

Powstaje Data Lake

#Astronautyka #Cyberprzestrzeń #Nowe technologie 4 kwietnia 2023

Rozpoczęto pierwszy etap prac nad budową Data Lake – ogromnego repozytorium danych, które jest kluczowym elementem przełomowej inicjatywy Unii Europejskiej – Destination Earth (DestinE). Ma ona na celu opracowanie cyfrowego modelu naszej planety m.in. na bazie danych satelitarnych, służącego do prowadzenia zaawansowanych badań i analiz ekosystemów Ziemi z wykorzystaniem sztucznej inteligencji i uczenia maszynowego w celu prowadzenia bardziej zrównoważonej gospodarki. Rolę głównego wykonawcy (prime contractor) projektu Data Lake powierzono polskiemu CloudFerro, a zakończenie pierwszego etapu prac planowane jest na czerwiec br.



/ Ilustracja: CloudFerro

Gwałtowne zmiany klimatu są powodem coraz większych obciążeń w zakresie rozwoju społecznego, gospodarczego i ochrony zdrowia w całej Europie. W odpowiedzi na te wyzwania, Komisja Europejska wraz z organizacjami partnerskimi – EUMETSAT (Europejską Organizacją Eksploatacji Satelitów Meteorologicznych), ESA (Europejską Agencją Kosmiczną) i ECMWF (Europejskim Centrum Prognoz Średnioterminowych), uruchomiła przełomową inicjatywę – Destination Earth (DestinE). Jej celem jest opracowanie cyfrowej repliki Ziemi, która umożliwi bardzo dokładne modelowanie wszystkich ekosystemów naszej planety, co pozwoli precyzyjnie monitorować zmiany klimatu i skutki działań człowieka, przewidywać zjawiska pogodowe i klimatyczne, a także testować różne scenariusze. Kluczowymi elementami inicjatywy DestinE będą:

W ramach tej inicjatywy, EUMETSAT jest odpowiedzialny za budowę wielo-chmurowego tzw. jeziora danych, czyli centralnego repozytorium danych, w tym zaprojektowanie, wdrożenie, testy i utrzymanie ogromnego repozytorium danych dostępnego online,

które stanowią podstawę Destination Earth (DestinE). Dane będą pochodzić z własnych systemów satelitarnych obserwacji Ziemi EUMETSAT, misji ESA, satelitów Copernicus Sentinel, ECMWF i innych europejskich dostawców danych.

Jako główny wykonawca dla EUMETSAT, polskie przedsiębiorstwo odpowiada za koordynację projektu Data Lake, dostawę infrastruktury chmurowej (Infrastructure-as-a-Service i Platform-as-a-Service), przetwarzanie Big Data oraz obsługę i utrzymanie usług. CloudFerro będzie współpracować z dwoma europejskimi partnerami – CS Group odpowiedzialnym za obszar odkrywania i dostępu do danych, oraz EODC odpowiedzialnym za przetwarzanie Big Data.

Budowane przez CloudFerro repozytorium danych Data Lake, którego będzie także operatorem, ma obejmować ok. 55 PB danych przechowywanych w prywatnych chmurach zlokalizowanych w centrach danych w całej Europie. W centrach tych znajdują się także potężne komputery zainstalowane w ramach projektu EuroHPC, na których przetwarzane są dane cyfrowego bliźniaka Ziemi. Jednym z takich centrów jest CSC Kajaani, gdzie znajduje się super komputer LUMI HPC. CSC - LUMI znacznie poszerzy zakres projektu DestinE.

Data Lake zapewni łatwy dostęp do wszystkich danych wymaganych przez użytkownika DestinE. Repozytorium będzie dostępne z dużej liczby zewnętrznych przestrzeni danych, cyfrowych bliźniaków Ziemi i aplikacji znajdujących się na platformie DestinE Core Service Platform, niezależnie od rodzaju danych i ich lokalizacji. Serwis Data Lake ma dostarczyć skalowalną strukturę usług, która umożliwi federację z różnymi przestrzeniami danych, a także świadczenie usług obliczeniowych i przechowywanie jak najbliżej danych DestinE.

DestinE jest częścią programu Komisji Europejskiej *Cyfrowa Europa* i ma przyspieszyć transformację Europy w zakresie ochrony klimatu i rozwoju cyfryzacji. DestinE jest również wspierany przez program *Horyzont Europa* oraz inne europejskie i krajowe inicjatywy w dziedzinie badań naukowych i innowacji. Komisja planuje dostarczyć pierwsze usługi DestinE do 2024, początkowo dla użytkowników z sektora publicznego. W pełni wdrożony system zostanie ukończony w ciągu 7-10 lat, aby służyć również naukowcom, badaczom, sektorowi prywatnemu i ogółowi społeczeństwa ([CloudFerro z kontraktem DLR](#) , 2019-11-12).

Powiązane wiadomości

[Powstaje Data Lake \(2023-04-04\)](#)

[CloudFerro z kontraktem DLR \(2019-11-12\)](#)

[Polski system dla Copernicusa \(2019-07-02\)](#)

[Copernicus coraz bardziej potrzebny \(2019-06-04\)](#)
