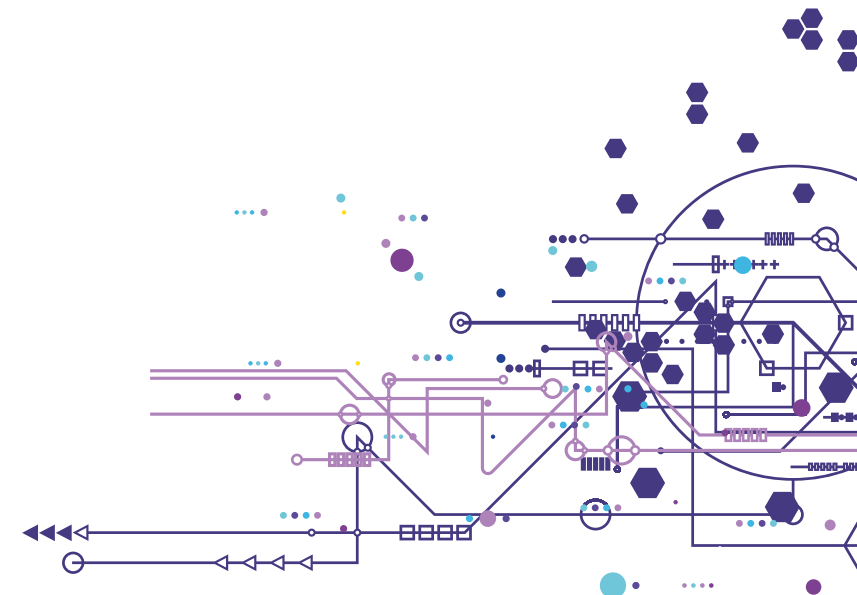
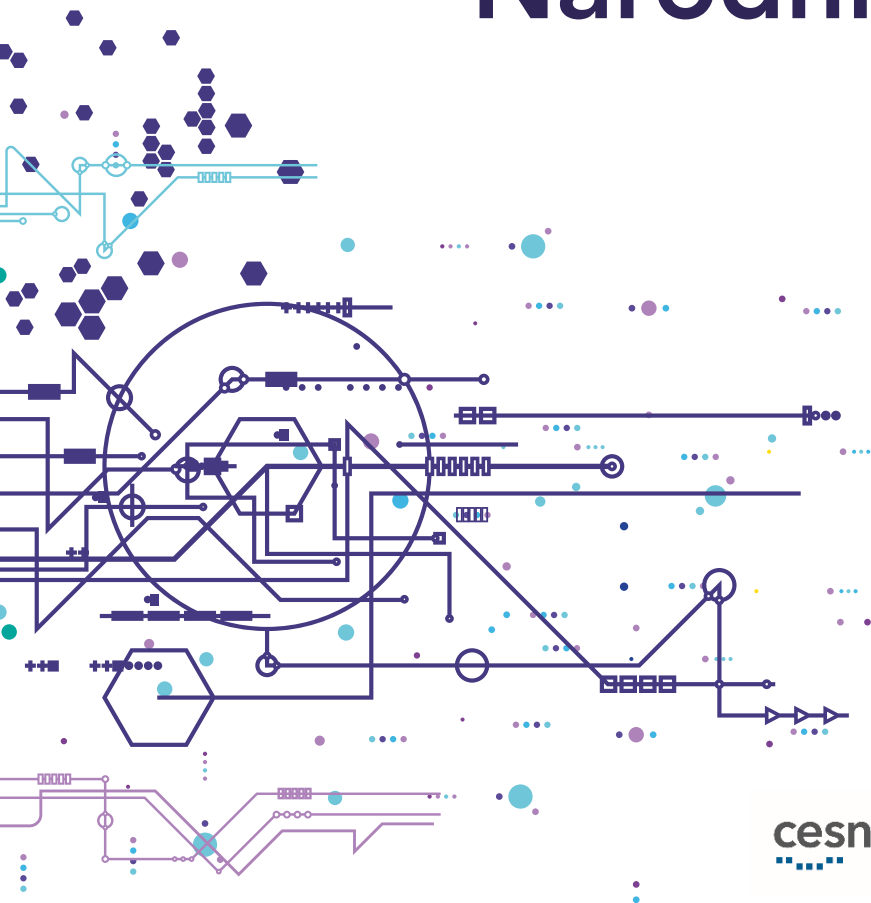


Příprava projektu Národní repozitářové platformy

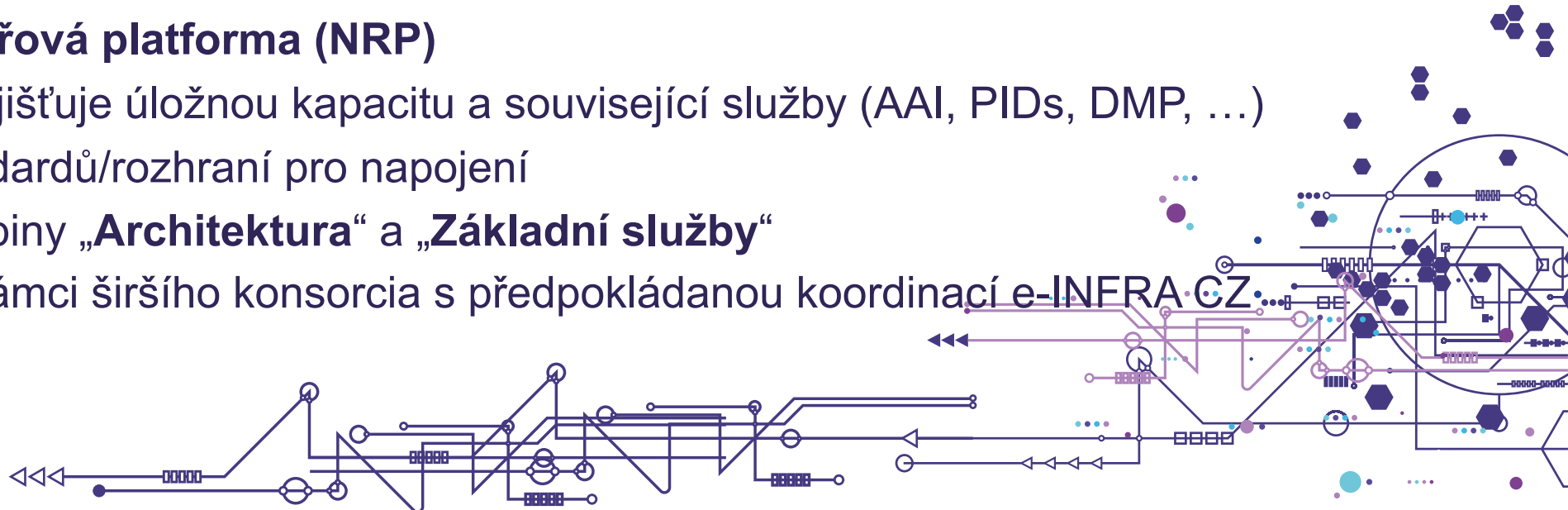
Luděk Matyska

e-INFRA CZ



Klíčové komponenty

- **Národní metadatový adresář (NMA)**
 - Klíčová komponenta, zajišťující vyhledatelnost a interoperabilitu dat
 - Pracovní skupina „**Metadata**“
 - Vazba na všechny ostatní aktivity v rámci NDI
 - Centrálně spravovaná komponenta
 - NTK (IPs CARDS) a e-INFRA CZ (IPs EOSC CZ)
- **Národní repozitářová platforma (NRP)**
 - Páteř NDI, zajišťuje úložnou kapacitu a související služby (AAI, PIDs, DMP, ...)
 - Definice standardů/rozhraní pro napojení
 - Pracovní skupiny „**Architektura**“ a „**Základní služby**“
 - Realizace v rámci širšího konsorcia s předpokládanou koordinací e-INFRA CZ (CESNET)



Národní datová infrastruktura (NDI)

Národní metadatový katalog (NMA)

Národní repozitářová
platforma (NRP)

Služby NDI

Tematické
repozitáře

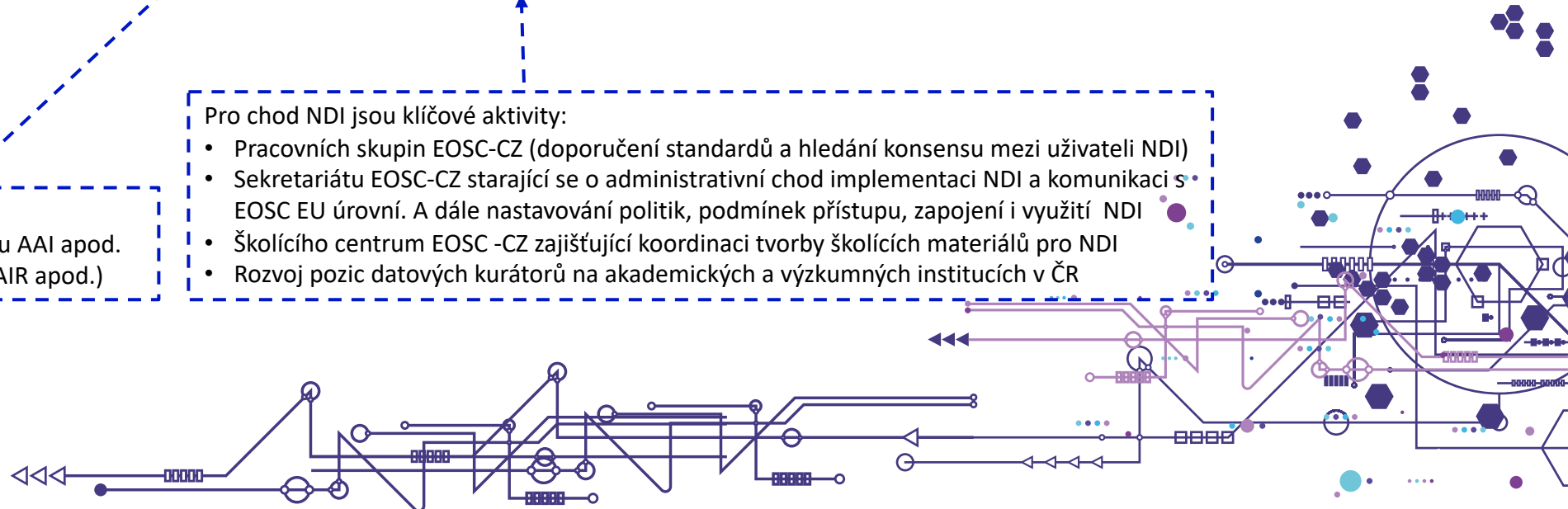
Ostatní
repozitáře

Služby NDI jsou dvojího typu:

- Základní technické služby typu AAI apod.
- EOSC-CZ Core služby (DMP, FAIR apod.)

Pro chod NDI jsou klíčové aktivity:

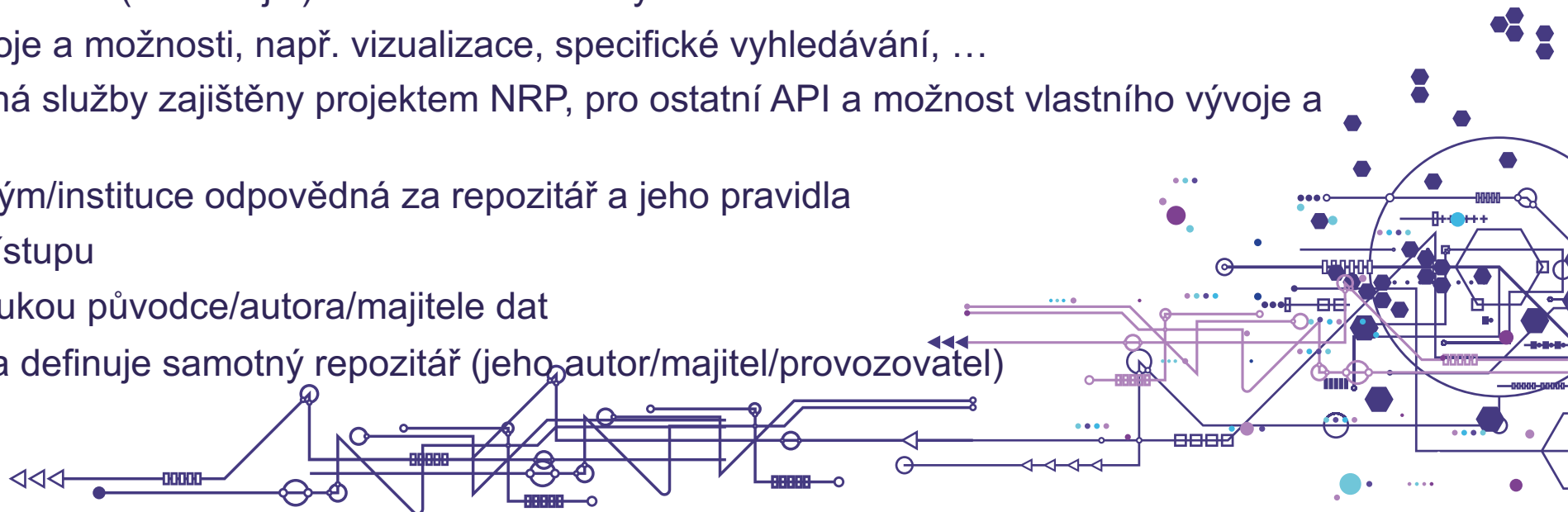
- Pracovních skupin EOSC-CZ (doporučení standardů a hledání konsensu mezi uživateli NDI)
- Sekretariátu EOSC-CZ starající se o administrativní chod implementaci NDI a komunikaci s EOSC EU úrovní. A dále nastavování politik, podmínek přístupu, zapojení i využití NDI
- Školícího centrum EOSC -CZ zajišťující koordinaci tvorby školících materiálů pro NDI
- Rozvoj pozic datových kurátorů na akademických a výzkumných institucích v ČR



Národní repozitářová platforma

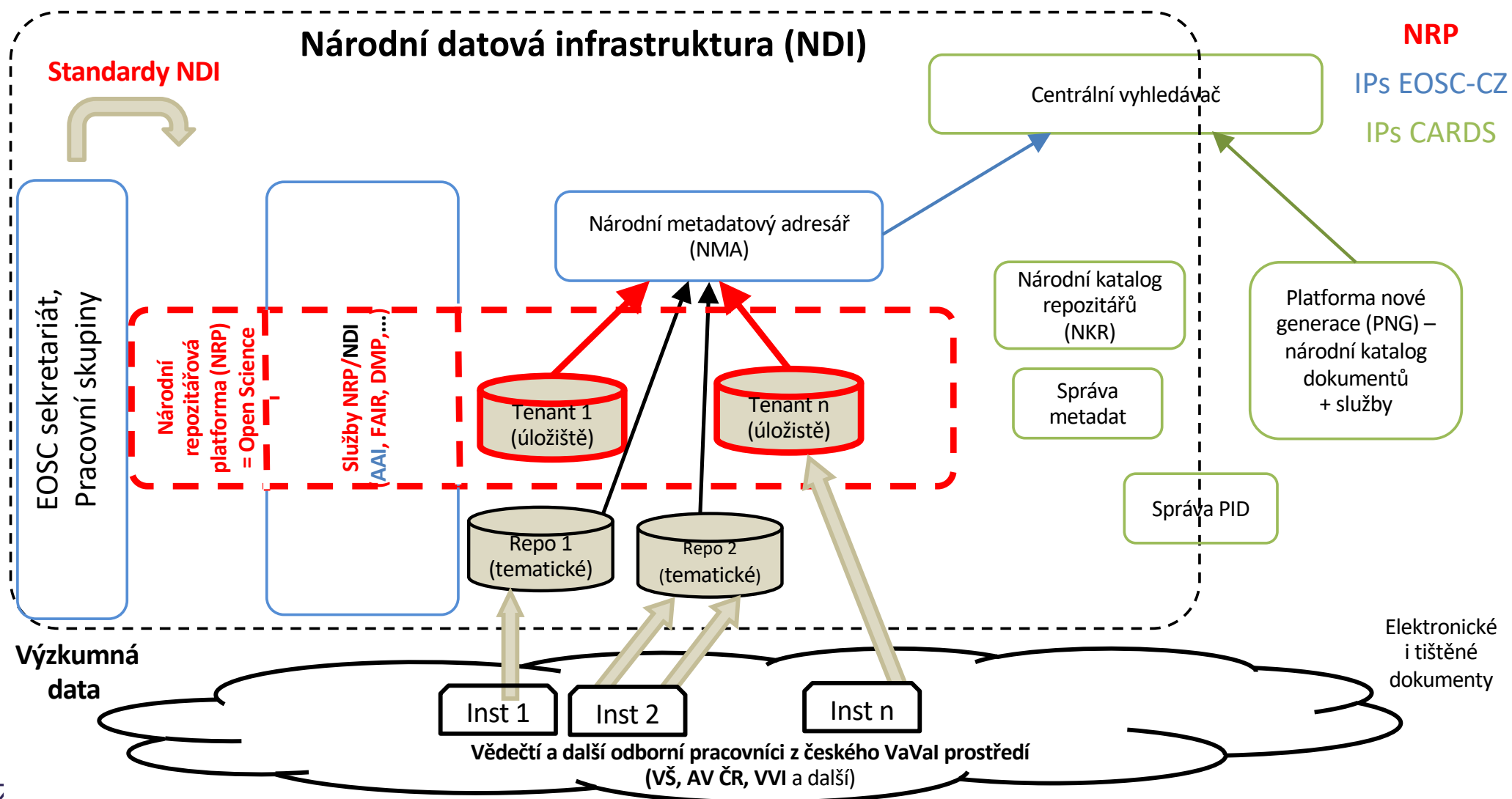
K architektuře NRP

- Systém pro vytváření instancí (tenantů) repozitářů
 - Distribuovaná mozaika (patchwork) fyzických úložišť
 - Vlastní úložiště budovaná vybranými institucemi, možnost rozšíření
 - Silné zajištění dat, s kontrolami binární konzistence, fyzická redundance uložených dat
- Tenanty (konkrétní repozitáře)
 - Branding, vlastní vzhled
 - Podpora pro specifická (detailnější) metadatové modely
 - Specifické nástroje a možnosti, např. vizualizace, specifické vyhledávání, ...
 - Jen společná služby zajištěny projektem NRP, pro ostatní API a možnost vlastního vývoje a integrace
 - Vždy konkrétní tým/instituce odpovědná za repozitář a jeho pravidla
- Klíčová role řízení přístupu
 - Plná kontrola v rukou původce/autora/majitele dat
 - Základní pravidla definuje samotný repozitář (jeho autor/majitel/provozovatel)



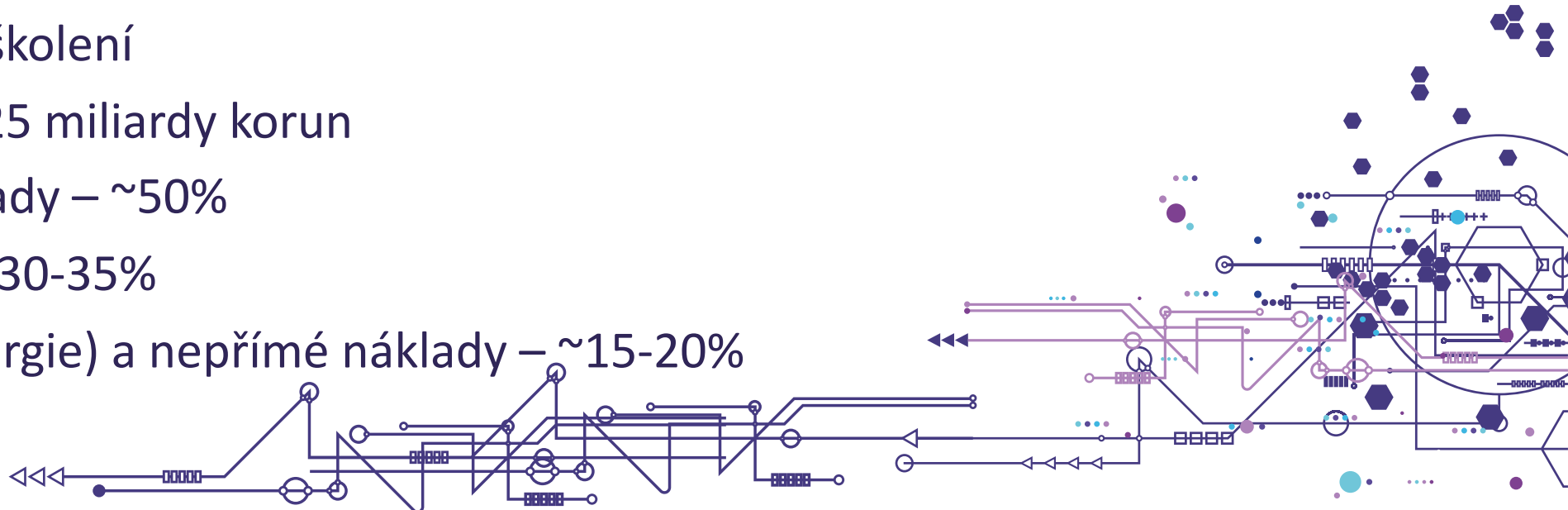
Národní repozitářová platforma (NRP)

Struktura intervence Open Science OP JAK (kontext **výzvy NRP**)



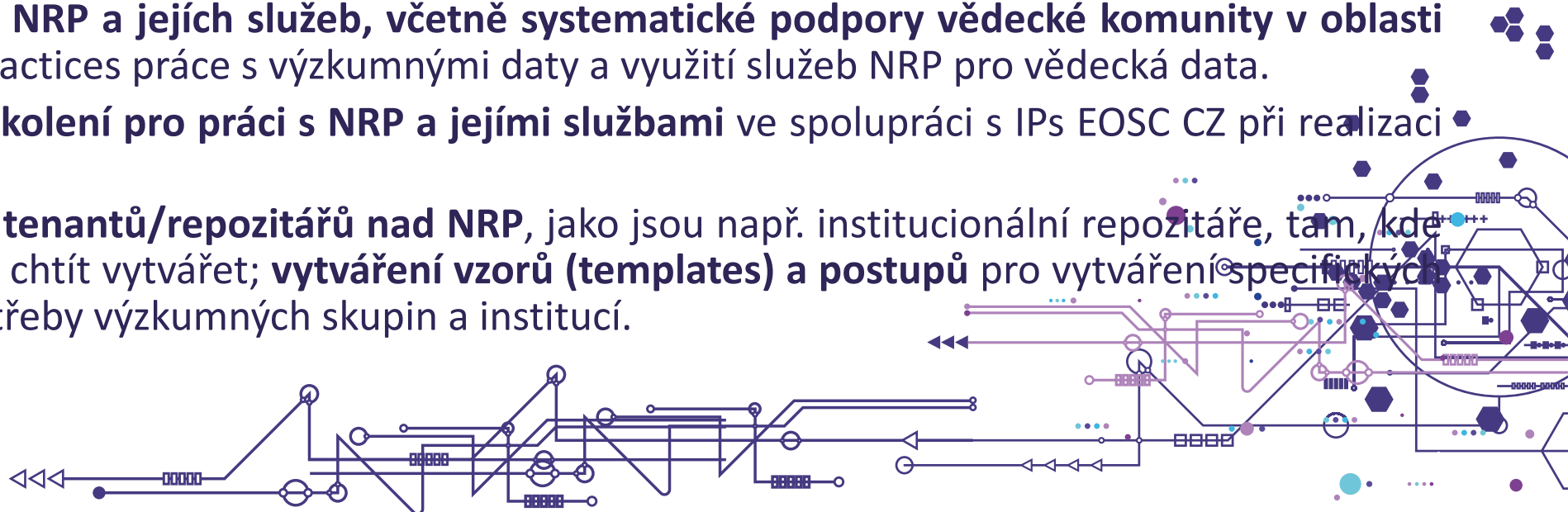
Společný projekt NRP

- Konsorciální projekt, cca 10+ partnerů
 - Koordinace Pověřenou organizací EOSC Asociace (CESNET)
- Povinná spolupráce s IPs EOSC CZ a CARDS
 - Napojení na NMA
 - Dodržování standardů
 - Koordinace školení
- Rozpočet cca 1,25 miliardy korun
 - Osobní náklady – ~50%
 - Investice – ~30-35%
 - Služby (+energie) a nepřímé náklady – ~15-20%



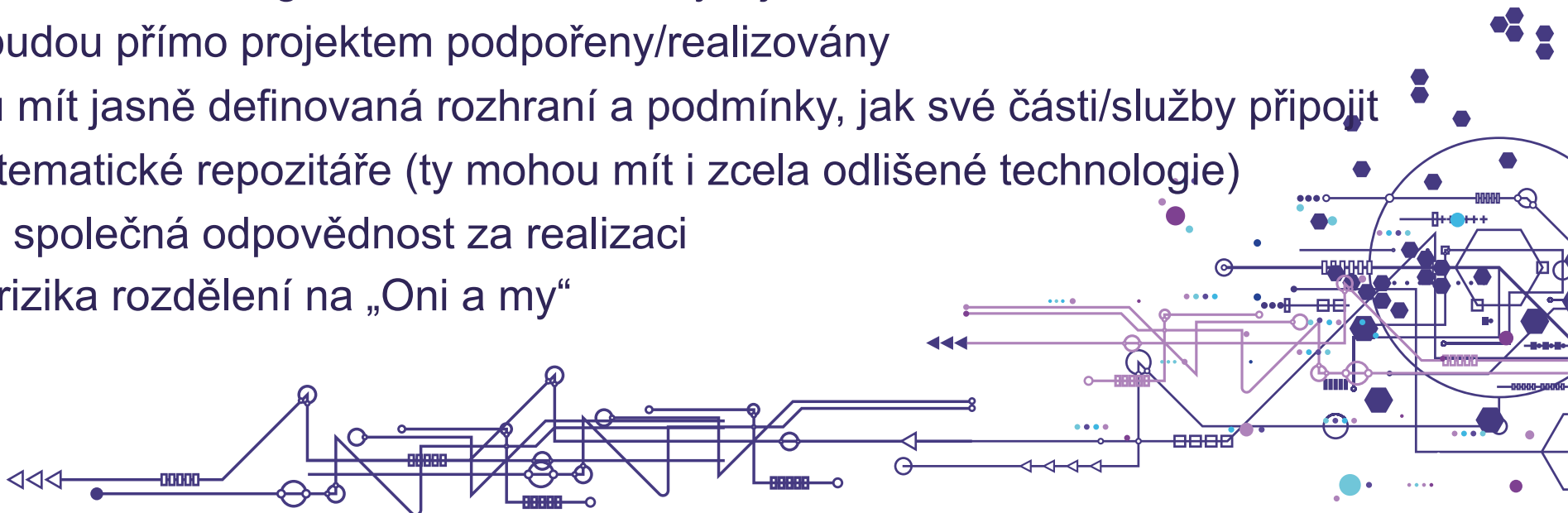
Projekt NRP – Očekávané výstupy

1. **Architektura Národní repozitářové platformy (NRP)** jako podklad pro její budování.
2. Kontinuálně rozvíjený **soubor služeb** s roadmapou jejich postupné implementace a zavádění do provozu.
3. **Národní repozitářová platforma** se základními i pokročilými službami, včetně odpovídající distribuované úložné kapacity a systému řízení přístupu k datům v NRP.
4. **Integrace klíčových aktérů datově-orientovaných akademických institucí a výzkumných týmů pro FAIR data** při budování jádra Národní datové Infrastruktury (NDI).
5. **Dlouhodobý provoz NRP a jejích služeb, včetně systematické podpory vědecké komunity v oblasti vzdělávání** o best-practices práce s výzkumnými daty a využití služeb NRP pro vědecká data.
6. **Školící materiály a školení pro práci s NRP a jejími službami** ve spolupráci s IPs EOSC CZ při realizaci školení.
7. **Tvorba konkrétních tenantů/repozitářů nad NRP**, jako jsou např. institucionální repozitáře, tam, kde si je instituce budou chtít vytvářet; **vytváření vzorů (templates) a postupů** pro vytváření specifických repozitářů podle potřeby výzkumných skupin a institucí.



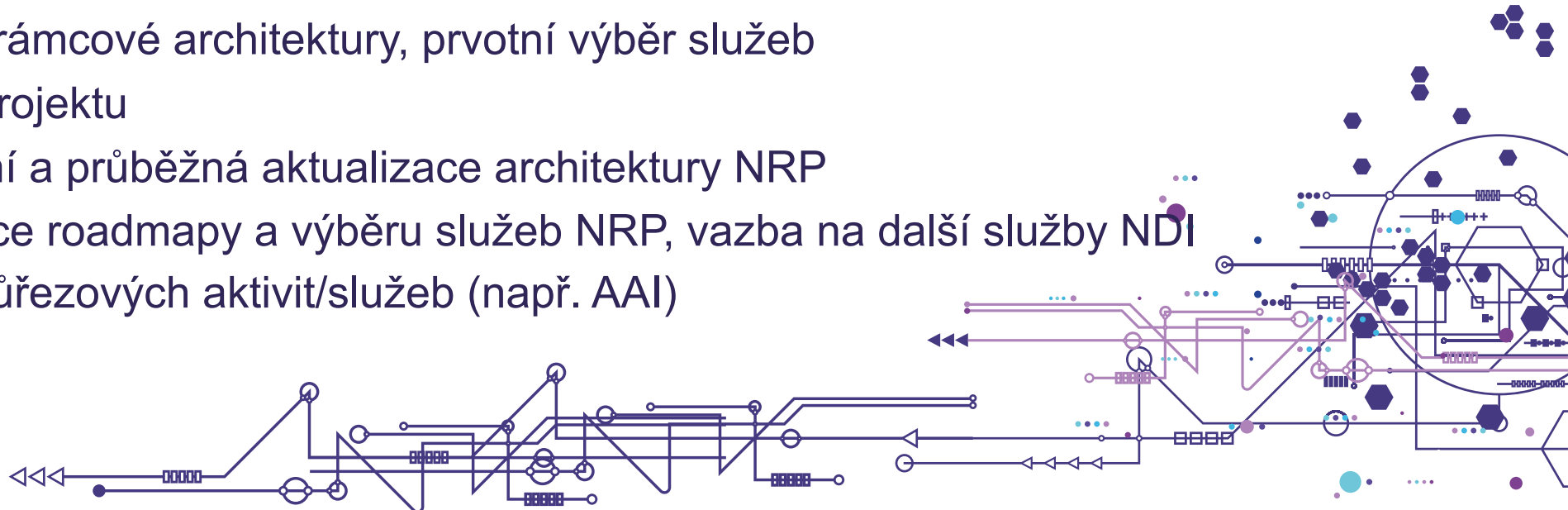
Základní principy projektu

- Definovat, implementovat a dlouhodobě provozovat NRP a její společné služby
- Spolupráce na definici akceptovatelné architektury a služeb NRP
 - Odpovědnost nejen za návrh, ale v rámci projektu i za realizaci
 - Výběr jen těch komponent/služeb, které mají široký/průřezový dopad
 - Tematicky specifické jsou úkolem komunit a souvisejících projektů
- Otevřený koncept, více technologií, možnost růstu a vývoje
 - Jen vybrané budou přímo projektem podpořeny/realizovány
 - Ostatní budou mít jasně definovaná rozhraní a podmínky, jak své části/služby připojit
 - Rozhraní pro tematické repozitáře (ty mohou mít i zcela odlišené technologie)
- Společná tvorba a společná odpovědnost za realizaci
 - Minimalizace rizika rozdělení na „Oni a my“



Projekt NRP a tematické pracovní skupiny

- Příprava na projekty do výzvy Open Science II
- NRP otevřený koncept, který umožní interoperabilitu s tematickými repozitáři
 - Nezbytný vstup z tematických PS
 - Ty přitom budou mít možnost podání projektů do další výzvy (2024)
- Spolupráce se zástupci tematických PS
 - Při přípravě projektu
 - vstup do rámcové architektury, prvotní výběr služeb
 - Při realizaci projektu
 - dokončení a průběžná aktualizace architektury NRP
 - aktualizace roadmapy a výběru služeb NRP, vazba na další služby NDI
 - využití průřezových aktivit/služeb (např. AAI)



Struktura projektu NRP – část 1

KA 1 Řízení projektu (koordinátor CESNET/e-INFRA CZ)

KA 1.1 Řízení a administrativa projektu

KA 1.2 Dlouhodobá udržitelnost NRP jako součást NDI

KA 2 Placeholder

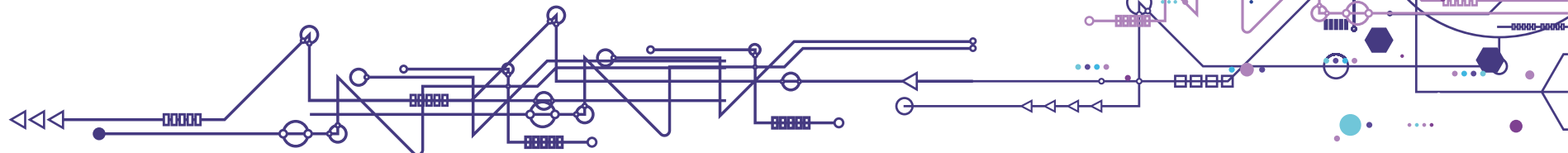
KA 3 Architektura, vývoj a provoz NRP (předpokládaná koordinace e-INFRA CZ) – *PS Architektura*

KA 3.1 Architektura NDI se specifickým důrazem na NRP

KA 3.2 Vývoj a provoz základních softwarových systémů pro NRP

KA 3.3 Pokročilá technická podpora („level 3“) pro vytváření tenantů NRP

KA 3.4 Pořizování hardware / veřejné zakázky



Struktura projektu NRP – část 2

KA 4 Integrace NRP (předpokládaná koordinace e-INFRA CZ) – spolupráce *PS Základní služby* s *PS Architektura* a *PS Metadata*

KA 4.1 Integrace s NMA

KA 4.2 Práce s PID

KA 4.3 Práce s licencemi

KA 4.4 Service desk včetně monitorování služeb a infrastruktury

KA 4.5 Řízení přístupu k datům

KA 5 Vývoj pokročilých služeb NRP – *PS Základní služby*, konzultace s *PS Architektura*

KA5.1 Podpora FAIRifikace vědeckých dat

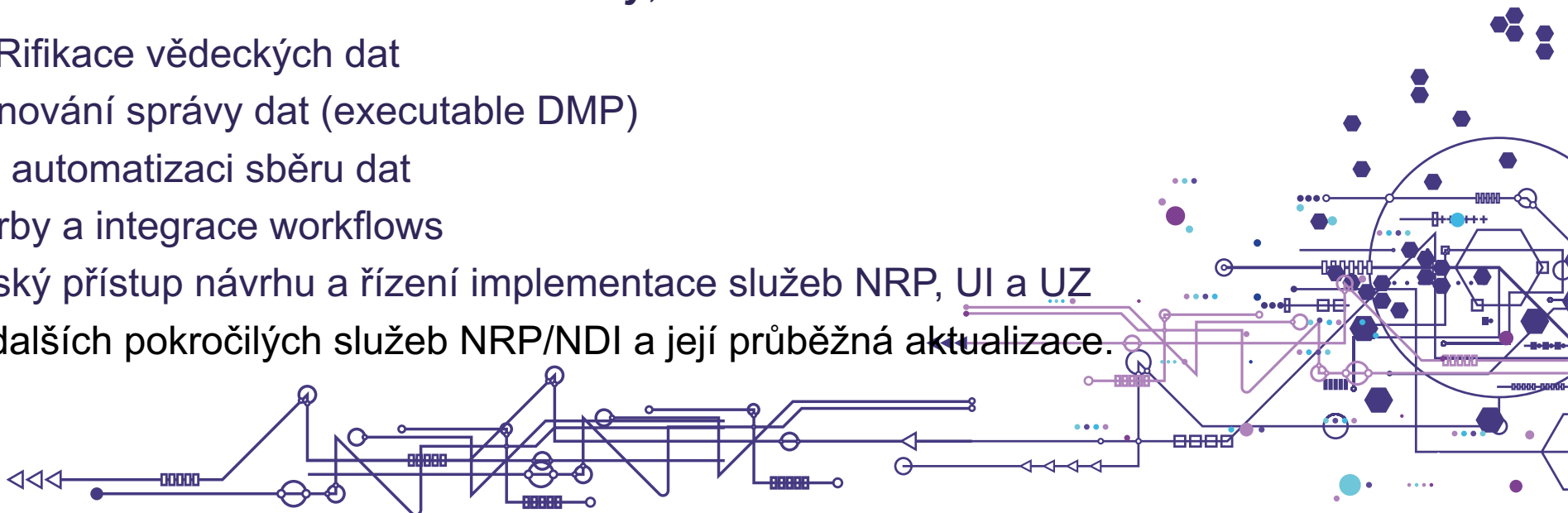
KA 5.2 Podpora plánování správy dat (executable DMP)

KA 5.3 Příprava pro automatizaci sběru dat

KA 5.4 Podpora tvorby a integrace workflows

KA 5.4 Pro-uživatelský přístup návrhu a řízení implementace služeb NRP, UI a UZ

KA 5.5 Roadmapa dalších pokročilých služeb NRP/NDI a její průběžná aktualizace.



Struktura projektu NRP – část 3

KA 6 Průřezové a specifické požadavky a služby

KA 6.1 Certifikace repozitářů

KA 6.2 Prostředí a podmínky pro citlivá data

KA 6.3 Kyberbezpečnost

KA 6.4 Právní podpora

KA 7 Podpora uživatelských komunit, trénink, vzdělávání a udržitelnost – *PS Vzdělávání*

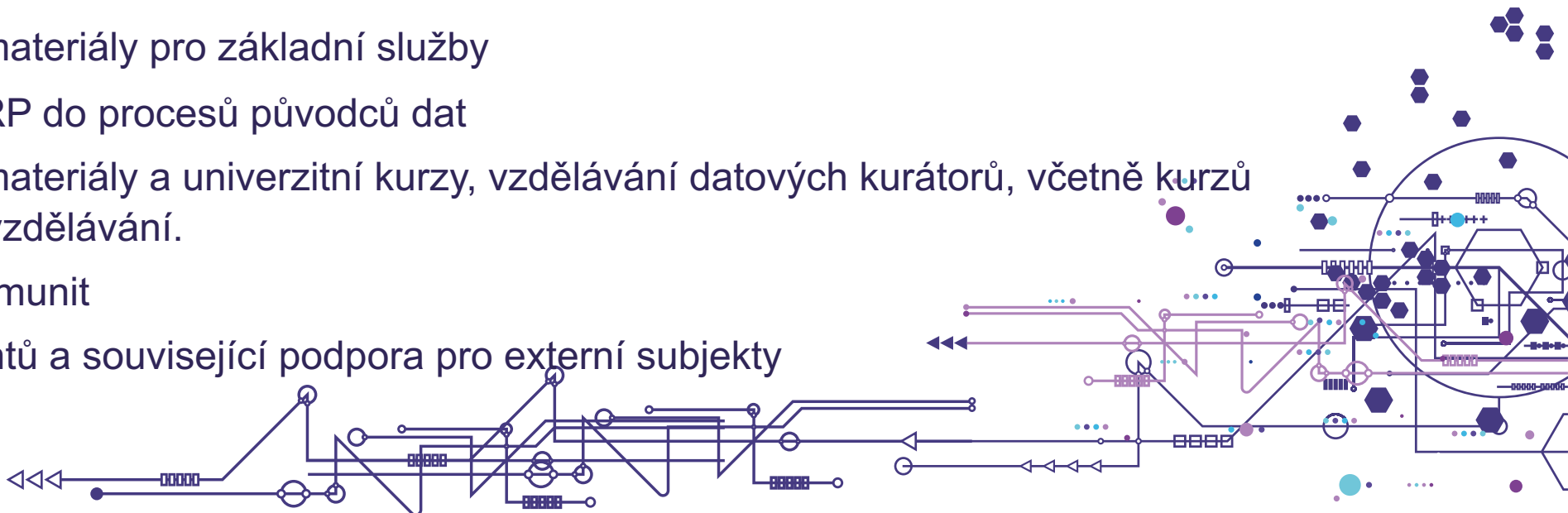
KA 7.1 Vzdělávací materiály pro základní služby

KA 7.2 Integrace NRP do procesů původců dat

KA 7.3 Tréninkové materiály a univerzitní kurzy, vzdělávání datových kurátorů, včetně kurzů celoživotního vzdělávání.

KA 7.4 Budování komunit

KA 7.5 Tvorba tenantů a související podpora pro externí subjekty



Proces přípravy NRP

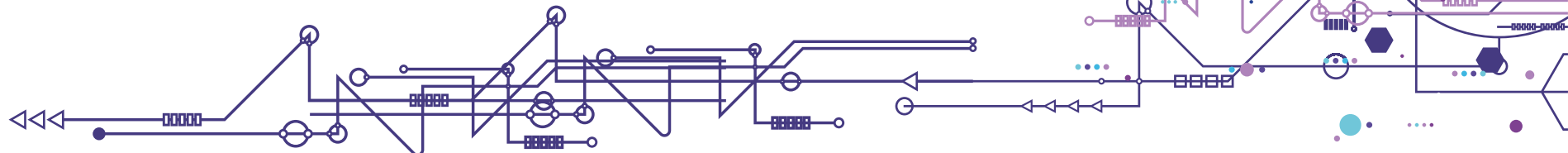
- **Přípravný výbor** (přelom 2022/2023)

- Oficiální zástupci přizvaných institucí
- Členové Asociace výzkumných univerzit a ZČU a VŠB-TUO (člen e-INFRA CZ)
- Zástupci AV ČR – KNAV, Středisko sdílených služeb, 4 ústavy AV ČR (koordinace KNAV)
- Členové E-INFRA CZ, CESNET jako koordinátor projektu NRP
- NTK (člen bez finančního příspěvku)

2 schůzky, diskuse záměru projektu, podmínek pro členy budoucího konsorcia, ...

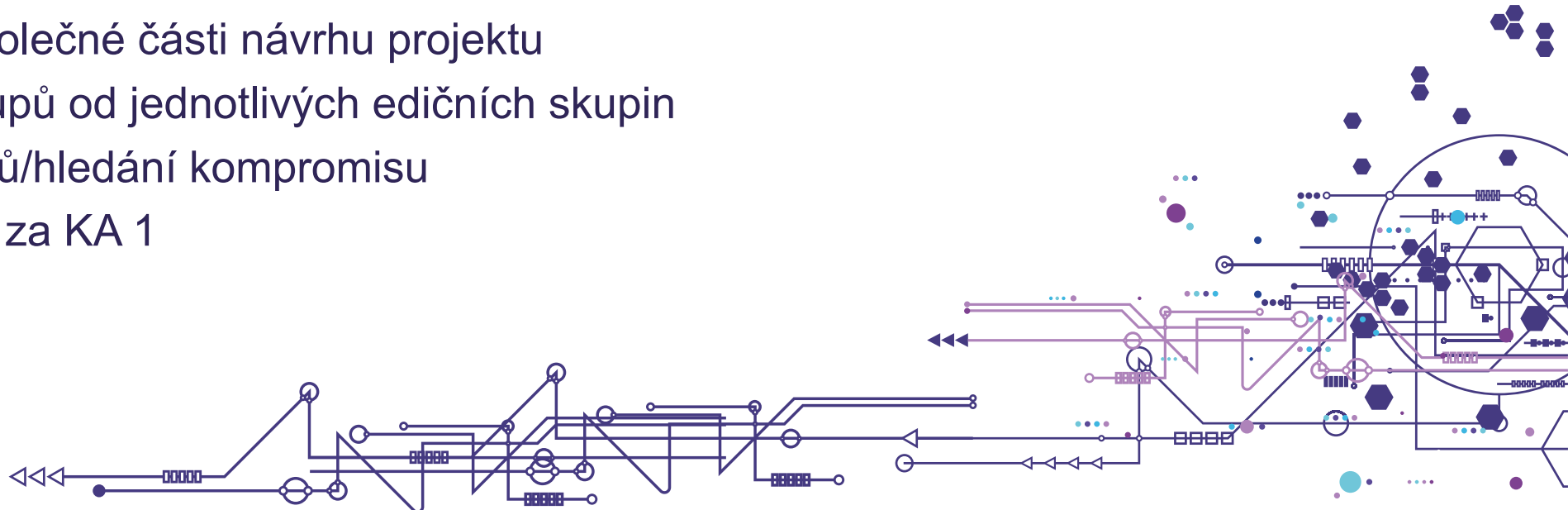
- **Ediční tým**

- Koordinace CESNET (LM), jádro z e-INFRA CZ
- **Jednotlivé ediční skupiny** ve vazbě na základní Pracovní skupiny EOSC
- Zapojení tematických pracovních skupin



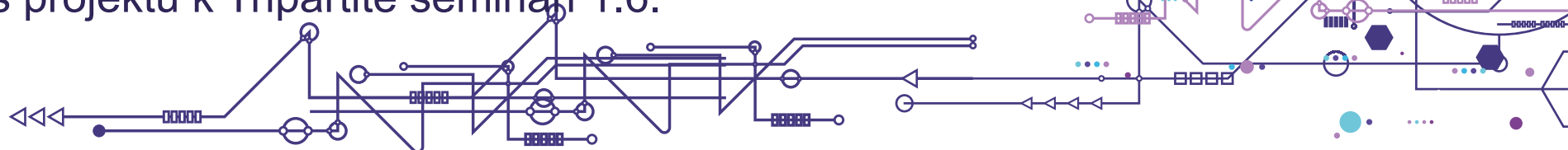
Úkoly edičních skupin

- Konkretizace jednotlivých Klíčových aktivit
 - Konkretizace jednotlivých výstupů
 - Identifikace odpovědných týmů/institucí
 - Požadavky na kapacitu (FTE, hardware, software, ...)
 - Vývoj v čase
- Hlavní ediční skupina
 - Základní a společné části návrhu projektu
 - Integrace vstupů od jednotlivých edičních skupin
 - Alokace zdrojů/hledání kompromisu
 - Odpovědnost za KA 1



Shrnutí

- NRP je (společně s NMA) klíčová součást NDI a implementace EOSC v ČR
- Nezbytná široká shoda přes republiku
 - Vstupy přes pracovní skupiny
 - Otevřený přístup přípravy projektu
 - (není připravován konkurenční projekt)
 - Širší konsorcium – realizace vyžaduje aktivní zapojení klíčových institucí
- Nejde o posílení zapojených institucí vůči ostatní
- ale o vybudování a provoz klíčové komponenty Národní datové infrastruktury
 - S tím je spojená i sdílená odpovědnost partnerů konsorcia z provoz a použitelnost
- Práce organizována v edičních skupinách
 - Optimální řešení včetně identifikace odpovědných institucí/týmů
- Konkrétnější popis projektu k Tripartite semináři 1.6.





Děkuji za pozornost

eosc-info@e-infra.cz

Další informace na <https://www.eosc.cz>

A circular logo consisting of two concentric circles, with the text 'e-infra.cz' centered between them.

e-infra.cz