



eosc.eu | #eoscsymposium2024



In cooperation with



## Zpráva ze zahraniční cesty

Mgr. Tereza Šorejsová  
Ústřední knihovna ČVUT

Datum a místo konání: 21. - 23. října 2024, Berlín, Německo  
Webová stránka: <https://eosc.eu/symposium2024/>  
Prezentace: <https://eosc.us6.list-manage.com/track/click?u=bd106f33ba5f588652c5ad1a4&id=0c34bd31b8&e=a3af491791>

V říjnu 2024 jsem díky Fondu podpory zahraničních cest AKVŠ dostala příležitost se zúčastnit EOSC Symposia 2024, které se konalo v Berlíně. Jde o nejvýznamnější setkání v rámci komunity EOSC (European Open Science Cloud), iniciativy Evropské unie, která vytváří digitální infrastrukturu pro sdílení vědeckých dat a služeb napříč Evropou. Cílem je podpořit otevřenou vědu, spolupráci a zjednodušit přístup k datům a nástrojům pro výzkumníky. Toto významné mezinárodní setkání přilákalo přibližně 460 účastníků osobně a dalších 900 diváků online z celkem 60 zemí, včetně 25 mimoevropských. Symposium se zaměřilo na rozvoj Evropského otevřeného vědeckého cloudu (EOSC Node) a jeho budoucí roli v evropském i globálním výzkumu a inovacích.

### Úvod

Úvodní příspěvky se zaměřily na proklamace o významu iniciativy EOSC, důležitosti principů FAIR (Findable, Accessible, Interoperable, Reusable), interoperability a dalších klíčových témat. Jak EOSC vstupuje do své provozní fáze, poselství je zřejmé: EOSC se stále více stává globálním projektem s klíčovým významem pro evropskou konkurenceschopnost.

Zasedání zahájila prostřednictvím videopřenosu německá ministryně pro vzdělávání a výzkum **Bettina Stark-Watzingerová**, která uznala EOSC jako nástroj pro výzkumné pracovníky. Úspěch jednotlivých národních iniciativ, jako je NFDI (německá iniciativa Nationale Forschungsdateninfrastruktur), a úspěch

EOSC jdou ruku v ruce, přičemž nyní je ten správný čas na to, aby se i ostatní státy zapojily. Vyzvala ke koordinaci a sdílení zdrojů napříč Evropou.

Rovněž prostřednictvím videopřenosu se za EOSC postavila evropská komisařka pro inovace, výzkum, kulturu, vzdělávání a mládež **Iliana Ivanová**, která důrazně potvrdila závazek Komise podpořit otevřenou vědu skrze interoperabilní infrastrukturu, která usnadní sdílení vědeckých dat a výsledků. Tuto myšlenku podpořila i **Signe Ratsová**, zástupkyně generálního ředitele Generálního ředitelství pro výzkum a inovace (DG RTD) Evropské komise, která označila EOSC za ústřední pilíř Evropského výzkumného prostoru. Přidal se k ní také bývalý italský premiér **Enrico Letta**, jenž zdůraznil, že EOSC hraje klíčovou roli pro tzv. **pátou svobodu**, kterou definoval ve své zprávě o budoucnosti vnitřního trhu Evropské unie<sup>1</sup> zadané Evropskou radou.

Cílem ustanovení páté svobody vnitřního trhu by měla být podpora volného pohybu výzkumu, inovací a vzdělávání. Letta potřebu ustanovení této páté svobody odůvodňuje tím, že existující čtyři svobody volného pohybu osob, zboží, služeb a kapitálu nejsou schopny plně pokrýt potřeby 21. století, jakými jsou například digitalizace či prolínání zboží a služeb. Tato svoboda má urychlit rozvoj v oblastech, jako jsou výzkum a vývoj, umělá inteligence, kvantové výpočty, biotechnologie, biorobotika či vesmírný průzkum. Záměrem je sjednotit rozptýlené znalosti a odstranit nerovnosti, čímž se vytvoří nové příležitosti pro inovace a inkluzivní růst. Klíčovým prvkem je také udržení talentů v EU. Letta věří, že pátá svoboda posílí evropskou integraci a konkurenceschopnost a zároveň umožní lépe využít talent a potenciál mladých lidí v EU.

## EOSC EU Node

Ústředním bodem letošního sympozia bylo představení **EOSC EU Node**<sup>2</sup>. Jeho spuštění oznámil **Gustav Kalbe**, ředitel jednotky *Digital Excellence and Science Infrastructure* v rámci Generálního ředitelství pro komunikační sítě, obsah a technologie Evropské komise (DG CONNECT). Jde o první node neboli uzel budoucí federace EOSC, který byl oficiálně spuštěn jen pár dní před konáním sympozia. Spuštění uzlu EOSC EU je vnímáno jako významný krok od teoretických proklamací a různých výzev k reálným a konkrétním činům. Tento krok představuje přechod od abstraktních diskusí o otevřené vědě a spolupráci k praktické implementaci, která umožní skutečné propojení evropských výzkumných komunit a efektivní sdílení vědeckých dat a nástrojů napříč Evropou.

Na sympoziu bylo předvedeno několik praktických ukázek, jak různé vědecké týmy mohou v rámci EOSC EU Node efektivně spolupracovat. Pro mě jako knihovníka to bylo velmi přínosné, protože jsem si dokázala mnohem lépe představit, jak taková spolupráce může vypadat. Vědecké týmy z různých evropských univerzit mohou využít službu pro synchronizaci a sdílení, která jim umožní jednoduše, bezpečně, a přitom efektivně sdílet obrovské soubory dat nebo celé databáze. Mohou použít interaktivní notebooky (např. Jupyter Notebooks) hostované na platformě EOSC pro analýzu dat nebo se připojit k výpočetním clusterům (HPC – High-performance computing), což umožňuje týmům spolupracovat na vývoji modelů a simulacích v reálném čase, přičemž každý člen týmu může přidávat své vlastní analýzy a komentáře, které jsou okamžitě dostupné pro ostatní. Procesy, které jsou roztržštěné mezi různé platformy, lze v EOSC Node soustředit do jednotného, bezpečného prostředí. Spolupráce napříč vědeckými týmy tak může být mnohem efektivnější.

---

<sup>1</sup> Enrico Letta - [Much more than a market](#)

<sup>2</sup> <https://open-science-cloud.ec.europa.eu/>

Když bych to měla přiblížit na jedné konkrétní ukázce, tak například vědci z oboru fyzikální geografie studují riziko přírodních pohrom, jako jsou zemětřesení, záplavy nebo sopečné erupce. Tento výzkum zahrnuje sběr a analýzu obrovského množství dat, geologických snímků, historických záznamů o pohybech zemské kůry nebo prediktivních modelů. Vědci z různých univerzit spolupracující na stejném projektu se přihlásí do EOSC Node (skrz domovskou instituci, nevytváří se nový účet, eduGAIN). Zde mohou pracovat s daty o seismické aktivitě z veřejné databáze, která je připojena k EOSC Node nebo sdílet svá vlastní měření. Každý si může stáhnout Jupyter Notebook, v něm vytvořit různé příkazy a skripty, následně se připojit k superpočítači (HPC), k čemuž se použijí kredity své univerzity. HPC provádí výpočty, simulace a modelování podle vědeckých požadavků. EOSC Node nabízí různé servery s různou výpočetní kapacitou. Některé servery jsou menší a spotřebovávají méně kreditů, respektive výpočetní síly, jiné jsou větší a náročnější na zdroje.

Vědci mohou vzájemně sdílet své notebooky, upravovat je a spouštět na nich výpočty. Pokud jeden vědec vytvoří model pro předpověď rizika zemětřesení, druhý ho může upravit nebo doplnit o nové analýzy. Všechny související soubory a výsledky jsou snadno dostupné a sdílené mezi členy týmu.

Po provedení analýzy mohou vědci převést své výsledky (grafy, tabulky, modely) do formátu PDF nebo jiného formátu a vystavit je jako open access. Tento PDF dokument obsahuje nejen výsledky, ale i kód, který vedl k těmto výsledkům. Kdokoli si tento dokument přečte, může vidět, jaké výpočty byly použity, a v případě potřeby se vrátit k původnímu kódu nebo datům na EOSC Node.

Do EOSC Node mají přístup zaměstnanci jejichž pracoviště je členem EOSC. Mohou se přihlásit údaji ze svých univerzit. Evropská komise poskytuje zdarma při prvním přihlášení každému 100 kreditů a automaticky je obnovuje každé tři měsíce. Kredity se spotřebovávají v závislosti na tom, jak moc výkonné stroje nebo služby jsou použity. Vědec je vždy dopředu informován o tom, kolik kreditů na určitý početní výkon bude potřebovat. Pokud vědec pracuje v týmu, počet kreditů se zdvojnásobuje. Studenti do EOSC Node přístup primárně nemají, ale je možné, aby je některý z vědců pozval do týmu a přístup dostanou. Pro testovací účely je také možné požádat o navýšení kreditů.

Spuštění EOSC Nodu mělo mezi vědci na sympoziu opravdu velký ohlas. Zazněly názory, že toto prostředí je přesně to, co vědci potřebují. Konečně je vidět nějaký hmatatelný výsledek a jeden z vědců dokonce uvedl, že to je poprvé, co si na nic nestěžuje. Do konce sympozia se zaregistrovalo zhruba 300 uživatelů, kteří chtěli prozkoumat nově dostupné služby, což svědčí o okamžitém zájmu a poptávce po uzlu.

The screenshot shows the user interface of the EOSC EU Node. At the top, it says "Hello Tereza Šorejsová" and "This is the overview of your EOSC EU Node account." Below this is a welcome message: "Welcome to the EOSC EU Node!". The main content area is titled "Services" and lists several services with their status and a "View Service" link:

- File Sync & Share**: Access enabled. View Service >
- Interactive Notebooks**: Access enabled. View Service >
- Large File Transfer**: Access enabled. View Service >
- Virtual Machines**: Not available. Please read our [User Access Policy](#) for more information. View Service >
- Cloud Container Platform**: Not available. Please read our [User Access Policy](#) for more information. View Service >
- Bulk Data Transfer**: Not available. Please read our [User Access Policy](#) for more information. View Service >

At the bottom, it says "Credits renewed 2025-01-21". On the left side, there is a navigation menu with options like "Overview", "Notifications", "Tools Hub", "SERVICES", "File Sync & Share", "Interactive Notebooks", "Large File Transfer", "Cloud Container Platform", "Virtual Machines", and "Bulk Data Transfer".

## EOSC Federation Handbook

EOSC Node je pouze jedním ze stavebních prvků vznikající federace uzlů. Cílem je, aby vznikly další uzly, a to na národních nebo komunitních/oborových úrovních. Některé uzly již mohou existovat, takže bude dobré, pokud se do federace zapojí. Pokud se jednomu uzlu něco stane, v Evropě budou další uzly, kde bude možné v rámci vědecké komunity spolupracovat. Pouze díky federaci bude vědecké prostředí odolné a flexibilní.

V plánu je tedy vybrat první skupinu kandidátských uzlů EOSC, které budou v rámci federace pilotně zavedeny v roce 2025. Jak se do federace zapojit je popsáno v jedné z kapitol právě vznikající příručky **EOSC Federation Handbook**<sup>3</sup>.

Příručka federace EOSC se bude komplexně zabývat účelem, strukturou, řízením, architekturou a fungováním federace EOSC. Plánem pro další roky je příručku průběžně obohacovat a revidovat, a to díky poznatkům a novým zkušenostem získaných v průběhu zapojování dalších uzlů. Draft příručky je možné komentovat do 17. listopadu 2024. Na konferenci tak byl návrh příručky detailněji představen a celá komunita byla vyzvána ke spolupráci.

Finální podoba příručky bude schvalována v rámci trojstranného řízení EOSC (EOSC Tripartite Governance). Což je koncepce strategické koordinace mezi EU, zastoupenou Komisí, sdružením EOSC a řídicím výborem EOSC, který zastupuje členské státy EU a přidružené země.

### Tripartite Collaboration

European Commission



Representing the European Union

EOSC Association



Representing the European research community

EOSC Steering Board



Representing EU Member States and Associated Countries\*

## AI a vědecká data

Často skloňovaným tématem i na této konferenci byla umělá inteligence. Jednu z klíčových přednášek na toto téma měla **Julia Lane**, profesorka New York University, renomovaná ekonomka a datová vědkyně, známá svými výzkumy v oblasti měření a analýzy dat o vědě, inovacích a pracovní síle. Zaměřuje se také na zlepšení dostupnosti a kvality dat pro analýzu veřejné politiky.

Byla zdůrazněna důležitost důvěryhodnosti a spolehlivosti dat pro AI modely. Zazněly myšlenky, že ChatGPT rozumí FAIR principům víc než někteří lidé, nebo že FAIR principy vědeckých dat jsou v podstatě obrácenou stranou AI (FAIR is the flipside of AI). Findable, Accessible, Interoperable, Reusable – tedy nalezitelnost, přístupnost, interoperabilita a znovupoužitelnost dat jsou klíčovými doplňkem a podmínkou pro efektivní využití umělé inteligence.

Všechna data splňující principy FAIR jsou v podstatě 'AI ready'. AI a FAIR jsou propojené, AI je efektivní pouze tehdy, když má přístup k FAIR datům, a FAIR principy maximalizují hodnotu dat právě pro AI aplikace.

Nejefektivnější uživatelé EOSCu budou nakonec algoritmy, takže je důležité, aby byly transparentní. Je potřeba upřít pozornost na GeoGPT, velký jazykový model, do něhož hodně investuje Čína, a který řeší geospatialní úlohy, analyzuje satelitní snímky, lokalizuje místa nebo generuje mapy s vizualizací dat.

<sup>3</sup> <https://eosc.eu/eosc-federation-handbook/>

Hodně se diskutovalo o potřebě určitých mantinelů a zásad při využívání AI. Uzavřené AI modely mohou být dokonale FAIR, nicméně je otázkou, zda mají být otevřené (open). Je důležité si uvědomit, že FAIR není to stejné jako open. Obavy ze zneužití AI jsou na místě.

## No data without software

Jedna z paralelních sekcí, které jsem se zúčastnila, se věnovala důležitosti softwaru ve vědeckém výzkumu.

Byla to velice interaktivní session, kde všichni zúčastnění odpovídali na různé otázky nebo hlasovali pomocí aplikace Slido. Díky tomu jsme získali jasnou představu, kdo jsou naši kolegové v sále. Velká část z nich byli samozřejmě vědci z různých oborů, plus také počítačovní vědci, ale v sále bylo i nezanedbatelné množství knihovníků a informačních pracovníků.

Software je dnes klíčovým nástrojem vědy. Bez něj by nebylo možné data získávat, analyzovat ani uchovávat. Každý vědec potřebuje software nejen k uchování dat, ale i k výpočtům, měřením či simulacím.

Bylo zdůrazněno, že vědci často potřebují programovat vlastní software pro zpracování dat. Diskutovalo se o problémech s udržitelností a dokumentací těchto softwarů. Je nezbytné, aby software byl kvalitní, tzn. dobře dokumentovaný, a to nejen pro uživatele, ale i programátory. Pravidelné aktualizace zajišťují jeho udržitelnost, což je zásadní, protože software, stejně jako lidé, stárne. Dozvěděla jsem se zde o iniciativě Software Heritage<sup>4</sup>, která za podpory UNESCO poskytuje služby pro archivaci a odkazování na historický i současný software se zaměřením na zdrojové kódy čitelné pro lidi. Sice je důležité uchovávat vědecká data, ale pokud k těmto datům neexistuje software, který je původně vytvořil, může být problém je znovu využít nebo ověřit. Mnoho vědců v sále tak potěšila zcela nová informace, že pokud uloží software spolu s daty do ZENODO, tak sw se automaticky uloží též do Software Heritage.

Zazněl apel na citování softwaru ve vědeckých pracích. Řešila se otázka, kdo je vlastníkem softwaru vyvinutého in-house. Obvykle je vlastníkem instituce, ve které vědec software vyvinul, zatímco vědec je jeho autorem. Je důležité znát a rozlišovat pojmy autorské právo (copyright), duševní vlastnictví (intellectual property) a licence. Je podstatné tyto pojmy definovat a určit, jak se s výzkumným softwarem nakládá, aby byla zajištěna ochrana práv a možnost sdílení softwaru či jeho dalšího využití.

## Závěr

Symposium EOSC 2024 přineslo významný milník v podobě oficiálního spuštění EOSC EU Node a potvrdilo důležitou roli tohoto uzlu jako základního stavebního kamene EOSC Federace, která směřuje k vytvoření jednotného digitálního výzkumného prostředí napříč Evropou. Téma dostupnosti dat, přirovnávané k ropě a zlatu 21. století, podtrhuje jejich strategický význam jako klíčového zdroje pro inovace a konkurenceschopnost.

Ačkoliv je diskuse o vědeckých datech v naší komunitě už nějaký čas přítomná, osobní účast na této konferenci mi umožnila uvědomit si, že jsme nejen svědky, ale přímo účastníky vzniku celého nového odvětví. Mění se také přístup ke správě a popisu nejen vědeckých dat, a to od tradiční bibliografické perspektivy k inženýrskému přístupu. Ta změna je opravdu zásadní a klade na všechny zúčastněné úplně nové nároky. Dnes už nestačí být pouze vědcem nebo knihovníkem – vzniká potřeba větší symbiózy mezi těmito obory.

---

<sup>4</sup> <https://www.softwareheritage.org/>