

Projekt DIR jako przykład praktycznej realizacji idei Open Access

Marek Niezgódka, Alek Tarkowski

ICM UW

marekn@icm.edu.pl

altar@icm.edu.pl

Konferencja EBIB, Toruń, 7 grudnia 2007



Rozwój napędzany przez przełomy technologiczne.

- nowa struktura kosztów dystrybucji
- konwergencja mediów elektronicznych
- nowe podejścia do własności intelektualnej
- działania oddolne w Internecie
- rosnąca świadomość zmian wśród naukowców

Obecne przemiany

- Eksplozja treści dostępnych w postaci cyfrowej
- Mechanizmy ochronne wydawców komercyjnych
- Instytucje działające w oparciu o nową wizję
 - Zapewnienie wolnych treści i narzędzi
 - Wykorzystanie sieciowej infrastruktury
 - Umożliwienie interakcji użytkowników wokół treści.
- Zagrożenia:
 - chaos, potencjalna utrata jakości
- Zyski:
 - oddolna mobilizacja
 - mechanizmy wzajemnej weryfikacji

Obecne przemiany: Otwarty dostęp (Open Access)

- Inicjatywa Budapesztańska (2001):
 - *„Tradycja naukowa wraz z nowymi technologiami umożliwiają stworzenie dobra wspólnego w postaci elektronicznej publikacji czasopism typu peer-reviewed oraz zapewnienie całkowicie darmowego i nieograniczonego dostępu do tego dobra”*

Obecne przemiany: Otwarty dostęp (Open Access)

- *“Literatura typu Open-access (OA) jest cyfrowa, dostępna w sieci i za darmo”.*

Peter Suber

Obecne przemiany: Otwarty dostęp (Open Access)

- OA w szerszej perspektywie
 - OA jako “rewolucja oparta na wiedzy”
 - Internet Archive: “Universal Access to All Knowledge”
 - Access to Knowledge (A2K) – globalny wymiar OA

Projekt DIR – Domena Internetowych Repozytoriów Wiedzy

Założenia projektu

- repozytorium w skali narodowej
- powszechny dostęp do materiałów naukowych
- dystrybucja internetowa jako rozwiązanie problemów z obiegiem wiedzy w Polsce

Projekt DIR – Domena Internetowych Repozytoriów Wiedzy

Założenia pomocnicze projektu

- Ochrona przed zacytaniem materiałów archiwalnych i trudno dostępnych
- Stworzenie standardów przygotowania i udostępniania tekstów w postaci elektronicznej
- Stworzenie prestiżowego *elektronicznego* wydawnictwa naukowego



Projekt DIR – Domena Internetowych Repozytoriów Wiedzy

Jednolita platforma dla:

- materiałów archiwalnych i aktualnych
- materiałów zastrzeżonych i dostępnych w modelu OA
- zasobów bibliograficznych i pełnotekstowych
- zasobów różnorodnej formy i w różnorodnych formatach
- zasobów polskich i zagranicznych

Projekt DRIVER

- Ogólno-europejska infrastruktura dla repozytoriów cyfrowych
- dowolne dokumenty, dane, obiekty
- wszystkie kraje europejskie
- wszystkie dyscypliny naukowe
- podstawowe zadanie: zapewnienie bezpośredniego dostępu do danych tekstowych

Projekt DRIVER

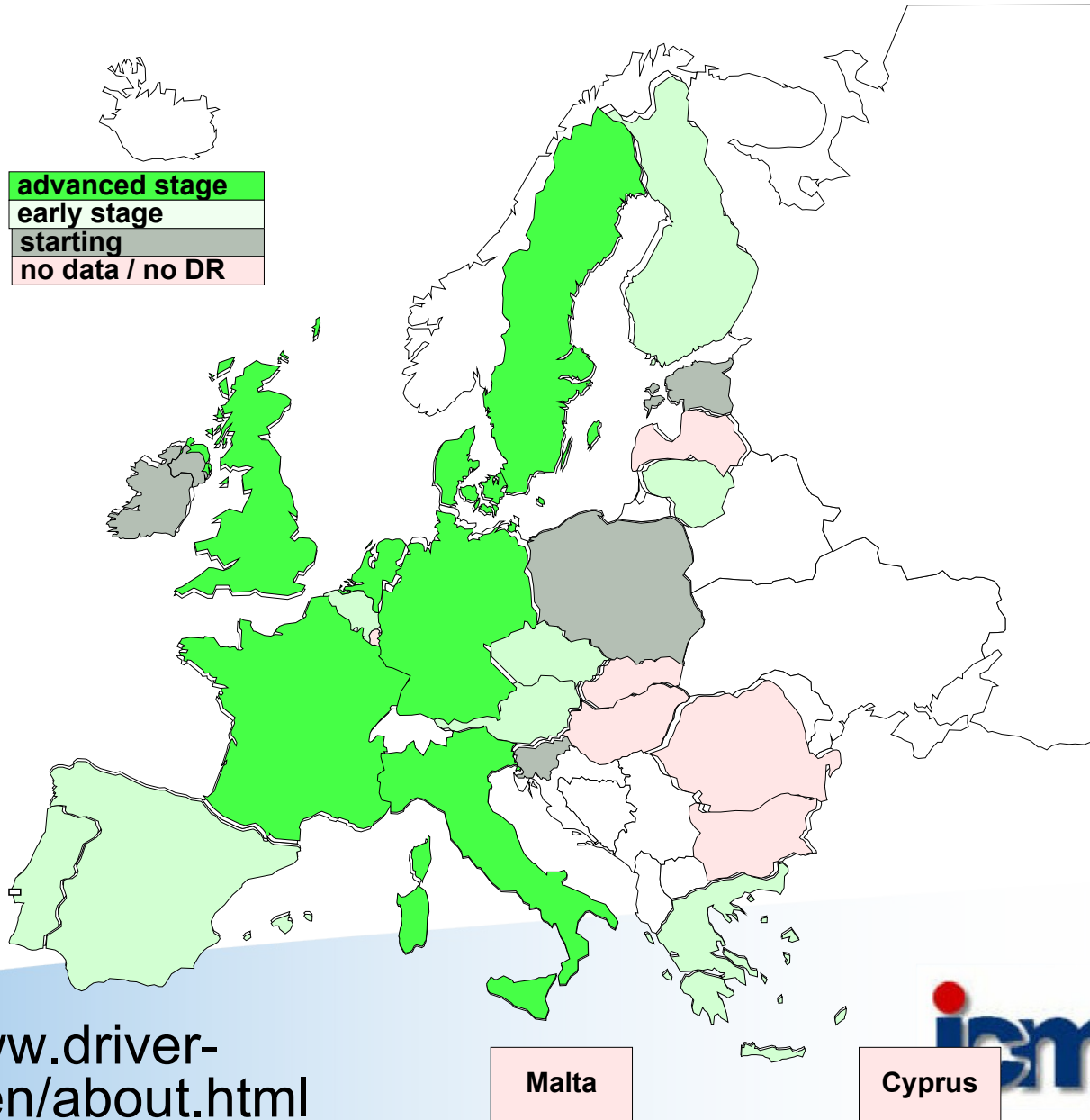
- wizja europejskiej infrastruktury wiedzy
- “Chcemy stworzyć i utrzymać europejskie środowisko, cechujące się długoterminową zdolnością trwania (*sustainable*), w którym treści i usługi mogą być otwarcie dzielone między sobą i zintegrowane, tak by mogła je wykorzystywać dowolna Aplikacja”

Projekt DRIVER

- inwentaryzacja repozytoriów europejskich
 - ok. 230 repozytoriów
 - w 7 krajach brak instytucji naukowych z repozytoriami
 - 5 krajów w fazie początkowej
 - w 15 krajach znacząca część instytucji posiada repozytoria (w 7 z nich, ponad połowa)

Projekt DRIVER

- inwentaryzacja repozytoriów europejskich
 - 1/3 wyników badawczych dostępna w repozytoriach
 - 90% danych to dane tekstowe
 - 68% danych to wyłącznie metadane



<http://www.driver-support.eu/en/about.html>



Zasoby DIR

- zbiory Biblioteki Wirtualnej Nauki (BWN)
- elektroniczne wersje czasopism polskich
- zbiory projektu dygitalizacyjnego
- projekt “Otwórz książkę”

Zasoby DIR: BWN

- Biblioteka Wirtualna Nauki (BWN)
- licencjonowane światowe bazy indeksowe oraz bibliograficzne
- licencjonowane kolekcje pełnotekstowe
- Kolekcje pełnotekstowe dostępne na zasadzie OA
 - Kolekcja matematyczno-fizyczna (14 tytułów)
 - Kolekcja nauk przyrodniczych (8 tytułów)

Zasoby DIR: projekt dygitalizacyjny

- kolekcja pełnotekstowa
- teksty w wersji PDF oraz TXT
- wyszukiwanie pełnotekstowe
- dobór pozycji w oparciu o rekomendacje ekspertów

Zasoby DIR – “Otwórz książkę”

- problem “utworów osieroconych”
- Lawrence Lessig: 1.7% książek z 1930 roku dostępnych w druku w 2001 roku
- Mark Lemley: 4% utworów z 1. połowy XX. wieku jest dostępnych w obiegu komercyjnym
- Landes i Posner: średni czas życia komercyjnego utworu: 25 lat dla muzyki, 11 dla książek, 7 dla grafiki

Zasoby DIR – “Otwórz książkę”

- Copyright Office
- British Library
- Ekspertki Komisji Europejskiej: niezbędne jest rozwiązanie problemu utworów osieroconych poprzez reformę systemu prawa autorskiego.

Zasoby DIR – “Otwórz książkę”

- Rozwiązanie problemu “utworów osieroconych”
- Udostępnianie książek naukowych w internecie na licencji CC.
- Możliwe dzięki specyfice polskiego rynku publikacji naukowych.

DIR – kwestie prawne

- Kwestie prawne mogą stanowić jedną z podstawowych barier dla OA w Polsce.
- Projekt Creative Commons Polska – promocja wolnych licencji w środowisku naukowym.
- Wolne licencjonowanie jako warunek pełnej realizacji modelu OA.
 - Mechanizm formalizujący swobodny dostęp.

DIR – kwestie techniczne

- YADDA – Yet Another Digital Document Archive:
- Baza zasilana z różnorodnych źródeł, również w sposób automatyczny
- Jednolita, zestandardyzowana platforma
- System archiwizacji i zarządzania treściami naukowymi

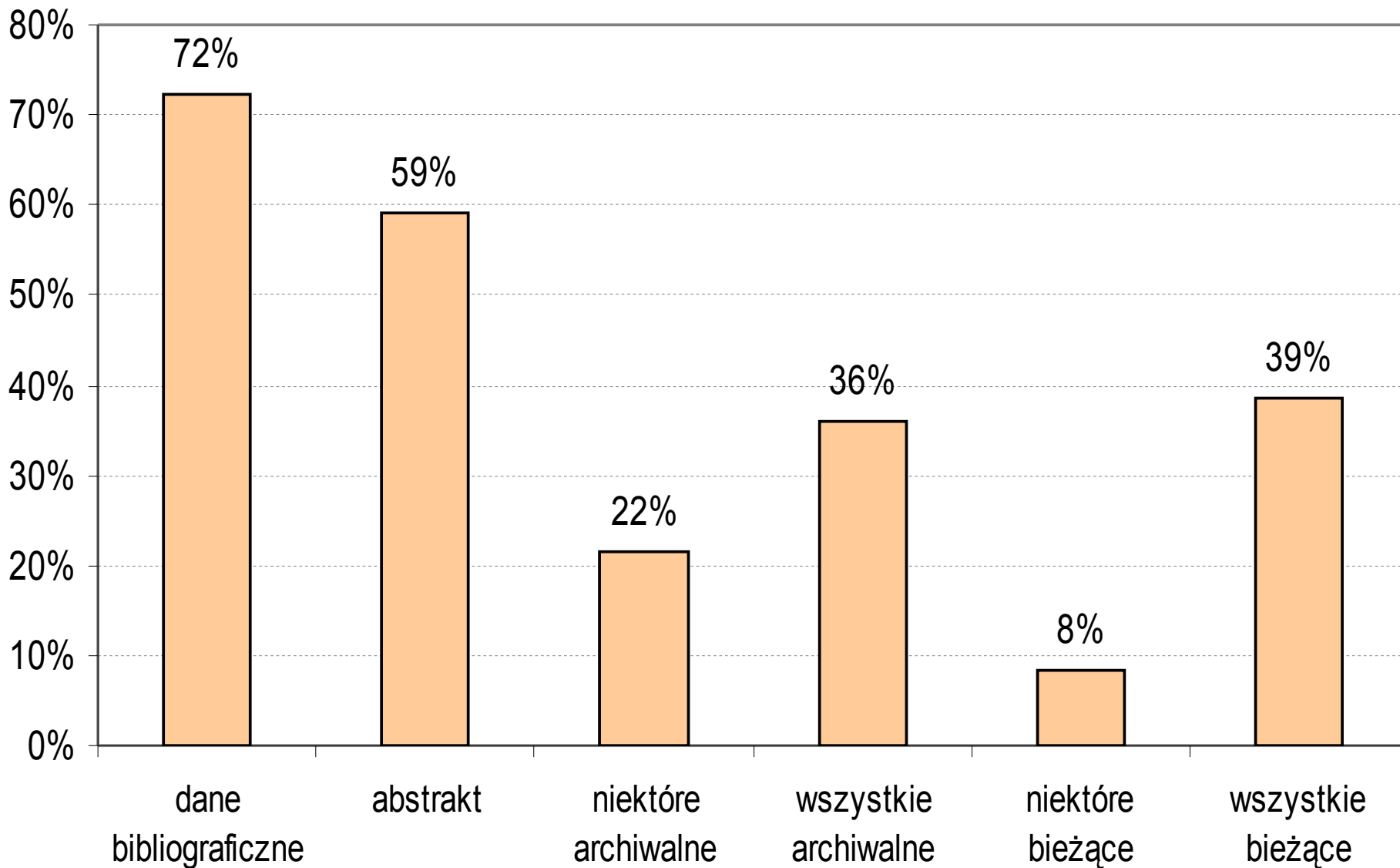
DIR jako projekt otwarty

- Nadrzędna zasada: swobodny, publiczny dostęp do materiałów finansowanych z pieniędzy publicznych
- Otwartość serwisu na różnych poziomach
 - teksty OA
 - otwarte dane
 - otwarte oprogramowanie

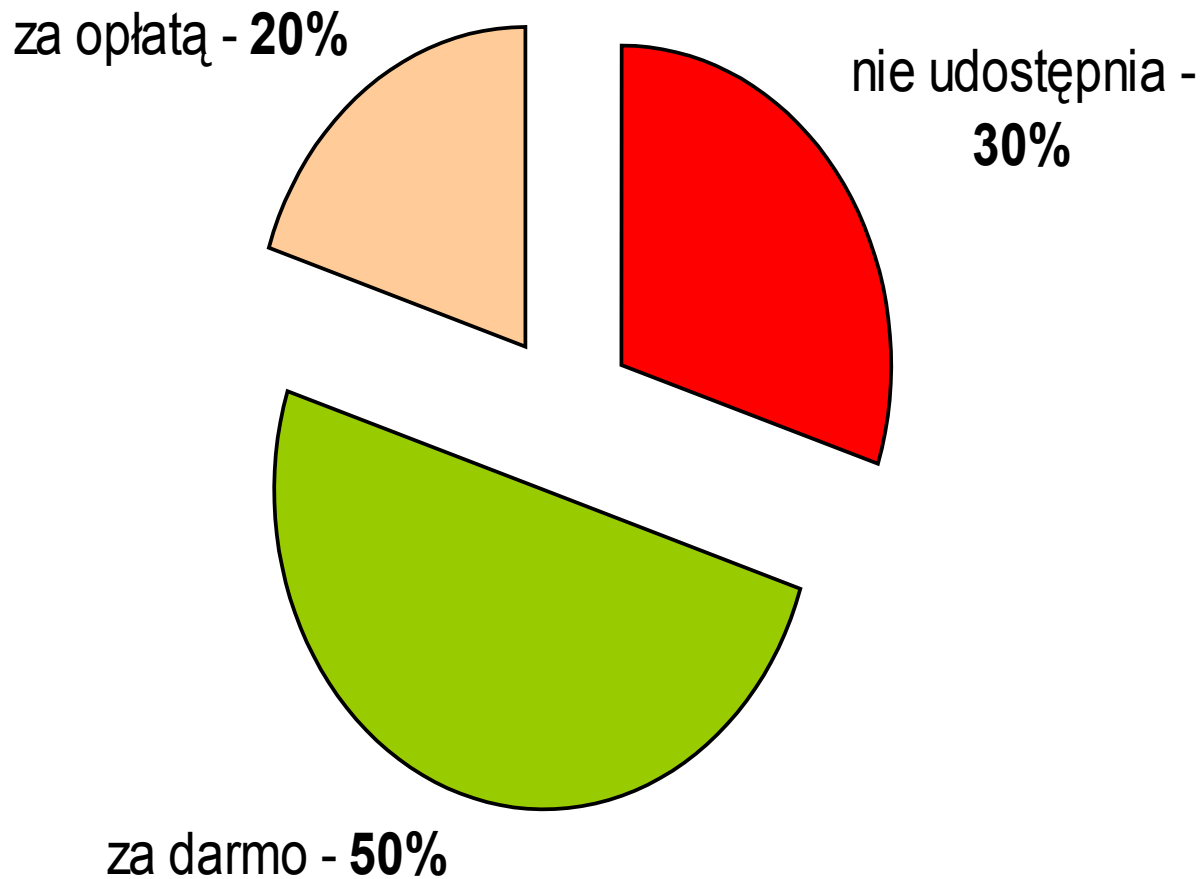
DIR – współpraca ze środowiskiem naukowym

- Badanie wydawców naukowych, wiosna 2007.
- Internet jest ciągle niewystarczająco wykorzystany w polskim środowisku naukowym.
- Niemal całkowity brak znajomości tematyki OA.

Materiały dostępne na stronie WWW czasopisma



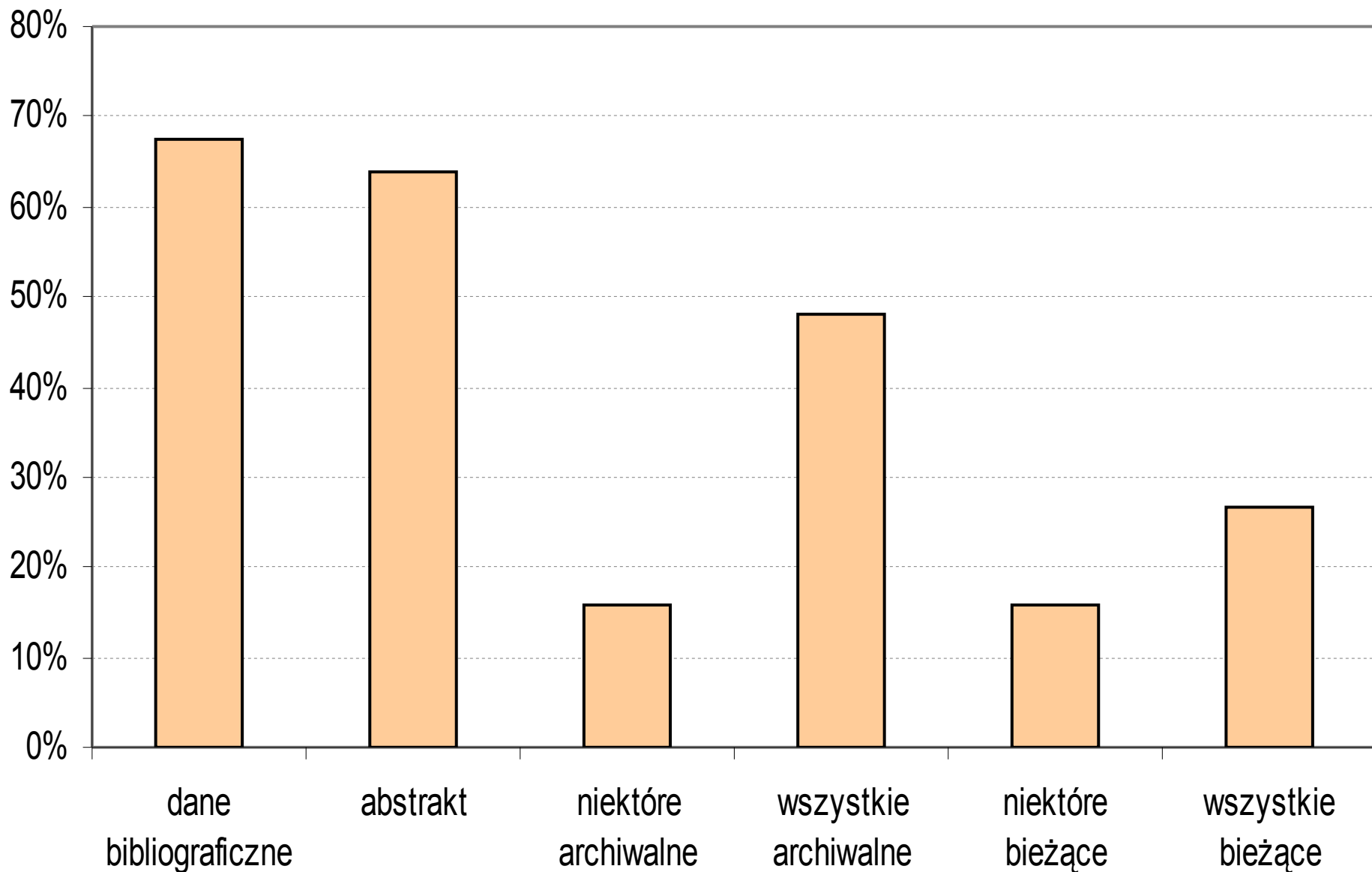
Udostępnianie artykułów w internecie



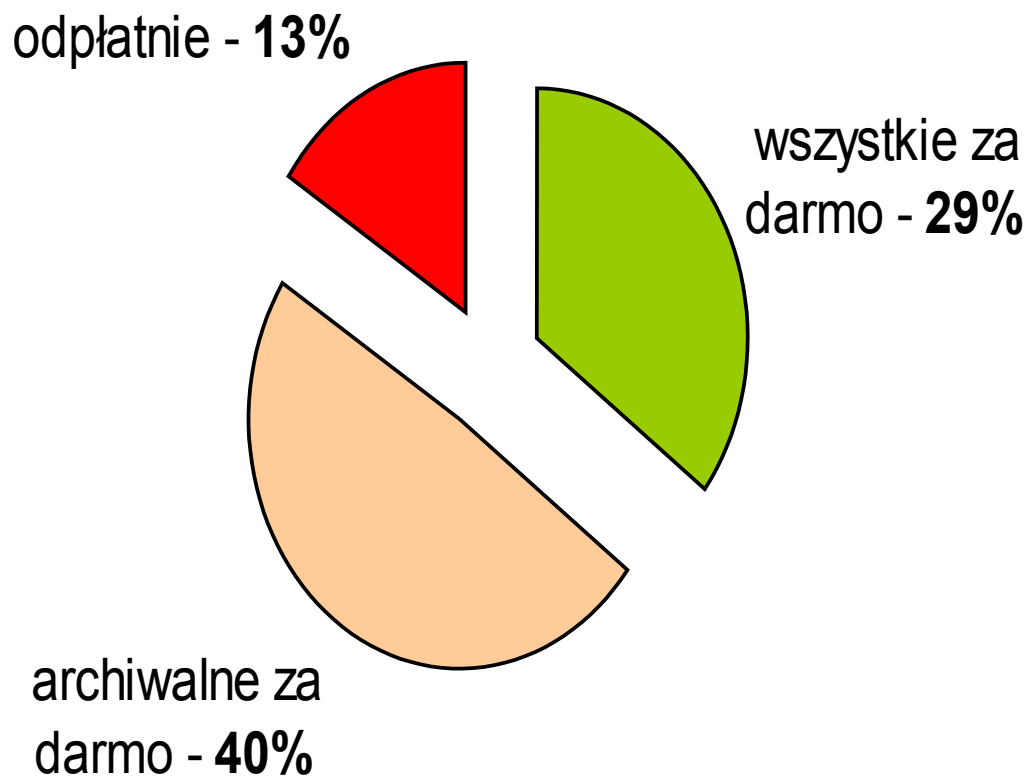
DIR – współpraca ze środowiskiem naukowym

- Repozytorium jako szansa dla dystrybucji wiedzy naukowej w Polsce.
- Trudności finansowe barierą dla dostępności.
- Zapewnienie infrastruktury jako istotny krok naprzód.
- 65% wydawnictw gotowe udostępnić materiały w repozytorium.

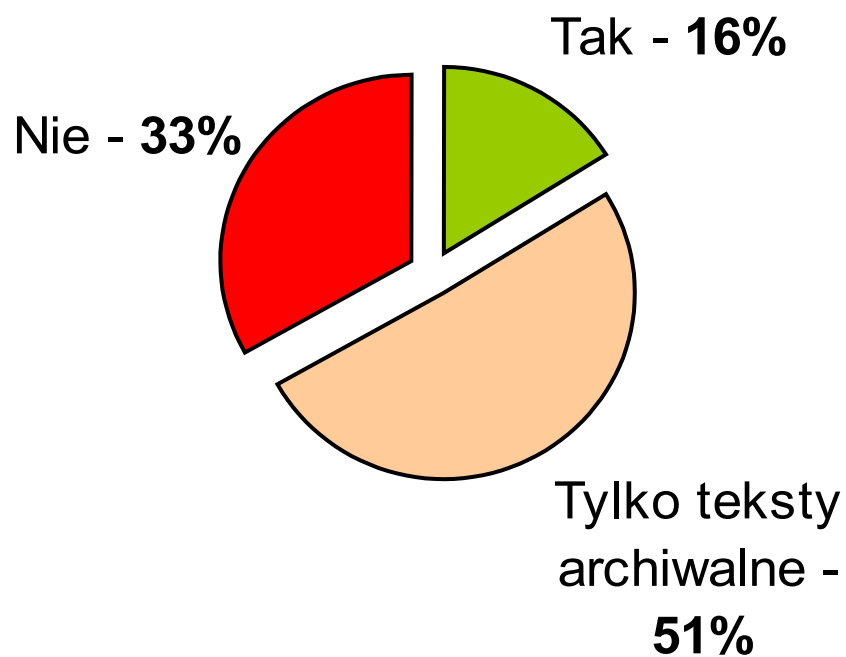
Materiały do przekazania do repozytorium



Warunki udostępniania materiałów w repozytorium



Udostępnianie tekstów na zasadzie Open Access



Open Data

- Idea otwartego dostępu do danych i wyników badań nie jest nowa – wynika z tradycji wymiany naukowej
- Nowe bariery dla dostępności: licencje, prawa autorskie, patenty, opłaty, techniczne ograniczenia
- European Database Directive – ochrona *sui generis*

Open Data

- 2004: deklaracja Ministrów nauki państw OECD: dane z badań finansowanych publicznie powinny być swobodnie dostępne
- 2007: rekomendacja “*OECD Principles and Guidelines for Access to Research Data from Public Funding*”

Open Data

- Science Commons: zintegrowany zestaw rozwiązań regulacyjnych, kontraktowych i technicznych, oparty na otwartych formatach
- projekt Neurocommons
 - data mining wykonywany na publicznie dostępnych danych w celu stworzenia sieci semantycznej
 - pilotaż: abstrakty, teksty OA oraz otwarte dane z zakresu neurologii

Przyszłość: od bibliotek cyfrowych do repozytoriów wiedzy

- Kluczowe kwestie:
- więcej niż proste przenoszenie treści do nowych formatów cyfrowych
- nowe struktury i paradygmaty wydawnicze
- zamiast wydania jako działania zamykającego proces wydawniczy – wydanie jako początek dynamicznego, wielowątkowego, kolektywnego procesu pracy z wiedzą