że wszystkie z nich z upływem czasu stały się elementem konsensu i doprecyzowania, wzbogacając poziom wspólnej wiedzy i języka opisu rzeczywistości. Problemy tego typu były doświadczeniem wielu państw wdrażających podobne systemy. Efektem tych, trwających niekiedy całe dekady, prac są wymienione w poprzednich rozdziałach modele i ontologie, takie jak np. CERIF.

2.3.2. Bezpieczeństwo i ochrona danych osobowych

Dane osobowe znajdujące się w systemie są w specjalny sposób chronione. Stosowane są zarówno specjalne środki techniczne, jak organizacyjne, które w znaczącym stopniu ograniczają ryzyko naruszenia danych przetwarzanych w systemie POL-on. Dodatkowo POL-on jest systemem teleinformatycznym w rozumieniu art. 3 pkt 3 ustawy o informatyzacji działalności podmiotów realizujących zadania publiczne [39] i spełnia wszystkie wytyczne sformułowanego w oparciu o zapisy tej ustawy rozporządzenia w sprawie Krajowych Ram Interoperacyjności [30]. Co roku system POL-on przechodził (z wynikiem pozytywnym) audyt bezpieczeństwa informacji w nim zawartych, co zapewniło bezpieczeństwo systemu oraz danych, które zostały do niego wprowadzone. Ponadto podmioty występujące w systemie szkolnictwa wyższego i nauki, które są zobowiązane do wprowadzania, aktualizowania, archiwizowania oraz usuwania danych, mają również obowiązek stosować w tym procesie przyjęte przez siebie wewnętrzne polityki bezpieczeństwa danych osobowych, z uwzględnieniem przepisów ustawy [37] oraz przepisów dotyczących ochrony danych osobowych [23].

Administratorem danych osobowych jest⁹⁴ Minister Edukacji i Nauki (MEiN). Administratorami danych są również instytucje⁹⁵, których dane znajdują się w systemie i które wprowadzają dane osobowe do systemu POL-on. Podmiotem przetwarzającym dane w systemie POL-on w imieniu administratora (MEiN) jest Ośrodek Przetwarzania Informacji – Państwowy Instytut Badawczy. OPI PIB jest instytutem badawczym nadzorowanym przez MEiN. Jednym z jego przedmiotów działania jest przedmiot związany ze świadczeniem usług w zakresie systemów informacyjnych, w tym baz danych. W oparciu o powyższe prowadzony jest system POL-on. Zadanie to jest realizowane na podstawie art. 353 ustawy [37] oraz w oparciu o umowę powierzenia przetwarzania danych osobowych zawartą na podstawie art. 28 RODO [23]. Dane w systemie POL-on są przetwarzane przez OPI PIB przez okres wskazany w rozporządzeniu w sprawie POL-on [32]. Dostęp do danych osobowych przetwarzanych w systemie POL-on przysługuje odbiorcom określonym w ustawie [37] i jest ograniczony tylko do tych danych, na które wskazują aktualne przepisy⁹⁶.



⁹⁴ w rozumieniu art. 4 pkt 7 RODO [23]

⁹⁵ w zakresie, w jakim zbierane są dane osobowe pracowników, doktorantów i studentów. MEiN jest administratorem danych dopiero od chwili ich wprowadzenia do systemu POL-on – do tego czasu administratorami są podmioty gromadzące dane

⁹⁶ w przypadku konkretnego odbiorcy

Zgodnie z zapisami ustawy „Prawo o szkolnictwie wyższym i nauce” ([37]) większa część danych wprowadzanych do systemu POL-on jest publicznie dostępna⁹⁷. Wcześniej raporty zawartości danych oparte były na systemie transakcyjnym – na tym samym źródle, na którym działała podstawowa aplikacja systemu. Później (obecnie) do przeglądania danych powstała hurtownia danych zawierająca wszystkie dane, które istnieją w systemie POL-on, a powinny zostać (zgodnie z ustawą [37]) publicznie udostępnione. System raportów ma służyć zarówno odbiorcom nauki (przedsiębiorcom, dziennikarzom, osobom decydującym o kształcie polityki naukowej w kraju), jak osobom, których dane znajdują się w systemie (pracownikom naukowym, studentom i doktorantom) jako podstawowe źródło informacji o nauce w Polsce, zapewniając równocześnie dostęp publiczny (opisany w ustawie [37]) do zamieszczonych w systemie danych. Pełen dostęp do wszystkich wprowadzonych danych mają poszczególne instytucje, których dane znajdują się w POL-onie i które są zobowiązane do wprowadzania danych poprzez system kont, które zostały założone przez ministra (zgodnie z § 17 rozporządzenia o POL-on [32]).

Z przyczyn oczywistych nie zostaną tu omówione dokładnie zastosowane w systemie POL-on mechanizmy ochrony danych osobowych. Mechanizmy te są wystarczające i adekwatne do danych, które mają chronić. Warto nadmienić, że system działa w architekturze rozproszonej geograficznie obejmujących kilka węzłów w trybie *disaster recovery*⁹⁸, czyli utrzymania w ruchu systemu nawet w sytuacji, gdyby fizycznie zniszczony został jeden z nich w wyniku np. klęski żywiołowej.

2.3.3. Rozwiązywanie konfliktów danych

Ponieważ część danych była współdzielona (te same dane były edytowane przez pracowników różnych jednostek naukowych), pojawił się problem istnienia konfliktów danych. Konflikty danych nie można było rozwiązać arbitralnie (jak proponuje np. [15] czy [2]) ze względu na to, że użytkownicy mieli dokumenty uzasadniające dane, które wprowadzili do systemu – wprowadzenie takich, a nie innych danych najczęściej nie wynikało z niedbalstwa, niedokładności lub niedoinformowania użytkowników, ale z dokładnego wprowadzenia udokumentowanych, posiadanych przez użytkowników danych. Przy czym nie istniał żaden zbiór danych⁹⁹, z którego można by było skorzystać w celu weryfikacji wprowadzonych informacji i rozstrzygnięcia, które z nich są poprawne. Poza tym dane wprowadzone do systemu nie były „jednorazowe”, ale miały służyć do identyfikacji obiektów, dla których miały zostać wprowadzone kolejne dane. W szczególnych wypadkach jednostki naukowe mogły również skorzystać finansowo na nieprawidłowo wprowadzonych do systemu danych (np. przez przyznanie dodatkowego finansowania działalności naukowej nieistniejących osób). Dlatego też w systemie



⁹⁷ między innymi po to właśnie powstał system RAD-on: radon.nauka.gov.pl (dostęp: 18.06.2021)

⁹⁸ dosłownie: odtwarzanie po katastrofie

⁹⁹ lub system POL-on nie miał do niego dostępu

zastosowano specyficzne rozwiązywania konfliktów danych, które zależały przede wszystkim od ich rodzaju.

Problem współdzielenia danych można tu wytłumaczyć na przykładzie danych osobowych studenta, który studiuje na więcej niż jednym kierunku studiów prowadzonym przez różne szkoły wyższe. Wszystkie instytucje mają pełen dostęp do danych osobowych tej osoby. Student zmienia nazwisko i przedstawia informacje o tej zmianie tylko w jednej z instytucji. Instytucja ta zmienia dane w POL-on na nowe. Pozostałe szkoły wyższe widzą zmianę, ale w posiadanych dokumentach mają inne (stare?) dane – pracownicy tych instytucji nie wiedzą, czy zmiana została wprowadzona dla odpowiedniej osoby i czy u nich również pojawią się dokumenty uzasadniające wprowadzenie tej zmiany.

W omawianym przypadku występuje ewidentny konflikt danych – dane znajdujące się w systemie nie odpowiadają danym znajdującym się w posiadaniu uczelni. Szkoła wyższa (inna niż ta, w której wprowadzono nowe dane) może w przedstawionym powyżej przypadku postąpić w dwojaki sposób: przywrócić stare dane (oparte na dokumentach posiadanych przez uczelnię – uwaga! nadpiszą one nowe informacje wprowadzone w zmienionych danych) albo przyjąć, że wprowadzona zmiana dotyczyła wskazanej osoby, która w przyszłości przedstawi stosowne dokumenty. Uwaga! Wszystkie dane (wprowadzane przez użytkowników) są równie ważne i nie można tu wskazać użytkownika, który ma najwyższy priorytet. W efekcie troskę o wprowadzenie do systemu poprawnych danych pozostawiono użytkownikom.

Opisany tu sposób postępowania dotyczy większości współdzielonych danych: danych osobowych, danych o realizowanych projektach badawczych, danych na temat wykonanych wdrożeń, otrzymanych patentów, osiągniętych przez jednostkę efektów praktycznych itp. W tym przypadku pełna odpowiedzialność za poprawność wprowadzonych danych spoczywała wyłącznie na użytkownikach systemu, którymi byli pracownicy wskazanych jednostek naukowych, a więc osoby nieprzypadkowe i umocowane prawnie. System był i jest w pełni rozliczalny, a więc w każdej chwili możliwa jest identyfikacja osoby, z której konta zostały wprowadzone konkretne zmiany. To był pierwszy sposób rozwiązania problemu konfliktu danych.

Druga możliwość, to niepozwalanie na wprowadzenie nadmiarowych lub niepoprawnych danych. W takim przypadku system nie pozwala na wprowadzenie danych, dopóki inne dane nie zostaną usunięte lub nie zostanie zakończony okres ich obowiązywania. W takim przypadku, wynikającym z konkretnych zapisów w obowiązujących przepisach prawa, POL-on był gwarantem, że dane nie zostaną wprowadzone nieprawidłowo. Jednak system mógł w ten sposób wypaczyć rzeczywistość, zapobiegając wprowadzeniu, niezgodnych z prawem, ale istniejących faktycznie zdarzeń. W tym przypadku pracownicy instytucji, które miały skonfliktowane dane musieli kontaktować się z pracownikami jednostki, która pierwsza wprowadziła nieprawidłowe dane, i we własnym zakresie rozwiązać istniejące problemy.

Istniała również trzecia możliwość: w niektórych przypadkach pojawiało się ostrzeżenie o niedopuszczalności pewnych danych i wszystkie kolidujące dane były zawieszane

– w efekcie nikt nie mógł mieć z nich żadnego pożytku. Wymagało to oczywiście uruchomienia robota sprawdzającego poprawność danych znajdujących się w systemie. Tak postąpiono w przypadku minimum kadrowego – pracownik mógł być dopisany do dwóch minimów, a jeżeli został dopisany do trzeciego to wszystkie wpisy dotyczące danego okresu były zawieszane i szkoły wyższe musiały się porozumieć w celu określenia tych minimów do których pracownik mógł być przypisany.

Oczywiście poprawność wprowadzanych danych była kontrolowana wszędzie. Nie pozwalano na wprowadzanie danych nieprawidłowych – tylko dane poprawnie wprowadzone były rejestrowane w systemie.

Obecnie, w wersji 2 systemu, wprowadzono jednolity mechanizm prezentacji najbardziej aktualnych danych dotyczących każdego elementu systemu dla danej instytucji. W ten sposób, każda z instytucji, może zamieścić te dane, które jej zdaniem są najbardziej aktualne dla danego obiektu (np. osoby). Dane te są następnie uszeregowane (nie są nigdzie zmieniane) na poziomie aplikacji. W ten sposób pracownicy instytucji nadzorujących system (MEiN, CKdsST, GUS i innych) widzą w systemie najbardziej aktualną wersję danych opisujących obiekt. Podstawą przeprowadzanego łączenia opisów są dane identyfikacyjne danego obiektu (np. dla osoby będzie to numer ORCID, PESEL, numer dowodu osobistego, numer paszportu itp.), które pozwolą na jednoznaczną identyfikację danego obiektu. Przy czym dane obiektu dostępne w systemie nie muszą się pokrywać z danymi opisującymi obiekt w rzeczywistości. Na przykład dla osoby, czym innym są dane identyfikacyjne, a czym innym nazwisko – dane te mogą być niezależnie brane w oparciu o najświeższą wersję danych dostępnych w systemie, lecz będą prezentowane jako dane najbardziej aktualne dla danego obiektu (osoby). Liczy się najbardziej aktualna wersja danych opisujących dany obiekt zarejestrowana dla danej instytucji, następnie wybierana jest najbardziej aktualna wersja danych w całym systemie – jeżeli występuje konflikt danych, tj. w systemie są dwie wersje danych opatrzonej tą samą datą, to decyduje data techniczna wprowadzenia danych – jako najświeższa wersja danych traktowana jest wersja wprowadzona do systemu z najpóźniejszą datą i te dane są prezentowane zarówno instytucjom nadzorującym system, jak i w serwisach internetowych.

Oprócz problemów merytorycznych w aplikacji pojawiły się też problemy techniczne. O ile poprzednio¹⁰⁰ spójność danych była zapewniona na poziomie bazy danych poprzez istnienie indeksów i więzów integralności na samej bazie, o tyle w podejściu mikrouslugowym spójność danych musi zostać zapewniona przez aplikację obsługującą dany fragment systemu. Rodzi to problemy z zapewnieniem jednoznacznej identyfikacji obiektu, która jest podstawą kontroli poprawności danych, a tym samym określeniem wystąpienia potencjalnych konfliktów wprowadzonych do systemu danych. Identyfikacja obiektu pomiędzy poszczególnymi modułami (np. identyfikacja osoby znajdującej się w module „studenci” z danymi osoby znajdującej się w module „pracownicy”) wymaga wykonania dodatkowej warstwy przeszukującej wszystkie moduły, gdzie dane mogą wystąpić.



¹⁰⁰ system w wersji 1 był oparty na relacyjnej bazie danych

SŁOWNIK POJĘĆ

C

CRIS <ang. Current Research Information Systems> Kategoria systemów do zbierania dorobku naukowego oparta o model danych CERIF

E

ERD <ang. Entity-Relationship Diagram> diagram związków encji – rodzaj graficznego przedstawienia związków pomiędzy encjami używany w projektowaniu systemów informacyjnych do zademonstrowania conceptualnych modeli danych używanych w systemie

O

OAI-PMH <ang. Open Archives Initiative Protocol for Metadata Harvesting> protokół komunikacyjny mający na celu zwiększenie interoperacyjności repozytoriów (archiwów, bibliotek). Jest to zestaw sześciu zleceń lub usług wywoływanych w ramach protokołu HTTP. Dzieli użytkowników protokołu na dwie grupy: dostawców danych i dostawców usług. Dostawcy danych to repozytoria, które udostępniają ustrukturyzowane metadane za pośrednictwem OAI-PMH. Dostawcy usług to konsumenci danych udostępnianych przez dostawców danych

ORCID <ang. Open Researcher and Contributor ID> kod alfanumeryczny stworzony w celu identyfikacji autorów publikacji naukowych

R

RESTful API <ang. Representational State Transfer, Application Programming Interface> styl architektury oprogramowania, opierający się o zbiór wcześniej określonych reguł opisujących jak definiowane są zasoby, a także umożliwiających dostęp do nich (REST). Połączony z zestawem reguł definiujących komunikację pomiędzy systemami komputerowymi oraz pomiędzy systemem komputerowym a człowiekiem (API)

U

UML Notacja UML <ang. Unified Modeling Language> umożliwia tworzenie diagramów przedstawiających różne punkty widzenia systemu. Jest przeznaczona do specyfikowania modelu obiektowego systemu informatycznego oraz kontekstu, w jakim system będzie używany na wybranym poziomie szczegółowości. Według definicji UML jest językiem służącym do zapisywania projektu systemu i może być stosowany do

graficznego opracowania, tworzenia, dokumentowania artefaktów podczas procesu budowy oprogramowania

X

XSD-XML Schema Definition Opracowany przez W3C standard służący do definiowania struktury dokumentu XML. XML Schema stanowi alternatywę dla DTD, przy czym posiada znacznie większe możliwości. XML Schema jest strukturą XML, w odróżnieniu od DTD niebędącego częścią standardu XML

SPIS ILUSTRACJI

1.1.	Kluczowe fazy tworzenia i rozbudowy systemu	12
1.2.	Struktura ekosystemu POL-on	17
1.3.	Statystyka aktywnych użytkowników systemu POL-on	19
2.1.	Weryfikacja ilości zarejestrowanej pomocy materialnej typu „zapomoga”	38
2.2.	Weryfikacja ilości zarejestrowanej pomocy materialnej typu „stypendium”	38
2.3.	Ewaluacja podmiotów w zakresie jakości kształcenia, szkół doktorskich i działalności naukowej	40
2.4.	Proces sprawozdawczości w systemie POL-on	41
2.5.	Proces sprawozdawczości GUS w systemie POL-on	42
2.6.	Deponowanie prac dyplomowych w ORPPD 2.0	43
2.7.	Proces ubiegania się o stopień doktora jako przykład procesu awansowego, którego dane gromadzone są w systemie POL-on	44
2.8.	Proces zbierania danych wykorzystywanych w systemie ELA	45
2.9.	Prezentacja różnic w zakresach danych na poziomie aktów prawnych implementowanych w systemie	48

SPIS TABEL

1.1. Statystyki POL-on według stanu na 30.09.2020 – wykaz instytucji.....	19
1.2. Statystyki POL-on według stanu na 30.09.2020 – wykaz pracowników.....	20
1.3. Statystyki POL-on według stanu na 30.09.2020 – wykaz studentów i doktorantów.....	20
1.4. Statystyki POL-on według stanu na 30.09.2020 – osiągnięcia naukowe.....	21
1.5. Statystyki POL-on według stanu na 30.09.2020 – baza dokumentów planistyczno-sprawozdawczych.....	21

BIBLIOGRAFIA

- [1] *Informatyzacja uczelni wyższych w świetle sprawozdawczości na rzecz Ministerstwa Nauki i Szkolnictwa Wyższego*, usos.edu.pl/sites/default/files/pl-cpi-2013-wadowice-polon.pdf (dostęp: 18.06.2021).
- [2] M. M. Beizer, D. Berg, R. Scullard, P. R. Simha, M. A. Solomon, *Method of resolving data conflicts in a shared data environment*, Patent US 6,240,414 B1.
- [3] S. Berglund, *Consortium Governance-Experiences of the Swedish LADOK Consortium*, *Higher Education Management*, 1998, 10(1), s. 7–12.
- [4] I. Blümel, S. Dietze, L. Heller, R. Jäschke, M. Mehlberg, *The Quest for Research Information*, *Procedia Computer Science*, 2014, 33, s. 253–260, DOI: 10.1016/j.procs.2014.06.040, 12th International Conference on Current Research Information Systems, CRIS 2014.
- [5] M. M. Bojko, M. Feldy, A. Knapińska, B. Kowalczyk, M. Ostaszewski, A. Tomczyńska, *Systemy publicznego finansowania nauki w ujęciu międzynarodowym*, Ośrodek Przetwarzania Informacji – Państwowy Instytut Badawczy, 2016.
- [6] D. Celjak, Z. Bekić, M. Cundeković, L. Jertec, M. Milinović, A. Zubić, *DABAR-the national infrastructure for digital repositories*, *EUNIS 23rd annual congress: Shaping the digital future of universities: Book of proceedings*, 2017, s. 16–24.
- [7] M. Czerniak, J. Mincer-Daszkiewicz, K. Sprawnik, *USOS–Uniwersytecki System Obsługi Studiów. Współpraca z IBM*.
- [8] D. Gagliardi, D. Cox, Y. Li, *What are the factors driving and hindering the adoption of open science? An exploratory study*, *Manchester Institute of Innovation Research working paper series*, 2014, DOI: 10.13140/2.1.3016.9602.
- [9] D. Ivanović, D. Surla, M. Racković, *A CERIF data model extension for evaluation and quantitative expression of scientific research results*, *Scientometrics*, 2011, 86(1), s. 155–172, DOI: 10.1007/s11192-010-0228-2.
- [10] R. Jeż, *Proces Boloński a system szkolnictwa wyższego w Polsce*, *Studia Ekonomiczne*, 2013, 131, s. 91–99.
- [11] E. S. Karlsen, G. C. Lingjærde, *CRISTin in the Norwegian research landscape*, hdl.handle.net/11366/975 (dostęp: 18.06.2021).
- [12] P. Korytkowski, E. Kulczycki, *Publication counting methods for a national research evaluation exercise*, *Journal of Informetrics*, 2019, 13(3), s. 804–816, DOI: 10.1016/j.joi.2019.07.001.

- [13] M. Marques, J. J. Powell, M. Zapp, G. Biesta, *How does research evaluation impact educational research? Exploring intended and unintended consequences of research assessment in the United Kingdom, 1986–2014*, *European Educational Research Journal*, 2017, 16(6), s. 820–842, DOI: 10.1177/1474904117730159.
- [14] M. Michajłowicz, M. Niemczyk, J. Protasiewicz, K. Mroczkowska, *POL-on: The Information System of Science and Higher Education in Poland*, *EUNIS 2018 Congress Book of Proceedings*, 2018.
- [15] F. Naumann, M. Häussler, *Declarative data merging with conflict resolution*, *Proceedings of the Seventh International Conference on Information Quality (ICIQ-02)*, 2002, s. 212–224.
- [16] S. V. Nordgård, *Integration Governance-A NTNU case study, how Enterprise modeling and Architecture improves Integration Governance*, Praca magisterska, NTNU, 2015.
- [17] C. S. Pinto, C. Simões, L. Amaral, *CERIF – Is the Standard Helping to Improve CRIS?*, *Procedia Computer Science*, 2014, 33, s. 80–85, ISSN 1877-0509, DOI: 10.1016/j.procs.2014.06.013.
- [18] J. Protasiewicz, S. Rosiak, I. Kucharska, E. Podwysocki, M. Niemczyk, Ł. Błaszczuk, M. Michajłowicz, *RAD-on: An integrated System of Services for Science – Online Elections for the Council of Scientific Excellence in Poland*, *European Journal of Higher Education IT*, 2019, 1.
- [19] M. L. Ramirez, J. T. Dalton, G. McMillan, M. Read, N. H. Seamans, *Do Open Access Electronic Theses and Dissertations Diminish Publishing Opportunities in the Social Sciences and Humanities? Findings from a 2011 Survey of Academic Publishers*, *College & Research Libraries*, 2013, 74(4), s. 368–380, DOI: 10.5860/crl-356.
- [20] L. Ribeiro, P. De-Castro, M. Mennielli, *Final report: EUNIS-EUROCRIS joint survey on CRIS and IR*, eunis.org/wp-content/uploads/2016/03/cris-report-ED.pdf (dostęp: 18.06.2021).
- [21] L. M. Ribeiro, R. H. Pereira, O. Pacheco, M. Bernardes, R. T. Martins, *Interoperability between Information Systems of Portuguese Higher Education Institutions*, *EUNIS 22nd Annual Congress Book of Proceedings*, 2016, s. 203–214.
- [22] M. Richards, *Software architecture patterns*, vol. 4, O'Reilly Media, 2015, ISBN 978-1491924242.
- [23] *Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE*, Dz. Urz. UE. L Nr 119, 4.5.2016.

- [24] *Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 1 września 2011 r. w sprawie tytułów zawodowych nadawanych absolwentom studiów, warunków wydawania oraz niezbędnych elementów dyplomów ukończenia studiów i świadectw ukończenia studiów podyplomowych oraz wzoru suplementu do dyplomu*, Dz.U. 2011 nr 196 poz. 1167.
- [25] *Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 22 września 2011 r. w sprawie danych zamieszczanych w ogólnopolskim wykazie studentów*, Dz.U. 2011 nr 204 poz. 1201.
- [26] *Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 29 września 2011 r. w sprawie centralnego wykazu nauczycieli akademickich i pracowników naukowych*, Dz.U. 2011 nr 207 poz. 1236.
- [27] *Rozporządzenie Prezesa Rady Ministrów z dnia 14 września 2011 r. w sprawie liczby członków Centralnej Komisji do Spraw Stopni i Tytułów oraz warunków i trybów ich wyboru*, Dz.U. 2011, poz. 1185.
- [28] *Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 23 sierpnia 2012 r. zmieniające rozporządzenie w sprawie warunków prowadzenia studiów na określonym kierunku i poziomie kształcenia*, Dz.U. 2012 poz. 983.
- [29] *Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 29 czerwca 2015 r. w sprawie Systemu Informacji o Nauce*, Dz.U. 2015 poz. 944.
- [30] *Rozporządzenie Rady Ministrów z dnia 12 kwietnia 2012 r. w sprawie Krajowych Ram Interoperacyjności, minimalnych wymagań dla rejestrów publicznych i wymiany informacji w postaci elektronicznej oraz minimalnych wymagań dla systemów teleinformatycznych*, Dz.U. 2012 poz. 526 z późn. zm.
- [31] *Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 22 lutego 2019 r. w sprawie ewaluacji jakości działalności naukowej*, Dz.U. 2019 poz. 392.
- [32] *Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 6 marca 2019 r. w sprawie danych przetwarzanych w Zintegrowanym Systemie Informacji o Szkolnictwie Wyższym i Nauce POL-on*, Dz.U. 2019 poz. 496.
- [33] P. Spyns, R. Meersman, M. Jarrar, *Data modelling versus ontology engineering*, *ACM SIGMod Record*, 2002, 31(4), s. 12–17, DOI: 10.1145/637411.637413.
- [34] W. Steinhoff, *Mapping the NARCIS aggregated domain to the OpenAIRE CERIF model*, hdl.handle.net/11366/989 (dostęp: 18.06.2021).
- [35] *Ustawa z dnia 18 marca 2011 r. o zmianie ustawy – Prawo o szkolnictwie wyższym, ustawy o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki oraz o zmianie niektórych innych ustaw*, Dz.U. 2011 nr 84 poz. 455.
- [36] *Ustawa z dnia 15 stycznia 2015 r. o zmianie ustawy o zasadach finansowania nauki oraz niektórych innych ustaw*, Dz.U. 2015 poz. 249.

- [37] *Ustawa z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce*, Dz.U. 2018 poz. 1668.
- [38] *Ustawa z dnia 3 lipca 2018 r. – Przepisy wprowadzające ustawę – Prawo o szkolnictwie wyższym i nauce*, Dz.U. 2018, poz. 1669.
- [39] *Ustawa z dnia 17 lutego 2005 r. o informatyzacji działalności podmiotów realizujących zadania publiczne*, Dz.U. 2005 nr 64 poz. 565 z późn. zm.
- [40] R. Vogl, H. Angenent, D. Rudolph, A. Thoring, C. Schild, S. Stieglitz, C. Meske, “*sciebo—theCampuscloud*” for NRW, *EUNIS Journal of Higher Education*, 2015.
- [41] J. Wilkin, *Ocena parametryczna czasopism naukowych w Polsce – podstawy metodologiczne, znaczenie praktyczne, trudności realizacji i perspektywy*, Nauka, 2013, 1.