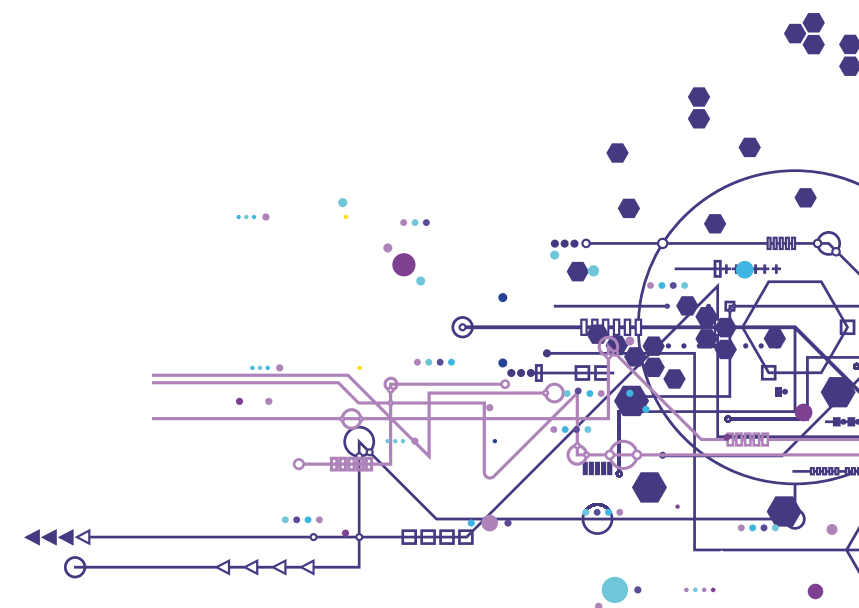
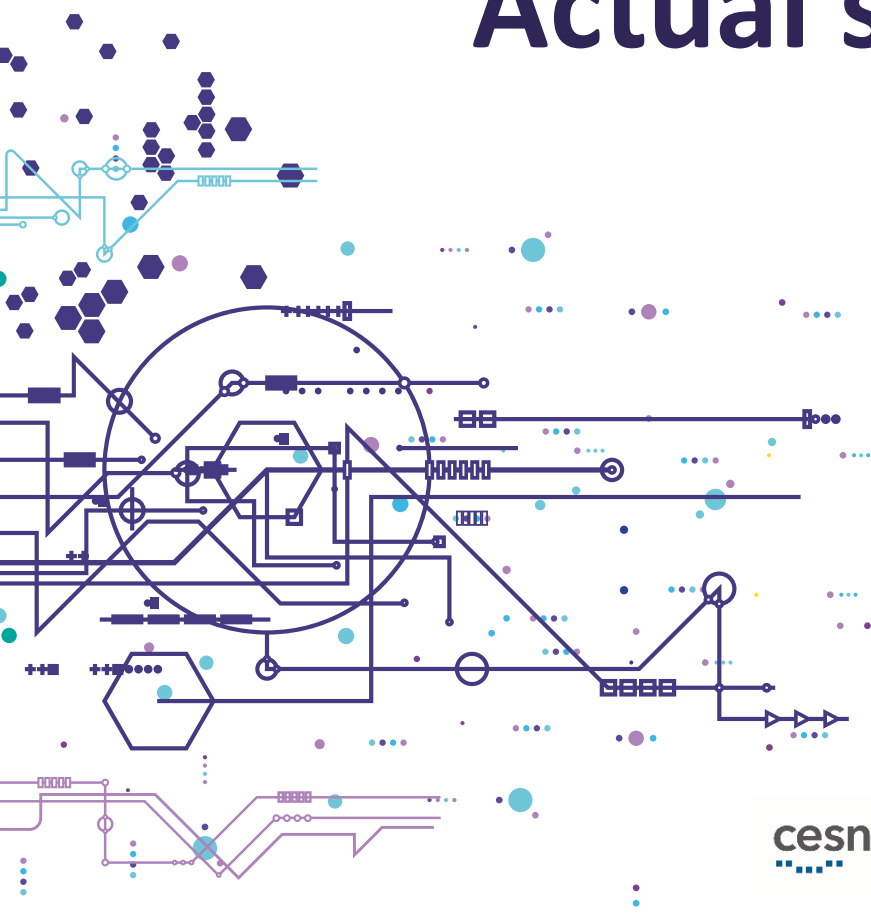


EOSC implementation in Czechia

Actual situation

Luděk Matyska

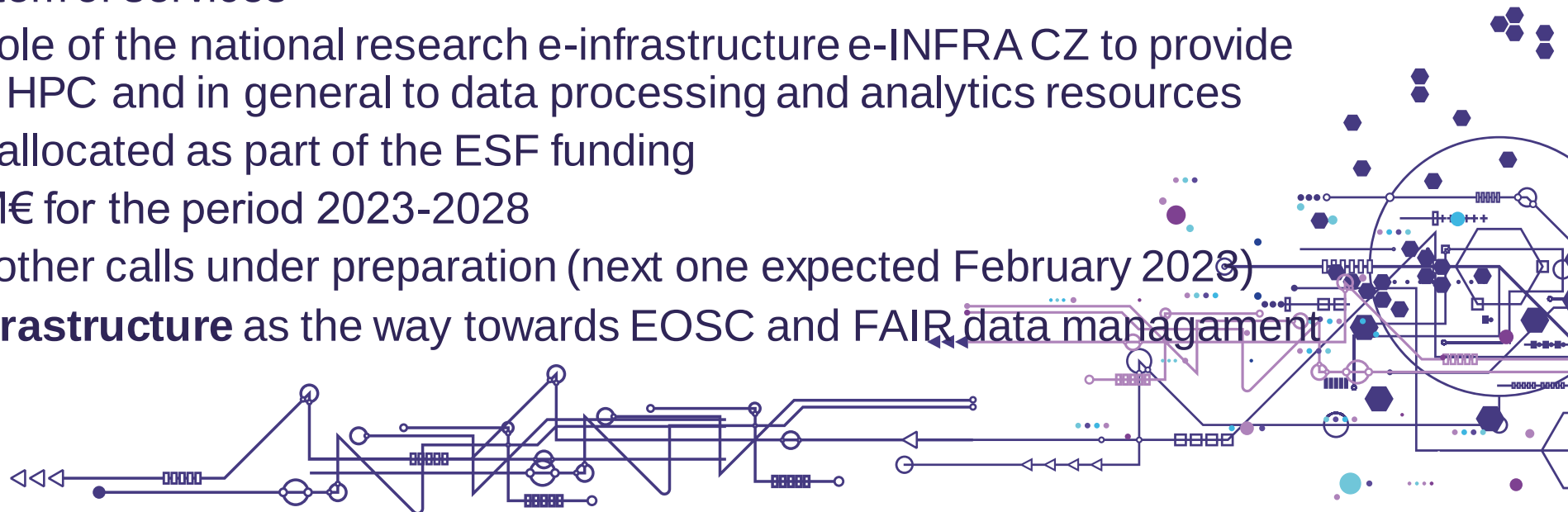
e-INFRA CZ



General overview



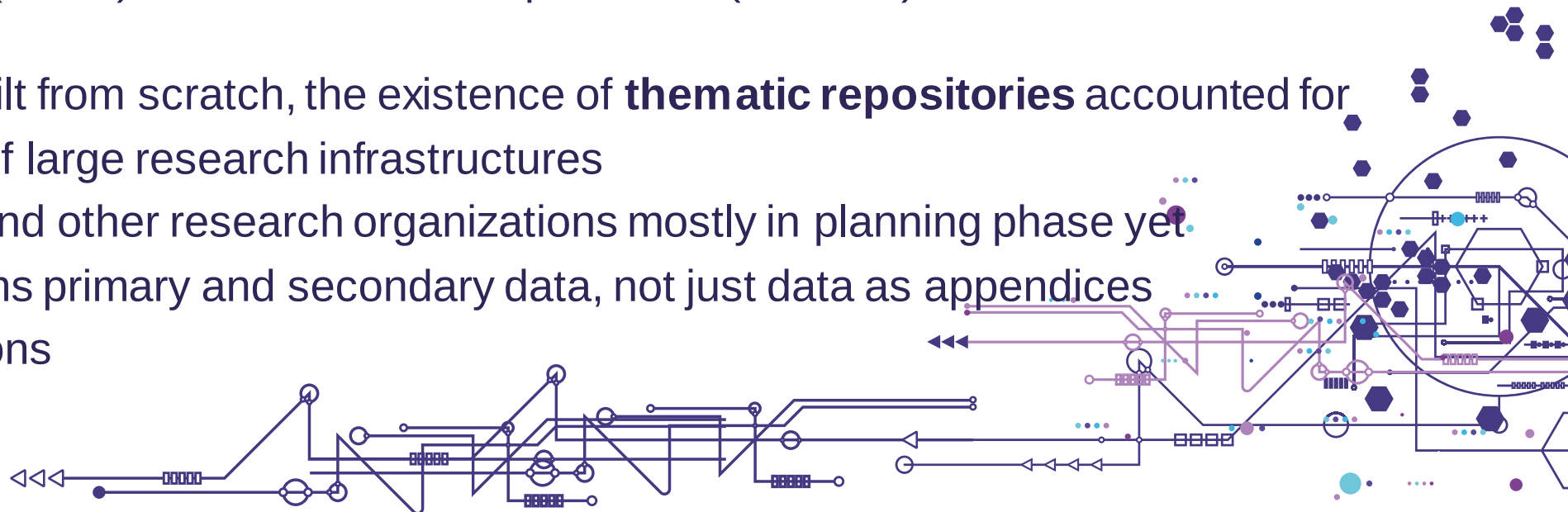
- **Architecture of EOSC implementation in Czechia**
 - Prepared during Spring 2021, accepted by the Ministry in June 2021, approved by the **EOSC Coordination Board** in December 2021
- **Focuses on the data management aspects** as the core for EOSC and Open Science
- EOSC understood as a federation of FAIR data and related services
 - Federation of existing and future data resources (FAIR data repositories)
 - Open ecosystem of services
 - Expects key role of the national research e-infrastructure e-INFRA CZ to provide connection to HPC and in general to data processing and analytics resources
- Financial support allocated as part of the ESF funding
 - Around 120 M€ for the period 2023-2028
 - First call out, other calls under preparation (next one expected February 2023)
- **National Data Infrastructure** as the way towards EOSC and FAIR data management



National Data Infrastructure

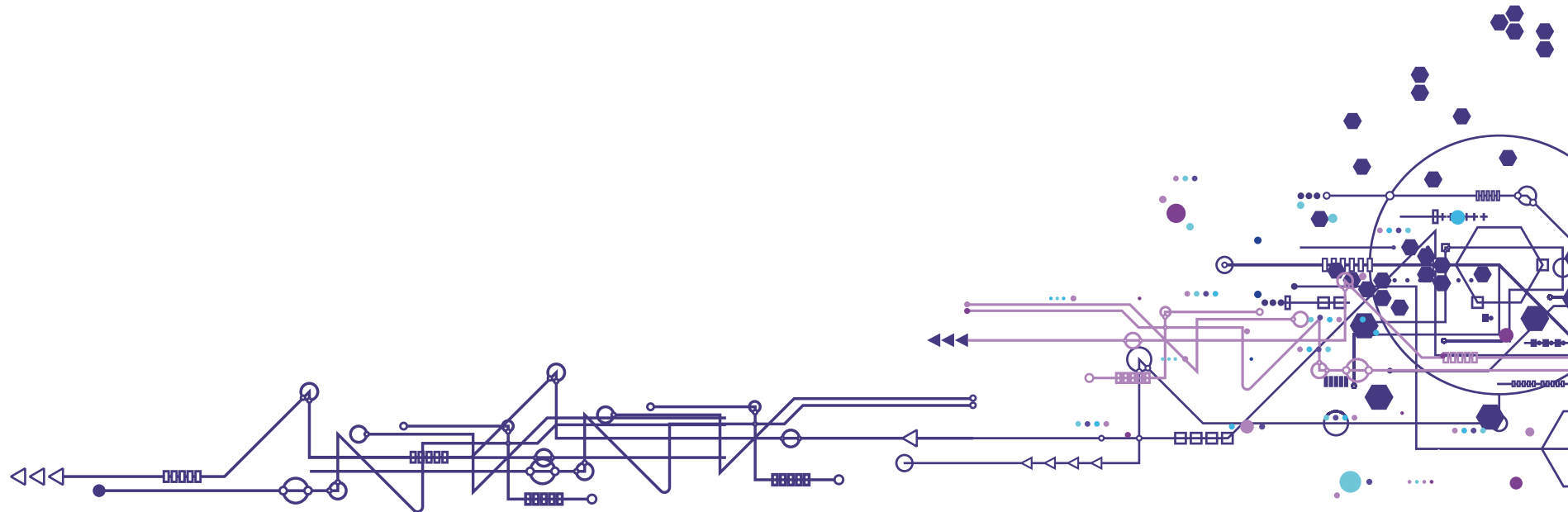


- Complex system with the following aims:
 - Basic **infrastructure** with **sufficient capacity** to store research data in Czechia
 - Open set of **key and other services** for data manipulation
 - Interaction with processing resources (HPC, clouds, ...)
- With parallel support of **human resources**
 - Education of (future) scientists, new competencies (curation), data stewards, data scientists
- Not everything built from scratch, the existence of **thematic repositories** accounted for
 - Usually part of large research infrastructures
 - Universities and other research organizations mostly in planning phase yet
- Data-centric means primary and secondary data, not just data as appendices to publications



4 Key components of NDI

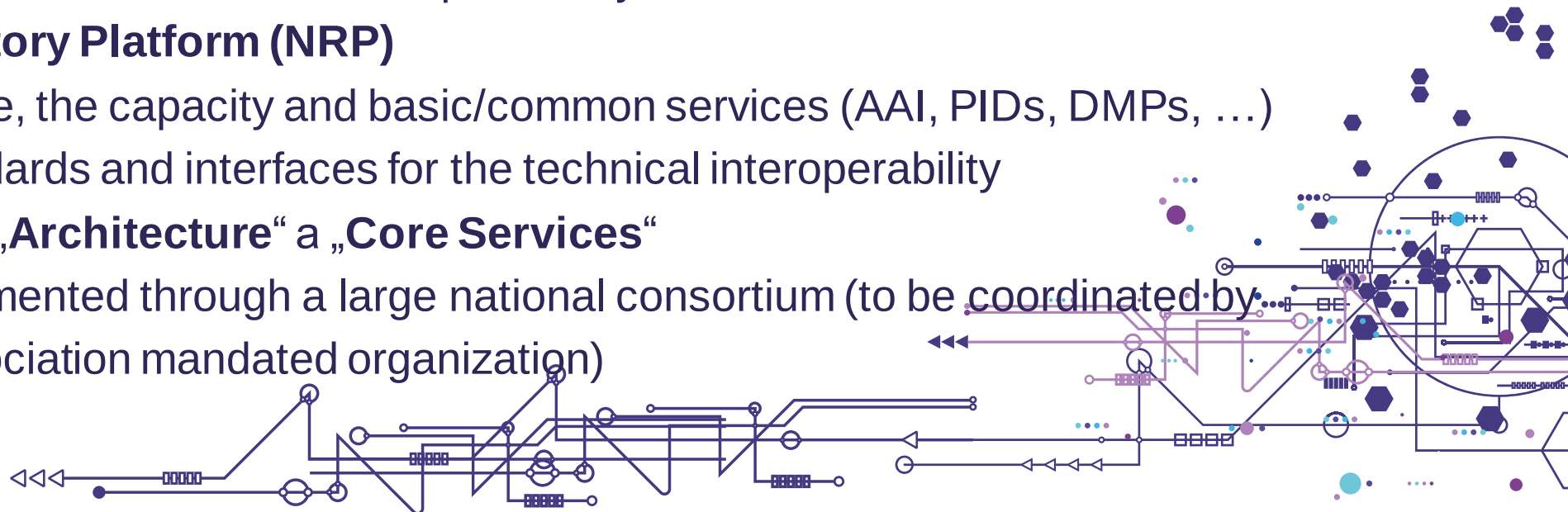
- **National Metadata Directory (NMA)**
- **National Repository Platform (NRP)**
- **Thematic repositories**
- **Education and training (human resources)**



National Data Infrastructure

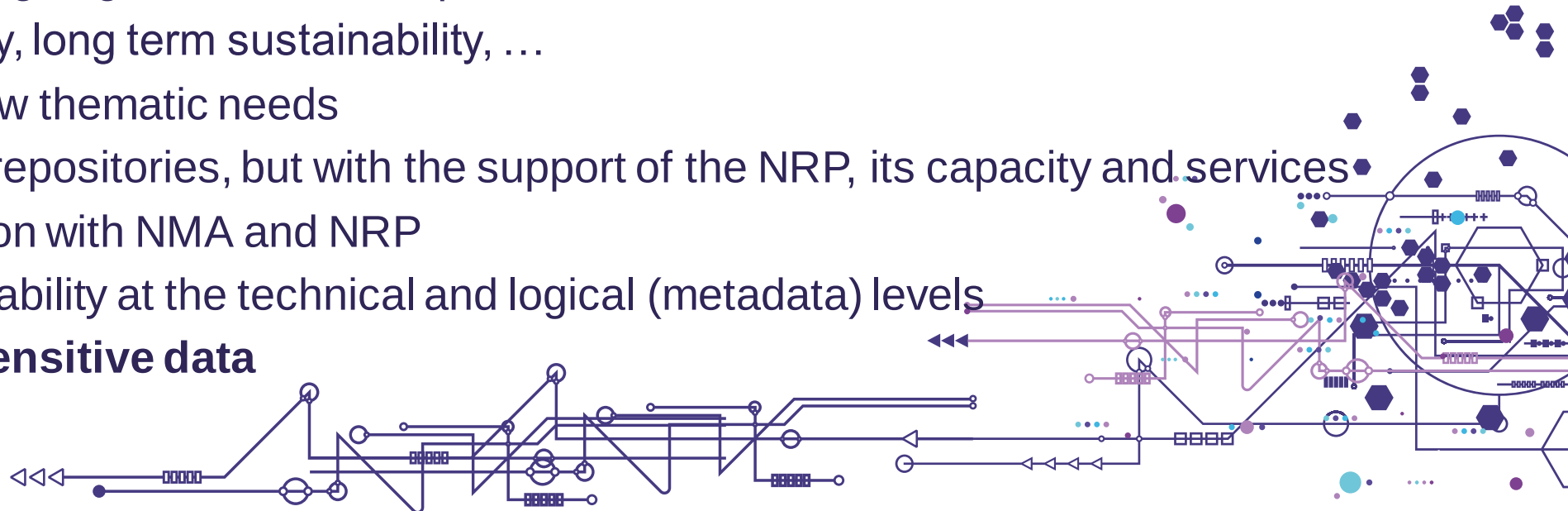
Key components

- **National Metadata Catalogue (NMA)**
 - Key component, guaranteeing findability and interoperability
 - Work group „**Metadata**“
 - Interacting directly with all other NDI components
 - Centrally managed NDI component
 - Still distributed; e-INFRA CZ responsibility
- **National Repository Platform (NRP)**
 - NDI backbone, the capacity and basic/common services (AAI, PIDs, DMPs, ...)
 - Defines standards and interfaces for the technical interoperability
 - Work groups „**Architecture**“ a „**Core Services**“
 - Will be implemented through a large national consortium (to be coordinated by EOSC Association mandated organization)



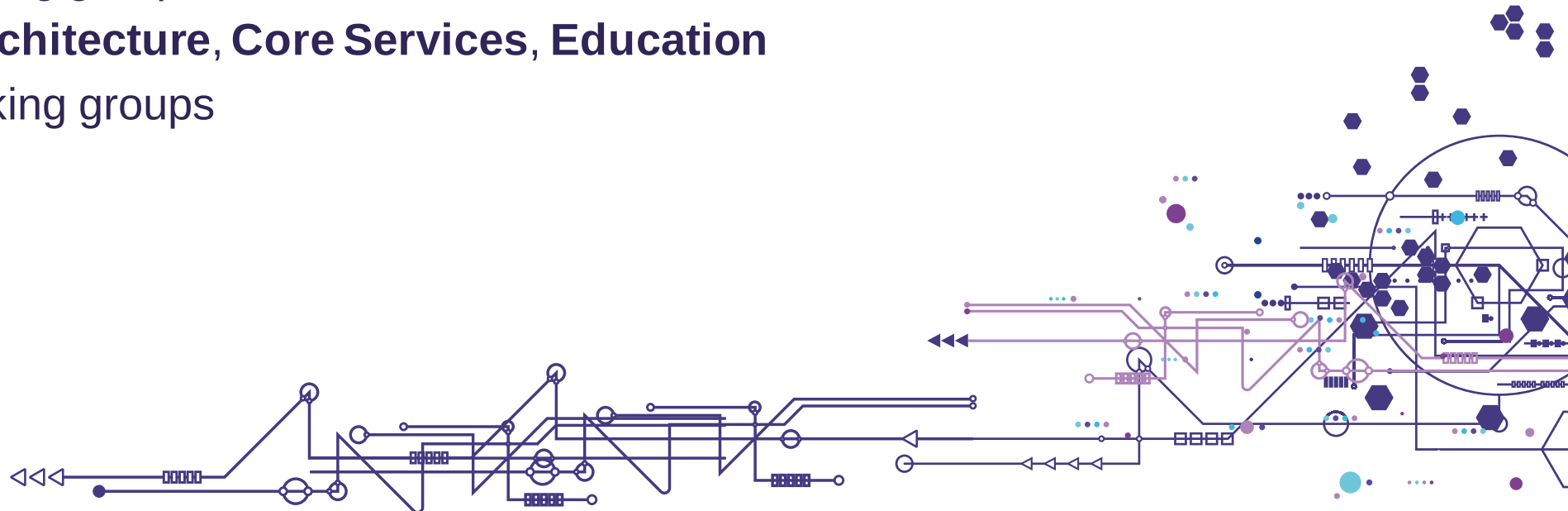
Other parts

- **Human Resources**
 - Horizontal Activity
 - Work group „**Education**“
 - To be supported through the whole implementation of NDI
- **Existing thematic repositories**
 - Support of on-going activities and repositories
 - Continuity, long term sustainability, ...
 - Support of new thematic needs
 - Not new repositories, but with the support of the NRP, its capacity and services
 - Interconnection with NMA and NRP
 - Interoperability at the technical and logical (metadata) levels
- Specific care of **sensitive data**



Preparatory steps

- Working groups established since Autumn 2021
 - Open platform, with fluid membership
 - Discussion of basic principles of all aspects of the EOSC CZ implementation
 - What is needed by researchers and their communities
 - And how to best fulfill these expectations
- 4 foundation working groups
 - **Metadata, Architecture, Core Services, Education**
- 7+1 thematic working groups



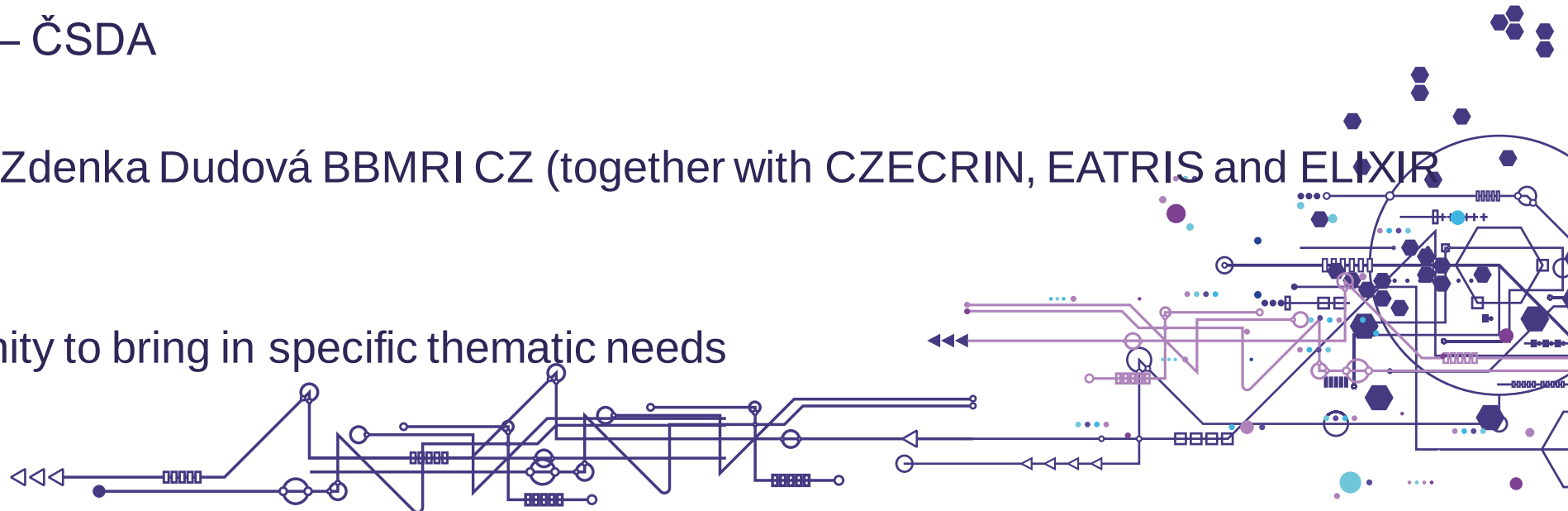
Thematic Working Groups

The leading RI mentioned, if any

- **Bio/Health/Food** – ELIXIR CZ
- **ENVRI** – RECETOX a EIRENE
- **Physics** – CERN
- **Material Sciences and Technology**
- **AI and Digital Science**
- **Humanities and Arts** – LINDAT/CLARIAH-CZ
- **Social Sciences** – ČSDA

and

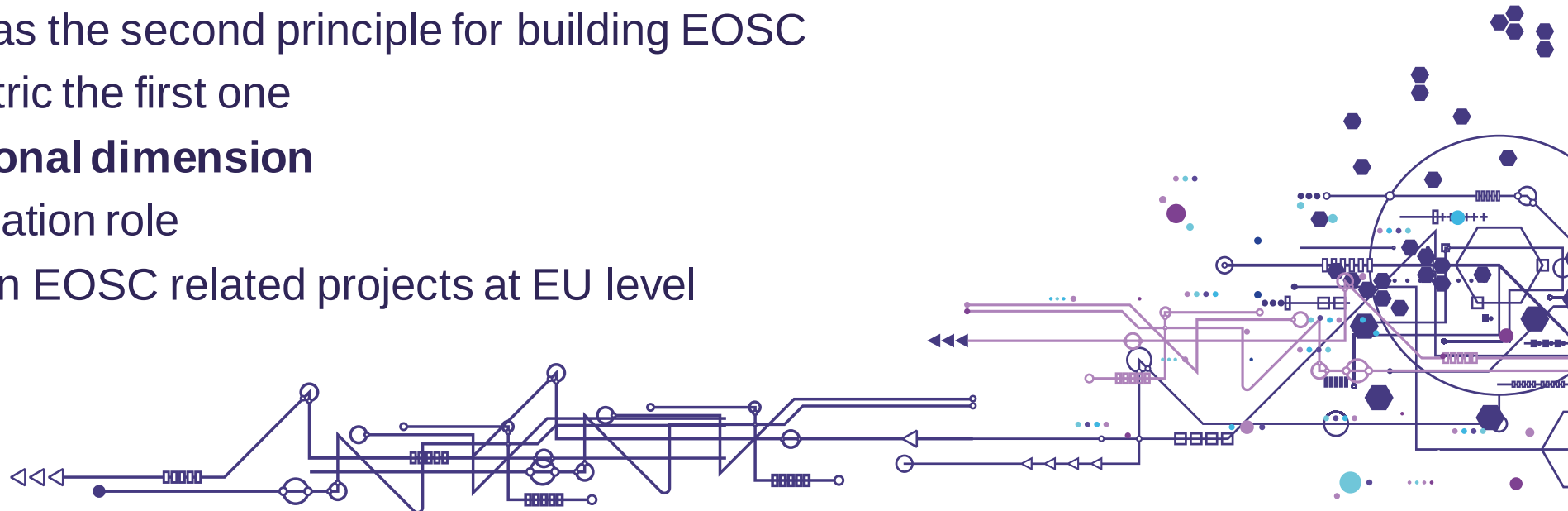
- **Sensitive Data** – Zdenka Dudová BBMRI CZ (together with CZECRIN, EATRIS and ELIXIR CZ)
- Provides opportunity to bring in specific thematic needs



EOSC CZ General Principles

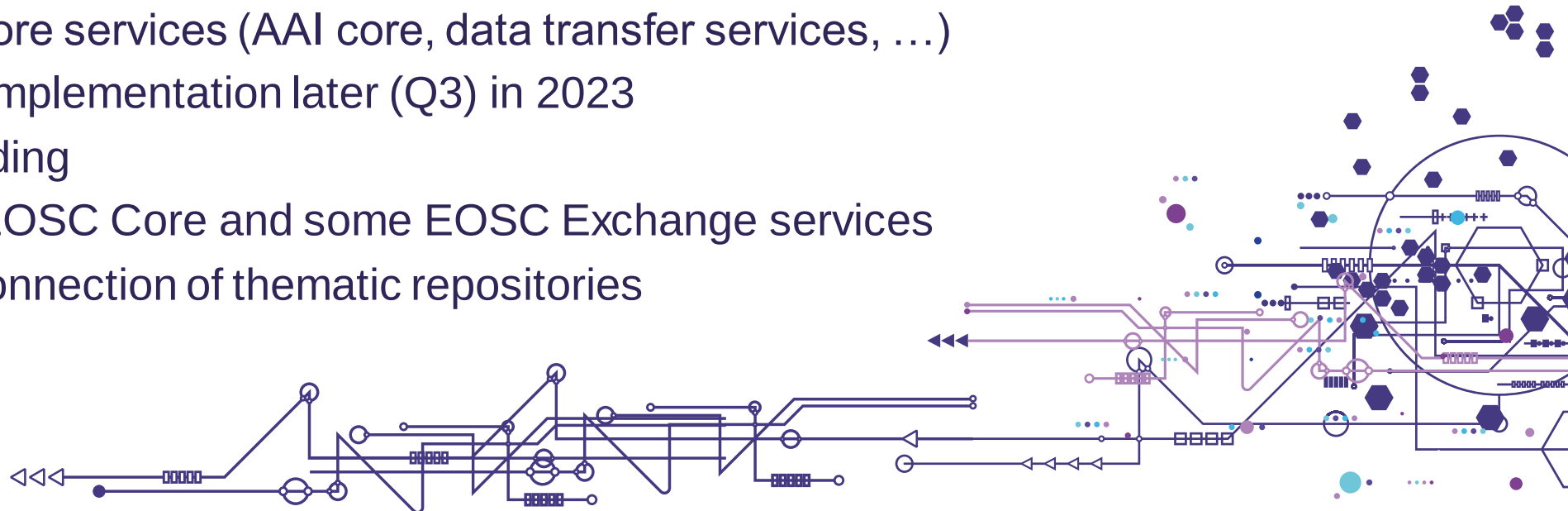


- Believe that **data-centric** approach is the core of the next generation federated computed and data infrastructures
 - Federated principles
 - Cloud computing in a distributed environments
 - Data and compute synergy (at logical and technical levels)
- The collaboration with research communities and large research infrastructures
 - **Partnership** as the second principle for building EOSC
 - Data centric the first one
- And the **international dimension**
 - EOSC Association role
 - Participation in EOSC related projects at EU level



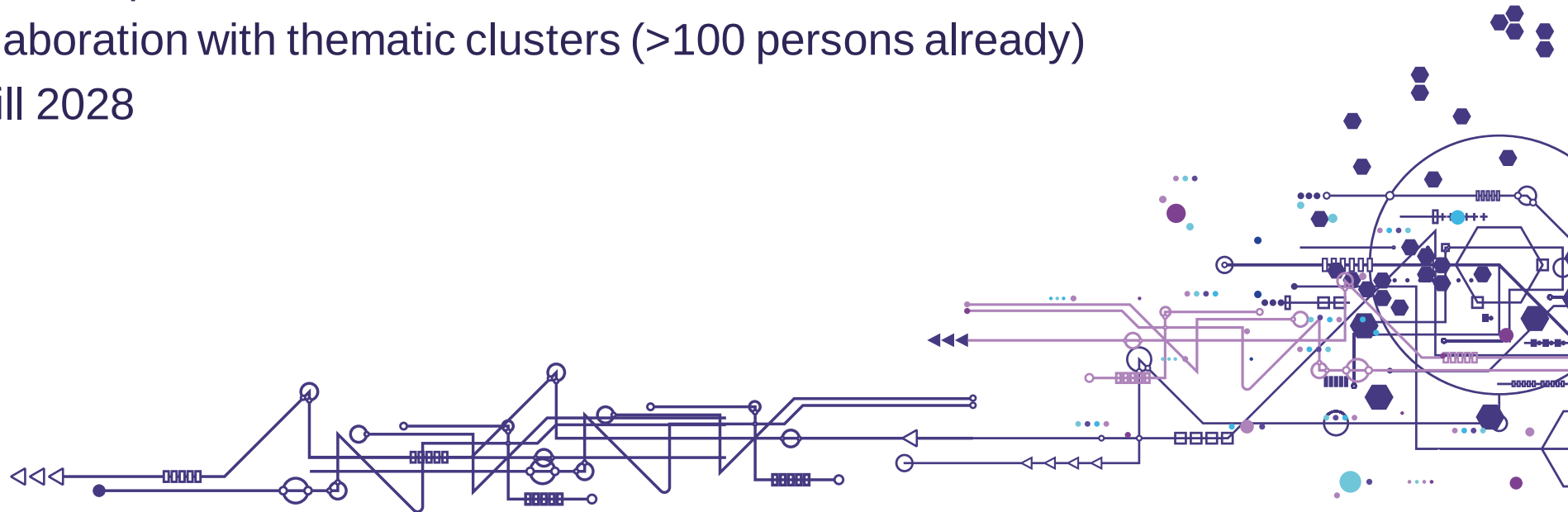
Steps for 2023

- Projects **EOSC CZ** (EOSC Secretariat, NMA, training coordination) and **CARDS** (Metadata, PIDs) to start in January 2023
 - EOSC CZ Secretariat
 - National EOSC contact point/Hub, EOCC CZ Portal, ...
 - NMA implementation
 - Metadata reference to the digital objects
 - First EOSC core services (AAI core, data transfer services, ...)
- Start of the NRP implementation later (Q3) in 2023
 - Capacity building
 - Fuller set of EOSC Core and some EOSC Exchange services
 - Preparing the connection of thematic repositories



Summary

- EOSC implementation as an opportunity to evolve country towards Open Science principles
- From compute to data-centric focus
 - Reshaping some of the core business of not only the e-INFRA CZ
- Open Science as a partnership – not a single institutions/infrastructure to **build EOSC**
 - Collaboration and partnership a key
 - Large consortium expected for NRP
 - Extensive collaboration with thematic clusters (>100 persons already)
- Secured funding till 2028





Thanks and questions?

info@e-infra.cz, eosc-info@infra.cz

More information on

<https://www.e-infra.cz>

<https://www.e-infra.cz/eosc>

e-infra.cz