OPI
OŚRODEK PRZETWARZANIA INFORMACJI
PAŃSTWOWY INSTYTUT BADAWCZY

Integracja ORPPD z uczelnianymi systemami antyplagiatowymi

Jarosław Protasiewicz
Ośrodek Przetwarzania Informacji - Państwowy Instytut Badawczy

Europejskie Forum Antyplagiatowe - EFA 2015
Warszawa, 19 października 2015

Plan wystąpienia

1. Podstawy prawne
2. Podstawy organizacyjne
3. Schemat integracji
4. Procedura dostępu
5. Techniczna integracja (opcjonalnie)
6. Zasilanie ORPPD (opcjonalnie)
7. Podsumowanie

1. Podstawy prawne

Ustawa „Prawo o szkolnictwie wyższym”

Art. 167b.

1. Minister właściwy do spraw szkolnictwa wyższego prowadzi ogólnopolskie repozytorium pisemnych prac dyplomowych.

4. Dane, o których mowa w ust. 2, wprowadzają do Systemu POL-on rektorzy uczelni. Dostęp do danych przysługuje promotorowi pracy dyplomowej oraz Komisji.

Art. 167a.

4. Uczelnia jest obowiązana do sprawdzania pisemnych prac dyplomowych przed egzaminem dyplomowym z wykorzystaniem programów antyplagiatowych współpracujących z ogólnopolskim repozytorium pisemnych prac dyplomowych.

1. Podstawy prawne

Obowiązki uczelni

1. **Sprawdzanie pisemnych prac dyplomowych** przed egzaminem dyplomowym z wykorzystaniem programów antyplagiatowych współpracujących z ORPPD.

2. **Zamieszczanie pisemnych prac dyplomowych w ORPPD** - uczelnia (rektor) ma obowiązek wprowadzenia pracy dyplomowej niezwłocznie po zdaniu egzaminu dyplomowego.

Pisemne prace dyplomowe, których obrona (pozytywna) odbyła się po dniu **30 września 2009 r.,** muszą być przekazane do ORPPD w terminie do dnia **31 grudnia 2016 r.**

2. Podstawy organizacyjne

Rola OPI PIB

Ogólnopolskie Repozytorium Pisemnych Prac Dyplomowych zostało powierzone OPI PIB przez MNiSW.

Odpowiedzialność OPI PIB jest następująca:

1. Rozwój oprogramowania ORPPD.
2. Utrzymanie bazy danych i systemu w centrum danych OPI PIB.
3. Świadczenie pomocy użytkownikom.
4. Integracja ORPPD z systemami antyplagiatowymi.

2. Podstawy organizacyjne

Współpraca OPI PIB z systemami antyplagiatowymi

1. Zapewniamy jednakowy dostęp do informacji oraz jednakowe warunki dostępu do ORPPD dla wszystkich podmiotów dostarczających rozwiązania antyplagiatowe na rynku polskim.
2. Nie wypowiadamy się na temat jakości systemów antyplagiatowych.
3. Nie wytwarzamy własnych systemów antyplagiatowych.
4. Nie wspieramy żadnych konkretnych rozwiązań.

Uwaga: W konferencji uczestniczy przedsiębiorstwo Turnitin, natomiast brak jest pozostałych dostawców systemów antyplagiatowych. Zdarzenie to jest niezależne od OPI PIB. Nie wspieramy tego przedsiębiorstwa jak też nie wyrażamy żadnych opinii na temat jego produktów.

2. Podstawy organizacyjne

OPI
OŚRODEK PRZETWARZANIA INFORMACJI
PAŃSTWOWY INSTYTUT BADAWCZY

Historia

1. Wersja „pre-beta” została wykonana przez ICM UW w ramach projektu „System informacji o szkolnictwie wyższym POL-on.”
2. Kod źródłowy systemu został przejęty przez OPI PIB i gruntownie przebudowany; system nadal jest rozwijany.
3. ORPPD zostało uruchomione 11.2014 w obszarze zasilania pracami dyplomowymi; obecnie zgromadzono **313 069** prac.
4. Trwają prace nad integracją z systemami antyplagiatowymi.

Pomoc: https://polon.nauka.gov.pl/repozytorium-orpd

Logowanie: https://polon.nauka.gov.pl/orpd/login

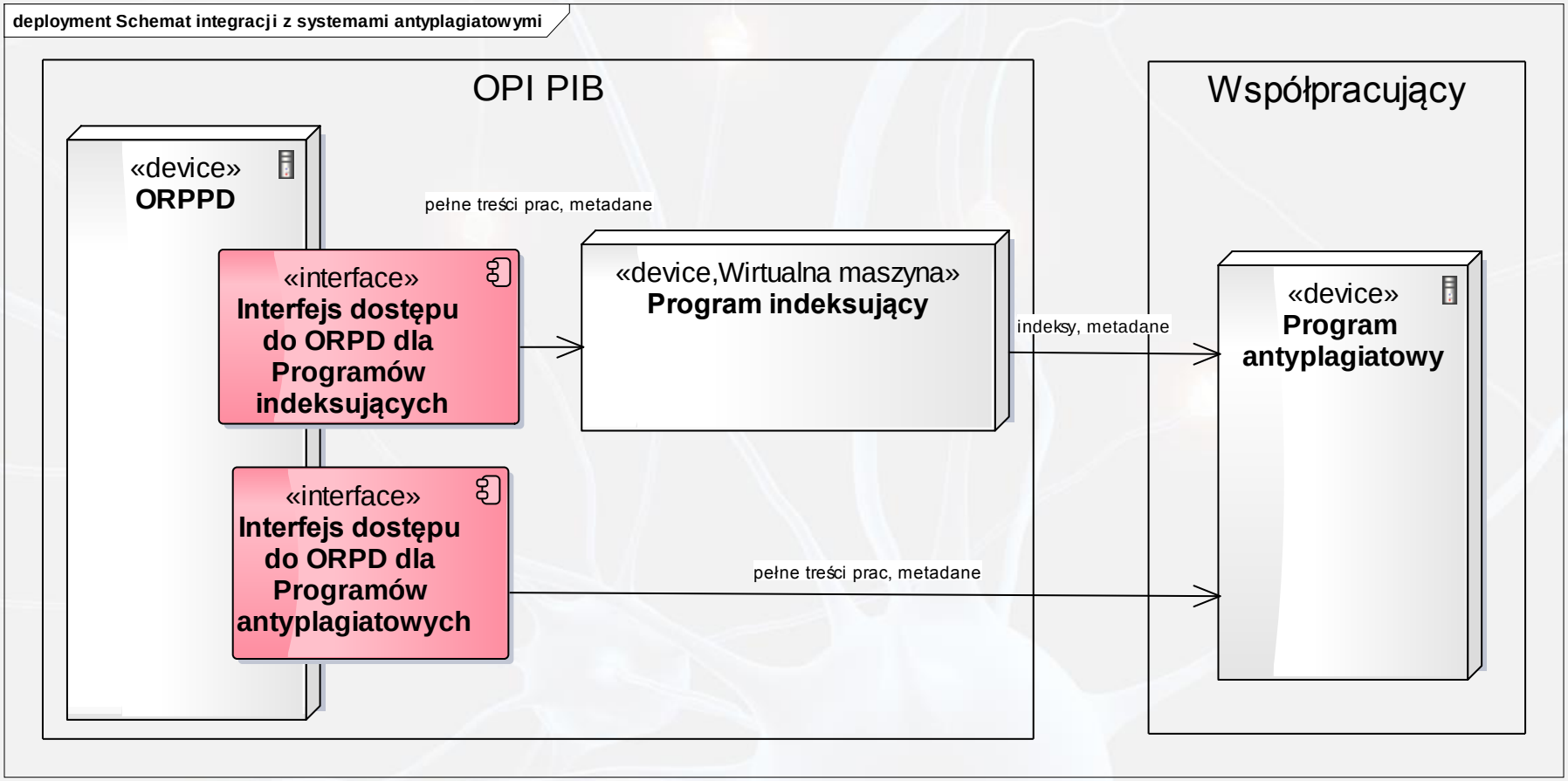
3. Schemat integracji

OPI
OŚRODEK PRZETWARZANIA INFORMACJI
PAŃSTWOWY INSTYTUT BADAWCZY

Podstawowe założenia

1. Prace dyplomowe nie są udostępniane poza serwerownię ORPPD za wyjątkiem promotorów i Polskiej Komisji Akredytacyjnej.
2. Każdy dostawca systemu antyplagiatowego otrzyma wirtualne maszyny, na których umieści oprogramowanie współpracujące z ORPD (nie cały system, lecz tylko program indeksujący).
3. Każdy dostawca będzie miał zagwarantowane takie same parametry techniczne (pamięć, procesor, sieć) w ORPPD.
4. Każdy dostawca zapewnia w swoim zakresie wszystkie niezbędne licencje do działania swojego oprogramowania.

3. Schemat integracji



3. Schemat integracji

Tryb B

- **Program indeksujący** znajdujący się w wirtualnej maszynie ma dostęp do pełnych treści prac i ich metadanych w ORPPD poprzez **Interfejs dostępu do ORPPD dla Programów indeksujących**

- **Program antyplagiatowy** ma dostęp do metadanych z ORPPD oraz indeksów wytworzonych przez Program indeksujący.

3. Schemat integracji

Tryb A

- Zawiera Tryb B.
- **Program antyplagiatowy** ma dostęp do pełnych treści prac, wobec których istnieje podejrzenie plagiatu poprzez **Interfejs dostępu do ORPPD dla Programów antyplagiatowych.**

Współpracujący zapewnia, że prace są udostępnianie wyłącznie uprawnionym promotorom, tzn. promotor otrzymuje dostęp wyłącznie do tych prac wobec których zachodzi podejrzenie plagiatu w stosunku do sprawdzanej i promowanej przez niego pracy.

3. Schemat integracji

OPI
OŚRODEK PRZETWARZANIA INFORMACJI
PAŃSTWOWY INSTYTUT BADAWCZY

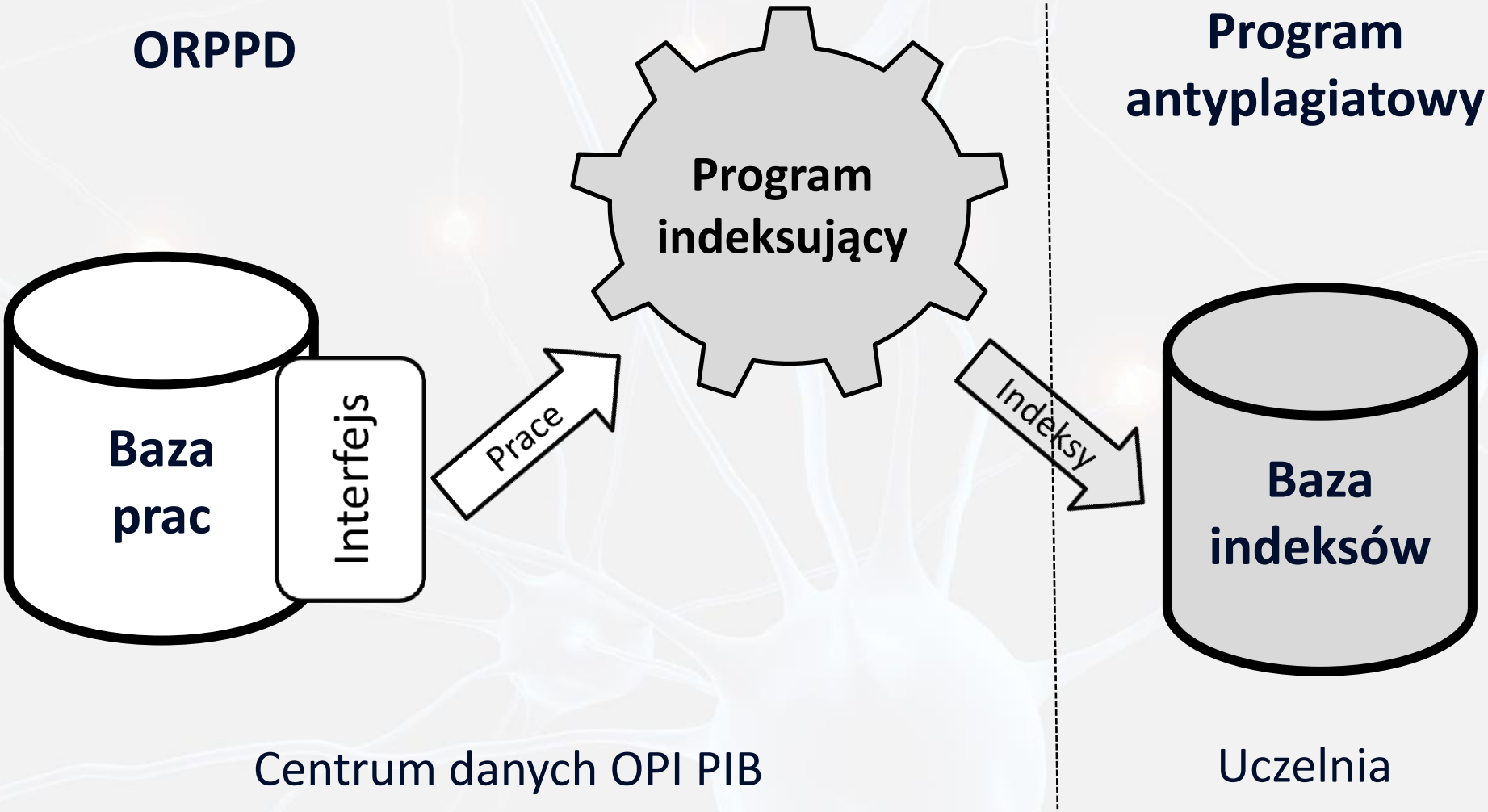
Tryb A

- Indeksacja prac przez program indeksujący umieszczony w ORPPD
- Przesłanie indeksów poza ORPPD.
- Odstęp do pełnych treści podejrzanych prac dla promotorów bezpośrednio poprzez program antyplagiatowy.

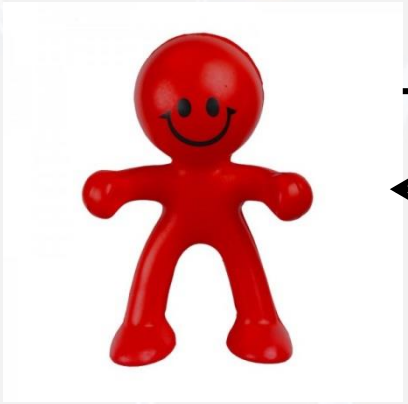
Tryb B

- Indeksacja prac przez program indeksujący umieszczony w ORPPD
- Przesłanie indeksów poza ORPPD.
- Odstęp do pełnych treści podejrzanych prac dla promotorów poprzez ORPPD (program antyplagiatowy wyświetli tylko metadane prac).

3. Schemat integracji



3. Schemat integracji



1. Sprawdź pracę

6. Raport z podejrzanymi częściami prac

ORPPD

Interfejs

4. Daj treści prac podejrzanych o plagiat

5. Treści prac podejrzanych o plagiat

Program anty-plagiatowy

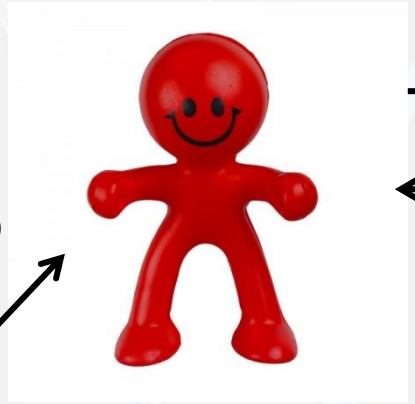
3.

2.

Baza indeksów

TRYB A

3. Schemat integracji



1. Sprawdź pracę

4. Raport z metadanymi podejrzanych prac

6. Dostęp do prac

ORPPD

Interfejs

Program anty-plagiatowy

5. Daj dostęp do podejrzanych prac

4. Raport z metadanymi podejrzanych prac

3.

2.

Baza indeksów

TRYB B

4. Procedura dostępu

Warunki przystąpienia do fazy wstępnej

- Przedstawienie struktury indeksu.
- Złożenie oświadczenia o niemożliwości odtworzenia treści pracy z indeksu.
- Stosowanie odpowiednich środków organizacyjnych i technicznych stawianych systemom teleinformatycznym.
- Przedstawienie technicznego i biznesowego opisu procesu udostępniania promotorom pełnych treści prac dyplomowych w przypadku wyboru trybu A.

OPI PIB zachowuje prawo audytu Współpracującego.

4. Procedura dostępu

Fazy wstępna

- Instalacja na Programu indeksującego Wirtualnej maszynie w centrum danych OPI PIB.
- Integracja Programu indeksujący z ORPPD wersja DEMO poprzez Interfejs dostępu do ORPPD dla Programów indeksujących.
- W przypadku wyboru Trybu A - integracja Programu antyplagiatowego z ORPPD wersja DEMO poprzez Interfejs dostępu do ORPPD dla Programów antyplagiatowych.
- Testy funkcjonalne na wersji DEMO i publikacja wyników OPI PIB (czas indeksacji prac dyplomowych, struktura indeksów).

4. Procedura dostępu

Fazy produkcyjna

- Integracja Programu indeksujący z ORPPD wersja PRODUKCYJNA poprzez Interfejs dostępu do ORPPD dla Programów indeksujących.
- W przypadku wyboru Trybu A - integracja Programu antyplagiatowego z ORPPD wersja PRODUKCYJNA poprzez Interfejs dostępu do ORPPD dla Programów antyplagiatowych.
- Testy funkcjonalne na wersja PRODUKCYJNA i publikacja wyników OPI PIB (czas indeksacji prac dyplomowych, struktura indeksów).

5. Techniczna integracja

Interfejs dostępu do ORPPD dla Programów indeksujących

1. Program indeksujący przesyła zapytanie o listy z pracami dyplomowymi (resorce list – wszystkie prace, change list – prace, które zostały zmienione) .
2. Autoryzacja Programu indeksującego: sprawdzenie zgodności nazwy użytkownika, hasła, IP z danymi przechowywanymi w ORPPD.
3. Wynik autoryzacji pozytywny - odczytanie list wygenerowanych w ORPPD i przesłanie do Programu indeksującego.
4. Wynik autoryzacji negatywny – przesłanie komunikatu informującego o negatywnej autoryzacji.

5. Techniczna integracja

(Tryb A) Interfejs dostępu do ORPPD dla Programów antyplagiatowych

1. Program antyplagiatowy przesyła zapytanie o pełną treść pracy dyplomowej z parametrami:
   - login i hasło Programu antyplagiatowego (login, hasło),
   - identyfikator pracy (metadateId),
   - dane promotora odpytującego:
     - ✓ kod jednostki (institutionCode),
     - ✓ e-mail promotora (e-mail),
     - ✓ imiona (ferstName),
     - ✓ nazwisko (lastName)

5. Techniczna integracja

(Tryb A) Interfejs dostępu do ORPPD dla Programów antyplagiatowych

2. Autoryzacja Programu antyplagiatowego:
   - sprawdzenie zgodności użytkownika, hasła, IP z danymi przechowywanymi w ORPPD.
3. Wynik autoryzacji pozytywny:
   - odczytanie i przesłanie pliku zip z metadanymi i plikami z treścią pracy dyplomowej.
4. Wynik autoryzacji negatywny:
   - przesłanie komunikatu informującego o negatywnej autoryzacji.

5. Techniczna integracja

(TYB B) Dostęp do podejrzanych prac za pomocą tokenów.

1. Użytkownik z rolą INST_ORPD_EDYTOR_ADM generuje token dostępu do pełnych treści prac (potencjalnych plagiatów) z interfejsu www, podając następujące dane:
   - identyfikatory prac (metadateId),
   - dane promotora:
     - kod jednostki (institutionCode),
     - e-mail (e-mail),
     - imiona i nazwisko.
2. Token zostaje przesłany na adres promotora.
3. Zapisanie informacji o wykonanej operacji.

6. Zasilanie ORPPD

OPI
OŚRODEK PRZETWARZANIA INFORMACJI
PAŃSTWOWY INSTYTUT BADAWCZY

Wysyłanie do ORPPD pliku XML

1. Przesłanie pliku zip z: plik **thesisMetadata.xml** z metadanymi pracy, plikami pracy (tekstowe lub binarne)
2. Autoryzacja jednostki zasilającej – sprawdzenie zgodności użytkownika i hasła z danymi przechowywanymi w ORPPD.
3. Sprawdzenie powiązania jednostki obrony pracy z jednostka zasilającą: musi istnieć powiązanie pomiędzy jednostką zasilającą a jednostką obrony wynikające ze struktury albo z przejęcia.
4. Walidacja metadanych pracy.
5. Zapianie lub uaktualnienie pracy.

6. Zasilanie ORPPD

Integracja z Archiwum Prac Dyplomowych

1. Udostępnienie danych prac przez uczelniane archiwa w formacie ETD-MS v1.1: an Interoperability Metadata Standard for Electronic Theses and Dissertations - http://www.ndltd.org/standards/metadata.
2. Pobranie udostępnionych prac przez ORPPD (raz dziennie).
3. Walidacja pobranych danych i zapisanie w ORPPD.
4. Zapisanie raportów z importu.

7. Podsumowanie

1. Obecnie ORPPD jest zasilane pracami.
2. Wzór umowy pomiędzy OPI PIB a dostawcami programów antyplagiatowych czeka na akceptację MNiSW. Z tego względu zaproponujemy tryb testowy.
3. Tryb A integracji – program antyplagiatowy otrzymuje z ORPPD tylko metadane podejrzanych prac; dostęp do pełnych treści pracy przez ORPPD.
4. Tryb B integracji – program antyplagiatowy otrzymuje z ORPPD pełne treści podejrzanych prac pod warunkiem dostarczenia ich tylko do promotora i braku zapisu na nośniku trwałym.

OŚRODEK PRZETWARZANIA INFORMACJI
PAŃSTWOWY INSTYTUT BADAWCZY

Ośrodek Przetwarzania Informacji
Państwowy Instytut Badawczy

al. Niepodległości 188 b
00-608 Warszawa

e-mail: jaroslaw.protasiewicz@opi.org.pl