



MORSKI
INSTYTUT
RYBACKI
PAŃSTWOWY
INSTYTUT
BADAWCZY



Elektroniczne Centrum Udostępniania Danych Oceanograficznych eCUDO.pl

platforma do gromadzenia i udostępniania danych
oceanograficznych

Marcin Wichorowski

Mirosława Ostrowska

Instytut Oceanologii Polskiej Akademii Nauk

II Sat4ENVI - Dane satelitarne w monitorowaniu środowiska Polski

Warszawa, 19-20 listopada, 2019



Fundusze
Europejskie
Polska Cyfrowa

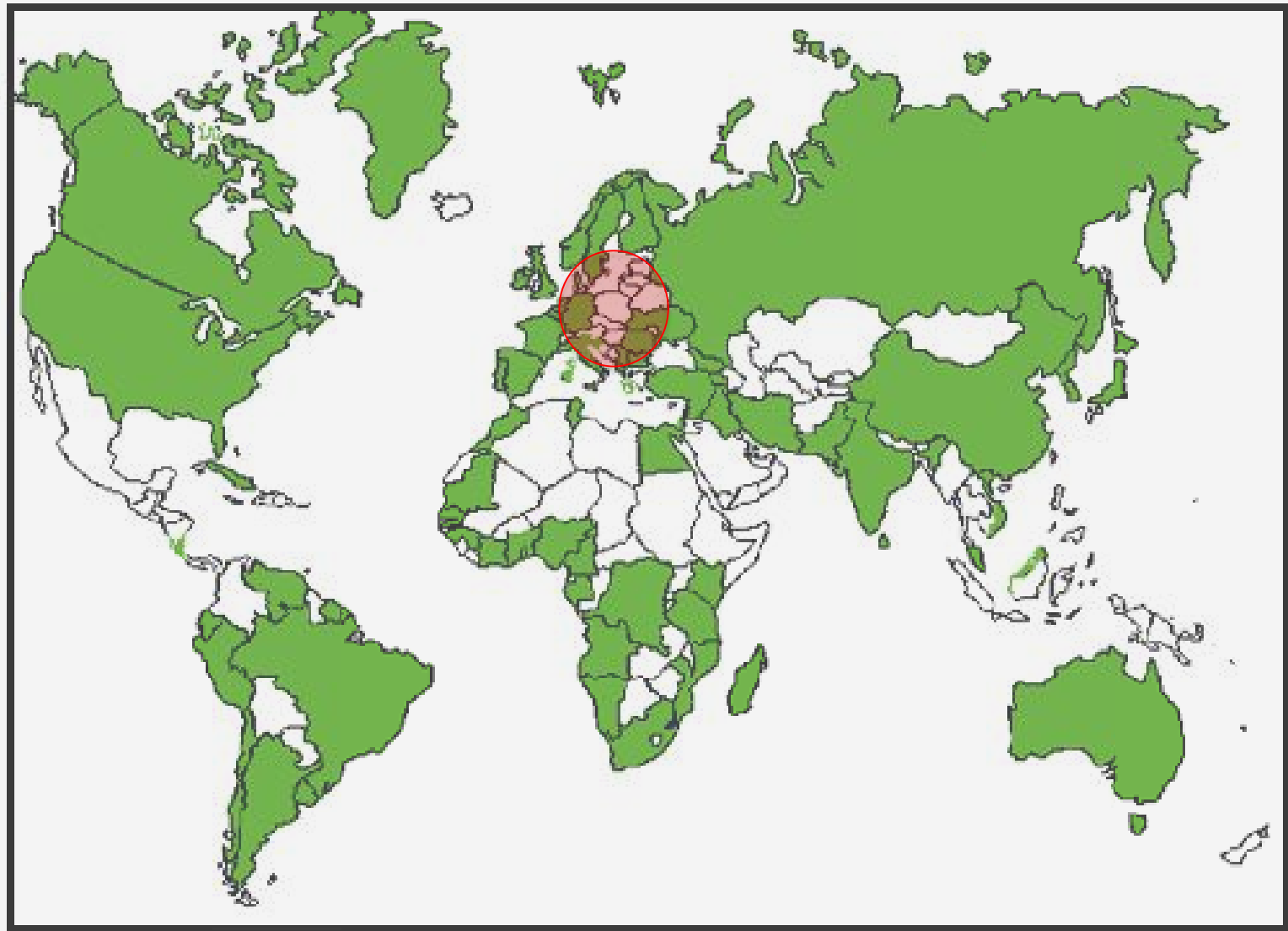


Rzeczpospolita
Polska

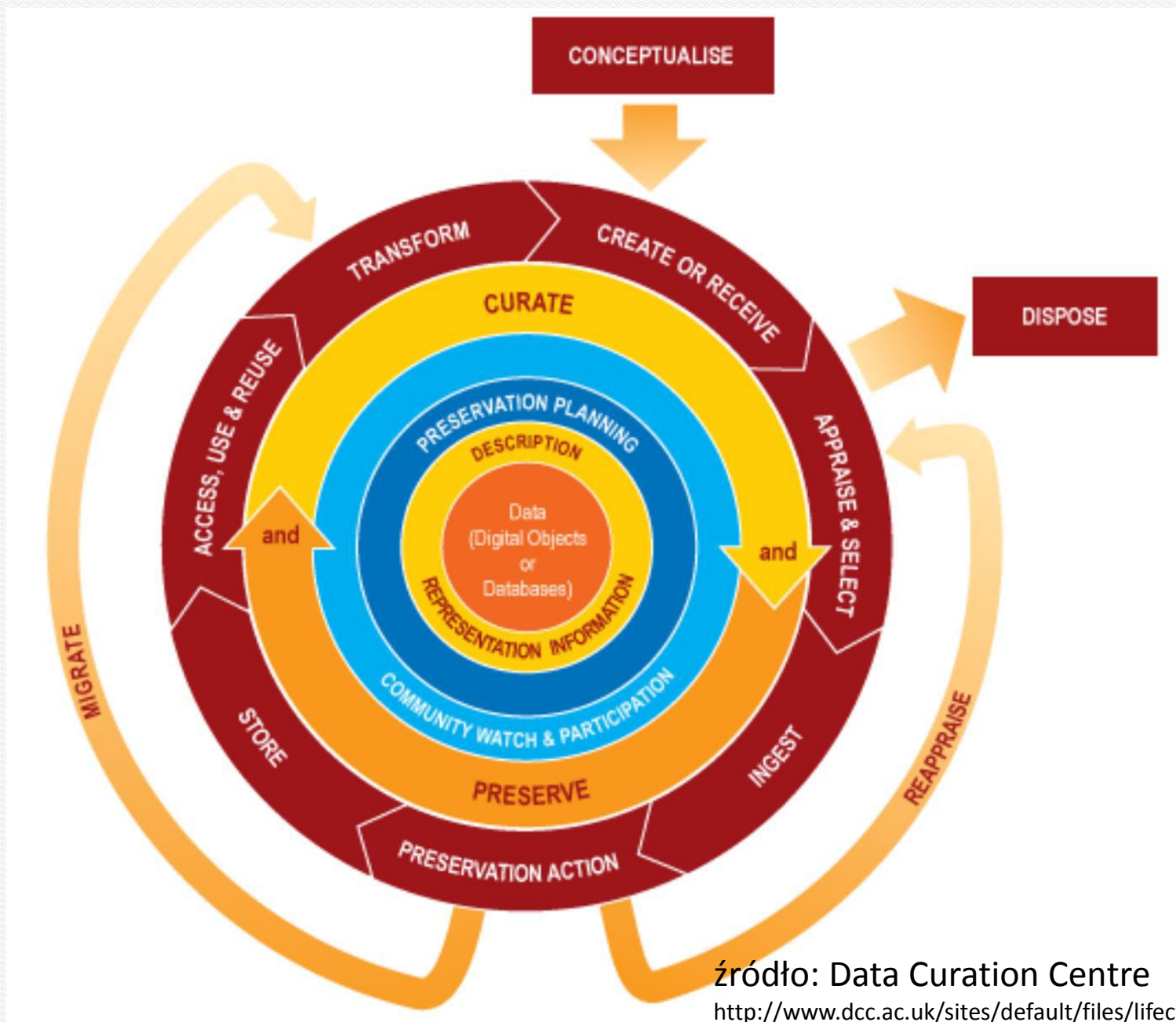


Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego





- Polska jest jedynym morskim krajem Unii Europejskiej i jednym z nielicznych na świecie, w którym nie zorganizowano jeszcze struktur pełniących funkcję Narodowego Centrum Danych Oceanograficznych (ang. National Oceanographic Data Centre, NODC),
- Brak przepływu informacji o prowadzonych badaniach i gromadzonych danych pomiędzy instytucjami zajmującymi się badaniami i eksploracją środowiska morskiego, finansowanymi niejednokrotnie z publicznych środków,
- Degradacja zgromadzonych zasobów danych,
- Ryzyko duplikowania prac badawczych,
- Utrudnienie dostępu do danych i informacji,
- Niewykorzystywanie potencjału informacyjnego gromadzonych danych, również na arenie międzynarodowej.



źródło: Data Curation Centre

http://www.dcc.ac.uk/sites/default/files/lifecycle_web.png

- **Findability**

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

- **Accessability**

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
 - A1.1 the protocol is open, free, and universally implementable
 - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

- **Interoperability**

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

- **Reuse**

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
 - R1.1. (meta)data are released with a clear and accessible data usage license
 - R1.2. (meta)data are associated with detailed provenance
 - R1.3. (meta)data meet domain-relevant community standards

źródło: **The FAIR Guiding Principles for scientific data management and stewardship**
Mark D. Wilkinson et. al, *Nature, Scientific Data* **volume 3**, Article number: 160018 (2016)

- **Findability**

udostępnienie danych jako “linked data” wraz z powiązanymi metadanymi propagowanymi do innych systemów (SeaDataCloud, EMODNet, BOOS)

- **Accessibility**

rozwój interfejsów M2M przy użyciu otwartych protokołów komunikacyjnych i formatów danych, dobrze udokumentowane API, konstrukcja systemu gwarantująca odporność na awarie i stałą dostępność danych i metadanych

- **Interoperability**

unikalne identyfikatory zestawów danych, standardy metadanych oparte o słowniki metainformacji kontrolowane przez SeaDataCloud (BODC), formaty danych i metadane oparte o standardy powszechnie używane w badaniach oceanograficznych (INSPIRE, SeaDataCloud CDI, Darwin Core, SeaBASS, NetCDF, GRIB)

- **Reuse**

bezpłatne udostępnianie metadanych i danych w oparciu o otwarte licencje z rodziny Creative Commons



Instytut Morski w Gdańsku



Państwowy Instytut Geologiczny
Państwowy Instytut Badawczy



Morski Instytut Rybacki
Państwowy Instytut Badawczy



Instytut Oceanologii
Polskiej Akademii Nauk



UNIwersytet GDAŃSKI

Uniwersytet Gdański



Akademia Pomorska w Słupsku



Uniwersytet Szczeciński



„Elektroniczne Centrum Udostępniania Danych Oceanograficznych eCUDO.pl”

Poddziałanie 2.3.1 Cyfrowe udostępnienie informacji sektora publicznego (ISP) ze źródeł administracyjnych i zasobów nauki (typ II projektu: cyfrowe udostępnienie zasobów nauki)

Działanie 2.3. Cyfrowa dostępność i użyteczność informacji sektora publicznego
II osi priorytetowej E-administracja i otwarty rząd

Program Operacyjny Polska Cyfrowa na lata 2014-2020

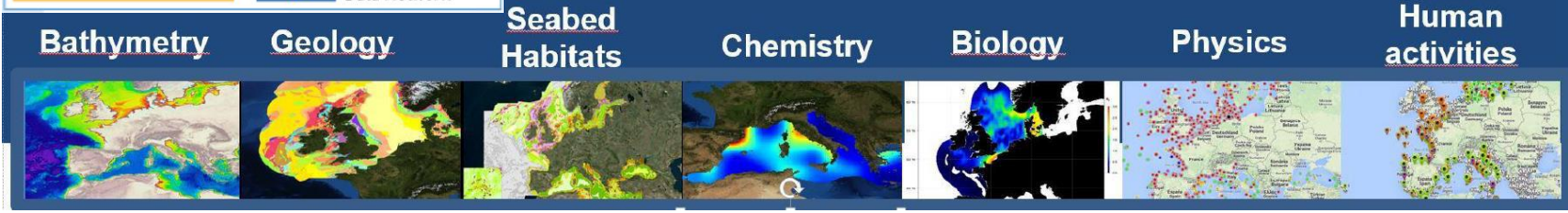
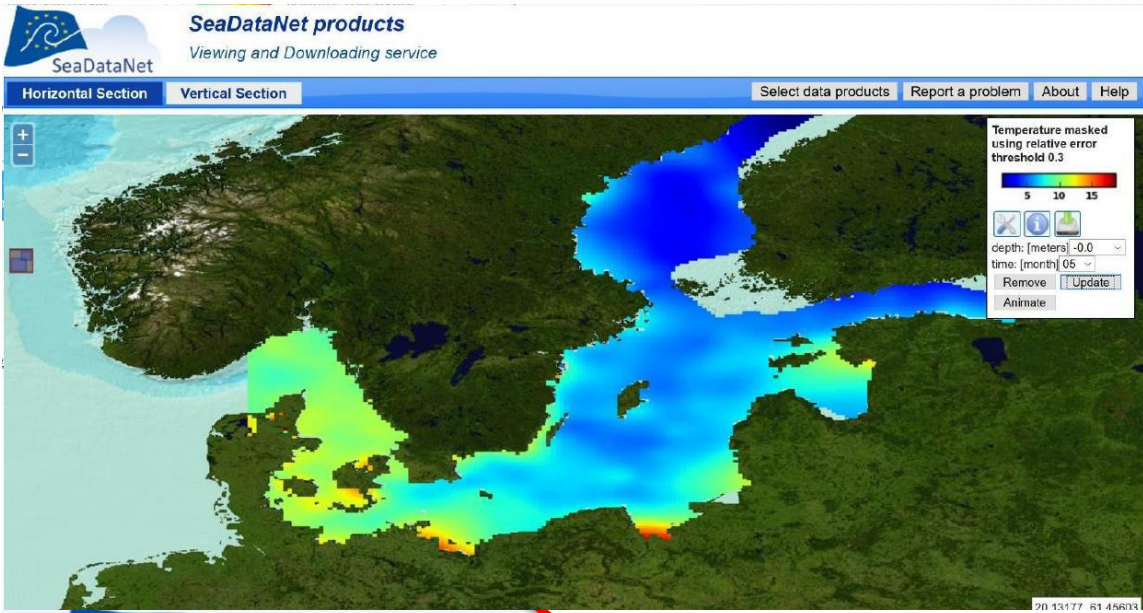
Wartość projektu: **15 261 546, 00 zł** (84,63% –ERDF, 15,37% środki krajowe)

- **zapewnienie zunifikowanego dostępu do danych i usług oraz opieka nad krajowymi zasobami danych oceanograficznych,**
- pozyskiwanie danych i informacji o danych oceanograficznych z realizowanych przez polskie instytucje projektów i programów badawczych, krajowych i międzynarodowych,
- weryfikacja i oznaczanie jakości danych na podstawie międzynarodowych standardów używanych przez organizacje zaangażowane w eksplorację zasobów morskich,
- zapewnienie długoterminowej archiwizacji danych wraz z informacjami niezbędnymi dla ich prawidłowej interpretacji,
- rozwijanie usług udostępniania informacji o danych oraz danych użytkownikom krajowym i międzynarodowym, rozwój narzędzi wizualizacji informacji
- pełnienie roli centrum przetwarzania danych o środowisku morskim dla międzynarodowych eksperymentów badawczych

Celem głównym Projektu jest zwiększenie cyfrowej dostępności i użyteczności informacji sektora publicznego dzięki sprowadzeniu zasobów nauki z zakresu oceanologii do ich postaci cyfrowej poprzez ich utrwalenie w postaci elektronicznej i opatrzenie metadanymi a także udostępnienie w Internecie, w formatach dostosowanych do przetwarzania maszynowego.

Odpowiada on na zidentyfikowany w toku analiz problem główny, jakim jest blokada społeczno-gospodarczego potencjału (użyteczności) zasobów nauki z zakresu oceanologii spowodowana brakiem ich pełnej cyfrowej dostępności.

Zidentyfikowane grupy użytkowników: administracja rządowa i samorządowa wraz z podległymi im jednostkami, jednostki naukowo-badawcze, przedsiębiorcy, organizacje pozarządowe i osoby prywatne



- Brak ogólnego systemu zarządzania i gromadzenia metadanych (część metadanych jest już w strukturach bazodanowych opisujących m.in. pobór prób)
- Słowniki, struktury danych i metadanych wymagają harmonizacji między instytutami
- Część danych znajdujących się w zasobach partnerów wymaga digitalizacji (sprowadzenie do postaci cyfrowej, powiązanie z metadanymi)
- Aktualny status danych wymaga weryfikacji prawnej
- Zbudowanie mechanizmów udostępniania danych i metadanych w ramach wspólnej platformy
- Udostępnianie części przetworzonych danych w formie map (serwisy mapowe, WxS)

Harmonogram realizacji

Kamienie milowe	Data rozpoczęcia	Data zakończenia	Data punktu krytycznego	Data punktu ostatecznego
Projekt systemu	01.03.2019	30.08.2019	27.09.2019	10.12.2019
Prace nad prototypem systemu eCUDO.pl	01.03.2019	26.02.2020	18.04.2020	27.01.2021
Dostawy sprzętu do digitalizacji danych	27.04.2019	06.10.2020	27.06.2020	30.03.2021
Testy walidacyjne systemów sfederowanych	01.03.2019	26.02.2021	30.03.2021	14.04.2021
Testy walidacyjne systemu centralnego	01.03.2019	28.04.2021	05.12.2021	17.05.2021
Zakończenie testów integracyjnych systemu eCUDO.pl	01.03.2021	17.08.2021	27.08.2021	27.02.2022
Testy akceptacyjne systemu eCUDO.pl i analiza działania systemu udostępnionego użytkownikom docelowym	01.03.2019	29.05.2021	27.08.2021	27.09.2021
Wdrożenie eksploatacyjnej instancji systemu eCUDO.pl	01.03.2019	28.09.2021	27.01.2022	27.02.2022

- 7 instytucji naukowych i badawczych
- 15 261 546, 00 zł (84,63% –ERDF, 15,37% środki krajowe)
- prawie 4 000 000 dokumentów, które zostaną zdigitalizowane i udostępnione
- prawie 500TB danych
- wdrożenie 5 * Open Data Model i polityki FAIR
- interfejs zgodny z WCAG 2.0
- uruchomienie interfejsów M2M do wymiany danych i informacji
- dostarczenie systemu: koniec roku 2021



Elektroniczne Centrum Udostępniania Danych Oceanograficznych

Beneficjent: Instytut Oceanologii Polskiej Akademii Nauk
realizuje **Projekt eCUDO.pl**
dofinansowany z Funduszy Europejskich

Celem głównym Projektu jest zwiększenie cyfrowej dostępności i użyteczności informacji sektora publicznego dzięki cyfrowemu udostępnieniu zasobów nauki z zakresu oceanologii

Okres realizacji Projektu: od 01.03.2019 do 27.02.2022

Partnerzy:



DZIĘKUJĘ



Fundusze
Europejskie
Polska Cyfrowa



Rzeczpospolita
Polska



Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego

