

# Matematyka czasu Horyzontu i poza nim

Marek Niezgódka

[marekn@icm.edu.pl](mailto:marekn@icm.edu.pl)

ICM, Uniwersytet Warszawski

*KWUMI, Olsztyn 29.05.2015*

# Otwarta Nauka

- **Matematyka jako dyscyplina Otwartej Nauki**
- **Otwarty Dostęp:**
  - tylko jeden z atrybutów Otwartej Nauki,
  - jej narzędzie a nie cel
- **Podstawy rozwoju Otwartej Nauki:**
  - głęboka transformacja systemu ocen naukowych i ewaluacji
  - tworzenie warunków i ram technicznych, społecznych i prawnych dla wtórnego użytku

# MNiSW: Strategia Otwartego Dostępu

- Rok 2012, 17.07: *Zalecenie Komisji Europejskiej w sprawie dostępu do informacji naukowej oraz jej ochrony*
- Termin wprowadzenia narodowych strategii Otwartego Dostępu: **grudzień 2012**
- Polska – **obecnie przygotowwany projekt** takiej strategii:
  - Pragmatyczny model Otwartego Dostępu, z miękkimi zaleceniami
  - Przewidywane zatwierdzenie: czerwiec 2015

## ZALECENIE KOMISJI

z dnia 17 lipca 2012 r.

w sprawie dostępu do informacji naukowej oraz jej ochrony

(2012/417/UE)



NINIEJSZYM ZALECA, ABY PAŃSTWA CZŁONKOWSKIE:

### Otwarty dostęp do publikacji naukowych

1. określiły jasne strategie rozpowszechniania i otwartego dostępu do publikacji naukowych, które powstają w wyniku badań finansowanych ze środków publicznych. Strategie te powinny uwzględniać:
  - konkretne cele i wskaźniki osiągniętych postępów,
  - plany wdrażania, w tym podział kompetencji,
  - odpowiednie plany finansowania.

# OpenAIRE: infrastruktura Otwartego Dostępu w Europie

- Infrastruktura OpenAIRE:  
wiodąca rola w realizacji otwartego dostępu do treści naukowych:
  - Publikacje
  - Dane naukowe
- Agregacja infrastruktur i zasobów krajowych UE
- Sieć integrująca działania krajowe
- **Polska: partner technologiczny i kontaktowy: ICM (także operator systemu)**

Kraje uczestniczące:



# Sprzeciw wobec nauki zamkniętej

- **Tim Govers** (matematyk, laureat medalu Fieldsa): skuteczny apel o wielopoziomowy bojkot głównych wydawnictw: publikowanie, ocenianie, edycja, ewaluacje
- **Harvard University**: rekomendacja jego uczonym publikowania w czasopismach Open Access, jako znak protestu przeciw nadmiernemu wzrostowi cen czołowych czasopism naukowych
- **Negatywna ocena zasady “*publish or perish*”** obowiązującej w systemie nauki, napędzanej przez instytucje zarządzające nauką
- **Prawdziwy poziom nieodkrytych nadużyć** w nauce jest nieznany; brak wsparcia remedium polegającego na docenieniu studiów replikacyjnych
- Do r. 2012: **52 noblistów** poparło apel do Kongresu USA o powszechne wprowadzenie zasad Otwartego Dostępu do wyników badań naukowych

# Otwarty dostęp do zasobów edukacyjnych – kolejny krok w stronę otwartości

- W serwisie Komisji Europejskiej, **Open Education Europa**:
  - prawie 1400 Massive Open Online Courses (MOOC) z krajów UE
  - oferta z prawie 500 instytucji
- Francja: Francuski Uniwersytet Cyfrowy, utworzona ze środków publicznych platforma oferująca kursy online
- Dynamiczny rozwój MOOC na amerykańskich uniwersytetach:
  - Duża, dynamicznie rosnąca liczba studentów z całego świata
- **Wizja: swoisty uniwersytet globalny**
- **Szansa i wyzwanie dla matematyki**

# Wirtualna Biblioteka Nauki: aktualności

## Użytkowanie WBN, 2014:

liczba pobranych książek (rozdziały lub całe książki), „testowy” dostęp do kolekcji Springer’a:  
**3.5 mln pobrań (I kwartał 2015: 1.5 mln)**

- liczba pobranych artykułów:
  - Elsevier 6.2 mln
  - Springer 1.2 mln
  - Wiley 1.2 mln
- Bazy bibliograficzne: relatywnie niższy poziom użytkowania



# WBN: infrastruktura Otwartego Dostępu

- Platforma [Otwartanauka.pl](http://Otwartanauka.pl)
- Repozytorium:
  - Polskie czasopisma naukowe:
    - ok. 1000 tytułów –metadane,
    - ok. 500 tytułów – pełne treści
  - Polskie książki naukowe:
    - Ok. 500 tytułów
  - Polska składowa EU–DML (Europejskiej Cyfrowej Biblioteki Matematyki)
- Polskie bazy bibliograficzne

# Otwarte dane naukowe: dobre praktyki

- Publikowane dane i kody oprogramowania dostępne na licencjach otwartych
- Udostępniana dokumentacja procesu badawczego – dostępne dane:
  - Kontrolowana wersja
  - Dostępne są dane surowe
  - Kody i metody są dostępne (kontrolowane wersje)
  - Wszystkie zewnętrzne (3rd party) kody i dane są cytowane
  - Podane są reguły zewnętrzne obowiązujące w badaniu (np. reguły fundatora)

# Matematyka a HORIZON 2020

- **Połączenie:** **Matematyka & Technologie Informacyjne i Komunikacyjne**, w tym High Performance Computing
- **Matematyka dla potrzeb modelowania** złożonych układów i procesów w nich zachodzących
- **Matematyka** dla pełnego zakresu problematyki związanej z przetwarzaniem **danych** (szczególnie tych bardzo dużych – Big Data)

# HORIZON 2020 – WORK PROGRAMME 2016-2017

*European research infrastructures (including e-Infrastructures), z projektu dokumentu*

- *Research infrastructures are facilities, resources and services that are used by the research communities to conduct research and foster innovation in their fields.*
- *They include: major scientific equipment (or sets of instruments); knowledge-based resources such as collections, archives or scientific data; e-infrastructures, such as data and computing systems and communication networks; and ...*
- *The Research Infrastructures Work Programme 2016–2017 will put wide emphasis on maximising the exploitation of data produced and/or collected by research infrastructures.*

# HORIZON 2020: *Computing e-infrastructure with extreme large datasets*

*Scientific communities are increasingly aware of the rising tide of data which **requires new approaches to data discoverability, management, quality assurance and trust, preservation, access and sharing.***

*e-Infrastructures incorporating advanced computing resources and software **based on up-to-date mathematical methods** are essential in order to increase the capacity to manage and store extremely large, heterogeneous and complex datasets and to enable intelligent and timely analyses, to extract accurate knowledge and information from huge datasets.*

# *Computing e-infrastructure with extreme large datasets (WP2016-17)*

*Proposals shall address all points below:*

- Develop service prototypes to cope with very large data resources. It should include the basis software layers supporting applications such as modelling, simulation, pattern recognition, visualisation, etc.
- The developments should be supported by robust mathematical methods and tools.
- Prototypes should follow an open source approach and aim at common interfaces to access and analyse underlying data collected/stored in different platforms, formats, locations and e-infrastructures and be tested against requirements of very large or highly heterogeneous research data sets.
- Clean slate approaches to high-performance computing and data management (e.g. HPC-through-the cloud, support of most innovative server's architectures for distributed computing in particular high Memory/Cores ratios allowing "in memory" processing) targeting 2020+ 'data factory' requirements of research communities and large scale facilities (e.g. ESFRI projects) are encouraged.

# OSIRIS: Polska Mapa Drogowa Infrastruktur Badawczych

- **Otwarta Infrastruktura** Zasobów Informacji Naukowej i Danych
- Jedyna pozycja w Programie umożliwiająca projekty łączące badania stosowane i rozwój nowych e-infrastruktur związanych z danymi informacją naukową
- **Ogłoszony nabór projektów**
- Lider OSIRISa: ICM
- Partnerzy pierwotni: 16 instytucji (głównie uczelnie, instytuty badawcze)

# Konsultacje online w KE: tematy

1. Rola matematyki w kontekście Big Data
2. Rola matematyki w HPC, w tym w obliczeniach exaskalowych
3. Rola e-infrastruktur w matematyce
4. Wpływ matematyki stosowanej i przemysłowej na innowacje
5. Przygotowanie z perspektywy *Excellent Science* w Programie Horizon 2020 Programów Roboczych 2016–2017 w zakresie FET Proactive (HPC) i/lub e-Infrastruktur



Matematyka jako dyscyplina wspomagająca rozwój badań stosowanych i projektów rozwojowych; główny obszar zaangażowania: **Modelowanie, symulacja i optymalizacja (MSO):**

- 1. Niezbędny rozwój współpracy interdyscyplinarnej matematyków, inżynierów, reprezentantów nauk ścisłych, technicznych i nauk o życiu, oraz liderów przemysłu*
- 2. Rozwój podejść na bazie MSO do zastosowań o wymiarze europejskim: m.in. Rozwój i zarządzanie wielkimi sieciami operacyjnymi (prąd, gaz, dane)*

# Główne konkluzje, 2

3. Matematyka wspomagająca rozwój kompetencji w głównych obszarach techniki:
- Multifizyka i układy wieloskalowe
  - Hybrydowe układy nieliniowe (dyskretno–ciągłe)
  - Procesy stochastyczne
  - Schematy redukcji modeli dla układów wysokowymiarowych
  - Metody traktowania niejednorodnych niepewności w modelach
  - Sprzężenia modeli
  - Metody optymalizacji wielokryterialnej

# Komisja Europejska, 10.2014: Mathematics for Digital Science

Raport: *Opportunities and Challenges at the Interface of Big Data, High-Performance Computing and Mathematics*

Tezy:

- *W wyniku rozwoju technologii obliczeniowych i komunikacyjnych, w naukach ścisłych i matematyce pojawiają się całkowicie **nowe wyzwania**: jak traktować i użyć zalew danych w układach społecznych, technicznych i ekonomicznych*
- *Dwukierunkowa natura wpływu: nauka i technika <-> matematyka*

# Raport: Rozwój metod wydobywania informacji/struktury z Big Data (BD)

Rozwój analityki wydobywającej sens z danych i układów złożonych (m.in. z uwzględnieniem topologii):

- **Rozwój analityki** dla wykrywania struktur i korelacji w Big Data (wspólne projekty matematyki i BD)
- **Doskonalenie algorytmów** wykrywających struktury: chemia obl., nauki materiałowe, prognozowanie pogody, wieloskalowa fizyka z upscaling'iem (integracja matematyków z reprezentantami odp. nauk)
- **Wykorzystanie metod statystycznych i adaptacyjnych** w obszarze BD:
  - redukcja,
  - Interakcja statystyków matematycznych z ekspertami w zakresie BD

# Raport: Matematyka – stymulator innowacji

- Tworzenie infrastruktury informacyjnej służącej współdziałaniu matematyki i przemysłu/gospodarki:
  - Istnieją rozwinięte powiązania w dużych firmach, jest ich mało lub brak w małych i średnich firmach
  - Rola konsorcjum EU-MATH-IN
- Rola szeregu obszarów matematyki teoretycznej w kontekście danych ekstremalnych wielkich

# Raport: Rozwiązania wyzwań efektywnych obliczeń exaskalowych

- Rozwój efektywnego matematycznego oprogramowania dla obliczeń exaskalowych:
  - Wysokoskalowalne solwery i biblioteki numeryczne
  - Algorytmy efektywne energetycznie
  - Hierarchiczne struktury algorytmów wykorzystujące wszystkie poziomy równoleglenia komputerów
  - Wysoka powtarzalność obliczeń
  - Rozwiązania dla obliczeń intensywnie korzystających z danych (zarządzanie danymi i ich analityka)
  - Algorytmy i oprogramowanie optymalizowane dla nowych generacji architektur komputerowych
- Rozwój statystycznych rozwiązań dla ilościowej charakteryzacji niepewności w zastosowaniach HPC

# Rozwój e-infrastruktur

- Wsparcie interakcji matematyki, nauki i gospodarki (przemysłu)
- Szerokie udostępnienie wyników matematycznych:
  - Otwarty Dostęp,
  - EU-DML ([eudml.org](http://eudml.org)): największa kolekcja matematycznych publikacji w Otwartym Dostępie

# Matematyka: wybór obszarów wyzwań badawczych wg KE

- Przykładowe problemy związane z Big Data i HPC:
  - Analizy ryzyka w kontekście Big Data
  - Prywatność i anonimizacja w kontekście BD
  - Analiza problemów społecznych



# Rola matematyki: referencje międzynarodowe

- STEM (Science, Technology, Engineering & Mathematics (STEM) for our Future), Australia
- Nauki ścisłe i matematyka wnoszą rocznie bezpośrednio \$145 mld do gospodarki Australii (raport, 2014)
- Odkrycia i innowacje w obszarach fizyki, chemii, nauk o Ziemi i matematyki wniosły bezpośrednio \$145 mld do gospodarki, a po uwzględnieniu oddziaływań pośrednich prawie \$300 mld

# Referencje:

- *The importance of advanced physical and mathematical sciences to the Australian economy*, prepared by the Centre for International Economics for the Office of the Chief Scientist and the Australian Academy of Science, March 2015.
- *Mathematics and Industry*. ESF Forward Look Report 2010

- Program finansuje matematyczne prace badawcze o podstawowym znaczeniu dla tworzenia i doskonalenia algorytmów:
  - Metody numeryczne rozwiązywania równań różniczkowych zwyczajnych i cząstkowych
  - Zastosowania tych metod w obliczeniowej dynamice przepływów
  - Użycie równań różniczkowych cząstkowych
  - Siatki obliczeniowe dla złożonych konfiguracji geometrycznych; wzajemna interakcja modeli dyskretnych i ciągłych
  - Metody numeryczne rozwiązywania wielkich układów równań liniowych i nieliniowych
  - Metody optymalizacji na poziomie problemów o wysokiej złożoności

# DOE, priorytety badań matematycznych dot. układów złożonych

- Metody analizy modeli i algorytmy dla układów wieloskalowych, sprzężonych i hybrydowych: dyskretno–ciągłych, stochastyczno–deterministycznych
- Metody redukcji modeli, metodologie walidacji
- Metody analizy i algorytmy optymalizacji stochastycznej i odpornej dla układów nieliniowych i zawierających wewnętrzne struktury
- Metody traktowania propagacji niepewności poprzez skale w modelach wieloskalowych i poprzez składowe modeli sprzężonych
- Optymalizacja algorytmów asymilacji danych i fuzji modeli/danych dla potrzeb identyfikacji, sterowania i ograniczania niepewności modelu

# Badania DOE, kont.

- Metody i algorytmy obliczeń wieloskalowych
- Obliczenia symulacyjne na poziomie multifizyki, symulujące w szerokim zakresie różnorodne procesy badawcze, np. powierzchniowe reakcje chemiczne
- Matematyczny software i biblioteki- kody modularne które można wpisywać w adaptowalne programy wywodzące się z różnych obszarów nauki

Zasady: Otwarty Dostęp, otwarte źródła software'owe

# DARPA: wyzwania matematyczne XXI wieku (wybór)

- *I: Matematyka mózgu:* teoria matematyczna, która umożliwi stworzenie modelu funkcjonalnymi, spójnego matematycznie
- *II: Dynamika sieci:* wysokowymiarowa teoria matematyczna jako podstawa modelowania i predykcji procesów w wielkoskalowych sieciach przestrzennych związanych z szeroką grupą zastosowań
- *III: Stochastyczność w przyrodzie:* Rozwój metod odnoszących się do manifestu Mumforda dot. nowej matematyki XXI w.
- *IV: Płyny XXI w.:* Nowe metody opisu złożonych płynów, w tym pian, żeli, ciekłych kryształów.
- *V: Biologiczna kwantowa teoria pola:* modelowanie bakterii, kontrolowana ewolucja patogenów
- *VI: Dualność obliczeniowa:* rozwój nowych klas algorytmów wykorzystujących dualność i cechy geometryczne
- *VII: Poza optymalizację wypukłą: na bazie systematycznej zamiany algebry liniowej przez geometrię algebraiczną*

*IX: Konsekwencje dowodu Perelmana twierdzenia geometrycznego Thurstona; W szczególności w kontekście projektowania nowych materiałów*

*X: Optimalne Nanostruktury: metody upscalingu*

*XI: Matematyka i algorytmika obliczeń kwantowych*

*XI: Wieloskalowa Matematyka: wymiana rrcz, wieloskalowa teoria gier*

*Xlii: Teoria informacji w opisie ewolucji wirusów*

*V: Biologiczna kwantowa teoria pola: modelowanie bakterii, kontrolowana ewolucja patogenów*

*VI: Dualność obliczeniowa:*

rozwój nowych klas algorytmów wykorzystujących dualność i cechy geometryczne

*VII: Poza optymalizację wypukłą: na bazie systematycznej zamiany algebry liniowej przez geometrię algebraiczną*

*Viii: Biologiczna kwantowa teoria pola: modelowanie bakterii, kontrolowana ewolucja patogenów*

*IX: Dualność obliczeniowa:*

rozwój nowych klas algorytmów wykorzystujących dualność i cechy geometryczne

*X: Poza optymalizację wypukłą: na bazie systematycznej zamiany algebry liniowej przez geometrię algebraiczną*

# Konsorcjum EU-MATH-IN

- *European Service Network of Mathematics for Industry and Innovation*
- Inicjatywa wokół europejskiego konsorcjum matematyki przemysłowej
- Cel: wspólne działanie w ramach Programu HORIZON 2020
- *Tworzenie innowacyjnych rozwiązań technologicznych wymaga szerokiego metod modelowania matematycznego, symulacji i optymalizacji, wspomaganych przez techniki HPC..., te podejścia o wysokim poziomie złożoności wymagają kształcenia ekspertów o nowych kompetencjach*



*The process of developing new products/processes increasingly requires a holistic approach, that considers the complete life cycle, including the impact and risks on society, environment etc. To realize this visionary and ambitious goal, it is necessary to accompany development and design with appropriate virtual products/processes that allow long-term risk analysis, improvements or even optimization in an early innovation phase. The key technology in this is the use of mathematical Modelling, Simulation and Optimisation (MSO) methods, processed via high performance computing (HPC). These methods are highly complex and require skilled experts, which often are not available, in particular in SMEs.*

*To improve this situation, the European network EU-MATHS-IN has been founded. The major objective of this proposal, organized by EU-MATHS-IN and its national networks, is to construct a virtual research infrastructure that provides, in a user-driven, integrative way, tailored access to the necessary services and resources. It will support industry, in particular SMEs, with easy access to external expertise, skilled staff, adequate MSO software, models, and benchmarks, and their integration, as well as a one-stop-shop for European industries and an MSO cloud that links MSO tools with HPC and Big Data. It will also support (interdisciplinary) teams of scientists and engineers in academia working in high technology areas with the necessary tools and expertise needed for their research, as well as appropriate electronic bibliographic services in the areas of literature and software.*

*This proposal is focusing on integrating and scaling up existing national initiatives and expertise to the European scale, including the recent European initiatives of a digital mathematical library.*

# Dane zgłoszonego projektu

- Lider: EU-MATH-IN [www.eu-maths-in.eu](http://www.eu-maths-in.eu) (z autoryzacją EMS)
- Koordynator: Mario Primicerio
- Partnerzy: krajowe sieci z ustanowionymi reprezentantami
- Cele: *zwiększyć wpływ MSO na innowacje w przemyśle, nauce i społeczeństwie poprzez rozwój wiedzy i wymianę informacji między zaangażowanymi stronami (przemysł, SME, środowiska akademickie) w skali europejskiej.*
- Zakres: *konstrukcja e-infrastruktury zapewniającej zasoby i realizującej serwisy na rzecz środowiska R&D w przemyśle, nauce i społeczeństwie, a także m.in. dla matematyków, informatyków pracujących w przemyśle.*
- i ... dalej w podobnym eurostylu

- ...it will support (interdisciplinary) teams of scientists and engineers working in high technology areas and using mathematical MSO with the necessary tools and expertise needed for their research. The e-infrastructure will also provide appropriate electronic bibliographic services in the areas of literature and software. Particular emphasis is put on the integrated service aspect, e.g. by providing together the MSO tools, the experts that can help in modelling and use of software, the cooperation with low and high performance computing facilities, and the corresponding bibliographic information.
- ...the first step in the creation of a strong culture of pan-European cooperation. The proposed e-infrastructure will help to balance the very different maturity in the creation and transfer of MSO methods, and the cooperation of academia with industrial partners (in particular SMEs) in the different European countries. It will provide not only the MSO software and benchmarking access facilities needed in industries but, and this is essential, also a European database of experts and a one-stop-shop that can support industry in the use of MSO tools. This is in particular important for SMEs, which usually do not have large R&D departments and thus are partially excluded from innovation. At the same time the infrastructure will provide creators of MSO tools the access to users and will facilitate and accelerate the transfer.

# Wynik ewaluacji projektu w Programie Horizon 2020

- Ocena sumaryczna: 8.5 (threshold 10)
- W tym:
  - Excellence: 3.0 (0–5, threshold 3.0)
  - Impact: 2.5 (0–5, threshold 3.0)
  - Quality and efficiency of the implementation 3.0 (0–5, threshold 3.0)
  - Operational capacity: not provided
- **Project rejected**

# Jedna ze składowych projektu: rozwój EU-DML

[eudml.org](http://eudml.org) (European Digital Mathematics Library)

## Statystyki:

- **242,136** przeszukiwalnych dokumentów, w tym:
  - 223,807 pełnych tekstów indeksowanych
  - 13,278 czytelnych dla słabowidzących
- **236,835** artykułów z **288** czasopism
- **3,233** monografii
- **747** tomów z **295** dzieł wielotomowych

# Partnerzy inicjatywy EuDML

Inicjatywa ustanowiona: luty 2014,

12 partnerów:

- [European Mathematical Society](#)
- [FIZ Karlsruhe - Leibniz Institut für Informationsinfrastruktur GmbH](#)
- [Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw](#)
- [Université Joseph-Fourier, Grenoble 1](#)
- [University of Birmingham](#)
- [Institute of Mathematics and Informatics BAS, Sofia](#)
- [Institute of Mathematics AS CR, Praha](#)
- [Ionian University, Corfu](#)
- [Società Italiana per la Matematica Applicata e Industriale](#)
- [Unione Matematica Italiana](#)
- [Niedersächsische Staats- und Universitätsbibliothek Göttingen](#)
- [Masaryk University, Brno](#)

Marek Niezgódka:

[marekn@icm.edu.pl](mailto:marekn@icm.edu.pl)

Załączniki:



# Wybrane kierunki reform systemu nauki

Zmiana paradygmatów ocen naukowych:

- **Otwarte peer review:** kilkuetapowe
- **Publikacje: dynamiczny status** – od technicznego, poprzez dyskusyjny, do ostatecznego – certyfikowanego poprzez publikowane referencyjne oceny

**Premiowanie otwartego udostępniania wyników** badań naukowych: publikacje, dane, oprogramowanie, dokumentacje, metody, ...

Wyjście poza zasadę oceny parametrycznej

# Otwarty dostęp w Polsce – dane

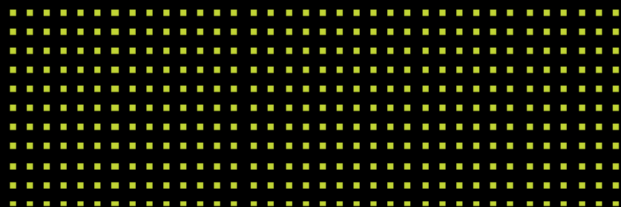


- 49% spośród niemal 2000 czasopism punktowanych z wykazu MNiSW udostępnia swoje bieżące numery bezpłatnie w Internecie.
- 59% przebadanych naukowców przynajmniej raz udostępniło w sposób otwarty swoją pracę, choć tylko 12% z nich robi to regularnie.
- 74% respondentów popiera udostępnianie w sposób otwarty wyników wszystkich badań naukowych finansowanych ze środków publicznych.
- Powstają pierwsze repozytoria instytucjonalne i dziedzinowe, w których naukowcy mogą umieszczać swoje prace (artykuły, książki, rozprawy doktorskie czy raporty) – obecnie 22+



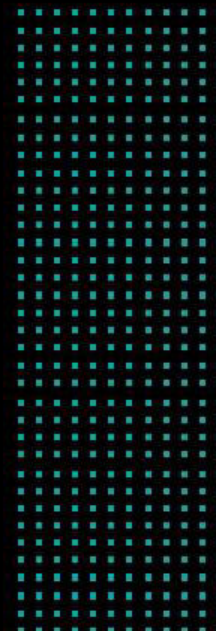
# Otwarta nauka w Polsce 2014

Diagnoza



Bo-Christer Björk, Wojtek Sylwestrzak, Jakub Szprot

## Analysis of Economic Issues Related to Open Access to Scientific Publications



Ref.: [Otwartanauka.pl](http://Otwartanauka.pl)

# Serwis Otwartanauka.pl



- Serwis o całości problematyki otwartości w badaniach naukowych i edukacji
- **Otwarte repozytorium**
- **Kolekcja otwartych polskich e-książek naukowych**
- Artykuły, analizy, wywiady z istotnymi postaciami otwartej nauki w Polsce i w świecie
- Seria wywiadów Otwartej Nauki: liderzy nauki w Polsce, liderzy światowej Otwartej Nauki

