



**EUROPEAN OPEN
SCIENCE CLOUD**

Doświadczenia EOSC i jego przyszłość

European Open Science Cloud (EOSC)



**EUROPEAN OPEN
SCIENCE CLOUD**

Główny cel

stworzenie godnego zaufania, wirtualnego, federacyjnego internetowego środowiska, wykraczającego poza granice i dyscypliny naukowe, służącego do przechowywania, udostępniania, przetwarzania i ponownego użycia obiektów cyfrowych wykorzystywanych w badaniach naukowych zgodnie z **zasadami FAIR (ang. Findable, Accessible, Interoperable and Reusable)**

European Open Science Cloud (EOSC)



**EUROPEAN OPEN
SCIENCE CLOUD**

2015-2018

2018-2020

2021-2030

Horizon 2020

Horizon Europe (2021-2027)

Faza koncepcyjna

Faza przejściowa

Faza implementacyjna

2016:
EC: 'Open Innovation, Open Science
and Open to the World – a vision for
Europe' –

- European Science Cloud
Initiative
- EOSC Initiative

Digital Single Market in Europe

Zarząd: Executive Board
Governance Board
Stakeholder Forum

Zarząd trójstronny:
KE+ EOSC Association + MS/AC

Środki na granty

Partnerstwo współprogramowane

Mapa Drogowa EOSC

Strategic Research and Innovation
Agenda (SRIA) of the European
Open Science Cloud



- Członkowie założyciele: CESAER, GEANT, GARR, CSIC
- EOSC AISBL zarejestrowana w lipcu 2020 r. (jako międzynarodowa organizacja non-profit)
- Pierwsze posiedzenie generalne 17.12.2020 – wybór Przewodniczącego i Zarządu
- **Członkami są:** uczelnie, instytucje finansujące badania, organizacje dostarczające usług
- Członkostwo w Stowarzyszeniu – uczestnictwo w Partnerstwie
- W ramach Stowarzyszenia działa 5 Advisory Groups oraz 13 Task Forces

EOSC Association – Advisory Groups i Task Forces



Implementation of EOSC

- PID Policy and Implementation
- Researcher Engagement and Adoption
- Rules of Participation Compliance Monitoring

Technical Challenges on EOSC

- AAI Architecture
- Infrastructure for Quality Research Software
- Technical Interoperability of Data and Services

Metadata and Data Quality

- FAIR Metrics and Data Quality
- Semantic Interoperability

Research Careers and Curricula

- Data Stewardship Curricula and Career Paths
- Research Careers, Recognition, and Credit
- Upskilling Countries to Engage in EOSC

Sustaining EOSC

- Defining Funding Models for EOSC
- Long-Term Data Preservation

EOSC Association - członkowie



**EUROPEAN OPEN
SCIENCE CLOUD**

- Obecnie ok. 150 członków i 60 obserwatorów
- Przyjmowane są zgłoszenia do uczestnictwa w Stowarzyszeniu
- Członkowie z Polski



NARODOWE CENTRUM NAUKI



**Instytut Geofizyki
Polskiej Akademii Nauk**

MoU Cele ogólne i Kluczowe Wskaźniki Efektywności (przykłady)



EUROPEAN OPEN
SCIENCE CLOUD

G01

nagradzanie i uczenie praktyk oraz umiejętności związanych z Otwartą Nauką – „new normal”

Percentage of publications from EOSC Association research-performing members that become immediate open access	70% do 2023
Number of national education systems that recognise European curricula for data stewardship	5 do 2025
Percentage of RPOs that are EOSC Association members that have data stewards to support their research	50% do 2025

G02

definicja standardów, rozwój narzędzi i usług pozwalających naukowcom na znajdowanie, dostęp, wykorzystywanie i kompilację wyników

Percentage of EOSC Association members that have policies which require FAIR to be implemented in project design via Data Management Plans	70% do 2023
Percentage of research data from EOSC Association members which is deposited in repositories that is made as open as possible	50% do 2025
A first generation of pan-European federation of infrastructures for preservation, management and sharing of research software is available	1 do 2025

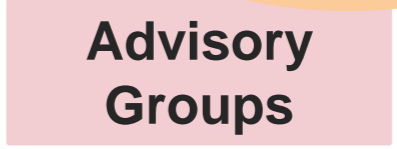
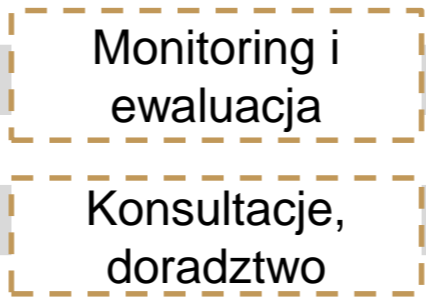
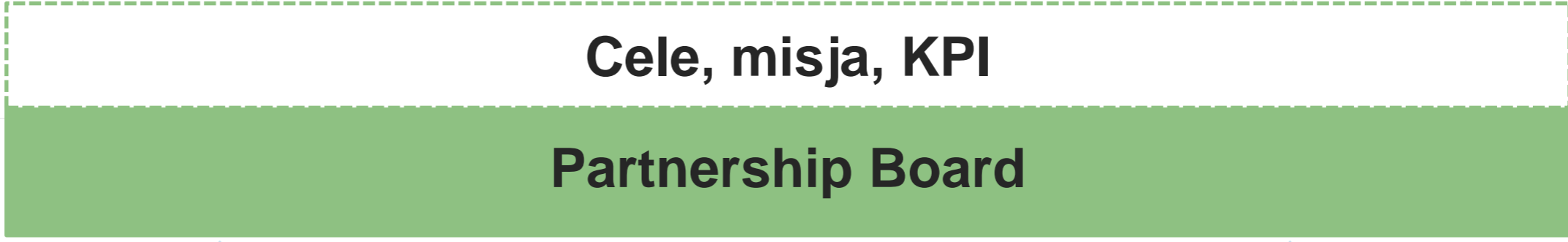




G03

utworzenie zrównoważonej i federacyjnej infrastruktury pozwalającej na otwartą wymianę rezultatów badawczych

Percentage of the repositories in EOSC that will have a certification such as CoreTrustSeal	30% do 2025
Percentage of research disciplines that have documented standards and protocols for data sharing and reuse	60% do 2023
Percentage of the metadata related to publicly funded research datasets which are defined as Open Data that are discoverable through EOSC federated infrastructure	70% do 2025





Science Europe

cOAlition S

Zarządzanie Danymi
Badawczymi

Otwarty dostęp do publikacji

EOSC Steering Board

EOSC Association

Polityka Otwartego Dostępu do Publikacji



Publikacje

z wyjątkiem: monografii, rozdziałów w monografiach i recenzowanych utworów zebranych

Dane powiązane (podstawowy zestaw danych)

Czasopisma full OA

Czasopisma
hybrydowe/subskrypcyjne
AAM lub VoR w repozytorium

Czasopisma
transformacyjne

Rights Retention
Strategy

CC-BY

CC-SA

Wiarygodne i zaufane **repozytorium**

Metadane - OpenAIRE

CC0

Plan Zarządzania Danymi / Co to jest Plan Zarządzania Danymi

- to **część planu badań**
- jego celem jest uzupełnienie planu badań o **techniczny opis** sposobu zarządzania danymi
- w wypełnieniu poniższej części formularza powinni pomóc pracownicy biblioteki/odpowiedniego repozytorium/działu IT/specjaliści ds. RODO

NCN dopuszcza, że w ramach niektórych projektów nie będą wytwarzane, na nowo wykorzystywane, ani poddawane analizie żadne dane badawcze ani inne podobne materiały. W takich wypadkach wymagane jest jednak krótkie uzasadnienie.



Plan Zarządzania Danymi / Dlaczego jest ważny?

- zabezpieczenie danych przed ich utratą
- pomaga monitorować postępy w realizacji badań
- ułatwia współpracę pomiędzy naukowcami pracującymi w danym projekcie
- wzmacnia kontrolę jakości badań
- pomaga zorganizować i skoordynować realizację badań
- zabezpieczenie wrażliwych danych np. przed ich nieuprawnionym wykorzystaniem
- wspiera organizację długotrwałego przechowywania danych



Plan Zarządzania Danymi / Ogólne cechy

- stosowany dla wszystkich konkursów i naborów
- oparty na stałym formacie (ale DMP musi być dostosowane do projektu!)
- w formie kwestionariusza
- dostępny we wniosku w systemie OSF i w ogłoszeniu konkursowym
- to dokument dynamiczny – może zmieniać się w trakcie projektu
- podlega ocenie merytorycznej





1. Opis danych oraz pozyskiwanie lub ponowne wykorzystanie dostępnych danych

Sposób pozyskiwania i opracowywania nowych danych i/lub ponownego wykorzystania dostępnych danych

Pozyskiwane lub opracowywane dane (np. rodzaj, format, ilość)

2. Dokumentacja i jakość danych

Metadane i dokumenty (np. metodologia lub pozyskiwanie danych oraz sposób porządkowania danych) towarzyszące danym

Stosowane środki kontroli jakości danych



3. Przechowywanie i tworzenie kopii zapasowych podczas badań

Przechowywanie i tworzenie kopii zapasowych danych i metadanych podczas badań

Sposób zapewnienia bezpieczeństwa danych oraz ochrony danych wrażliwych podczas badań

4. Wymogi prawne, kodeks postępowania

Sposób zapewnienia zgodności z przepisami dotyczącymi danych osobowych i bezpieczeństwa danych w przypadku przetwarzania danych osobowych

Sposób zarządzania innymi kwestiami prawnymi, np. prawami własności intelektualnej lub własnością.
Obowiązujące przepisy



5. Udostępnianie i długotrwałe przechowywanie danych

Sposób i termin udostępnienia danych. Ewentualne ograniczenia w udostępnianiu danych lub przyczyny embarga

Sposób wyboru danych przeznaczonych do przechowania oraz miejsce długotrwałego przechowywania danych (np. repozytorium lub archiwum danych)

Metody lub narzędzia programowe umożliwiające dostęp do danych i korzystanie z danych

Sposób zapewniający stosowanie unikalnego i trwałego identyfikatora (np. cyfrowego identyfikatora obiektu (DOI)) dla każdego zestawu danych



6. Zadania związane z zarządzaniem danymi oraz zasoby

Osoba (np. funkcja, stanowisko i instytucja) odpowiedzialna za zarządzanie danymi (np. data steward)

Środki (np. finansowe i czasowe) przeznaczone do zarządzania danymi i zapewnienia możliwości odnalezienia, dostępu, interoperacyjności i ponownego wykorzystania danych

Dziękuję za uwagę

Kontakt:

aneta.pazik-aybar@ncn.gov.pl

Narodowe Centrum Nauki
ul. Twardowskiego 16, 30-312 Kraków



EOSC Association – Advisory Groups i Task Forces



**EUROPEAN OPEN
SCIENCE CLOUD**

Implementation of EOSC

- PID Policy and Implementation
- Researcher Engagement and Adoption
- Rules of Participation Compliance Monitoring

Technical Challenges on EOSC

- AAI Architecture
- Infrastructure for Quality Research Software
- Technical Interoperability of Data and Services

Metadata and Data Quality

- FAIR Metrics and Data Quality
- Semantic Interoperability

Research Careers and Curricula

- Data Stewardship Curricula and Career Paths
- Research Careers, Recognition, and Credit
- Upskilling Countries to Engage in EOSC

Sustaining EOSC

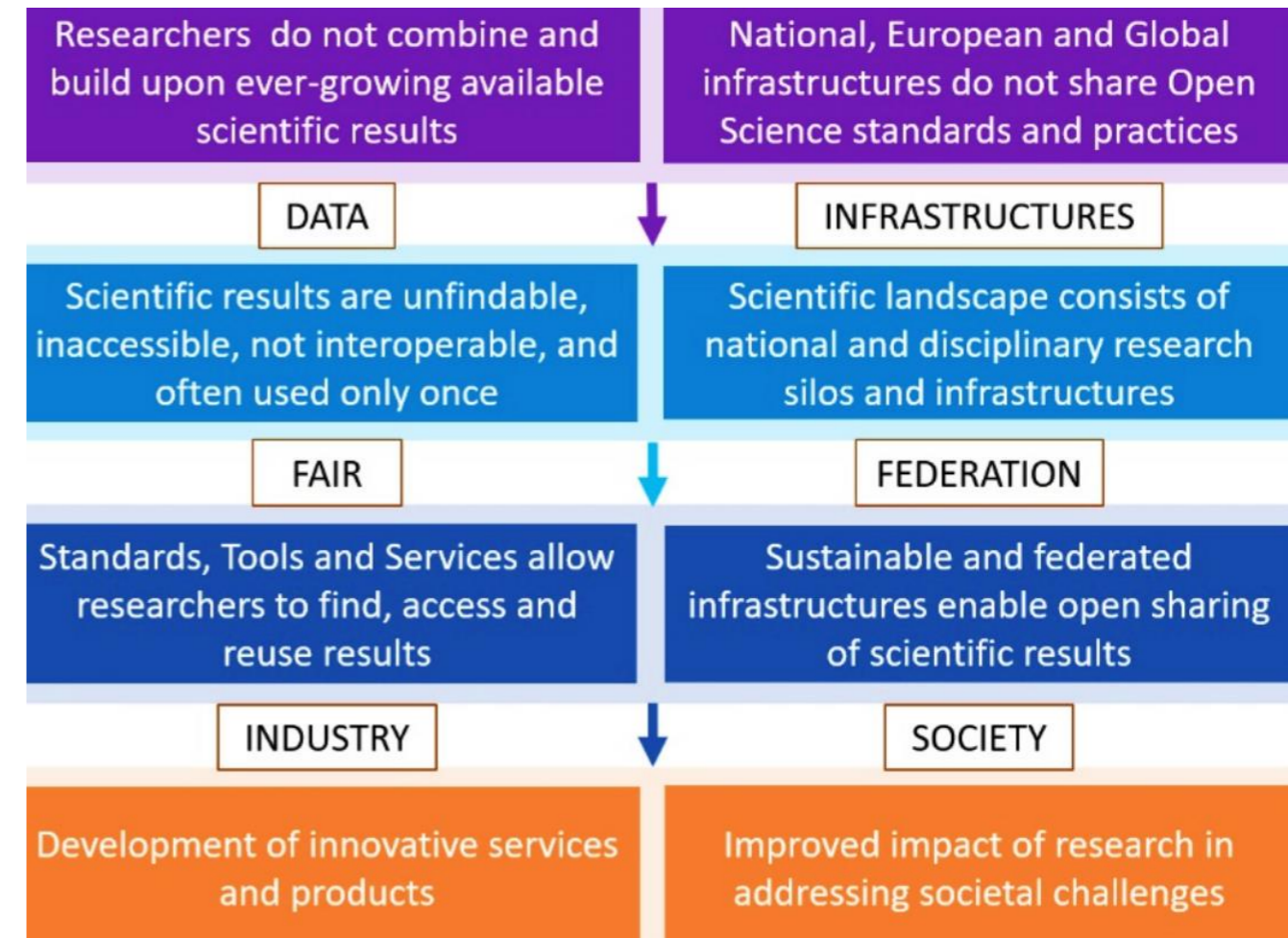
- Defining Funding Models for EOSC
- Long-Term Data Preservation

Zasady FAIR w EOSC – nowe wyzwania



**EUROPEAN OPEN
SCIENCE CLOUD**

- Zasady FAIR **nie ograniczają się już tylko** do zagadnień otwartych danych
- Zasady FAIR zostały **rozszerzone i obejmują cały cykl życia badań naukowych w EOSC**
- Zasady FAIR odnoszą się do całych, często skomplikowanych i czasochłonnych procesów badań, które **wymagają wsparcia ze strony infrastruktury informatycznej nauki**
- Niezależnie, obserwujemy **znaczący wzrost wolumenów danych** generowanych i wykorzystywanych w badaniach naukowych

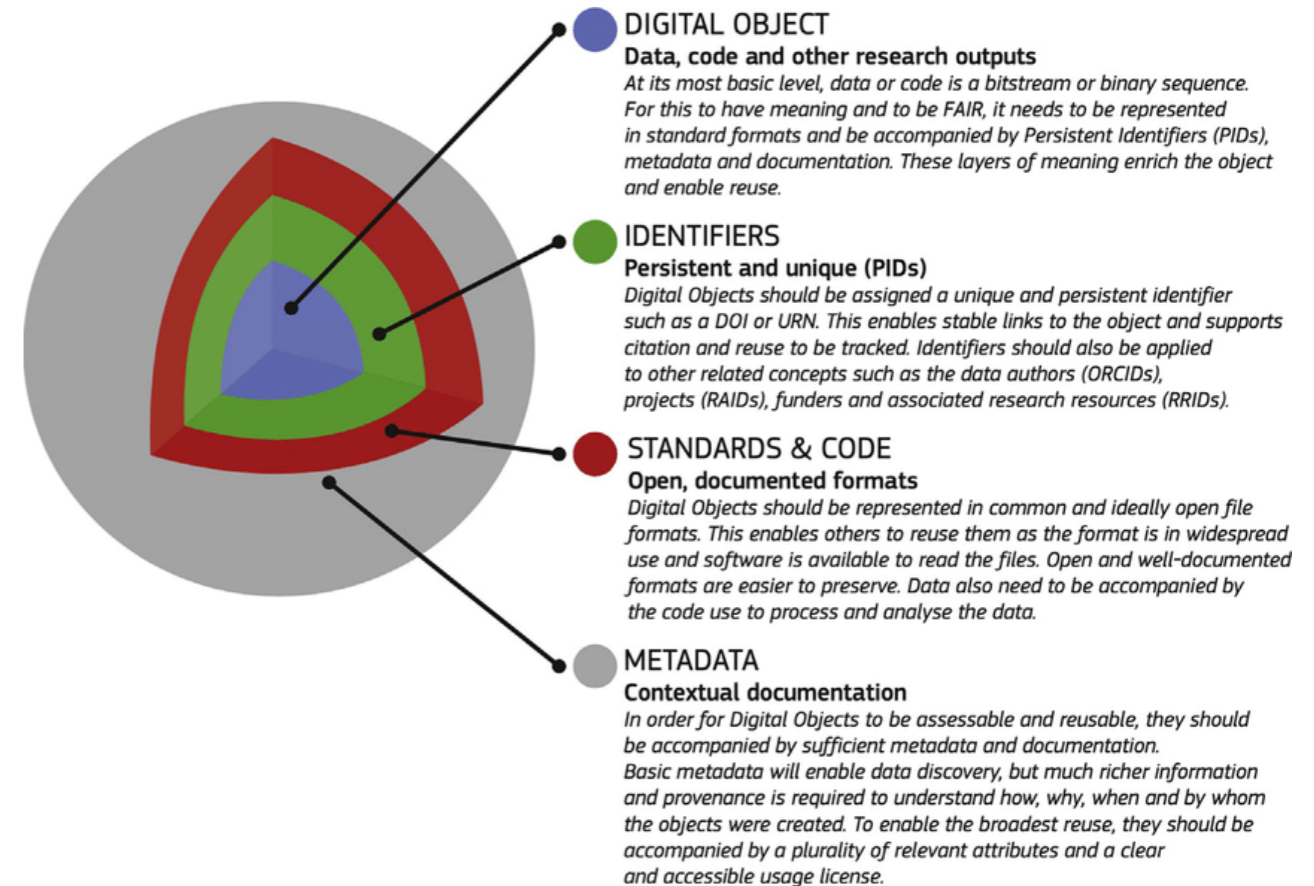


Zasady FAIR w EOSC – nowe wymagania



**EUROPEAN OPEN
SCIENCE CLOUD**

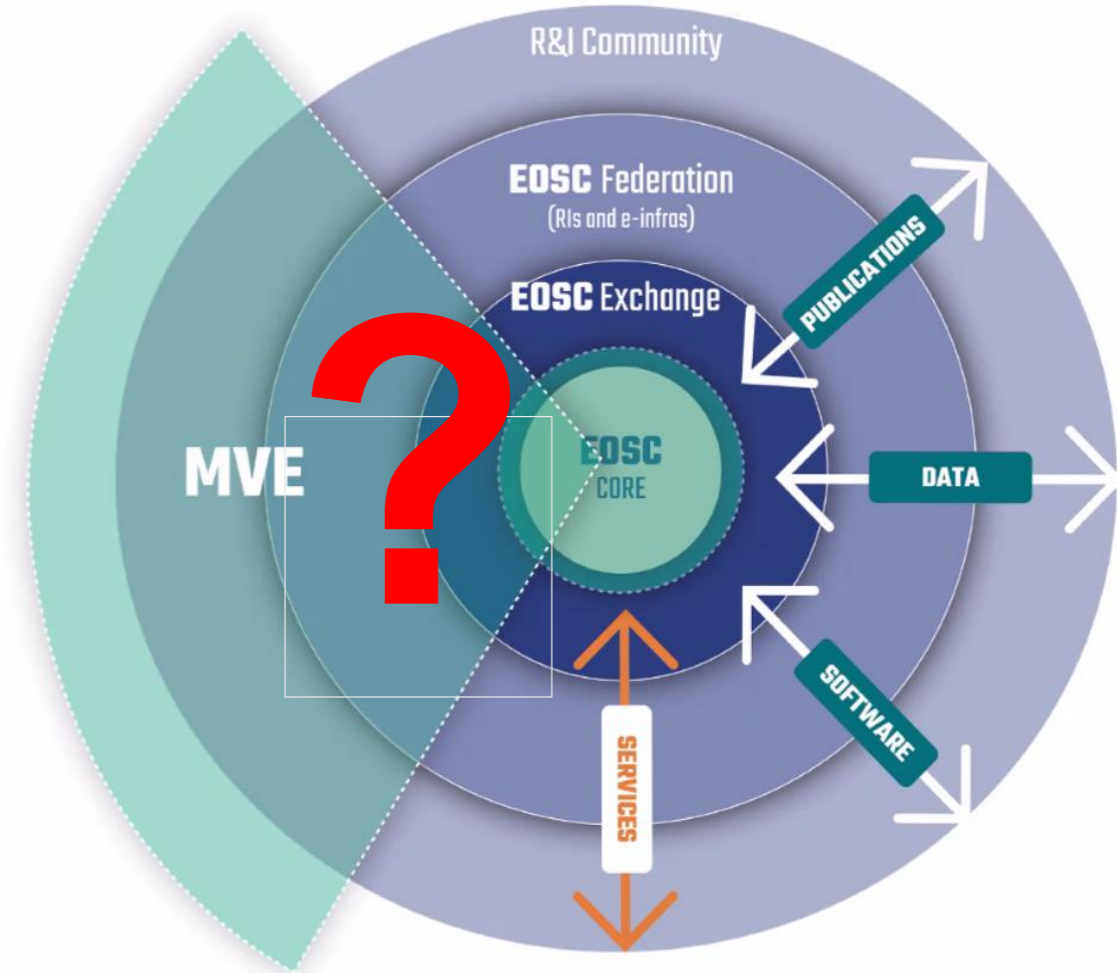
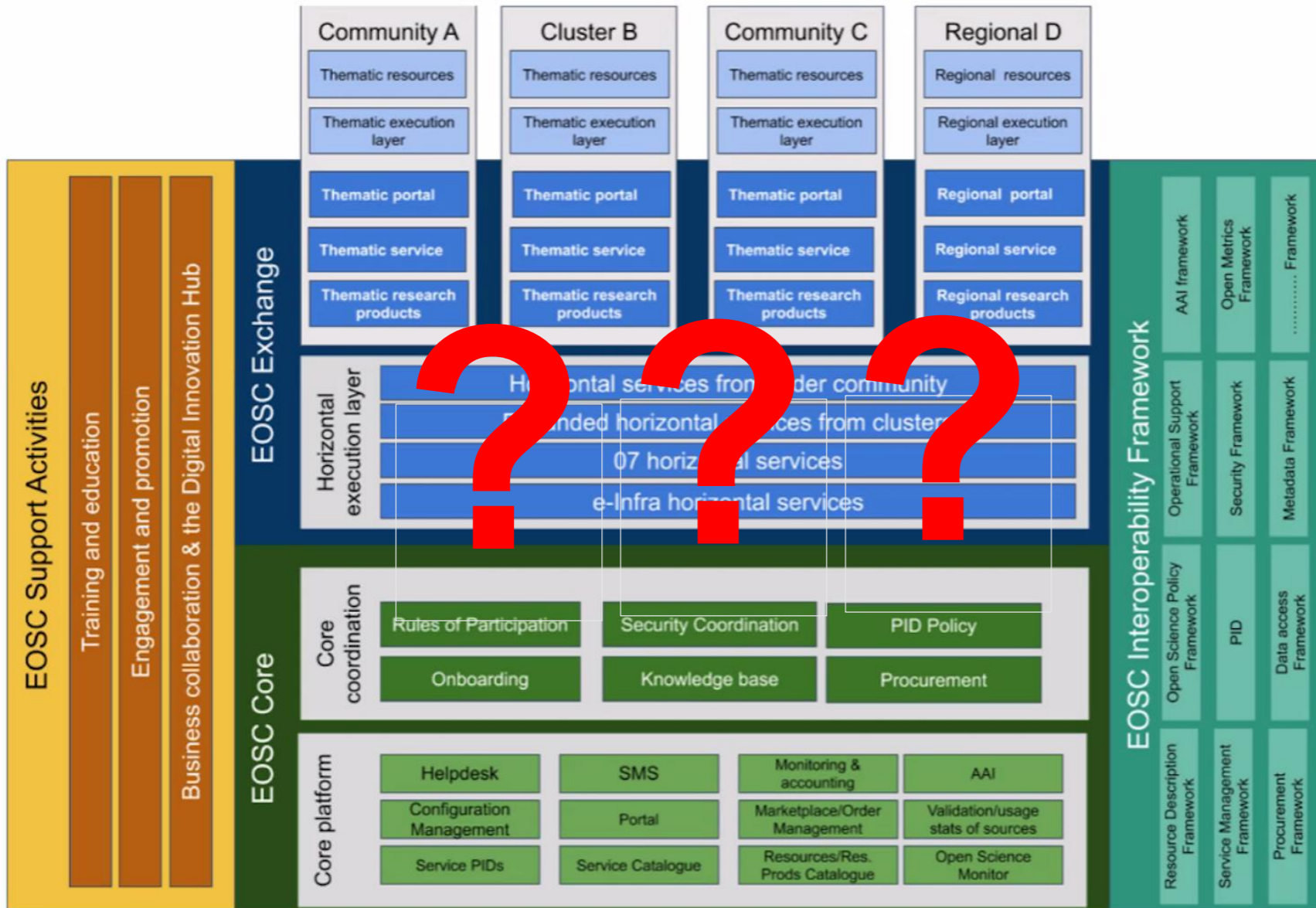
- Dodatkowe wymagania związane z **opisami semantycznymi** danych (*ang. metadata*) z różnych dziedzin nauki
- Regulacje w zakresie **powiązań i zależności** w procesach przetwarzania danych naukowych (*ang. workflows/pipelines*)
- Wymagania **długoterminowego składowania, archiwizacji i udostępniania** danych naukowych
- Wymagania **federacyjnego i kontrolowanego dostępu** oraz **zabezpieczeń**
- Zapewnienie i wsparcie **interoperacyjności** na poziomie **technicznym, semantycznym, prawnym** oraz **organizacyjnym**



Architektura EOSC



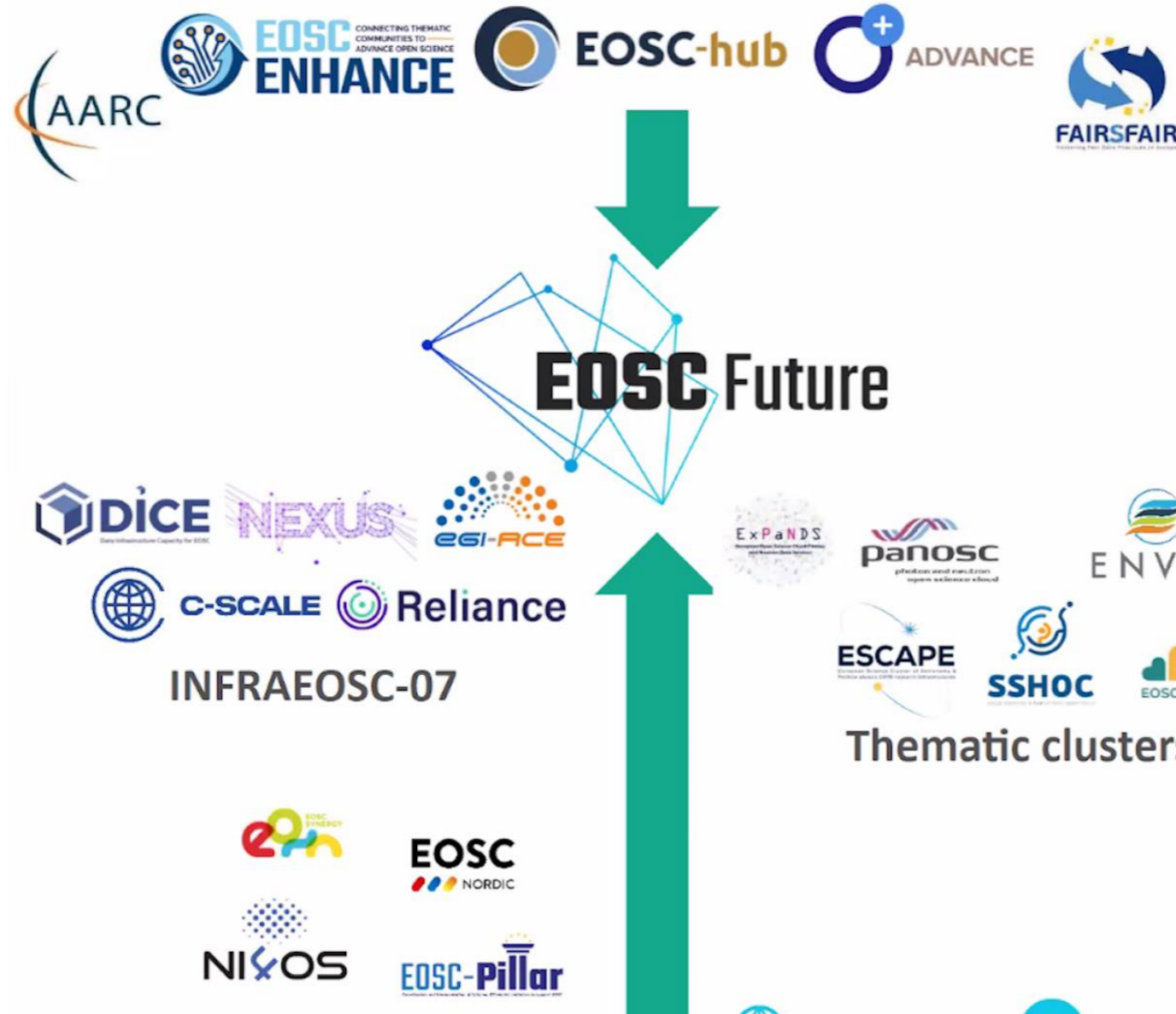
EUROPEAN OPEN SCIENCE CLOUD



EOSC na poziomie EU



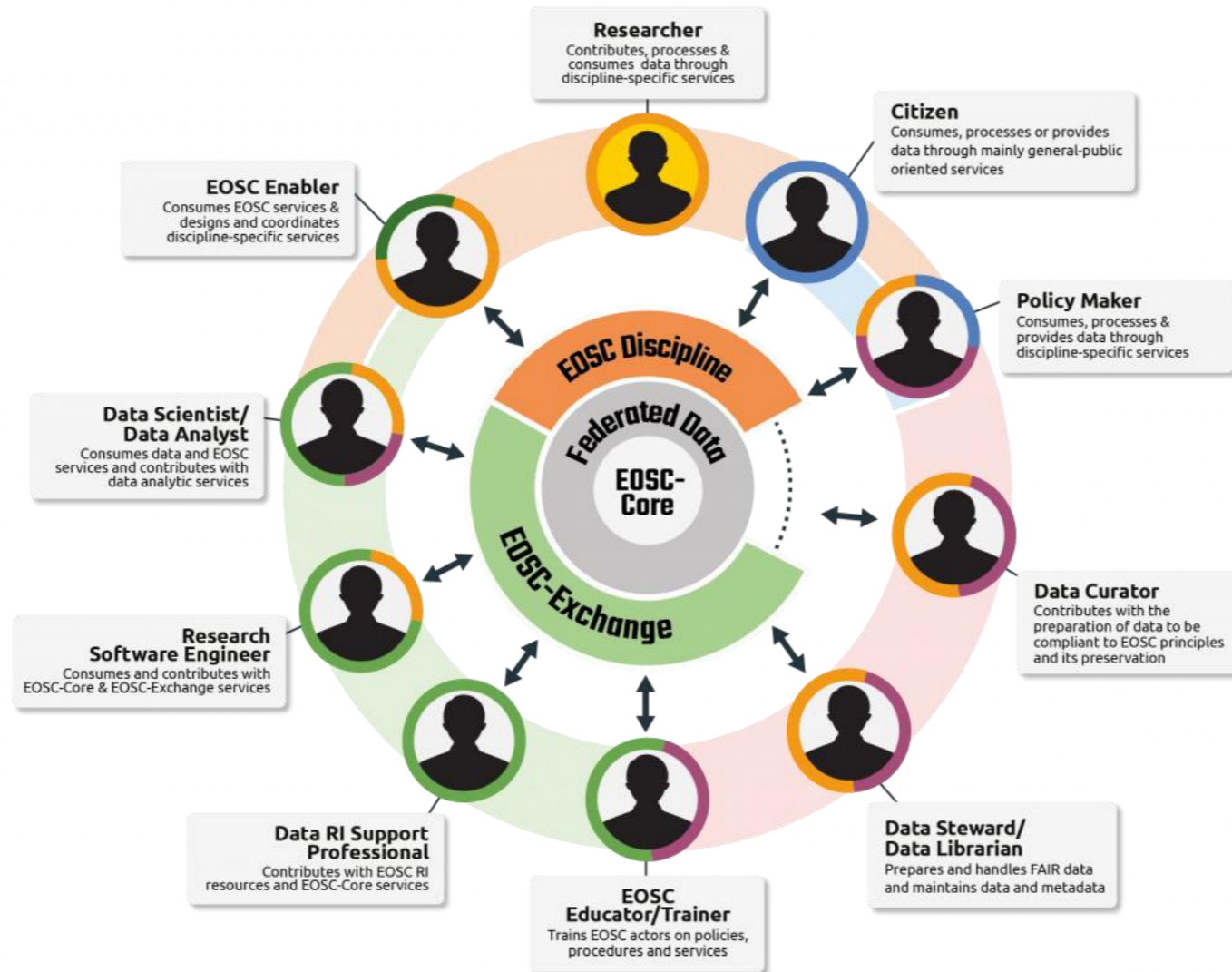
**EUROPEAN OPEN
SCIENCE CLOUD**



Główni aktorzy w ekosystemie EOOSC



**EUROPEAN OPEN
SCIENCE CLOUD**



ICT-Specific
Developing Software



Library & Information Science
Understanding Data



Discipline Specific
Conducting Research



General Public



NARODOWE CENTRUM NAUKI



Polska Mapa Infrastruktury Badawczej



**EUROPEAN OPEN
SCIENCE CLOUD**

Polska Mapa Infrastruktury Badawczej

Europejskie Forum Strategiczne Infrastruktur Badawczych (ESFRI)

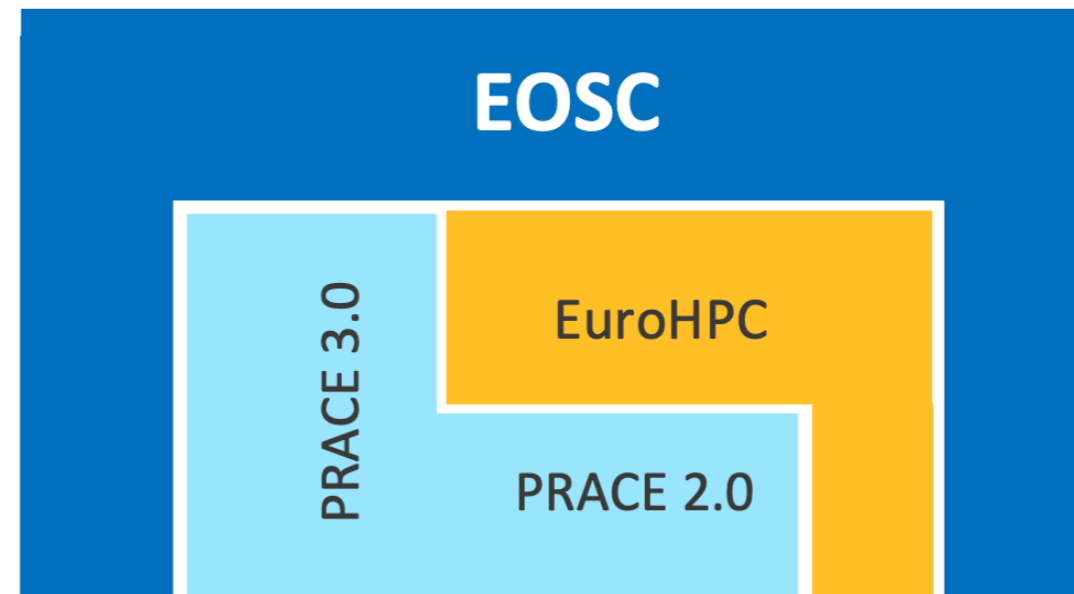
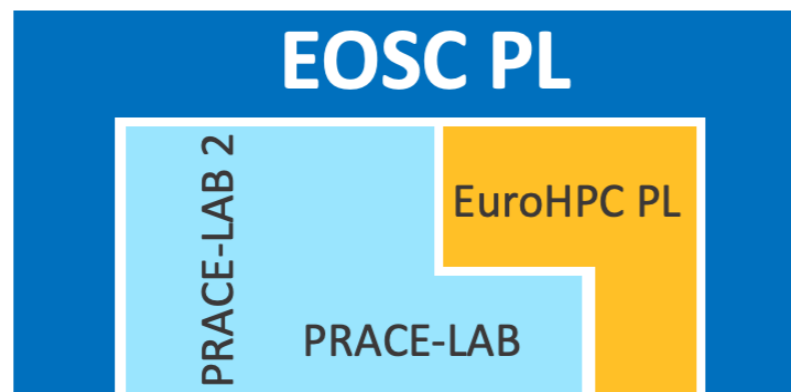
Obszar aplikacyjny i dziedzinowy

Centers of Excellence (Hildago, EoCoE, EoCoE II, Coegss),
Deep, LOFAR, ...
EOSC-HUB, EOSC- Synergy, ...

E-infrastruktura informatyczna nauki w Polsce

E-infrastruktura informatyczna nauki w UE

Krajowy Magazyn Danych

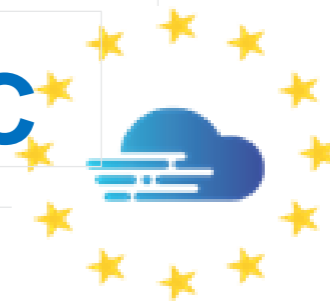


PIONIER-LAB (ośrodki MAN i KDM)

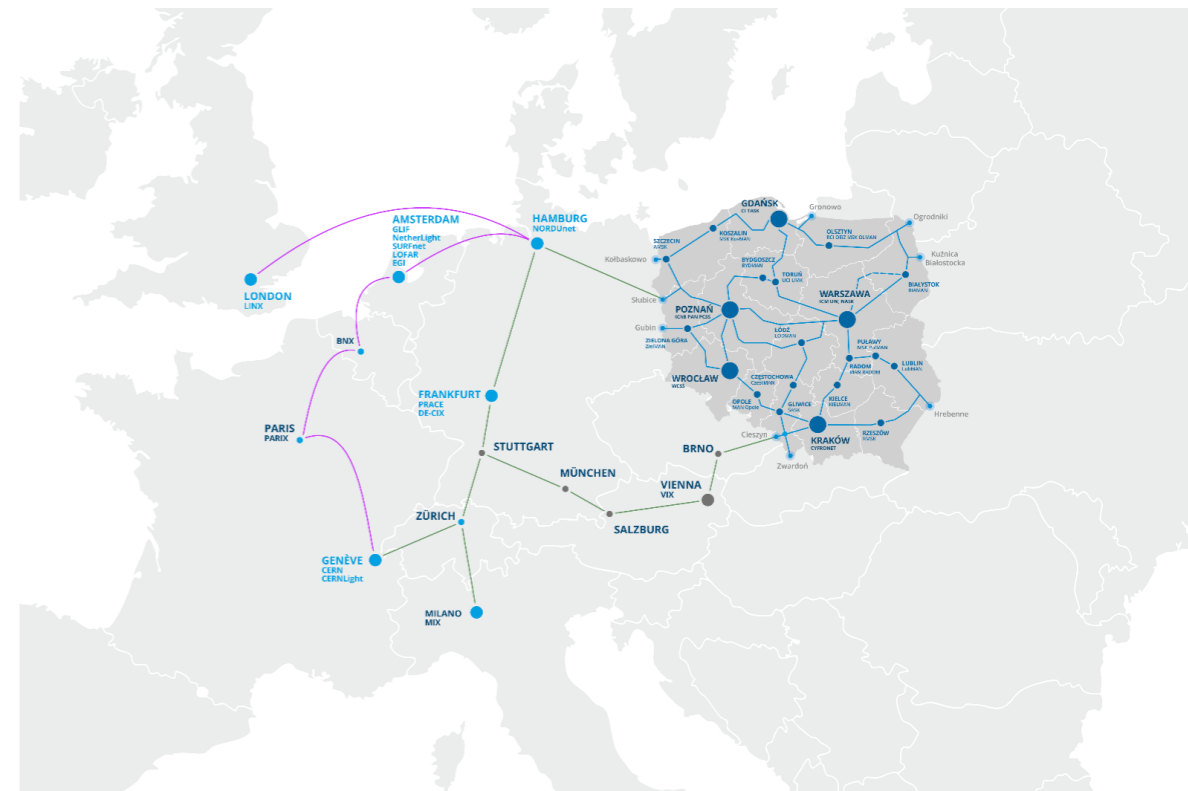
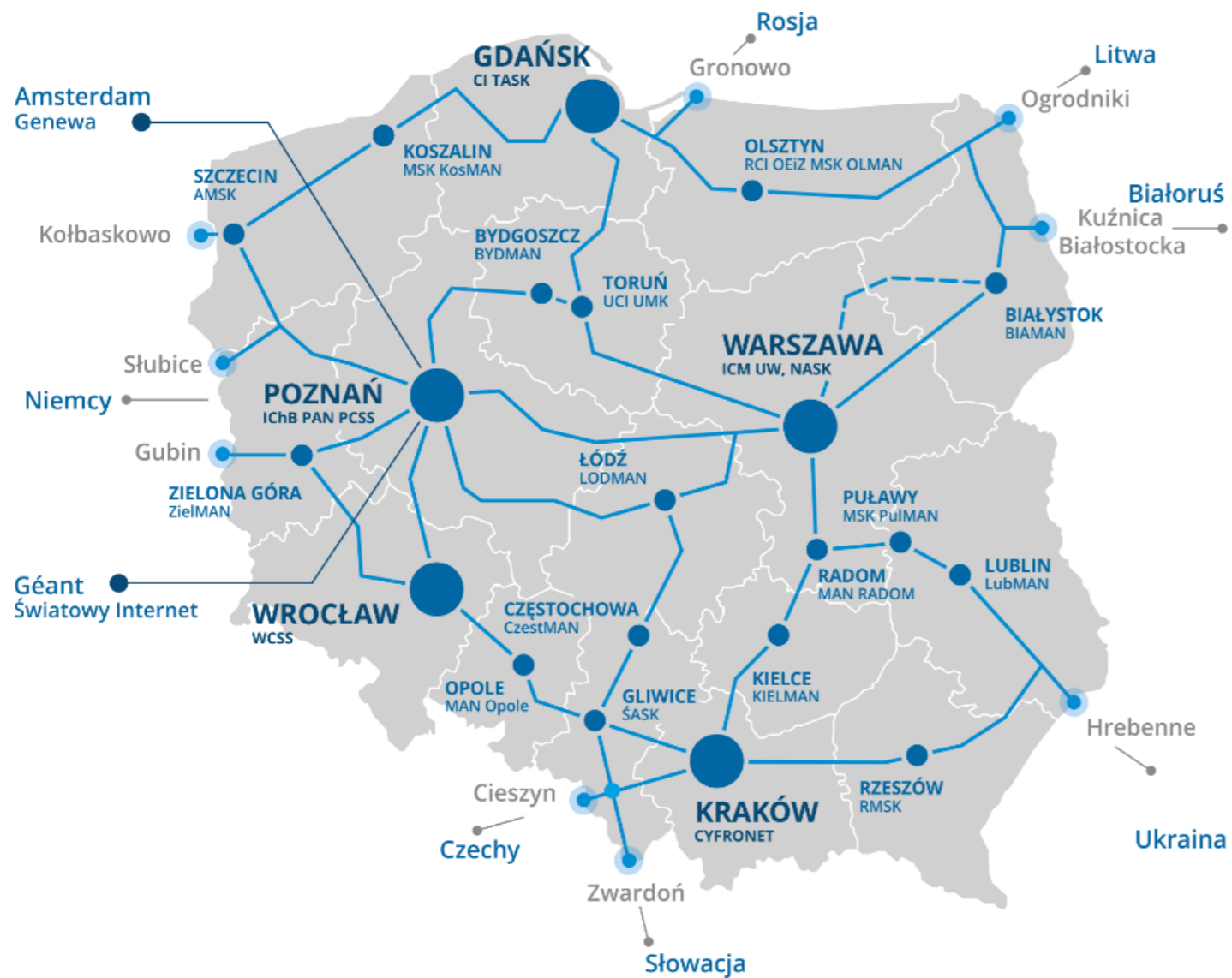
Sieć krajowa PIONIER i sieci miejskie MAN

Sieć GEANT i NREN

Sieć naukowa PIONIER i sieci MAN dla EOSC



**EUROPEAN OPEN
SCIENCE CLOUD**

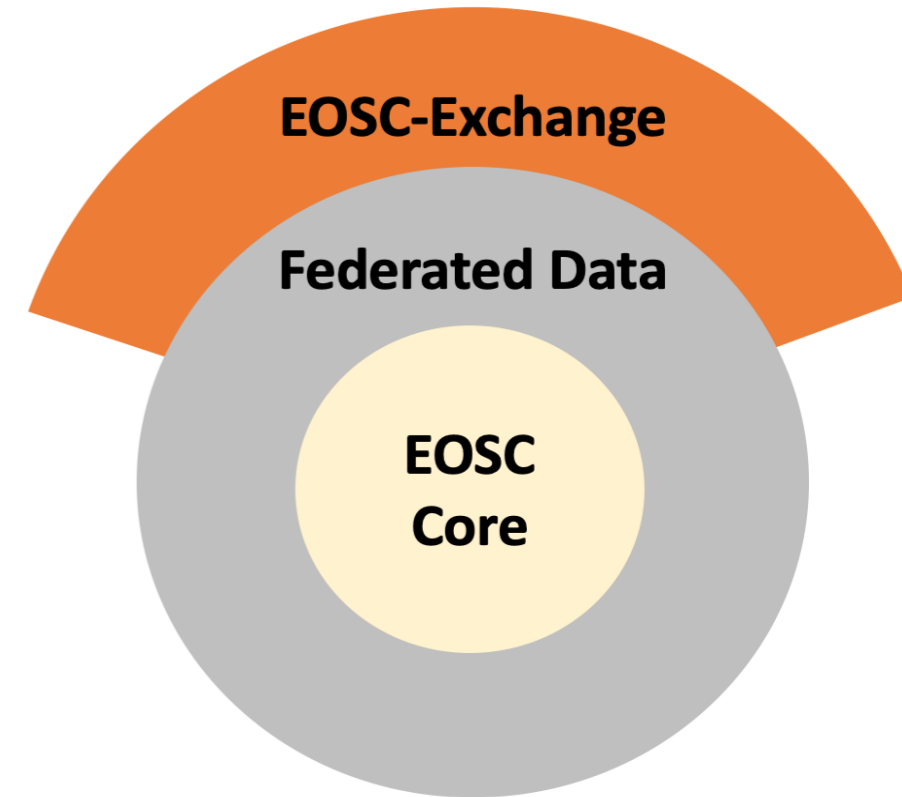


EOSC-Core – podstawowy zakres



**EUROPEAN OPEN
SCIENCE CLOUD**

- Ramowa **polityka otwartej nauki**, dotycząca sposobu w jaki dane naukowe oraz elementy składowe (np. otwarte oprogramowanie) w procesie naukowym są publikowane, udostępniane i ponownie wykorzystywane w EOSC
- Ramowa **infrastruktura informatyczna dla uwierzytelniania i autoryzacji** (*ang. Authentication Authorization Infrastructure - AAI*) - usługi zarządzania tożsamością naukowców umożliwiające im bezproblemowy dostęp do wszystkich otwartych zasobów w EOSC
- Nowe zasady dostępu i współdzielenia danych naukowych, których podstawową rolą jest oferowanie danych jako usługi online, w tym **trwałych identyfikatorów obiektów cyfrowych** (*ang. Persistent Identifier*) wraz z usługami do generowania unikalnych identyfikatorów oraz ich walidacji

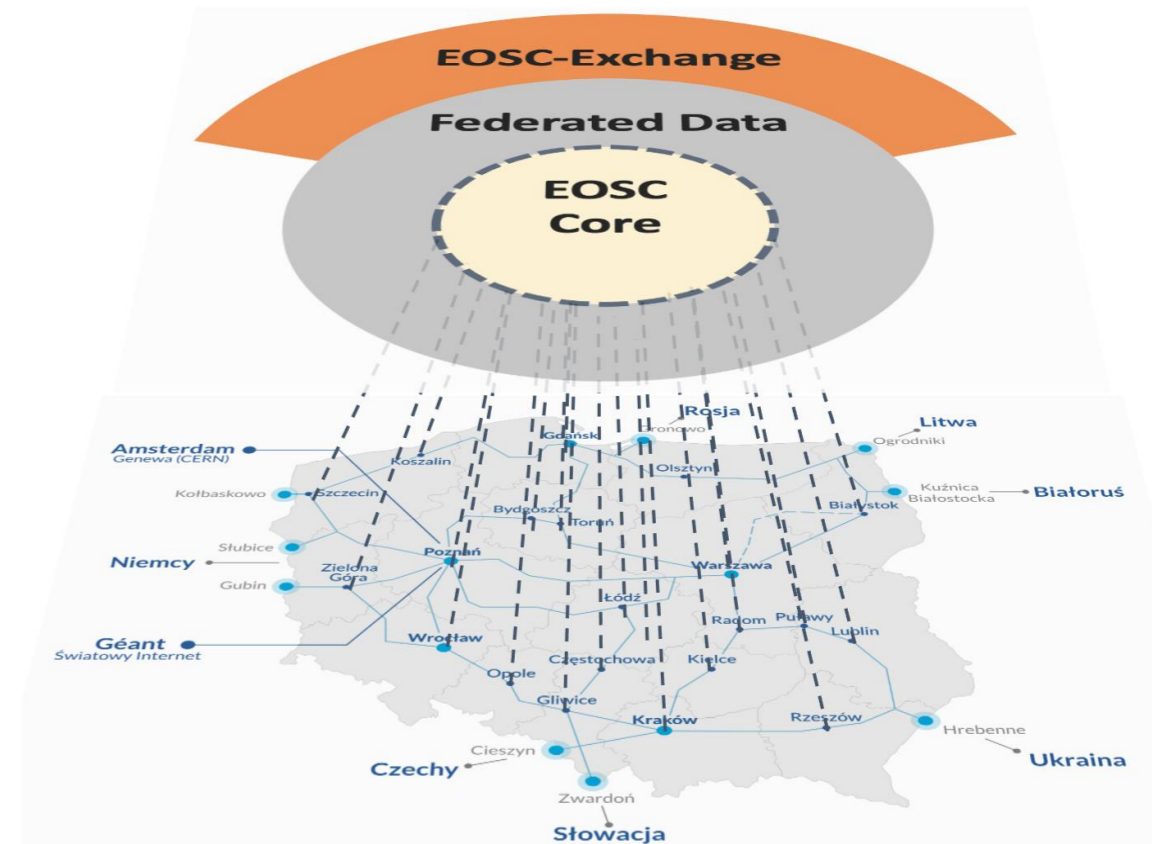


PIONIER dla EOSC-Core (2021)

- **Polska Federacja Zarządzania Tożsamością** dla użytkowników naukowych oraz kontroli dostępu w oparciu o otwarte standardy uwierzytelniania i autoryzacji (AAI) niezbędnych do realizacji założeń EOSC-Core dla federacji usług oraz repozytoriów
- **Usługi Data ID** dla rejestracji trwałych identyfikatorów obiektów cyfrowych w sieci PIONIER gwarantująca trwałą i jednoznaczny sposób dostępu do różnorodnych obiektów cyfrowych
- **Federacja Bibliotek Cyfrowych** dla wsparcia dalszego procesu gromadzenia, przetwarzania i udostępniania obiektów z polskich instytucji nauki oraz kultury w modelu otwartych danych i zgodnie z zasadami FAIR
- ... (wdrożenia projektów z POIR i POPC)



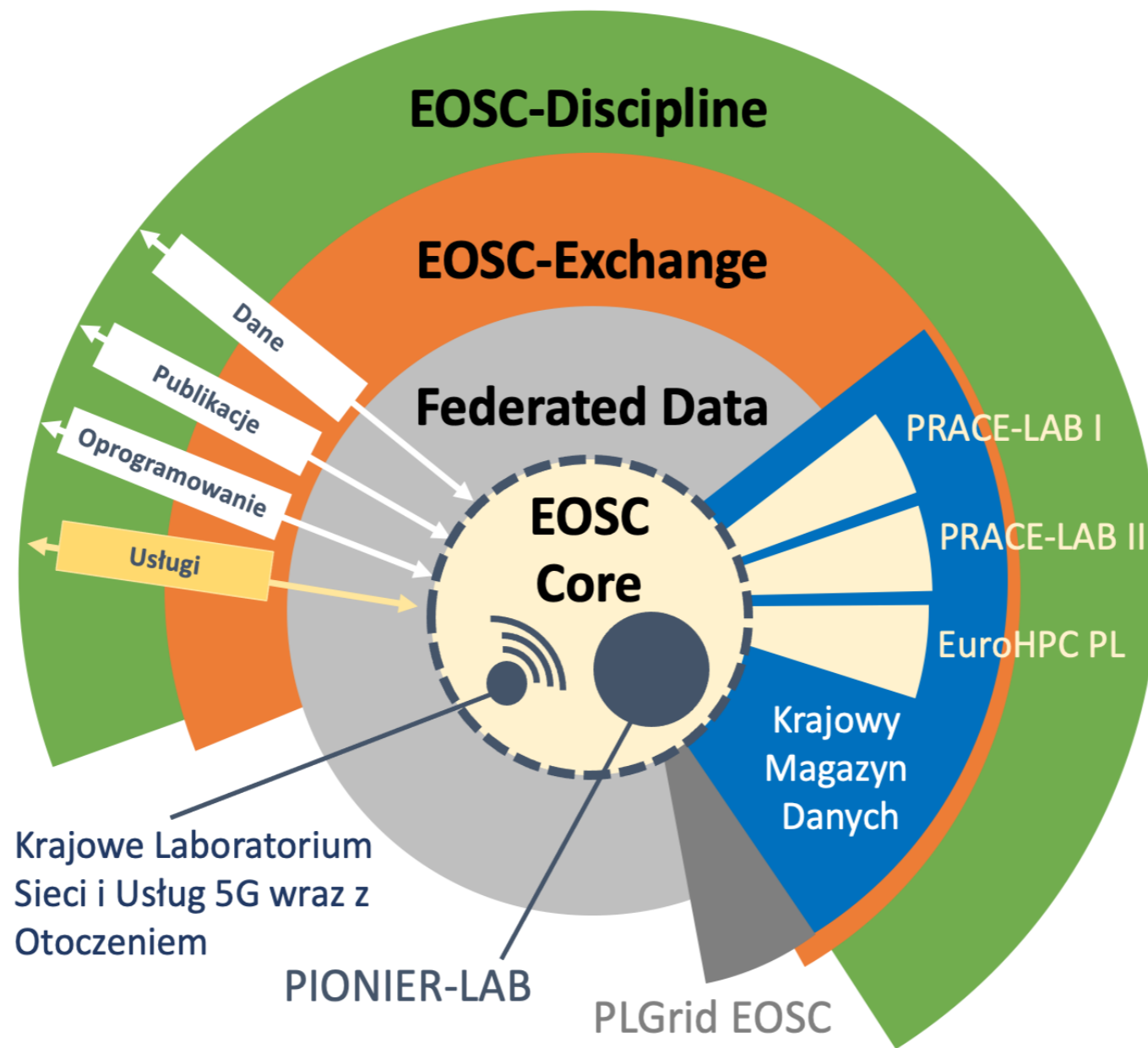
**EUROPEAN OPEN
SCIENCE CLOUD**



Informatyczne infrastruktury dla EOSC



**EUROPEAN OPEN
SCIENCE CLOUD**



PRACE-LAB dla EOSC (2021)



**EUROPEAN OPEN
SCIENCE CLOUD**

Partnerzy projektu PRACE-LAB:

Instytut Chemii Bioorganicznej
PAN, PCSS – Lider

Akademickie Centrum
Komputerowe CYFRONET AGH

Politechnika Białostocka

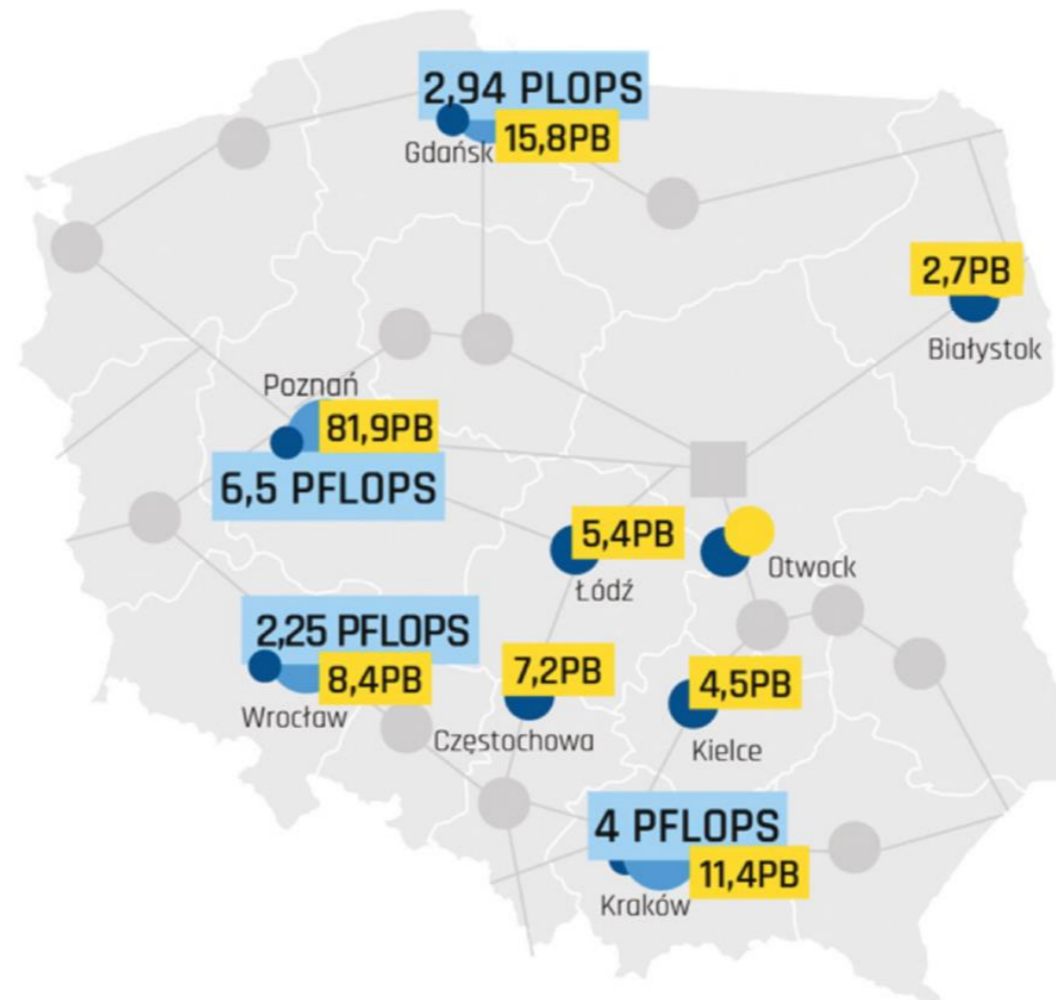
Politechnika Częstochowska

Politechnika Gdańska CI TASK

Politechnika Łódzka

Politechnika Świętokrzyska

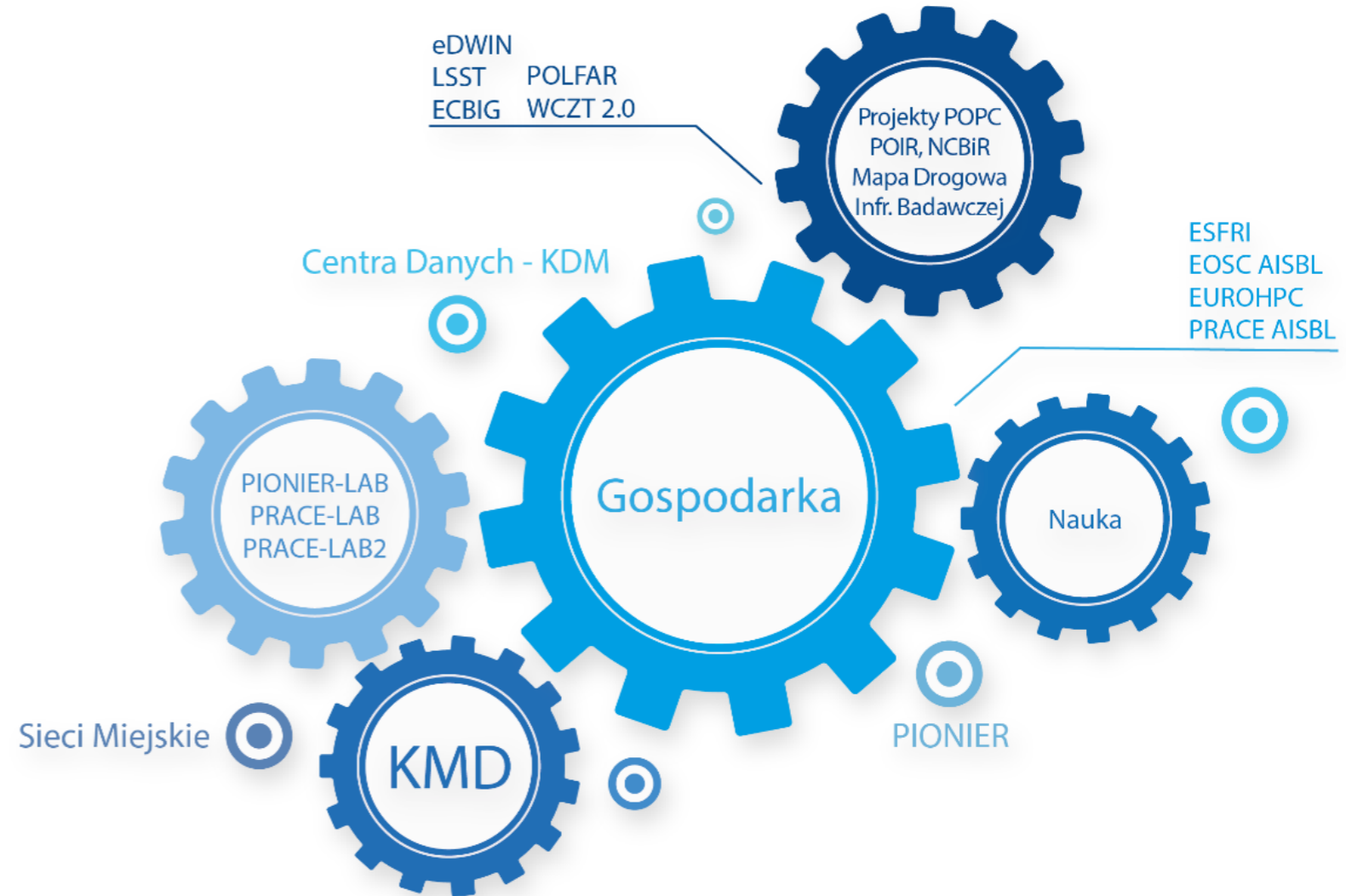
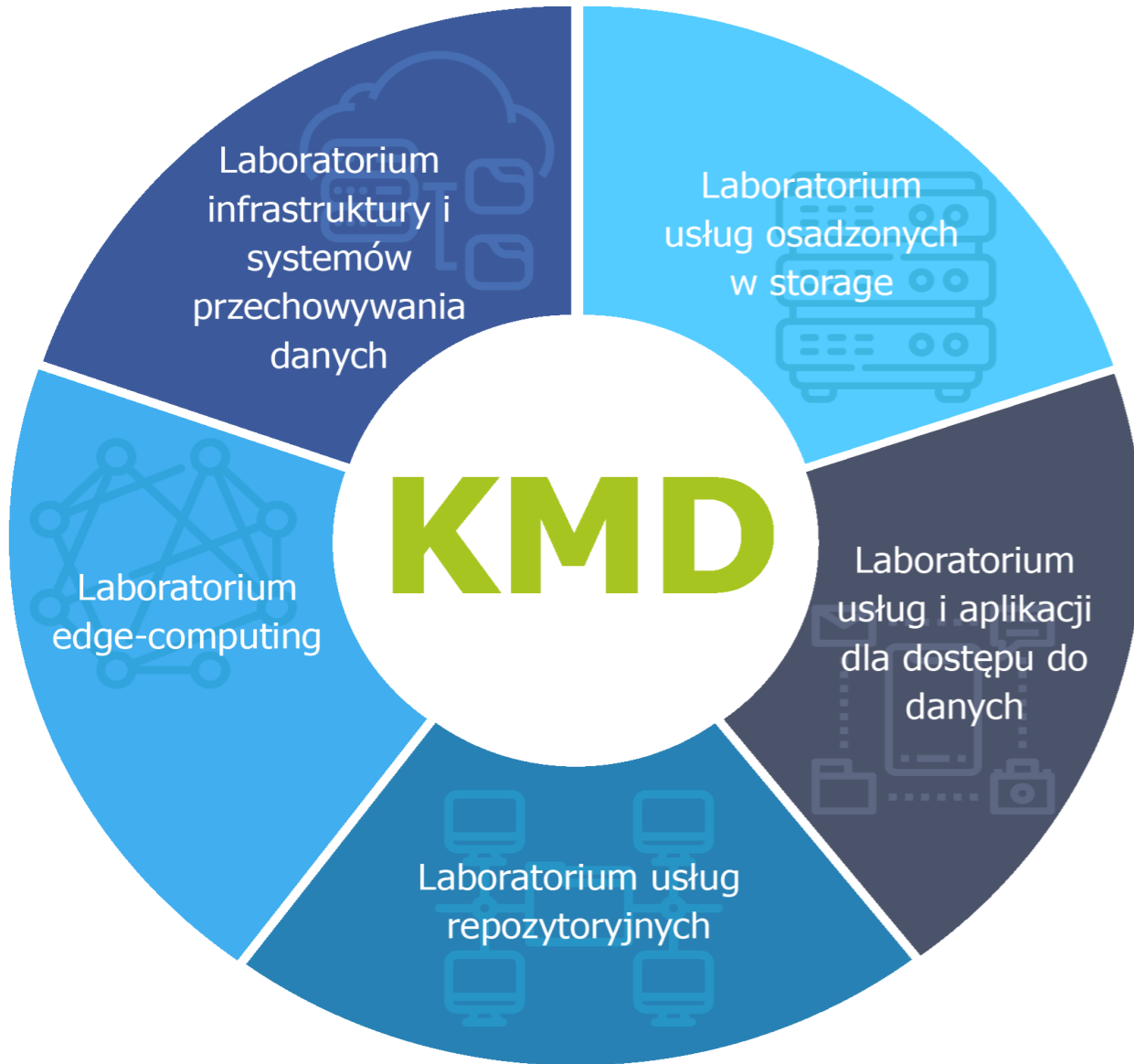
Politechnika Wrocławska
– WCSS



Krajowy Magazyn Danych dla EOSC (2022/3)



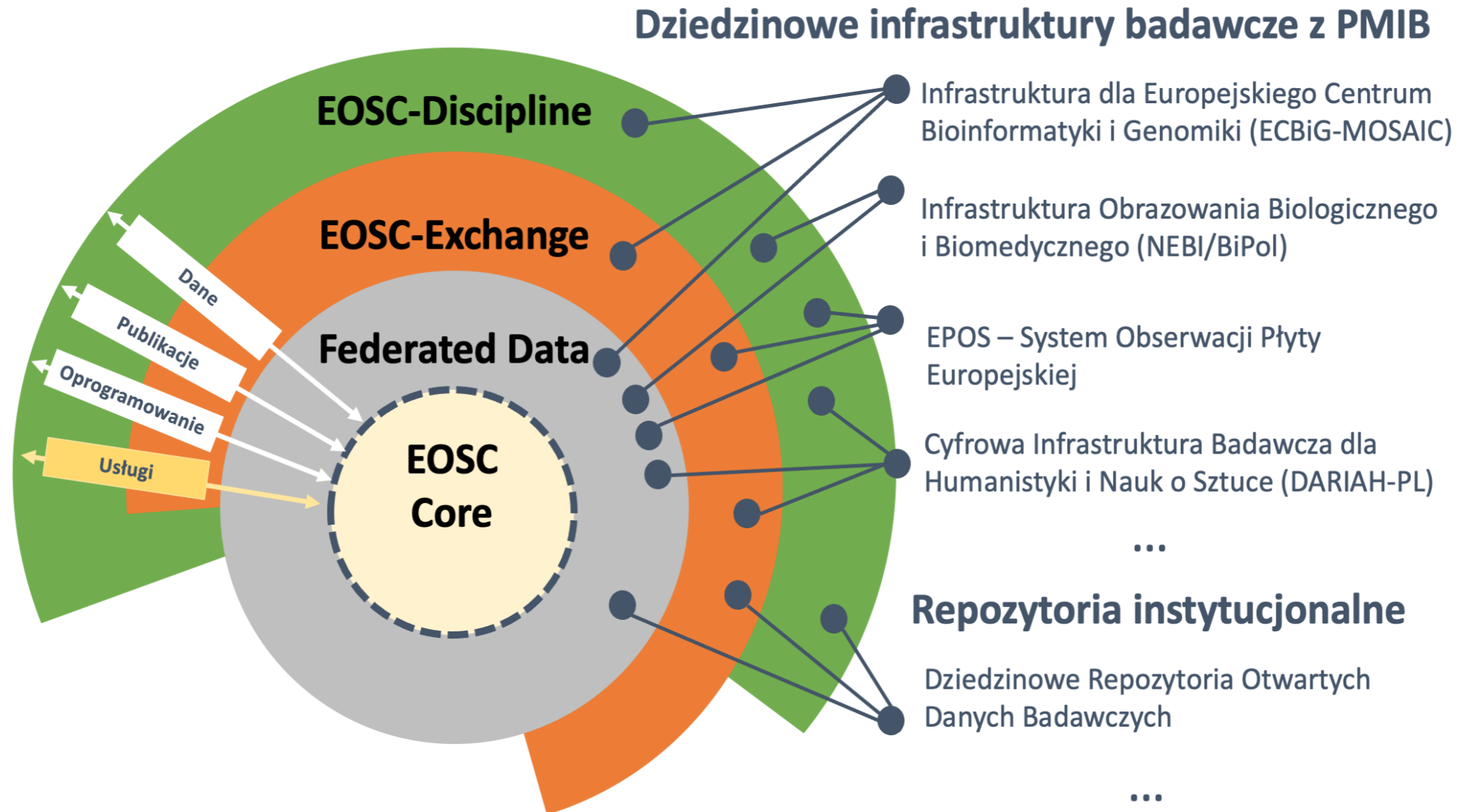
**EUROPEAN OPEN
SCIENCE CLOUD**



Dziedzinowe infrastruktury w EOSC



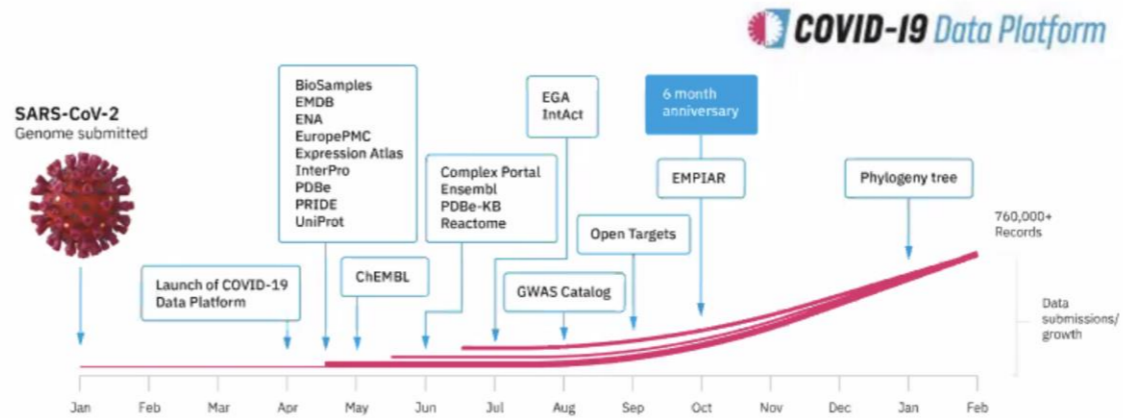
**EUROPEAN OPEN
SCIENCE CLOUD**





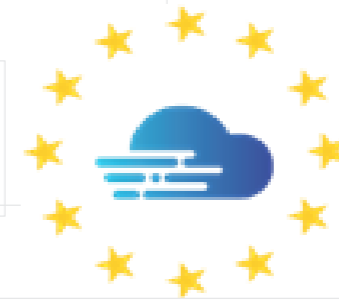
COVID-19 Data Portal

- Over 2,000,000 records across molecular platforms and literature
- Access to data resources and tools
- 70 linked “related” resources
- Web, API and download



Viral sequences		Host sequences		Expression		Protein	
All	1,719,048	All	16,527	All	97	All	1,709
Sequences	454,495	Human studies controlled access)	14	Gene expression	4	Protein sequences	106
Reference sequences	2	Human reads consented for full access)	8,804	Gene expression experiments	24	Protein families	315
Raw reads	645,456	Other species reads	7,669	Single cell expression	4	Protein structures - Knowledge Base	12
Sequenced samples	606,010	Association studies	40	Single cell expression experiments	22	Protein structures	737
Studies							523
Genes		Biochemistry		Imaging		Literature	
Browser		All	4,717	All	18	All	403,193
Variants		Pathways	16	Images	2	Coronaviruses	203,278
		Interactions	3,297	Electron microscopy public image archive	16	Diseases	189,782
		Complexes	31			Related viruses and diseases	2,501
		Compound document	10			Genes, receptors and antibodies	7,6
		Drug targets	1,361			Supplementary material	
		Metabolomics experiments	2				

Główne etapy wdrożenia EOSC



**EUROPEAN OPEN
SCIENCE CLOUD**

- **Etap 1 (okres 2021 – 2023)** Stworzenie bazowych mechanizmów dla EOSC-Core w celu zapewnienia infrastruktury uwierzytelniania i autoryzacji (AAI) dla użytkowników naukowych oraz zabezpieczenia elementów infrastruktury informatycznej zgodnie z założeniami modelu MVE oraz federacyjnej architektury EOSC
- **Etap 2 (okres 2024 – 2025)** Rozszerzenie modelu MVE dla EOSC o dostęp do usług o wartości dodanej, aplikacje i narzędzia (EOSC-Exchange) wspierające pełen cykl naukowy i związane z tym przepływy danych, obliczeń, współpracy naukowej, ...
- **Etap 3 (okres okres 2026 - 2027+)** Wdrażanie sieci danych zgodnie z rozszerzonymi zasadami FAIR i powiązanych usług, w tym EOSC-Core, EOSC-Exchange i EOSC-Discipline oraz innych ramowych warunków dotyczących pełnej federacyjności oraz interoperacyjności danych naukowych

Główne etapy wdrożenia EOSC (2021)



**EUROPEAN OPEN
SCIENCE CLOUD**

- Należy **wzmocnić działania informacyjno-promocyjne** dla modelu EOSC w Polsce, w szczególności skierowane do konsorcjów i partnerów zaangażowanych w realizację projektów infrastruktury informatycznej nauki na poziomie kraju oraz uczelni i instytucji naukowych odpowiedzialnych za udostępnienie oraz utrzymanie kluczowych repozytoriów danych
- Dla konkursów ogłaszanych przez **NCN** i **NCBiR** należy wypracować nowe procedury wraz z nowymi wymaganiami (np. poprzez rozbudowę zakresu istniejących **Planów Zarządzania Danymi**) związanymi ze współdzieleniem gromadzonych oraz przetwarzanych danych zgodnie z architekturą i modelem EOSC oraz rozszerzonymi **zasadami FAIR**
- Należy **aktywnie monitorować rozwój ekosystemu i modelu EOSC na poziomie europejskim** i odpowiednio synchronizować **działania podejmowane na poziomie krajowym**, w szczególności w ramach dużych projektów związanych z dalszym rozwojem infrastruktury informatycznej nauki w Polsce (udział polskich konsorcjantów w projektach EU - **EOSC Future**, **RELIANCE**, ...)



**EUROPEAN OPEN
SCIENCE CLOUD**

Dziękuję za uwagę

Kontakt:

krzysztof.kurowski@man.poznan.pl

Poznańskie Centrum Superkomputerowo-Sieciowe ICHB PAN

Od otwartych danych do otwartej nauki...



**EUROPEAN OPEN
SCIENCE CLOUD**

 Enabling **reproducible**, transparent research.

