

Szkolenie systemu POL-on

dr Piotr Rodzik
ekspert systemu POL-on

Ośrodek Przetwarzania Informacji - Państwowy Instytut Badawczy
Al. Niepodległości 188B, 00-608 Warszawa
Numer KRS: 0000127372 • Sąd Rejonowy dla m. st. Warszawy w Warszawie XII Wydział Gospodarczy KRS
REGON: 006746090 • NIP: 525-000-91-40 • <http://www.opi.org.pl/> • e-mail: opi@opi.org.pl

Projekt współfinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚĆ



OPI



INDEX COPERNICUS
INTERNATIONAL

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Masowy import danych

Projekt współfinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚĆ



OPI



INDEX COPERNICUS
INTERNATIONAL

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Jako **Masowy import danych** rozumiane jest przesłanie do systemu POL-on odpowiednio przygotowanego pliku w formacie XML (ang. Extensible Markup Language) zawierającego dane jednego lub wielu obiektów.

Moduł dostępny jest z menu głównego POL-on –
opcja **Administracja** -> **Lista importów**
-> **Lista eksportów**

Do obsługi modułu służy rola **INST_IMPORT**.

Rola może zostać nadana na dowolnym poziomie struktury instytucji.

Pozwala na wykonywanie operacji importu tylko dla jednostki, na poziomie której, przypisano rolę oraz dla jej jednostek podrzędnych.

Założenia zrealizowanej modyfikacji procesu importu

1. Minimalizacja nakładów związanych ze zmianą zawartości przygotowywanych plików - nowy model importu utrzymuje w dalszym ciągu większość atrybutów i rozwiązań z poprzedniej wersji.
2. Stworzenie możliwości ewolucyjnego przejścia w dłuższej perspektywie z dotychczasowej struktury importu na rozwiązania skierowane na model usługowy (w przyszłości).
3. Uwzględnienie możliwie jak największej liczby zgłoszonych przez Państwa propozycji dotyczących modyfikacji obecnej struktury plików XSD, mających na celu uproszczenie sposobu raportowania.

Wprowadzone zmiany

1. Zmiany merytoryczne wynikające z nowelizacji ustawy Prawo o szkolnictwie wyższym:
 - Rozszerzenie zakresu danych osobowych.
 - Usunięcie danych dotyczących odpłatności za studia.
 - Dodatkowe atrybuty studenta związane z zakończeniem studiów oraz uzyskaniem tytułu zawodowego - absolwenci.
 - Nowy rejestr studentów studiów trzeciego stopnia - doktoranci.
 - Dodatkowe informacje o liczbie godzin zajęć dydaktycznych nauczycieli akademickich.

2. Zmiany merytoryczne wynikające z nowelizacji Ustawy o zasadach finansowania nauki:
- **Rozszerzenie zakresu importowanych danych – zmiana definicji „pracownik naukowy”.**
 - **Rozszerzenie zakresu informacji o zatrudnieniu.**
 - **Rozszerzenie słowników określających stanowisko, na którym zatrudniony jest pracownik.**
 - **Dodanie informacji o złożeniu przez pracownika oświadczenia o zaliczeniu go do pracowników B+R jednostki.**
 - **Dodanie informacji o prowadzeniu prac B+R oraz dziedzinie /dyscyplinie, w której są prowadzone.**
 - **Dodanie informacji o instytucji prowadzącej postępowanie w sprawie nadania tytułu naukowego profesora.**

3. Zmiany merytoryczne wynikające ze zgłoszonych uwag i wprowadzonych udogodnień (wstrzymanych przez „zamrożenie” XSD):
- Zmiany w modelu zatrudnienia nauczycieli akademickich.
 - Przedstawienie zgody rektora na dodatkowe zatrudnienie pracownika jako informacji powiązanej z pracownikiem.
 - Dodanie możliwości wskazania dyscypliny realizowanej przez doktoranta w ramach danych studiów doktoranckich.
(obecnie na żądanie MNiSW nie wykorzystywane)
 - ...

4. Zmiany w architekturze procesu:
 - Import całodobowy.
 - Zarządzanie kolejką importu.
 - Zmiana mechanizmu blokowania danych.
5. Dodanie możliwości wywoływania usług z poziomu pliku XML.
6. Dodanie możliwości eksportowania danych z systemu zgodnie z nowym schematem XSD.

Szczegółowe informacje o:

- przebiegu masowego importu danych
- wymaganiach dotyczących składni pliku XML
- regułach walidacji danych wprowadzanych do systemu
- słownikach danych wykorzystywanych w imporcie
- dostępności środowiska testowego dla importu
- przykładowych plikach XML

oraz Instrukcję importu

można znaleźć na stronie:

<http://polon.nauka.gov.pl/dokumentacja-masowego-importu-danych>

Zakres importowanych danych

- dane o nauczycielach akademickich i pracownikach naukowych,
- dane o doktorantach,
- dane o studentach i absolwentach (dla szkół wyższych).

Dla pozostałych danych nie ma
(i prawdopodobnie nie będzie)
opcji importu masowego.

Przesłany plik musi:

- być zgodny z aktualnie obowiązującą, opublikowaną przez OPI PIB za pomocą pliku XSD, definicją schematów importu (obecnie wersja 6.0.1),
- zostać przygotowany w oparciu o aktualną „Instrukcję importu” opublikowaną przez OPI PIB w formie pliku PDF,
- zawierać tylko jeden typ obiektów lub listę operacji,
- odwoływać się do aktualnych słowników systemowych.

Przygotowanie importu

1. Nadanie w systemie roli INST_IMPORT w kontekście wybranej jednostki.
2. Pobranie z systemu aktualnych słowników (dla wybranej jednostki).
3. Wprowadzenie aktualnych identyfikatorów danych słownikowych do lokalnego systemu, z którego eksportowane będą dane.
4. Ustalenie zakresu danych i eksport wybranych danych do pliku XML.
5. Skompresowanie pliku do formatu ZIP.
6. Przesłanie pliku do systemu POL-on.

Usługi sieciowe – pobieranie słowników

1. W części dotyczącej słowników stałych z powszechnie dostępnej strony systemu
<http://polon.nauka.gov.pl/dokumentacja-masowego-importu-danych>
2. Wszystkich słowników (również specyficznych dla jednostki) po zalogowaniu:
Administracja -> Lista importów -> [Pobierz słowniki importu]

3. Wszystkich słowników za pomocą usług sieciowych (web services) za pomocą usług zdefiniowanych na stronie:
<https://polon.nauka.gov.pl/polon-ws/slownik?wsdl>
szczegółowy opis wywoływania poszczególnych operacji można znaleźć w pliku Usługi_Sieciowe_Slowniki.pdf dostępnym na stronie
<https://polon.nauka.gov.pl/dokumentacja-masowego-importu-danych>

Sprawdzenie poprawności plików XML

Sprawdzenie technicznej struktury pliku:

1. Plik po skompresowaniu do formatu ZIP nie może być większy niż 1.91 MB.
2. Przesłany plik ZIP musi zawierać tylko jeden plik o rozszerzeniu .XML.
3. Plik XML musi zawierać dane kodowane zgodnie ze standardem UTF-8.
4. Struktura pliku musi być zgodna z opublikowanym na stronie POL-on schematem XSD.

Weryfikacja wstępna:

1. Plik musi zawierać dane: ***studentów/absolwentów albo pracowników albo doktorantów albo listę operacji.***
2. Plik musi zawierać dane wykorzystujące aktualną wersję słowników systemowych.
3. Plik musi zawierać przynajmniej jeden identyfikowalny obiekt deklarowanego typu.
4. Plik musi zawierać dane odpowiadające zakresowi uprawnień użytkownika, który importuje plik.

Pliki XSD

dostępne schematy XML

Projekt współfinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



KAPITAŁ LUDZKI
NARODOWA STRATEGICZNA SPÓŁNOŚĆ



OPI



INDEX COPERNICUS
INTERNATIONAL

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



W nowej wersji importu dostępna jest:

- **tylko jedna wersja pliku XSD dla pliku zawierającego dane** (obecnie jest to wersja 6.0.1)
- **jedna wersja pliku zawierającego operacje** (obecnie jest to wersja 4.0.0)

Szczegółowy opis i zawartość aktualnych wersji plików XSD zawsze można znaleźć na stronie

<http://polon.nauka.gov.pl/dokumentacja-masowego-importu-danych>

Znajdują się tam również przykłady plików XML zgodnych z tą specyfikacją.

Poprzednie wersje XSD: 2.0, 2.1 i 4.0 nie będą już przez system wspierane.

Aktualna struktura plików XSD:

Dane:

- plikImp.xsd
- element
 - doktorant.xsd
 - osoba.xsd
 - pracownik.xsd
 - student.xsd
- wspólne
 - wspólne.xsd

Usługi (lista operacji):

- _biblioteka.xsd
- _operacje.xsd
- uslugi.xsd

Ważne zmiany w XSD:

Student:

- rozszerzenie opisu osoby: płeć, rok urodzenia, obywatelstwa, kraj pochodzenia
- rozszerzenie informacji o studiach: miejsce zamieszkania przed studiami, data ukończenia, uzyskany tytuł, numer dyplomu
- zmiana danych o punktach ECTS: podaje się punkty uzyskane i uznane jako efekt kształcenia
- Przeniesienie sekcji <pomocMat> na zewnątrz sekcji <semestr> - jest ona obecnie elementem sekcji <daneDotyczaceStudiow>

Doktorant:

- całkowicie nowa sekcja
- podaje się dane:
 - dotyczące osoby
 - przyporządkowujące do studiów doktoranckich
 - o pobieranej pomocy materialnej
 - opis studiów doktoranckich został rozszerzony o możliwość wskazania dziedziny i dyscypliny opisujących te studia - **aktualnie przekazywane w sekcji wartości są ignorowane. Prosimy nie wypełniać!**

Pracownik:

- zamiast elementu <pracownicyDydaktycznoNaukowi> mamy teraz po prostu element <pracownicy>,
- rozszerzenie opisu osoby: płeć, rok urodzenia, kraj pochodzenia,
- uniwersalny element <zatrudnienie> zastępuje elementy istniejące poprzednio: <nauczycielAkademickiZatrudnienie> i <pracownikNaukowyZatrudnienie>,
- zmiana sensu sekcji <zatrudnienie>
- wprowadzenie sekcji <warunkiZatrudnienia>

- wprowadzenie sekcji <dziedzinaDyscyplinaBr>
- wprowadzenie sekcji <oswiadczenieBr>
- w sekcji <warunkiZatrudnienia> rozszerzony zostaje zestaw pól opisujących stanowisko: pola: <charakterWykPracy>, <grupaStanowisk>, <stanowiskoKod>, <stanowiskoInne>, <realizacjaPracBr>
- zmianie/rozszerzeniu ulegają słowniki związane z tymi polami
- element <tytulNaukowy> rozszerzony jest o pola wskazujące na jednostkę w której przeprowadzono postępowanie w sprawie nadania tytułu naukowego: <instytucjaKod>, <instytucjaNazwa> i <instytucjaKrajKod>

- sekcja <zgodaDodZatr> została umiejscowiona w konkretnym zatrudnieniu i określona w czasie
- w sekcji <zatrudnienie> dodano sekcję <rocznyWymiarZajecDydakt>
- w sekcji <funkcja> dodano elementy <email> i <telefon>

Narzędzie wstępnej weryfikacji pliku XML

Dla wygody użytkowników wprowadzono narzędzie wstępnej, strukturalnej weryfikacji zgodności przygotowanego pliku z aktualną wersją schematu XSD.

Narzędzie wykonane jest jako kod programu napisany w języku Java i może być uruchomione na każdym komputerze, gdzie zainstalowano środowisko uruchomieniowe języka Java.

Szczegóły dotyczące narzędzi oraz opis jego stosowania można znaleźć na stronie:

<http://polon.nauka.gov.pl/dokumentacja-masowego-importu-danych>

Przebieg importu

Projekt współfinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚĆ



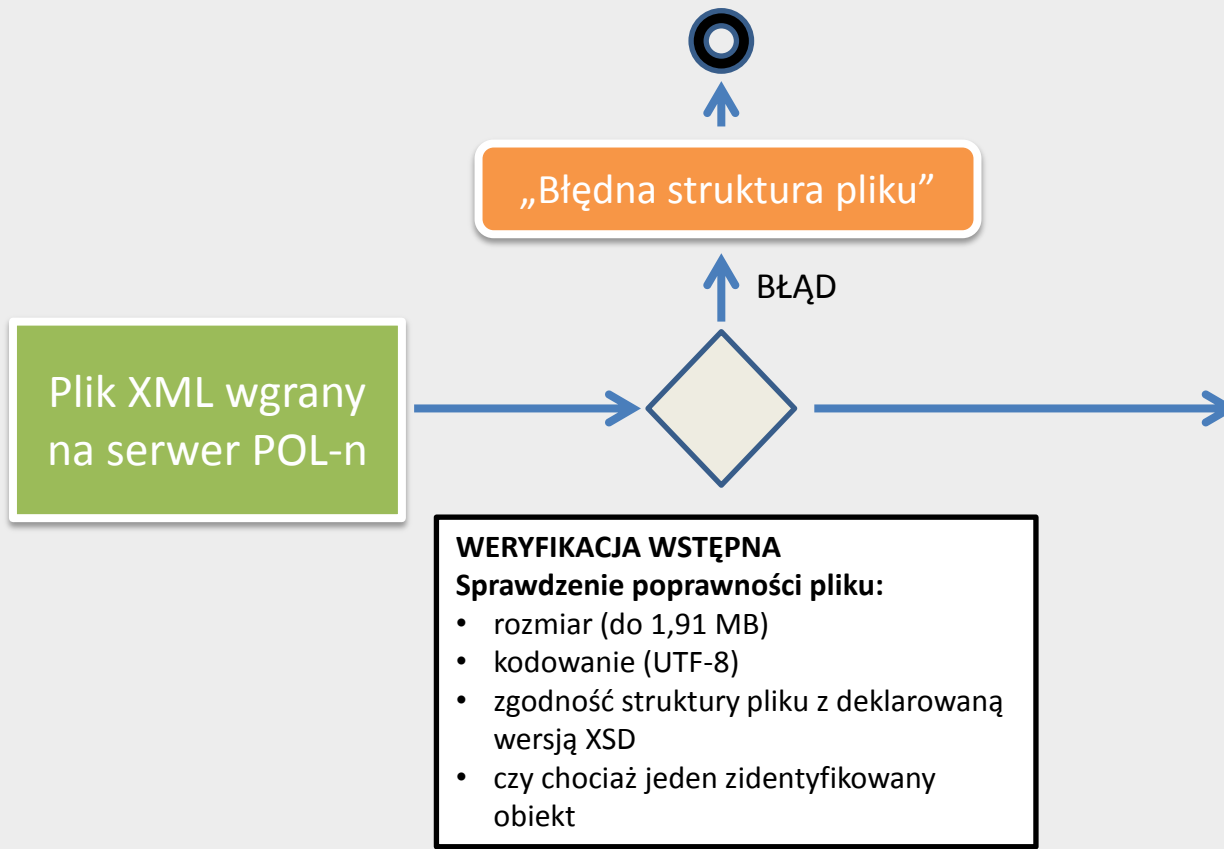
OPI

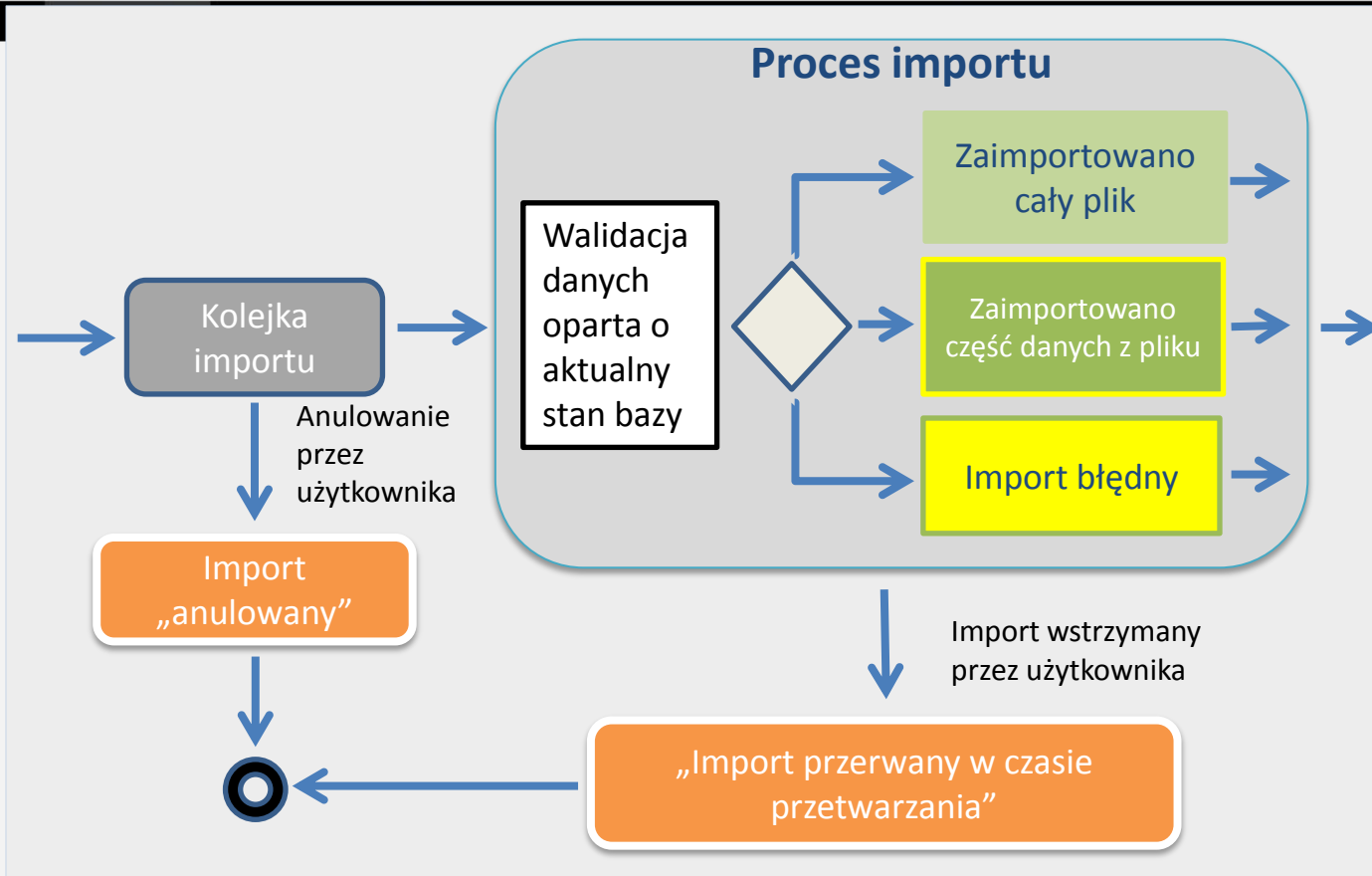


INDEX COPERNICUS
INTERNATIONAL

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY







Projekt współfinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



Projekt współfinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚĆ



OPI



INDEX COPERNICUS
INTERNATIONAL

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Reguły walidacji danych

Szczegółowy opis aktualnie stosowanych podczas importu reguł walidacji można obecnie znaleźć w pliku: [Reguly 6.0.1.pdf](#).

Projekt współfinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚĆ



OPI



INDEX COPERNICUS
INTERNATIONAL

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Reguły zostały podzielone logicznie na:

1. zakresy danych, których dotyczą:

- Globalne
- Administracja
- Dane osobowe
- Doktoranci
(Pomoc materialna, Studia)
- Pracownicy
(Funkcje kierownicze, Minimum kadrowe, Roczny wymiar zajęć dydaktycznych, Warunki zatrudnienia, Zatrudnienie)
- Studenci
(Ogólne, Pomoc materialna, Semestry, Studia)

2. stopień komplikacji danego warunku:

- Pola
- Proste
- Złożone

i informują szczegółowo o występujących w pliku danych nieprawidłowościach danych.

Identyfikacja osób w systemie

1. Obywatel Polski:

- < cudzoziemiec > N < / cudzoziemiec >
- pierwsze imię
- PESEL

2. Cudzoziemiec:

- < cudzoziemiec > T < / cudzoziemiec >
- pierwsze imię
- PESEL (jeżeli cudzoziemiec posiada numer PESEL)
- rodzaj, numer i kraj wydania dokumentu tożsamości (jeżeli cudzoziemiec nie posiada numeru PESEL)

Identyfikacja innych obiektów w systemie

1. Studia: <kierunekId> + <dataRozpoczecia>
2. Semestr: [studia] + <rokAkademicki> + <semestrNr>
3. Studia doktoranckie: <kierunekId> + <dataRozpoczecia>
4. Pomoc materialna: jest za każdym razem nadpisywana wartościami przekazanymi w pliku, w przypadku jej wykrycia w systemie w tym samym interwale miesięcznym.
5. Zatrudnienie: <instytucjaKod> + <dataOd>
6. Warunki zatrudnienia: [Zatrudnienie] + <dataOd>
7. Minimum kadrowe: [Zatrudnienie] + <kierunekInstancja01Id> + <rokAkademicki>

Usługi wywoływane z poziomu pliku XML

Opcja ta stanowi otwarty mechanizm pozwalający na wykonywanie w systemie zdefiniowanych operacji (usług).

Lista usług będzie rozwijana i dostosowywana do bieżących potrzeb użytkowników bez konieczności zmiany wersji obowiązującego XSD.

Mechanizm ten w zamyśle stanowi uzupełnienie dla standardowych operacji odpowiadających procesom importu danych i ma umożliwić wykonanie operacji niedostępnych podczas „zwykłego” importu.

Plik XML zawierający listę operacji musi zostać przesłany do systemu jako oddzielny plik opisany przez oddzielny schemat XSD. Mechanizm ten został tak pomyślany, aby definiowanie nowych operacji nie wiązało się z koniecznością zmiany wersji XSD.

Opis aktualnej wersji XSD można znaleźć w pliku [xsd_uslugi_4.0.0.zip](#)

Natomiast opis dostępnych operacje, sposób ich użycia i format w pliku [Usługi wywoływane z poziomu pliku XML 4.0.2.pdf](#)

Obecna wersja stanowi początek procesu rozwoju tego mechanizmu, w związku z czym zdefiniowane operacje trzeba traktować jako przykłady wykorzystania tej technologii w systemie.

Eksport danych z systemu POL-on

W nowej wersji systemu dodano opcję umożliwiającą eksport danych znajdujących się w systemie w takim zakresie i formacie jaki obejmuje import masowy danych do systemu.

Eksport wykonywany jest na żądanie użytkownika posiadającego w systemie rolę INST_IMPORT.

Eksport wykonywany jest asynchronicznie – istnieje kolejka eksportu i zlecone pliki są generowane w miarę dostępnych mocy obliczeniowych systemu.

Po zakończeniu eksportu wygenerowany plik udostępniany jest użytkownikowi do pobrania.

Możliwe jest wyeksportowanie tylko wszystkich danych z określonego zakresu:

- aktywnych studentów
- wszystkich studentów
- aktywnych pracowników
- wszystkich pracowników

Dziękuję za uwagę.

Koniec części „teoretycznej”

dr Piotr Rodzik

POLON_szkolenia@opi.org.pl

Ośrodek Przetwarzania Informacji - Państwowy Instytut Badawczy
Al. Niepodległości 188B, 00-608 Warszawa
Numer KRS: 0000127372 • Sąd Rejonowy dla m. st. Warszawy w Warszawie XII Wydział Gospodarczy KRS
REGON: 006746090 • NIP: 525-000-91-40 • <http://www.opi.org.pl/> • e-mail: opi@opi.org.pl

Projekt współfinansowany przez Unię Europejską w ramach środków Europejskiego Funduszu Społecznego



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚĆ



OPI



INDEX COPERNICUS
INTERNATIONAL

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

