

Paweł Szczęsny

OTWARTA NAUKA

czyli dobre praktyki uczonych

Toruń 2013

Stowarzyszenie EBIB

Paweł Szczęsny

OTWARTA NAUKA

czyli dobre praktyki uczonych

Wydawnictwa Zwarte Stowarzyszenia EBIB
pod red. Bożeny Bednarek-Michalskiej

Korekta

Izabella Milewska-Warta, Sylwia Wołek-Biernat

Skład i opracowanie graficzne

Emanuel Kulczycki

Nie wszystkie prawa zastrzeżone.

Broszura opublikowana na licencji [Creative Commons Uznanie autorstwa-
-Na tych samych warunkach 3.0](#) (CC BY-SA 3.0).

Publikacja dotowana przez Ministerstwo Nauki i Szkolnictwa Wyższego ze
strumienia finansowego Działalności Upowszechniającej Naukę



Ministerstwo Nauki
i Szkolnictwa Wyższego

Wydawca

Stowarzyszenie EBIB
Toruń 2013

ISBN 978-83-63458-03-4

SPIS TREŚCI

1.	Czym jest otwartość w nauce?	5
2.	Jaki problem rozwiązuje otwarta nauka?	7
3.	Otwarte zasoby nauki	11
3.1.	Otwarty dostęp do literatury naukowej	11
3.2.	Otwarte dane badawcze	14
3.3.	Otwarta komunikacja i nauka 2.0	16
3.4.	Otwarty proces badawczy (czyli nauka obywatelska)	19
3.5.	Co jeszcze można uwolnić?	22
4.	Nie wszystko w nauce powinno być otwarte	23
5.	Polskie otwieranie nauki	24
6.	Konkretne korzyści z otwierania zasobów nauki	27
6.1.	Przewaga pierwszego gracza	27
6.2.	Korzyści z punktu widzenia badacza	28

6.3.	Korzyści z punktu widzenia instytucji	29
6.4.	Korzyści z punktu widzenia systemu finansowania badań	30
7.	Praktyczne aspekty otwierania nauki	31
7.1.	Infrastruktura techniczna dla otwartej nauki	31
7.2.	Aspekt prawny otwierania zasobów	33
8.	Podsumowanie	35
9.	Przykładowa strategia otwierania nauki – wersja dla badacza	36
	Bibliografia	38
	Spis ilustracji	42
	Biogram autora	43

1. Czym jest otwartość w nauce

Cyprian Kamil Norwid, polski poeta okresu romantyzmu, nie został doceniony za życia. Większość dorobku została wydana dopiero po jego śmierci, m.in. dzięki „odkryciu” go na nowo przez Zenona Przesmyckiego. Dopiero wtedy w oczach zarówno czytelników, jak i specjalistów Norwid doczekał się uznania go za jednego z czterech najważniejszych polskich poetów tego okresu. Jak się to ma do otwartej nauki? Otóż otwartość w nauce jest jak dorobek zapomnianego poety — była tu od dawna, ale wymaga odkrycia na nowo, bo zapomnieliśmy o jej istnieniu.

Otwartość w nauce odwołuje się do dwóch podstawowych zasad prowadzenia badań naukowych: do swobodnej wymiany myśli i wiedzy oraz do weryfikowalności doświadczeń naukowych. Jest integralnym elementem dobrych praktyk naukowych. Często nie nazywamy jej nawet po imieniu. Dla przykładu, wielokrotnie zalecenia otwartości i transparentności znalazły się w opublikowanym w 2001 r. przez Komisję ds. Etyki w Nauce PAN dokumencie pt. „Dobre obyczaje w nauce. Zbiór zasad i wytycznych”¹, natomiast słowo „otwartość” pojawia się sporadycznie:

Publikacja w czasopiśmie kontrolowanym przez Komitet Redakcyjny jest istotnym końcowym punktem

1 Wersja z przedmową prof. Kornela Gibińskiego opisującą proces tworzenia tego dokumentu dostępna jest na stronach PAN. [Dokument elektroniczny]. [Dostęp 30 lipca 2013]. Dostępny w: <http://www.ken.pan.pl/images/stories/pliki/pdf/down.pdf>.

procesu naukowego. Z chwilą publikacji treść jej staje się częścią własności publicznej, jaką jest wiedza naukowa i otwarta zostaje droga do jej wykorzystywania z uwzględnieniem poniższych zasad.

Autorzy „Dobrych obyczajów w nauce” zwrócili uwagę na fakt, że wiedza naukowa jest własnością publiczną — jej wykorzystanie (z zachowaniem zasady uwzględniania praw autorów do opublikowanych materiałów) nie powinno podlegać ograniczeniom w szczególności:

Pracownik nauki przeciwstawia się jako szkodliwym dla jej rozwoju:

- *cenzurowaniu wypowiedzi i publikacji naukowych;*
- *ograniczaniu dostępu do materiałów źródłowych;*
- *utrudnianiu swobodnego przepływu informacji naukowej; (...)*

Piętnuje się tu nie tylko system wydawania literatury naukowej, ale też specyficzne praktyki samych uczonych, jak chociażby zwlekanie z publikowaniem wyników:

Opóźnianie ogłoszenia wyników pracy naukowej może być usprawiedliwione jedynie dobrem samej nauki

lub, co gorsza, publikowanie wyników badań na modłę Galileusza (i niestety wielu jego następców, jak chociażby Huygensa czy Hooke’a), czyli w sposób, który uniemożliwia odtworzenie głównych tez lub doświadczeń wyłącznie z tekstu komunikatu czy książki z powodu np. pominięcia pewnych kluczowych informacji:

Ogłaszanie nierzetelnych wyników badań, a także publikacje nieudokumentowane, np. dla zapewnienia sobie priorytetu są niedopuszczalne. Zarówno publikacje, jak i sprawozdania z działalności naukowej muszą być odtwarzalne przez innych.

Termin „otwarta nauka” jest dość nowy i to samo się tyczy angielskiego odpowiednika „open science” (nawiasem mówiąc pierwsza bądź jedna z pierwszych konferencji nt. idei otwartej nauki, która odbyła się w 1999 r., nosiła tytuł „Open Source/Open Science”², odnosząc się do etosu wolnego oprogramowania). Natomiast sama **otwartość nauki to nie żadna nowa moda, tylko podstawa prowadzenia badań we współczesnym świecie.**

2. Jaki problem rozwiązuje otwarta nauka?

Nigdy nie miałem pomysłu, którego nie można byłoby ulepszyć dzięki opowiedzeniu o nim możliwie największej liczbie ludzi — i nie sądzę, żeby ktoś kiedykolwiek takowy miał³.

² Strona konferencji wciąż jest dostępna: <http://openscience.bnl.gov/>.

³ Cytat pochodzi z felietonu „The Future of Science is Open” opublikowanego na portalu 3Quarks Daily. [Dokument elektroniczny]. [Dostęp 30 lipca 2013]. Dostępny w: http://3quarksdaily.blogs.com/3quarksdaily/2006/10/the_future_of_s_1.html. Jest to pierwsza z trzech części cyklu napisanego przez biochemika Billa Hookera w 2006 r.

Otwartość i wynikająca z niej transparentność leżą u podstaw nieskrępowanej wymiany wiedzy pomiędzy uczonymi. Jednakże przez ostatnie kilkadziesiąt lat środowisko naukowe straciło w sporej części kontrolę nad procesem komunikacji naukowej, głównie poprzez systematyczne przekazywanie podmiotom, spoza struktur akademickich, wielu czynności, które historycznie wykonywane były w całości przez akademików. Dla przykładu, czasopisma, które niegdyś w większej części wydawane były przez uczelnie, towarzystwa naukowe czy inne instytucje badawcze, w tej chwili niemalże w całości kontrolowane są przez komercyjnych wydawców. Naukowe bazy danych w istotnym stopniu (zależnym jednak od dziedziny) podzieliły los czasopism. Ponadto, niegdyś oczywisty obowiązek popularyzacji nauki i etosu uczonego, „zrzucony” został na dziennikarzy i rzadko oczekuje się od akademika działalności popularyzacyjnej.

Oczywiście wszystkie te zmiany podyktowane były idącymi za nimi praktycznymi korzyściami, a ponadto wynikały z takiego, a nie innego sposobu finansowania nauki. Jednakże nieoczekiwanymi konsekwencjami stały się utracenie części kontroli na komunikacją naukową oraz części kontroli nad procesem ewaluacji nauki. Jednym z przejawów tych konsekwencji jest fakt, że wciąż używamy w procesie ewaluacji uczonych miary „Impact Factor” obliczanej przez firmę Thomson Reuters, mimo tego, że jest miarą niereprodukowalną⁴ i co gorsza, negocjowalną⁵ (pomińmy inne jej wady). A dlaczego? Ponieważ nie mamy dostępu do tekstów czasopism w taki sposób, żeby samodzielnie zebrać dane bibliograficzne.

Obecny system funkcjonowania nauki jest mało skuteczny w wielu warstwach, natomiast najważniejszą kwestią, wywołującą coraz więcej kontrowersji, są koszty związane z jego funkcjonowaniem. Przede wszystkim całe środowisko ponosi bezpośrednio gigantyczne koszty obsługi systemu wymiany wiedzy naukowej. Coraz częściej dochodzi do spięć na linii wydawnictwo — instytucje akademickie. Dla przykładu, w 2007 r. instytuty Maxa-Plancka w Niemczech zrezygnowały z dostępu do czasopism Springera.

4 Rossner, M., Van Epps, H. and Hill, E. Show me the data. *Journal of Cell Biology* 2007, 179:1091-1092.

5 The Impact Factor Game. *PLoS Medicine* 2006, 3(6): e291.

ger Verlag z powodu nadmiernych żądań finansowych wydawcy (po kilku miesiącach po cichu osiągnięto ugodę). W 2010 r. Nature Publishing Group (NPG) zostało zbojkotowane przez system uniwersytetów kalifornijskich — z powodu czterokrotnego podniesienia opłat za dostęp do czasopism, nie tylko nie wyrażono zgody na przedłużenie umowy z NPG, ale także uczeni uniwersyteccy odmówili recenzowania dla czasopism wydawanych przez tego wydawcę i zagrozili, że zrezygnują z publikowania w tychże czasopismach (po kilku miesiącach uzgodniono nowe warunki, które nie są publicznie znane). Natomiast w 2012 r. zapoczątkowany został międzynarodowy bojkot wydawnictwa Elsevier — w tej chwili ponad 13 tysięcy uczonych podpisało się pod zobowiązaniem, że nie będą publikować w czasopismach tego wydawcy lub nie będą recenzować artykułów tam wydawanych, bądź też odmówią/zrezygnują z wykonywania pracy edytora (www.thecostofknowledge.com). Kwestie kosztów dotyczą nie tylko literatury, ale także baz danych naukowych w niektórych dziedzinach. Dla przykładu, otwarte bazy danych biologicznych mają długą historię sięgającą kilkudziesięciu lat. Natomiast pierwsza duża publiczna baza danych związków chemicznych powstała dopiero w 2004 r., nawiasem mówiąc przy akompaniamencie protestów firm udostępniających komercyjne bazy⁶.

Poza bezpośrednimi kosztami, obecny system generuje również trudne do oszacowania koszty pośrednie. Nierównomierny dostęp do wiedzy naukowej, wynikający z różnego poziomu finansowania nauki w poszczególnych krajach i instytucjach powoduje dodatkowe koszty związane ze źle ukierunkowanymi badaniami (wynikającymi z niepełnego obrazu wiedzy naukowej) oraz powtarzaniem doświadczeń, które zostały opublikowane, ale badacz nie ma dostępu do publikacji.

Podtrzymuje również dysproporcje pomiędzy nauką w krajach rozwiniętych oraz w krajach rozwijających się. Ponadto, dochodzą do tego koszty pewnych praktyk wywołanych takim, a nie innym systemem ewaluacji pracy badacza. Dla przykładu, dość często dochodzi do sytuacji, w których istotna część badań klinicznych ma dwuletni poślizg pomiędzy wygenerowaniem

6 Protest American Chemical Society przeciwko powstaniu bazy PubChem opart się o amerykański Kongres.



i opracowaniem wyników badań a ich publikacją⁷, wynikający z „konieczności opublikowania wyników w lepszym czasopiśmie”. Jak łatwo sobie wyobrazić, powoduje to podobny poślizg przy wprowadzaniu leków do obiegu, a co za tym idzie, generuje koszty nieskutecznej pomocy lekarskiej i usuwa często doświadczonych pracowników z rynku pracy.

Otwarta nauka ma swoje źródło w fundamentalnych zasadach funkcjonowania nauki, natomiast **w dzisiejszych czasach otwartość ma rozwiązać najbardziej palący problem: wydatków, jakie ponosimy na funkcjonowanie systemu nauki**. To nie znaczy, że to jedyna korzyść z otwartej nauki — ale temat innych, wymiernych korzyści rozwinęty będzie później.

⁷ Sommer, J., The delay in sharing research data is costing lives. *Nat Med.* 2010, Jul;16(7): 744.

3. Otwarte zasoby nauki

Nauka to nie tylko publikacje czy dane — zasoby nauki mają wiele twarzy. Ich otwieranie jest nierównomierne i specyficzne dla każdego typu zasobu. Zależy od środowiska, dziedziny oraz dostępności infrastruktury prawnej czy technicznej. Przyjrzyjmy się najpierw najbardziej oczywistemu przykładowi, czyli publikacjom naukowym.

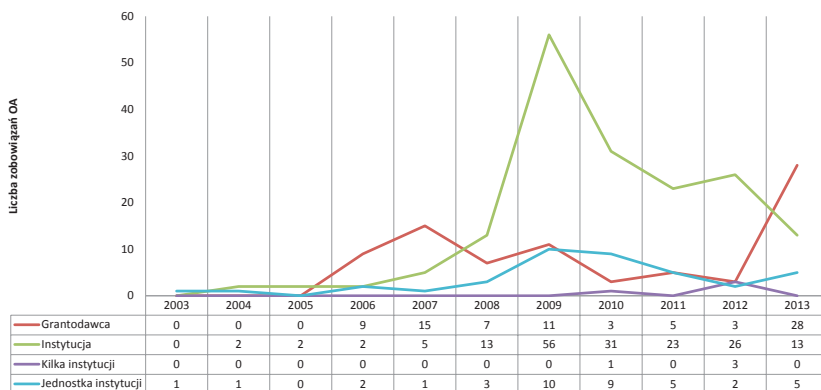
3.1. Otwarty dostęp do literatury naukowej

Idea otwartego dostępu do literatury naukowej (zwana z ang. Open Access, dalej: OA) jest jednym z kluczowych elementów procesu zwiększania dostępności zasobów nauki. Nie tylko dotyczy jednego z najbardziej istotnych elementów komunikacji naukowej, ale jest także najbardziej zaawansowanym pod względem prawnym i technicznym procesem systematycznego otwierania nauki.

Chociaż określenie OA ma dopiero nieco ponad 10 lat, to idea wolnego dostępu do publikacji ma dużo dłuższą historię. Próby współczesnej implementacji tego pomysłu są prawie tak stare jak Internet. Jedno z pierwszych otwartych czasopism naukowych powstało w 1989 r. („The Public Access Computer Systems Review”), natomiast w roku 1991 założono dwa repozytoria prac naukowych w odpowiedzi na zapotrzebowanie fizyków i matematyków na szybką wymianę wiedzy, mp_arc oraz arXiv (oba funkcjonują do dzisiaj, ale arXiv jest większy i bardziej znany, m.in. dzięki temu, że poszerzył katalog tematyczny o większość nauk przyrodniczych). Obecnie szacuje się, że około jedna czwarta wszystkich artykułów naukowych opublikowanych w 2012 r. jest dostępna w Internecie. Baza ROARMAP zbierająca dane

o repozytoriach oraz zobowiązaniach ze strony instytucji badawczych bądź fundujących badania określa liczbę tych ostatnich na świecie na ok. 300 i ta liczba dynamicznie rośnie.

Liczba zobowiązań OA w latach



Rys. 1. Liczba nowych zobowiązań OA w poszczególnych latach, podzielona na typy instytucji⁸. Warto zwrócić uwagę na skok liczby takich zobowiązań nakładanych na uczonych ze strony instytucji finansujących badania.

Istnieją dwa główne modele realizowania polityki otwartego dostępu do publikacji naukowych i każdy z nich ma swojego lidera. Pierwszy, czyli otwarte repozytoria oferujące pełne wersje opublikowanych artykułów (tzw. Green OA), wybrały Stany Zjednoczone. Badacze finansowani przez National Institutes of Health (NIH), a zgodnie z ostatnim memorandum Białego Domu (www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf), już wkrótce wszyscy badacze finansowani z budżetu

⁸ Dane pochodzą z serwisu ROARMAP. [Dokument elektroniczny]. [Dostęp 30 lipca 2013]. Dostępny w: <http://roarmap.eprints.org>.

tu państwa, będą zobowiązani do deponowania swoich prac w publicznym repozytorium, niezależnie od tego, gdzie ukazała się publikacja. Drugi model, czyli otwarte czasopisma (tzw. Gold OA) jest silnie wspierany przez Wielką Brytanię. Pokrywa się tam ze środków publicznych koszty zarówno udostępnienia publikacji, jak i koszty zatrzymania pełni praw przy autorze oraz udostępnienia ostatecznej wersji artykułu na wolnej licencji⁹.

Oba modele OA różnią się zarówno kosztami, jak i kwestiami praw autorskich. Bezwzględne uwolnienie artykułu w modelu Gold OA kosztuje zdecydowanie więcej, ponieważ jednostkowy koszt udostępnienia artykułu rekompensować musi utracone zyski wydawcy z publikowania w modelu zamkniętym (subskrypcja). Tych kosztów nie ma w modelu Green OA, a jednostkowy koszt udostępnienia artykułu jest znikomy, ponieważ koszty obsługi repozytorium są niskie. Model Green OA nie gwarantuje jednak możliwości swobodnego wykorzystania publikacji, ponieważ często nie mogą być udostępnione na wolnych licencjach, nawet przez ich autorów. Autorzy publikacji naukowych w znaczącej części nie są dysponentami praw autorskich — te, zazwyczaj, cedowane są na wydawcę.

Wybór modelu implementacji OA jest nietrywialny, aczkolwiek argument finansowy przechyla szalę w stronę modelu repozytoryjnego. Wprawdzie sumaryczne korzyści z przyjęcia modelu Gold OA są większe (wynikające np. ze spójnego modelu licencjonowania treści), to jednak na etapie przejściowym koszty finansowe modelu otwartych czasopism przewyższają model repozytoryjny co najmniej pięciokrotnie (<http://www.dlib.org/dlib/january13/houghton/01houghton.html>). Wyraźnych wskazań nie ma też na poziomie Unii Europejskiej. Komisja Europejska uznaje oba modele OA za akceptowalne. Komisja planuje też, że dla projektów finansowanych z programu Horizon 2020 wymagać się będzie albo zdeponowania kopii artykułu w repozytorium po upływie czasu embarga, postulowanego na co najwyżej 6 miesięcy,

⁹ Więcej informacji na stronach Research Councils UK. [Dokument elektroniczny]. [Dostęp 30 lipca 2013]. Dostępny w: <http://www.rcuk.ac.uk/research/Pages/outputs.aspx>.

albo publikowania w modelu Gold OA, przy jednoczesnym pokrywaniu kosztów takiego publikowania (europa.eu/rapid/press-release_IP-12-790_en.htm).

W obecnej chwili najrozsądniejszym rozwiązaniem wydaje się być przejście na model repozytoryjny i rozpoczęcie prac nad przejściem na model otwartych czasopism w następnym kroku. Spójność licencjonowania pozwoli na automatyczną analizę tekstu, która zostanie omówiona szerzej w następnej sekcji.

3.2. Otwarte dane badawcze

Kwestia otwartości danych pozyskanych w procesach badawczych, w przeciwieństwie do literatury naukowej, wydaje się być dziedzinowo specyficzna. Są obszary nauki, gdzie otwartych baz danych jeszcze nie ma, a ponadto w niektórych dziedzinach dane naukowe często nie są dostępne w ogóle (np. nie są w ogóle w formie cyfrowej) i nie zawsze podyktowane jest to racjonalnymi przesłankami. W większości przypadków, wolny dostęp do danych źródłowych wynika z umowy społecznej danego środowiska badaczy. W biologii strukturalnej otwarty dostęp do danych funkcjonował już w latach 70., a w latach 90. stał się obowiązkiem (nie można w tej chwili opublikować pracy o nowej strukturze białka, nie udostępniając tej struktury). Analogiczne umowy społeczne obowiązują od lat 90. w innych obszarach biologii. Rozwój dziedziny doprowadził do kolejnych kroków i skodyfikowania dotychczas obowiązującego „zwyczaju”. Udostępnianie danych dopiero w momencie publikacji przestało być w niektórych obszarach biologii warunkiem wystarczającym. Od kilku lat funkcjonują inicjatywy przyspieszające dostęp do surowych danych, to jest do danych zebranych, ale jeszcze nieprzeanalizowanych. Dla przykładu w genomice powszechnie już stosuje się tzw. Ft. Lauderdale Agreement (www.genome.gov/10506537), czyli umowę na udostępnienie danych, która gwarantuje autorom eksperymentu pierwszeństwo w opublikowaniu globalnych analiz genomu, natomiast pozwala na wykorzystanie zebranych danych do badań szczegółowych (pojedyncze geny, mutacje) bez żadnych ograniczeń.

Na ten trend nakładają się coraz częstsze decyzje o zmianie polityki czasopisma naukowego na taką, która nie dopuszcza do pojawienia się w artykule stwierdzeń, które poparte są niedostępnymi danymi (schowanymi za sformułowaniem: „data not shown”). Uzasadnienia tych zmian są oczywiste: takie praktyki utrudniają zarówno proces recenzji artykułu, jak i możliwość późniejszego odtworzenia opisanych w nim eksperymentów¹⁰. Ponadto coraz częściej podnosi się kwestię dostępności nie tylko wyników eksperymentów, ale także materiałów źródłowych, którymi mogą być zarówno skan niedostępnego wcześniej dokumentu poddanego później analizie, jak i materiał sienny roślin, na których wykonane zostały eksperymenty¹¹.

Otwartość danych naukowych ma olbrzymie znaczenie nie tylko wprost dla rozwoju danej dziedziny, ale także dla całej nauki i środowisk, które bezpośrednio z dorobku nauki korzystają. Od niedawna mówi się o erze „wielkich danych” (ang. „big data”) i o możliwościach, które mogą się kryć w ich analizie na dużą skalę. Zazwyczaj tego rodzaju analizy były domeną pracowników instytucji finansowych (analizy milionów transakcji w poszukiwaniu tych nietypowych, będących prawdopodobnie wynikiem kradzieży karty kredytowej), lub fizyków wysokich energii (analizy wyników zderzania cząstek). Obecnie dostępność danych naukowych, jak i ich semantycznego opisu w postaci publikacji pozwala dokonywać automatycznych odkryć — chodzi o wnioskowanie na podstawie zgromadzonej wiedzy przez odpowiednie algorytmy¹². I możliwości takich analiz dotyczą wszystkich dziedzin życia, nie tylko samych badań. Dla przykładu, w 2011 r. McKinsey Global Institute oszacował, że zdjęcie prawnych barier dotyczących automatycznych analiz

10 Data shown. *Nat Chem Biol*, 2008, 4:575.

11 Przykładem inicjatywy, która za cel obrała sobie standaryzację dostępu do materiału biologicznego jest Creative Commons Biological Material Transfer. [Dokument elektroniczny]. [Dostęp 30 lipca 2013]. Dostępny w: <http://sciencecommons.org/projects/licensing/>. Niestety, projekt nie jest już aktywny.

12 Dwa projekty, które można podać za przykład to LarkC (Large Knowledge Collider) oraz OpenPHACTS, z których ten pierwszy jest systemem do wnioskowania, a drugi bazą zarówno fizykochemicznych właściwości związków chemicznych, jak i opisów ich działania, interakcji, syntezy, itd.

treści i danych zmniejszyłoby sumaryczne wydatki rządów krajów europejskich o ok. 100 miliardów euro rocznie (www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation). W tym kontekście staje się jasne, dlaczego model Gold OA jest bardziej perspektywiczny — nieskrępowany dostęp do treści (w ramach licencji Creative Commons) pozwala na lepsze „zrozumienie” dołączonych do treści danych (kontekst, metodologia pozyskania).

Otwarte dane naukowe wpisują się też w szerszy trend otwartości danych publicznych (takich jak chociażby statystyki GUS, albo informacje o budżecie państwa). I podobnie jak w przypadku danych publicznych, ich otwartość (lub jej brak) może być kwestią interesu państwowego (więcej o tym w rozdziale „Nie wszystko w nauce powinno być otwarte”).

3.3. Otwarta komunikacja i nauka 2.0

Określenie *Sieć 2.0* opisuje zmianę, jaka zaszła w ostatnich kilku latach dzięki technologiom umożliwiającym tworzenie i edytowanie treści w Internecie bez zaawansowanej wiedzy technicznej. Po trzydziestu latach od pojawienia się poczty elektronicznej, Internet w końcu wydaje się realizować pokładane w nim nadzieje bycia międzynarodową platformą konwersacji, wymiany myśli, twórczości i informacji. Użyteczność tej platformy skusiła także wielu uczonych. Tradycyjne formy komunikacji naukowej sprowadzały się do konferencji oraz recenzowanych artykułów — Internet zaoferował kanały, które są bardziej otwarte, transparentne, znacznie szybsze i zachęcały do współpracy. Całość aktywności uczonych w sieci z czasem zaczęto nazywać *Nauką 2.0*.

Wciąż jesteśmy na etapie wielu eksperymentów z użyciem narzędzi Sieci 2.0 w nauce. Z jednej strony blogi naukowe już na dobre zagościły pod dachami instytucji akademickich. Ich zalety są oczywiste — służą do komentowania bieżącej literatury, dzielenia się odnośnikami do ciekawych

stron (także o nauce), jako dzienniki laboratoryjne, platformy do wspólnego pisania publikacji, a także do luźnej dyskusji, chociażby na temat przyszłości nauki. Z drugiej strony nie udały się próby stworzenia wspólnej platformy dla naukowych blogerów — porażką okazały się projekty zarówno profesjonalnych wydawców naukowych, jak i firm zajmujących się stricte mediami (platformę Nature Network zbudowaną przez Nature Publishing Group, znanego wydawcy prestiżowych czasopism naukowych, opuściło większość użytkowników; wielu czytelników oraz autorów straciła też platforma ScienceBlogs założona przez grupę Seed Media). Podobne fluktuacje dotyczą w zasadzie każdej innej technologii współczesnego Internetu. Twitter, platforma mikroblogowa, gdzie użytkownicy dzielą się krótkimi, 140-znakowymi komunikatami, cieszył się ogromnym zainteresowaniem uczonych chociażby przy okazji misji lądownika NASA na Marsa, gdzie był wykorzystywany jako alternatywny kanał komunikacji.

**Tradycyjne formy
komunikacji naukowej
sprowadzały się do
konferencji oraz recen-
zowanych artykułów**



W konsekwencji, pojawiły się prace, gdzie ten portal używany był w kontekście prawdziwych badań naukowych (np. wykorzystanie krótkich komunikatów do monitorowania połowów¹³). Nie przełożyło się to natomiast na stabilną obecność mikroblogów w nauce. Swego czasu sporym zainteresowaniem cieszyły się wirtualne konferencje i sesje posterowe na platformie Second Life, albo niszowe portale społecznościowe budowane wyłącznie dla naukowców. Analogicznie, w pewnym momencie wydawało się, że umożliwienie komentowania artykułów naukowych wyzwoli potencjał nieformalnej komunikacji i kilku wydawców (m.in. PLoS) udostępniło taką funkcjonalność na swoich stronach. Podobnie jak i w poprzednich przypadkach, kiedy minął pierwszy zachwyty, spora część pierwszych użytkowników odeszła, czasem po prostu rezygnując, a czasem do innych serwisów. Czy to oznacza, że nieformalna komunikacja naukowców w sieci jest skazana na porażkę? Bynajmniej.

Otwarty dostęp do publikacji jest procesem, który w tej chwili najwygodniej regulować i inicjować na poziomie instytucji lub fundatora. Otwarta nieformalna komunikacja z oczywistych względów nie podda się takim regulacjom i próby stworzenia centralnego systemu dla nauki, jak pokazuje krótka historia Nauki 2.0, tak czy inaczej nie mają szansy na długofalowy sukces. To, co stanowi o sile narzędzi współczesnego Internetu, to udostępnienie możliwości samoorganizacji środowisk naukowych wokół jednego czy kilku platform. Innymi słowy, potrzeba nieformalnej wymiany wiedzy odgrywa pierwszoplanową rolę — wybór narzędzi jest trochę kwestią przypadku. Dla przykładu, w dziedzinie autora, bioinformatyce, funkcjonuje kilka „miejsc”, wokół których skupieni są badacze. Dwa skrajne przypadki to forum Seqanswers oraz serwis „pytanie-odpowiedź” Biostar (z ang. Q&A, koncepcja mocno osadzona w informatyce — największy serwis tego typu, Stack Overflow ma w archiwum ponad pięć milionów różnych pytań <http://stackoverflow.com/questions>). Jeśli chodzi o sposób używania obu stron (zarówno czysto technicznie, jak i w warstwie „etykiety” portalu) to jest on stosunkowo różny. Natomiast tematyka dyskusji na obu portalach w sporej części się pokrywa, mimo tego, że

13 Baker, S. M., Oschenger, I., Description and Initial Evaluation of a Text Message Based Reporting Method for Marine Recreational Anglers. *Mar Coast Fish.* 2009, 1:143-154.

stosunkowo niewielu użytkowników jest mocno aktywnych na obu portalach jednocześnie.

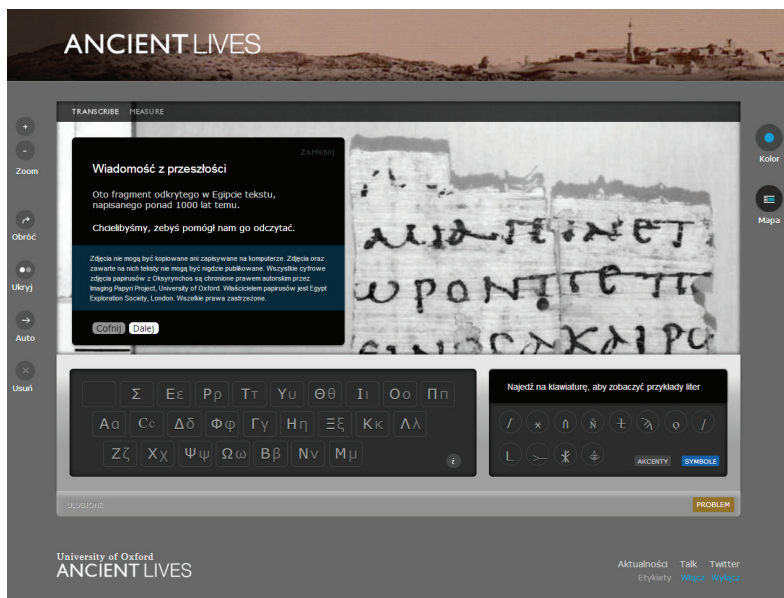
Podsumowując, o realnym i stabilnym otwarciu nieformalnej komunikacji decydują całe środowiska badaczy, a nie marketing czy dostępność narzędzi. *Nauka 2.0* szybciej rozwija się tam, gdzie tempo rozwoju dziedziny dyktuje wysokie potrzeby badaczy w zakresie wymiany wiedzy.

3.4. Otwarty proces badawczy (czyli nauka obywatelska)

Historycznie pod nazwą nauka obywatelska (ang. *citizen science*) rozumiano zaproszenie „niefachowców” do zbierania danych dla naukowców, np. poprzez liczenie ptaków lub odsyłanie kartek z butelek, które morze wyrzuciło na plażę (na początku XX w. w ten sposób badano prądy morskie). Internet, wydawać by się mogło, niewiele może zmienić w tym podejściu — pozwoli tylko na sięgnięcie do szerszego grona. Okazało się, że jest inaczej. Sieć, będąc nie tylko „oknem na świat”, ale także platformą współpracy pozwoliła na rozwinięcie współpracy pomiędzy naukowcami-amatorami a akademikami na niespotykaną dotąd skalę. Ramię w ramię uczeni i społeczeństwo odkrywają nowe ciała niebieskie, rozwiązują skomplikowane problemy matematyczne, przewidują strukturę białek lub dokonują interpretacji starożytnych tekstów. Nigdy do tej pory nie było łatwiej włączyć się w badania naukowe bez formalnych kwalifikacji akademickich.

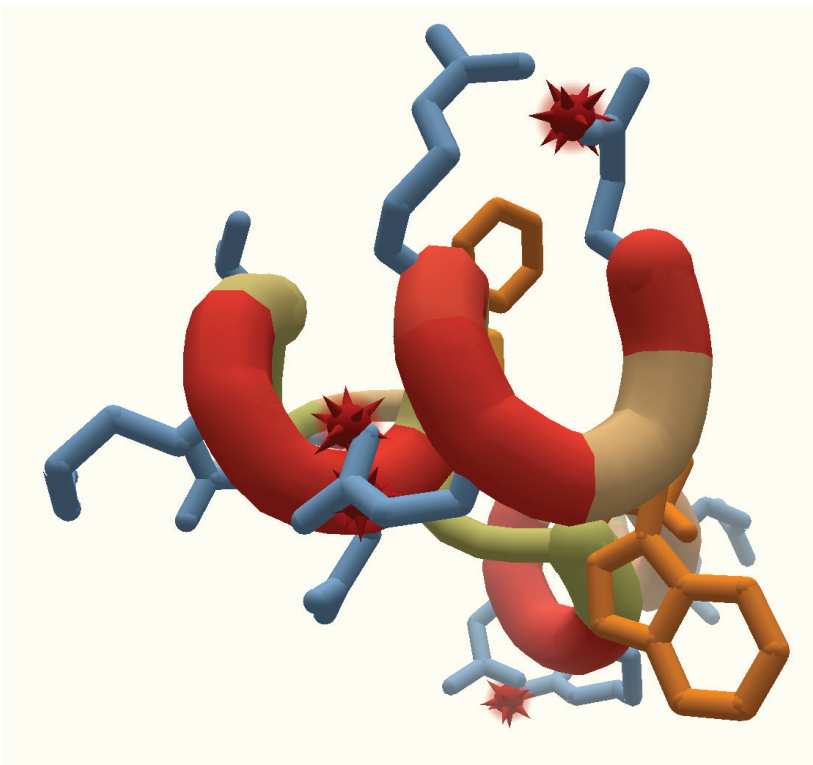
Metodyka tych projektów jest bardzo różna. Są to zarówno gry (Fold It, Phylo), quizzowe systemy adnotacji danych (Spectralgame, Annotathon), projekty stawiające na właściwe rozpoznanie obrazu (Ancient Lives, Galaxy Zoo) lub luźne projekty oparte na dyskusji (Polymath Project, Open Dinosaur Project), gdzie każdy uczestnik może zgłosić swój pomysł rozwikłania danej zagadki. Wydawać by się mogło, że odpowiedni poziom wiedzy będzie barierą w większości projektów nauki obywatelskiej. Okazuje się jednak, że nawet w bardzo hermetycznych obszarach nie wszyscy zainteresowani są zawo-

dowymi naukowcami. Przykładem może być Polymath Project, w którym w pierwszej edycji (michaelnielsen.org/polymath1/index.php?title=Polymath1) 27 osób w ciągu 37 dni pomogło rozwiązać problem matematyczny postawiony przez jednego z laureatów Medalu Fieldsa, Tima Gowersa. Nie wszyscy uczestnicy byli czynnymi naukowcami, aczkolwiek oczywiste jest, że głęboka wiedza matematyczna była niezbędna. Współpraca ta oczywiście bywa doceniona. W przypadku gry Fold It, wkład intelektualny graczy został oceniony tak wysoko, że gracze razem występują jako zbiorowy współautor na publikacjach twórców gry jako „Fold It Players”¹⁴.



Rys. 2. Widok portalu Ancient Lives. Widoczne jest powiększenie fragmentu z tekstem, oraz klawiatura do wprowadzania odczytanych znaków.

¹⁴ Publikacje graczy Fold It znajdują się w bazie PubMed. [Dokument elektroniczny]. [Dostęp 30 lipca 2013]. Dostępny w: <http://www.ncbi.nlm.nih.gov/pubmed/?term=foldit+players%5Bau%5D>.



Rys. 3. Widok „białkowego puzzla” programu Fold It. Na różne sposoby (tutaj na czerwono) program pokazuje problemy steryczne danej struktury i użytkownik musi je rozwiązać.

Należy też wspomnieć o jeszcze jednym aspekcie otwierania procesu naukowego. Jest to mianowicie bezpośrednie uczestnictwo amatorów nie tylko jako współpracowników, ale także jako źródło materiału badawczego lub danych. Przykładem takiej inicjatywy może być projekt Cancer Commons, gdzie chorzy na jeden z trzech rodzajów nowotworów proszeni są o dużą ilość informacji medycznych, które potem wykorzystywane są w sumarycznych statystykach do ulepszania metod leczenia tych chorób. Oczywiście uczestnicy programu informowani są o efektach oraz sugestjach terapii na

bieżąco. Współpraca w obszarze medycyny bywa także bardziej wielopoziomowa — w projekcie μ Biome uczestnicy nie tylko dostarczają próbki flory bakteryjnej z własnych ciał, ale także finansują całe badania. Należy się spodziewać, że tego rodzaju projektów będzie coraz więcej.

Otwieranie procesu naukowego dla innych społeczności zdarza się w większości dziedzin, niezależnie czy projekt dotyczy nauk ścisłych, medycznych czy humanistyki. Wprawdzie nie każdy projekt badawczy da się tak przeformułować, żeby naturalnie odpowiadał specyficznym warunkom prowadzenia projektów nauki obywatelskiej, ale na pewno nie dziedziną jest ograniczeniem.

3.5. Co jeszcze można uwolnić?

W zależności od dziedziny i umowy społecznej, na którą przystali badacze z konkretnego obszaru otworzyć można np. doktoraty, kod źródłowy, recenzje, konferencje, e-laboratoria. Wydaje się, że najmniej kontrowersji toczy się wokół kodu źródłowego oprogramowania tworzonego przez naukowców (zwłaszcza w tych obszarach nauk podstawowych, gdzie kod źródłowy jest elementem analizy danych eksperymentalnych, a nie głównym osiągnięciem badacza). Natomiast dyskusja o otwieraniu zarówno recenzji (publikacji, grantów), jak i doktoratów, niesie z sobą duży ładunek emocjonalny, niezależnie od tego, że przykładów na otwieranie obu zasobów znajdziemy mnóstwo za granicą.

Zasobem, który nie przynależy wprost do nauki, mimo tego, że tworzą go uczeni, są otwarte materiały edukacyjne. Tutaj na szczęście kontrowersji nie ma i mamy w kraju wiele świetnych przykładów otwierania uczelnianych materiałów dydaktycznych lub nawet tworzenia otwartych kursów e-learningowych.

4. Nie wszystko w nauce powinno być otwarte

Omówiliśmy wcześniej wiele przykładów otwierania zasobów nauki i zanim przejdziemy do praktycznych korzyści z otwartości, trzeba wspomnieć o ograniczeniach, które trzeba nałożyć na ten proces.

Jedną z głównych ról odgrywa tu kwestia prywatności danych osobowych, którą w dobie wspomnianych wcześniej „wielkich danych” łatwo naruszyć. Przykładem są badania nad identyfikacją nazwisk dawców próbek pozornie anonimowych danych genetycznych^{15, 16}. Znajomość sekwencji czyjegoś genomu może doprowadzić do sytuacji, w której przestępca „podrzuci” czyjeś DNA (zsyntetyzowane wcześniej w laboratorium na podstawie dostępnych informacji) na miejsce przestępstwa, tak aby przypisać je komuś innemu. Kolejnymi istotnymi elementami są: bezpieczeństwo kraju, ochrona przemysłu czy też utrzymanie przewagi strategicznej zarówno naukowej, jak i innowacyjnej. Przykładem mogą być dane geologiczne (krytyczne z punktu interesów państwa) lub publikacje kryminalistyczne (opisy procedur kryminalistycznych to prawie jak podręcznik dla potencjalnych przestępców). Jednakże problem otwartości danych dotyka dużo większej liczby wąskich dziedzin i często w obszarach, które na pierwszy rzut oka wydają się niekrytyczne. Dla przykładu, dane o bioróżnorodności wraz z identyfikacją rzadkich gatunków przyczyniają się do ich przetrzebień przez kłusowników, ponieważ na czarnym rynku rzadkie egzemplarze zwierząt osiągają zawrotne ceny¹⁷. Innym przykła-

15 Gymrek, M., McGuire, A. L., Golan, D., Halperin, E., Erlich, Y., Identifying Personal Genomes by Surname Inference., *Science*, 2013, (339) 6117, 321-324.

16 Bohannon, J., Genealogy Databases Enable Naming of Anonymous DNA Donors. *Science* 2013, (339) 6117, 262.

17 Hall, R. J., Milner-Gulland, E. J., Courchamp, F., Endangering the endangered: The effects of perceived rarity on species exploitation. *Conservation Letters* 2008 1:75–81.

dem są badania nad mutantami groźnych wirusów — w przypadku mutantu wirusa grypy H5N1 spora część danych, w tym opisy metodologii, zostały utajnione z obawy przed bioterroryzmem¹⁸, niezależnie od tego, że badania te mają fundamentalne znaczenie w obliczu potencjalnej epidemii grypy.

Radykalna otwartość zasobów nauki nie jest najlepszym pomysłem w naszych, bardzo złożonych, czasach. Mimo tego, że ogólne korzyści z otwierania nauki trudne są do przecenienia, w pewnych dziedzinach nauki powinno się ustalić sztywne granice dostępności informacji. W przeciwnym przypadku pojawi się ryzyko wygenerowania problemów na skalę również trudnej do przecenienia.

5. Polskie otwieranie nauki

Z powodów zarówno historycznych, jak i systemowych, stopień otwarcia polskich zasobów nauki nie czyni naszego kraju liderem. Z drugiej strony pozytywnie wyróżniamy się na tle krajów z podobnym bagażem historycznym, ponieważ proces otwierania tychże zasobów zdecydowanie nabrał rozpędu w ostatnich latach¹⁹.

Otwartość polskich zasobów nauki jest fragmentaryczna, ponieważ nigdy nie stworzono jakiegokolwiek spójnej strategii otwierania nauki na poziomie chociażby instytucji. Dostępne zasoby są wynikiem starań pojedynczych jednostek lub grup lobbujących na rzecz otwartości. Nie wydaje się, aby któraś z dziedzin nauki znacząco wybijała się na tle innych — pozytywne przykłady są obecne w prawie każdym obszarze. Infrastruktura dla otwartej nauki powstała dopiero w ostatnich dwóch, trzech latach, ciężko więc

18 Herfst, S. et al Airborne Transmission of Influenza A/H5N1 Virus Between Ferrets. *Science* 2012, (336) 6088, 1534-1541.

19 Świetne kalendarium rozwoju wolnego dostępu do literatury naukowej w Polsce przygotowały organizacje KOED i EBIB. [Dokument elektroniczny]. [Dostęp 30 lipca 2013]. Dostępny w: http://www.ebib.pl/?page_id=887 pod redakcją Bożeny Bednarek-Michalskiej.

spodziewać się, że jej obecność natychmiast zmieni obraz polskich zasobów otwartych.

Najbardziej zaawansowanym pod względem otwartości zasobem nauki są oczywiście publikacje. W Polsce funkcjonuje ponad 160 czasopism w modelu OA²⁰ (czyli całość czasopisma jest otwarta), mamy też ok. ok. 100 bibliotek cyfrowych²¹ zawierających prace naukowe i parę klasycznych repozytoriów naukowych, jak :

- ◀ **ECNIS** — Instytut Medycyny Pracy w Łodzi.
- ◀ **AMUR** — Uniwersytet A. Mickiewicza, Poznań
- ◀ **RUW** — Uniwersytet Warszawski
- ◀ **CEON** — ogólnopolskie repozytorium naukowe ICM
- ◀ **RUMAK** — Uniwersytet M. Kopernika, Toruń
- ◀ **SUW** — Politechnika Krakowska
- ◀ **IBB PAS Repository** — Polska Akademia Nauk, Warszawa
- ◀ **ENY** — Politechnika Wroclawska
- ◀ **RUŁ** — Uniwersytet Łódzki

Z powyższych repozytoriów tylko trzy objęte są zobowiązaniem, tj. uchwałą obligującą pracowników do deponowania wszystkich swoich publikacji naukowych w repozytorium instytucjonalnym (jedyne jak do tej pory zobowiązanie na poziomie całej instytucji przyjęte zostało w Instytucie Biochemii i Biofizyki PAN). Od kilku lat toczy się dyskusja na wszystkich szczeblach na temat otwierania polskich zasobów wydawniczych oraz polityki otwartego dostępu do publikacji wydawanych przez komercyjne podmioty. Wydaje się, że wkrótce zostaną zaproponowane rozwiązania na poziomie dużych instytucji, jeśli nie od razu na poziomie krajowym. Zapowiedzią takich decyzji jest np. wspólna uchwała PAN i KRASP z lipca 2013, dotycząca kierunku

20 Dane pochodzą z bazy DOAJ. [Dokument elektroniczny]. [Dostęp 30 lipca 2013]. Dostępny w: <http://http://www.doaj.org/doi?func=byCountry&uiLanguage=en>.

21 Dane pochodzą z bazy FBC dostępnej w: <http://fbc.pionier.net.pl/owoc/oai-hosts>.

prac nad otwartym dostępem do treści publikacji naukowych i edukacyjnych²². Inicjatorem prac nad tymi rozwiązaniami jest Komisja Europejska ze swoimi zapowiedziami wdrożenia polityki otwartego dostępu dla wszystkich projektów finansowanych z programu Horizon 2020.

Co do innych zasobów, to znajdziemy na stronach polskich instytucji naukowych (w tym także uczelni) przykłady otwartych danych, otwartego kodu źródłowego, czy otwartych zasobów edukacyjnych. Dobrym przykładem są Otwarte Zasoby Edukacyjne (OZE) Uniwersytetu Mikołaja Kopernika w Toruniu (<http://portal.umk.pl/web/otwarte-zasoby>), które stały się częścią konglomeratu zasobów, jakie ta uczelnia upowszechnia w modelu Open Access od lat:

1. czasopisma otwarte <http://wydawnictwoumk.pl/czasopisma/>,
2. repozytorium otwarte RUMAK <http://repozytorium.umk.pl/>,
3. biblioteka cyfrowa KPBC <http://kpbc.umk.pl/dlibra>

Paradoksalnie, otwartość nie zawsze idzie w parze z dostępnością — część z zasobów otwartych jest trudna do znalezienia. Brakuje ogólnopolskiego katalogu zasobów otwartych (z czasem tę lukę ma zapełnić baza Otwarte Zasoby (<http://otwartezasoby.pl/>) budowana przez Centrum Cyfrowe), więc trudno ocenić skalę tych inicjatyw. Czasami też się zdarza, że np. rekrutacja do projektu nauki obywatelskiej, w założeniu badacza — otwartego dla wszystkich zainteresowanych, ogłaszana jest zamkniętymi kanałami (tylko do szkół, brak informacji w Internecie). Natomiast kierunek zmian od wielu lat się nie zmienił — w znakomitej większości przypadków instytucje naukowe czy uczelnie ułatwiają dostęp do swoich zasobów, a nie utrudniają.

22 Uchwała została podjęta na wspólnym posiedzeniu Prezydium KRASP i Prezydium PAN. Treść uchwały: [Dokument elektroniczny]. [Dostęp 30 lipca 2013]. Dostępny w: http://www.aktualnosci.pan.pl/images/stories/pliki/2013/Wiadomo%C5%9Bci_biez/stanowisko_KRASP-PAN_open_access.pdf.

6. Konkretnie korzyści z otwierania zasobów nauki

Otwarta nauka mierzy się z rozwiązaniem fundamentalnych problemów funkcjonowania systemu prowadzenia badań naukowych. Systemowe korzyści z otwierania nauki takie, jak niższe koszty badań, wzrost współpracy naukowej, wzrost poziomu edukacji, czy też spadek różnic pomiędzy państwami rozwiniętymi i rozwijającymi się, są niekoniecznie najlepszą zachętą do otwarcia zasobów nauki z punktu widzenia pojedynczego badacza, instytucji badawczej czy instytucji finansującej badania. To jednak nie oznacza, że innych korzyści nie ma — są, ale częściej wynikają z dynamiki całego systemu nauki i jej otoczenia niż wprost z procesu otwierania.

6.1. Przewaga pierwszego gracza

Przewaga pierwszego gracza (ang. first mover advantage) jest to taka sytuacja, kiedy osoba lub instytucja jako pierwsza decydująca się na dany krok (może to być decyzja o otwarciu zasobów, lub też decyzja o nowym obszarze badawczym) zyskuje przewagę nad tymi, które w danym kierunku podążają z opóźnieniem. Jest to dość dobrze znany mechanizm z punktu widzenia strategii biznesowej. Co to ma jednak wspólnego z nauką? Okazuje się, że całkiem sporo.

Przewaga pierwszego gracza jest mechanizmem leżącym u podstaw wyników większości analiz sugerujących większą cytowalność otwartych zasobów naukowych. Im mniej otwartych zasobów do wyboru, tym większe znaczenie tych nielicznych dostępnych. Co za tym idzie, pozytywny efekt otwartości w zakresie cytowań maleje w miarę jak liczba otwartych zasobów rośnie.

Innymi słowy, w zakresie wzrostu cytowań korzyść dla badacza, czy instytucji może wynikać z ich bezpośredniego merytorycznego lub instytucjonalnego otoczenia.

6.2. Korzyści z punktu widzenia badacza

Naturalną konsekwencją dostępności zasobów naukowych jest możliwość ich odnalezienia przy za pomocą wyszukiwarek internetowych. W oczywisty sposób zwiększa to widoczność zarówno zasobu, jak i jego autora, co może (w zależności od okoliczności) przełożyć się na wzrost cytowań danego zasobu.

Za wzrostem widoczności w sieci idzie też kwestia rozpoznawalności. Budowanie wizerunku w sieci przestało już być fanaberią, a stało się koniecznością, zwłaszcza wśród młodych uczonych. Stosunkowo rzadko propozycje współpracy płyną poprzez unijne bazy ekspertów — pojawiają się natomiast oferty od osób, które znalazły kogoś w wyszukiwarce nie po nazwisku, a po interesujących je słowach kluczowych (czasami katalogów internetowych skończyły się przeszło 10 lat temu). Posiadanie profilu na portalach społecznościowych, zwłaszcza tych używanych przez uczonych, również przekłada się na większą widoczność i rozpoznawalność (autor tego opracowania nawiązał współpracę naukową m.in. poprzez portal FriendFeed) i wspomaga proces budowania wizerunku uczonego.

Obecność uczonego w sieci pomaga w budowaniu zaufania ze strony społeczeństwa, co jest nie bez znaczenia w kontekście krytycznych głosów dotyczących transparentności systemu nauki (zwłaszcza w kontekście badań medycznych lub związanych z bezpieczeństwem kraju). Potencjalnie może też być krytyczne w sytuacji, kiedy naukowiec zechce spróbować sfinansować swój projekt ze zbiórki publicznej (tzw. crowdfunding, coraz bardziej popularny na Zachodzie).

Za otwartością zasobów idą możliwości mierzenia ich wpływu, np. przez statystyki ściągnięć, liczby wzmianek w sieci, odnośników z naukowej blogos-

fery, komentarzy itp. Część wydawnictw (np. PLoS) już udostępnia tego rodzaju informacje jako tzw. article-level metrics, czyli metryki użycia indywidualnego artykułu. Te miary niekoniecznie mówią o naukowym znaczeniu danego zasobu (dlatego ostrożnie należy z nich korzystać w kontekście ewaluacji pracy badacza), ale pozwalają na poziomie całego systemu ocenić rolę i widoczność danych zasobów w świecie. W oczywisty sposób obecność uczonego w sieci przekłada się na większe wykorzystanie (w sensie odwiedzin/pobrań) zasobów jego autorstwa, dlatego niektórzy naukowcy umieszczają te statystyki w CV lub formularzach ewaluacyjnych.

6.3. Korzyści z punktu widzenia instytucji

Podstawowe korzyści dla instytucji pokrywają się z korzyściami indywidualnych badaczy. Jest to ponownie wzrost widoczności i rozpoznawalności instytucji, jeśli nie tylko obecna jest w sieci poprzez swoją stronę, ale także oferuje w niej otwarte zasoby naukowe. Natomiast kwestia wizerunku ma wówczas inny wymiar — instytucja, która dzięki swojej otwartości pozyska wysokie zaufanie społeczne, może być głównym ośrodkiem opiniotwórczym w swojej dziedzinie, może wywierać wpływ na politykę państwa w pewnym zakresie, może łatwiej pozyskiwać partnerów, w tym partnerów komercyjnych, oraz może cieszyć się większym zainteresowaniem np. studentów, czy potencjalnych pracowników.

W kontekście samych uczelni, otwieranie zasobów edukacyjnych (kursów czy podręczników) przez uczelnie może mieć wpływ na decyzję o wyborze miejsca studiowania przez przyszłych studentów. Z ich punktu widzenia, wolny dostęp do treści edukacyjnych ma przede wszystkim wymiar finansowy. W obliczu nadchodzącego kryzysu demograficznego otwartość zasobów edukacyjnych dla szkół wyższych może stać się jednym z elementów promocji.

Ponadto, idąca za otwartością transparentność pozwoli na łatwe wdrożenie systemów kontroli potencjalnych nieuczciwości. Dotyczy to zarówno plagiatów (nie tylko publikacji naukowych, ale także prac dyplomowych czy doktoratów), jak i przypadków fałszowania wyników badań. Sama możliwość łatwej weryfikacji wszystkich etapów procesu badawczego zmniejszy chęć podjęcia próby fałszowania jego wyników.

6.4. Korzyści z punktu widzenia systemu finansowania badań

W przypadku systemu finansowania badań (czyli m.in. z perspektywy fundatora), główne korzyści z otwartości istotnie różnią się od tych odnoszonych przez badacza czy instytucje. Fundator zainteresowany jest jak największą maksymalizacją efektywności finansowej wynikającej z nakładów na badania, co m.in. oznacza, że im większe ponowne wykorzystanie wyników badań, tym większa efektywność i większy wpływ tych badań na naukę. Otwartość zasobów wydaje się być warunkiem koniecznym (aczkolwiek oczywiście nie wystarczającym) do osiągnięcia tego celu.

Powyższe ma zastosowanie zwłaszcza, jeśli fundatorami są instytucje dysponujące pieniędzmi publicznymi, które przed podatnikami muszą wykazać się efektywnością wydawania tych funduszy. Otwieranie i upowszechnianie wyników badań nie pozostaje bez wpływu na ocenę społeczną zarówno instytucji dysponującej pieniędzmi, jak i samych uczonych.

Pośrednie korzyści z otwierania zasobów nauki, to także wspomniane wyżej wzrost innowacyjności, oszczędności w systemie obrotu wiedzą oraz widoczność stopy zwrotu inwestycji poniesionych na naukę.

7. Praktyczne aspekty otwierania nauki

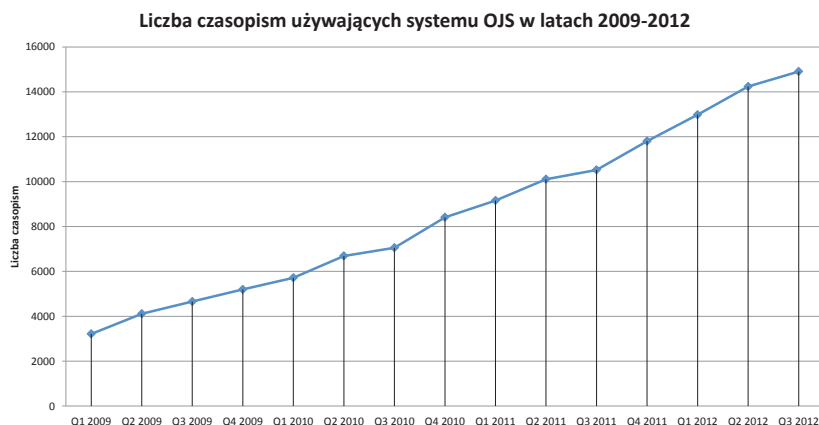
Podstawową kwestią, jeśli chodzi o praktyczne aspekty otwierania nauki jest założenie, że otwieranie jest długim procesem, a nie pojedynczym krokiem. Historia otwartego dostępu do literatury naukowej pokazuje, że mimo oczywistych podstaw (swobodny dostęp do tekstów) wielokrotnie modyfikowano szczegółowe warunki zarówno pod naciskiem lobby wydawniczego, jak i pod kątem nowych technologii, które pojawiły się w tym czasie. Dlatego w przypadku otwierania nauki najważniejsze jest określenie drogi (opracowanie strategii, nakreślenie planu, kierunku), nie ostatecznego celu, gdyż ten ostatni będzie się zmieniał w trakcie całego procesu.

7.1. Infrastruktura otwierania nauki

Otwarte zasoby naukowe nie wiszą w próżni — potrzebują infrastruktury (otoczenia technicznego i prawnego), która zapewni im maksymalne możliwe wykorzystanie.

W przypadku infrastruktury technicznej, otwieranie publikacji naukowych jest stosunkowo proste, ponieważ istnieje wiele dobrych i darmowych systemów repozytoryjnych, a ich utrzymanie nie jest kosztowne. Przykładami systemów repozytoryjnych są DSpace (1500 wdrożeń na świecie), EPrints

(500 wdrożeń) czy Fedora (51 wdrożeń)²³. Istnieją też otwarte systemy dla wydawców czasopism. Dla przykładu Open Journal Systems (OJS) zbudowany przez Public Knowledge Project jest zaimplementowany do ponad 11 500 czasopism naukowych²⁴ (w Polsce także w Polskiej Akademii Nauk, na Uniwersytecie Mikołaja Kopernika w Toruniu czy w Dolnośląskiej Szkole Wyższej, która zresztą podjęła się tłumaczenia systemu na język polski i w roku 2012 udostępniła go dla naszego środowiska naukowego za darmo).



Rys. 4. Wdrożenia OJS w czasie. Źródło: OJS User Numbers. <http://pkp.sfu.ca/>

W przypadku danych, sama ich ilość powoduje konieczność znaczących nakładów zarówno na oprogramowanie (dostosowanie istniejącego opro-

23 Więcej informacji o systemach repozytoryjnych można znaleźć na stronach ROAR lub DOAR.

24 Dane pochodzą ze stron OJS. [Dokument elektroniczny]. [Dostęp 30 lipca 2013]. Dostępny w: <http://pkp.sfu.ca/>.

gramowania do specyficznego rodzaju danych), jak i fizyczną infrastrukturę dostępu. Koszty te są niebagatelne i nawet w dobrze finansowanych instytucjach obserwuje się zamykanie otwartych baz danych²⁵.

Jeśli chodzi o infrastrukturę techniczną, dużym ułatwieniem dla indywidualnych uczonych są publiczne serwisy służące udostępnianiu zasobów nauki (lub ogólne serwisy służące udostępnianiu zasobu określonego rodzaju, np. kodu źródłowego). Z ciekawszych propozycji wymienić należy Figshare (udostępnianie danych cząstkowych, obrazków, wyników negatywnych), GitHub (udostępnianie kodu źródłowego oparte na systemie kontroli wersji Git), Academia.edu (budowanie publicznego profilu naukowca), czy Annotum (system tworzenia czasopism dostępny m.in. w ramach platformy Wordpress.com). W Polsce od niedawna funkcjonuje też ogólnodostępne repozytorium artykułów, prac dyplomowych i doktoratów CEON zbudowane przez ICM z Uniwersytetu Warszawskiego. Jest ono stworzone dla tych naukowców, którzy nie mają własnego repozytorium instytucjonalnego, a chcą lub muszą (wymóg fundatora) umieścić swoją pracę w otwartym Internecie.

7.2. Aspekt prawny otwierania zasobów

Tradycyjnie, istotną kwestią dla otwierania zasobów naukowych są kwestie legislacyjne. Niestety, międzynarodowa infrastruktura prawna dla otwartej nauki wciąż jest jeszcze w budowie.

Stosunkowo klarowna sytuacja dotyczy treści np. tekstów publikacji naukowych. Licencje Creative Commons (wraz z ich narodowymi wersjami) wprowadzają ład i kompatybilność pomiędzy prawodawstwami różnych krajów. Problemem są jedynie klauzule restrykcyjne (np. tylko użycie nieko-

25 Closure of the NCBI SRA and implications for the long-term future of genomics data storage, *Genome Biology* 2011, 12:402.

mercyjne, zakaz tworzenia utworów pochodnych, konieczność udostępniania materiału przetworzonego na tych samych warunkach co oryginał), które powodują, że nie wszystkie materiały można wykorzystywać w tym samym dziele pochodnym. W kontekście publikacji dużym ułatwieniem jest fakt, że większość wydawców bez żadnych dodatkowych formalności i umów zgadza się na archiwizację artykułów przez autorów na ich stronach domowych lub w repozytoriach instytucjonalnych²⁶ (niekoniecznie oznacza to zgodę na licencjonowanie na dowolnych warunkach, natomiast czyni tekst dostępnym chociażby do przeczytania, bez dodatkowych obciążeń kwestiami prawnymi). Po szczegółową analizę prawnych aspektów otwartego dostępu do publikacji naukowych z polskiej perspektywy warto sięgnąć do książki Krzysztofa Siewicza zatytułowanej „Otwarty dostęp do publikacji naukowych. Kwestie prawne” dostępnej w repozytorium ICM UW: <http://ceon.pl/pl/otwarta-nauka/kwestie-prawne>.

W przypadku surowych danych badawczych nie ma satysfakcjonujących rozwiązań prawnych, ale proces ich powszechnego otwierania jest stosunkowo nowy i dyskusje na temat szczegółów dopiero niedawno się rozpoczęły. Do pogodzenia są np. istotne różnice w prawodawstwie po obu stronach Atlantyku dotyczące stosowalności prawa autorskiego do danych.

Niestety (paradoksalnie) najgorzej jest w przypadku oprogramowania. Wprawdzie otwarte licencje na kod źródłowy programów są popularne i uczeni chętnie z nich korzystają, to jednak większość z nich jest wzajemnie niekompatybilna, wliczając w to nawet dwie kolejne wersje tej samej licencji (np. GPL v2 i GPL v3). Oznacza to, że spora część dostępnego oprogramowania (w tym biblioteki) nie może być swobodnie (tj. w sposób, który nie ogranicza ponownej dystrybucji) włączana do nowo tworzonych programów,

26 Około 68% wydawców na świecie pozwala na samo-archiwizację w pewnej postaci. Można to prześledzić w bazie Sherpa/Romeo. [Dokument elektroniczny]. [Dostęp 30 lipca 2013]. Dostępny w: <http://www.sherpa.ac.uk/romeo/statistics.php?la=en&flDnum=&mode=simple>. To przybliżenie potwierdzają wąskie analizy dziedzinowe takie, jak w tym artykule: <http://library.queensu.ca/ojs/index.php/IEE/article/view/4555>.

niezależnie od tego, czy czyni to uczoney w trakcie prowadzenia badań, czy też firma dla celów zarobkowych.

8. Podsumowanie

Otwarta nauka to długi proces reformowania systemu funkcjonowania nauki tak, aby z jednej strony wrócić do etycznych jej korzeni, a z drugiej strony, aby pozbyć się technicznych i finansowych barier w swobodnym jej rozwoju. Regulacje i eksperymenty na tym polu kształtować będą system nauki przez długie lata.

Z punktu widzenia pojedynczego badacza, zwłaszcza w takim kraju jak Polska, który nie jest w centrum światowej nauki, otwartość zasobów nauki bezpośrednio wpływa na widoczność i rozpoznawalność ich autora. Może mieć to olbrzymie znaczenie dla uczonych usiłujących przebić się ze swoimi pomysłami do świadomości mainstreamu naukowego.

Natomiast z perspektywy funkcjonowania całego systemu, zwiększenie transparentności funkcjonowania nauki ma szansę pomóc w odbudowaniu społecznego zaufania do uczonych, które zostało nadszarpnięte przez oczywiste przykłady nieuczciwości naukowej (tych nigdy nie brakowało, niestety) i przez nieścisłości w metodologii wyolbrzymione przez media do rozmiarów skandalu²⁷.

27 Patrz afera Climategate w Wikipedii. [Dokument elektroniczny]. [Dostęp 30 lipca 2013]. Dostępny w: pl.wikipedia.org/wiki/Climategate.

9. Przykładowa strategia otwierania nauki — wersja dla badacza

Założenia: badacz chciałby skorzystać bezpośrednio na procesie otwierania nauki, do tej pory cały jego dorobek dostępny jest w zamkniętym obiegu. Żadna z przedstawionych rekomendacji nie wpływa bezpośrednio na proces badawczy oraz nie zmienia przyzwyczajzeń dotyczących publikowania prac naukowych.

Co należy zrobić?

1. Zidentyfikować „liderów otwartości” tj. osoby lub instytucje, które otwierają zasoby nauki. Określić dobre praktyki, bariery, itp. oraz określić zakres otwierania zasobów nauki, który jak wspomniano w opracowaniu, bywa bardzo specyficzny w zależności od dziedziny.
2. Otworzyć i upowszechnić własne publikacje poprzez umieszczenie ich w instytucjonalnym lub dziedzinowym repozytorium, ewentualnie na własnej stronie domowej. Możliwość i warunki udostępnienia należy skonsultować z bazą Sherpa/Romeo (która zawiera polityki wydawców wobec modelu Open Access).
3. W przypadku braku strony WWW poświęconej tematyce badawczej konkretnej osoby na stronach instytucji/pracodawcy, stworzyć własną stronę domową podsumowującą dotychczasowe osiągnięcia naukowe (popularnym wyborem jest użycie platformy Wordpress). Umieścić tam listę publikacji z odnośnikami do pełnych tekstów.
4. Używać i budować otwarte oprogramowanie.
5. Używać sieci społecznościowych, repozytoriów dokumentów do dystrybuowania i przechowywania innych materiałów niż publikacje (np. materiały dydaktyczne).

6. Używać licencji Creative Commons do wszystkich treści publikowanych w Internecie (dla korzyści z poinformowania o zakresie praw dla autorów i użytkowników).
7. Wspierać tworzenie otwartych treści w Internecie (Wikipedia, Scholarpedia, blogi naukowe).
8. Wspierać alternatywne metryki oceny wykorzystania zasobów nauki (na poziomie systemowym, nie do ewaluacji pracy badacza).

Bibliografia

1. Zbiór zasad i wytycznych, pt. „Dobre obyczaje w nauce” z przedmową prof. Kornela Gibińskiego. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://www.ken.pan.pl/images/stories/pliki/pdf/down.pdf>.
2. Open source/open science. Międzynarodowa konferencja Brookhaven National Laboratory. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://openscience.bnl.gov/>.
3. HOOKER, Bill, 3quarksdaily: The Future of Science is Open, Part 1: Open Access. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: http://3quarksdaily.blogs.com/3quarksdaily/2006/10/the_future_of_s_1.html.
4. ROSSNER, M., EPPS, VAN E., and HILL, E., Show me the data. *Journal of Cell Biology* 2007, 179:1091-1092.
5. The Impact Factor Game. *PLoS Medicine* 2006, 3(6): e291.
6. The Cost of Knowledge. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://www.thecostofknowledge.com>
7. SOMMER J., The delay in sharing research data is costing lives. *Nat Med* 2010, July 16(7):744.
8. ROARMAP: Registry of Open Access Repositories Mandatory Archiving Policies [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://roarmap.eprints.org>
9. Memorandum for the heads of executive departments and agencies. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.

10. RCUK Policy on Open Access. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://www.rcuk.ac.uk/research/Pages/outputs.aspx>.
11. HOUGHTON J., Swan, A., Planting the Green Seeds for a Golden Harvest: Comments and Clarifications on “Going for Gold”. *D-Lib Magazine* 2013, 19:1/2.
12. Scientific data: open access to research results will boost Europe’s innovation capacity. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: http://europa.eu/rapid/press-release_IP-12-790_en.htm.
13. Reaffirmation and Extension of NHGRI Rapid Data Release Policies: Large-scale Sequencing and Other Community Resource Projects. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://www.genome.gov/10506537>.
14. Data shown. *Nat Chem Biol*, 2008, 4:575.
15. Creative Commons Biological Material Transfer Project [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://creativecommons.org/projects/licensing/>.
16. Big data: The next frontier for innovation, competition, and productivity [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation.
17. BAKER, S. M., Description and Initial Evaluation of a Text Message Based Reporting Method for Marine Recreational Anglers. *Mar Coast Fish*. 2009, 1:143-154.
18. StackOverflow. All questions. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://stackoverflow.com/questions>.
19. Polymath1. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://michaelnielsen.org/polymath1/index.php?title=Polymath1>
20. Fold it players[au] — PubMed — NCBI. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://www.ncbi.nlm.nih.gov/pubmed/?term=foldit+players%5Bau%5D>.
21. GYMREK, M., McGUIRE, A. L., GOLAN, D., HALPERIN E., ERLICH Y., Identifying Personal Genomes by Surname Inference. *Science* 2013, (339) 6117, 321-324.

22. BOHANNON, J., Genealogy Databases Enable Naming of Anonymous DNA Donors. *Science* 2013, (339) 6117-6262.
23. HALL, R. J., MILNER-GULLAND E. J., COURCHAMP, F., Endangering the endangered: The effects of perceived rarity on species exploitation. *Conservation Letters* 2008 1:75–81.
24. HERFST, S. et al, Airborne Transmission of Influenza A/H5N1 Virus Between Ferrets. *Science* 2012, (336) 6088-1534-1541.
25. Kalendarium działań na rzecz open access (otwartej nauki) w Polsce. Pod redakcją Bożeny Bednarek-Michalskiej. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: http://www.ebib.pl/?page_id=887.
26. DOAJ: Directory of Open Access Journals. Browse by Country. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://www.doaj.org/doaj?func=byCountry&uiLanguage=en>.
27. Federacja Bibliotek Cyfrowych. Zestawienie polskich bibliotek cyfrowych. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://fbc.pionier.net.pl/owoc/oai-hosts>.
28. Stanowisko Prezydium KRASP i Prezydium PAN z dnia 5 lipca 2013 r. w sprawie zasad otwartego dostępu do treści publikacji naukowych i edukacyjnych. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: http://www.aktualnosci.pan.pl/images/stories/pliki/2013/Wiadomo%C5%9Bci_biez/stanowisko_KRASP-PAN_open_access.pdf.
29. Portal edukacyjny Uniwersytetu Mikołaja Kopernika. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://portal.umk.pl/web/otwarte-zasoby>.
30. Otwarte Zasoby. Pod redakcją Kamila Śliwowskiego. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://otwartzasoby.pl/>.
31. Closure of the NCBI SRA and implications for the long-term future of genomics data storage, *Genome Biology* 2011, 12:402.
32. RoMEO Statistics. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://www.sherpa.ac.uk/romeo/statistics.php?la=en&fIDnum=1&mode=simple>.
33. HASSALL, C., „Going Green”: self-archiving as a means for dissemination of research output in ecology and evolution. *IEE* 2012, 5(2).

34. Climategate — Wikipedia, wolna encyklopedia. [Dokument elektroniczny]. [Dostęp 23 lipca 2013]. Dostępny w: <http://pl.wikipedia.org/wiki/Climategate>.

Najważniejsze publikacje i kursy polskie dotyczące otwartej nauki:

1. HOFMOKL, J., TARKOWSKI, A., BEDNAREK-MICHALSKA, B., i in. Przewodnik po otwartej nauce, ICM UW Warszawa 2009. [Dostęp 22.07.2013]. Dostępny w: <http://otwartanauka.pl/przewodnik-po-otwartej-nauce/>.
2. BEDNAREK-MICHALSKA, B., TARKOWSKI, A., GANICZ, T., GRODECKA, K., Rekomendacje Koalicji Otwartej Edukacji dla nauki polskiej, dotyczące reguł otwartości, Warszawa KOED 2010. [Dostęp 22.07.2013]. Dostępny w: <http://koed.org.pl/blog/2010/10/18/rekomendacje-otwartosci-dla-swiata-nauki/>.
3. NIEZGÓDKA, Marek M. i in. Wdrożenie i promocja otwartego dostępu do treści naukowych i edukacyjnych. Praktyki światowe a specyfika polska. Przewidywane koszty, narzędzia, zalety i wady, Warszawa: ICM [UW], 2011. [Dostęp 22.07.2013]. Raport dla MNiSW. Dostępny w: <https://depot.ceon.pl/handle/123456789/1545>.
4. Kurs e-learningowy Otwarta Nauka pod red. Bożeny Bednarek-Michalskiej i Karoliny Grodeckiej. Toruń, Kraków 2011. [Dostęp 22.07.2013]. Dostępny w: <http://otwartanauka.cel.agh.edu.pl/index.php?id=1&theme=1024>.
5. SIEWICZ, K., „Otwarty dostęp do publikacji naukowych. Kwestie prawne”. Wydawnictwo Uniwersytetu Warszawskiego, Warszawa 2012. ISBN 978-83-235-0959-2. [Dostęp 22.07.2013]. Dostępny w: <http://ceon.pl/pl/otwarta-nauka/kwestie-prawne>.

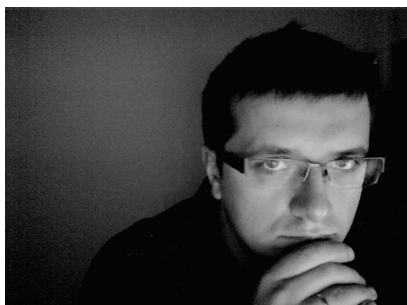
Spis ilustracji

Okładka – CC BY-SA 2.0 – <http://creativecommons.org/licenses/by-sa/2.0/> – Autor: SanFranAnnie

Strona 10 – CC BY 2.0 – <http://creativecommons.org/licenses/by/2.0/> – Autor: U.S. Fish and Wildlife Service - Midwest Regio

Strona 17 – CC BY-SA 2.0 – <http://creativecommons.org/licenses/by-sa/2.0/> – Autor: poptech

Biogram autora



Paweł Szczęsny – doktor biologii, bioinformatyk zatrudniony w Instytucie Biochemii i Biofizyki Polskiej Akademii Nauk i na Uniwersytecie Warszawskim. W wolnych od pracy badawczej chwilach zajmuje się kwestiami otwartości w nauce oraz rolą internetu w komunikacji naukowej.

Prowadzi warsztaty i wykłady dotyczące komunikacji naukowej i otwartej nauki zarówno w kraju, jak i za granicą. Współpracuje z międzynarodowymi instytucjami (Instytut Społeczeństwa Otwartego, Fundacja Otwartej Wiedzy, EIFL, UNESCO) w zakresie strategii rozwoju i promocji otwartej nauki. Jest członkiem Komisji ds. Open Access Polskiej Akademii Nauk.

Zarządza też pionem naukowym fundacji Instytut Systemów, która to instytucja wspiera organizacyjnie ruch Obywatele Nauki oraz buduje m.in. infrastrukturę techniczną dla otwartej nauki.

Zapoznaj się z innymi naszymi broszurami poświęconymi Otwartej Nauce



Emanuel Kulczycki

Otwarte czasopisma. Zakładanie czasopism naukowych oraz transformacja czasopism zamkniętych

Bożena Bednarek-Michalska

Modele biznesowe otwartego publikowania naukowego. Informator dla polskich wydawców uczelnianych

Karolina Grodecka

Studium przypadku: udane projekty Open Access w Polsce

Wszystkie broszury dostępne na stronie wydawcy:

<http://ebib.pl/>