The Digital Europe Programme and the Common European Language Data Space
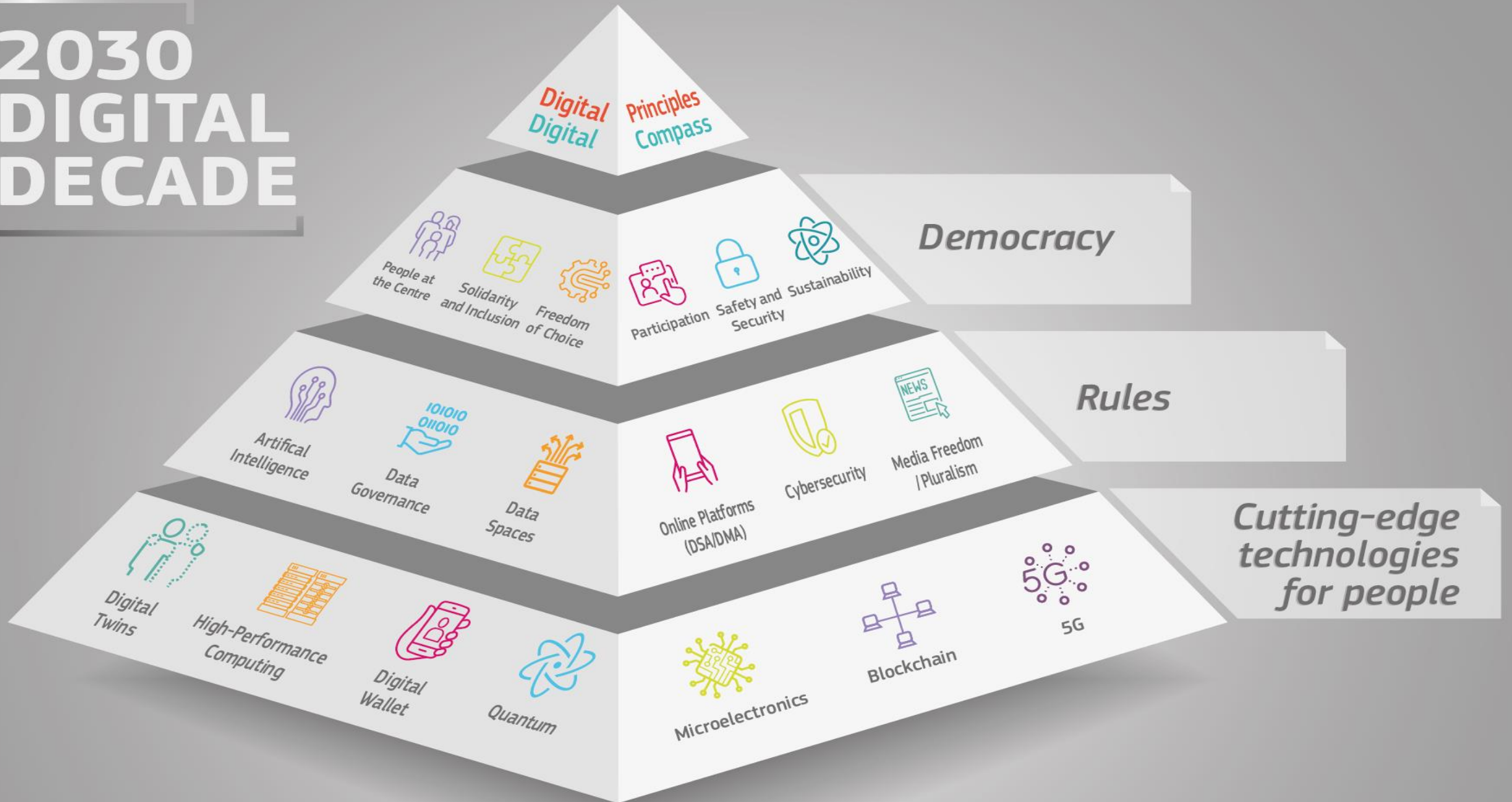
The European Language Data Space (LDS) Workshop in Slovakia

Mateusz SZOTURMA
Legal and Policy Officer
European Commission – DG CONNECT – G3

Revolution !

European Commission

2030 DIGITAL DECADE

Digital Digital Principles Compass

Democracy

People at the Centre
Solidarity and Inclusion
Freedom of Choice
Participation
Safety and Security
Sustainability

Rules

Artificial Intelligence
Data Governance
Data Spaces
Online Platforms (DSA/DMA)
Cybersecurity
Media Freedom / Pluralism

Cutting-edge technologies for people

Digital Twins
High-Performance Computing
Digital Wallet
Quantum
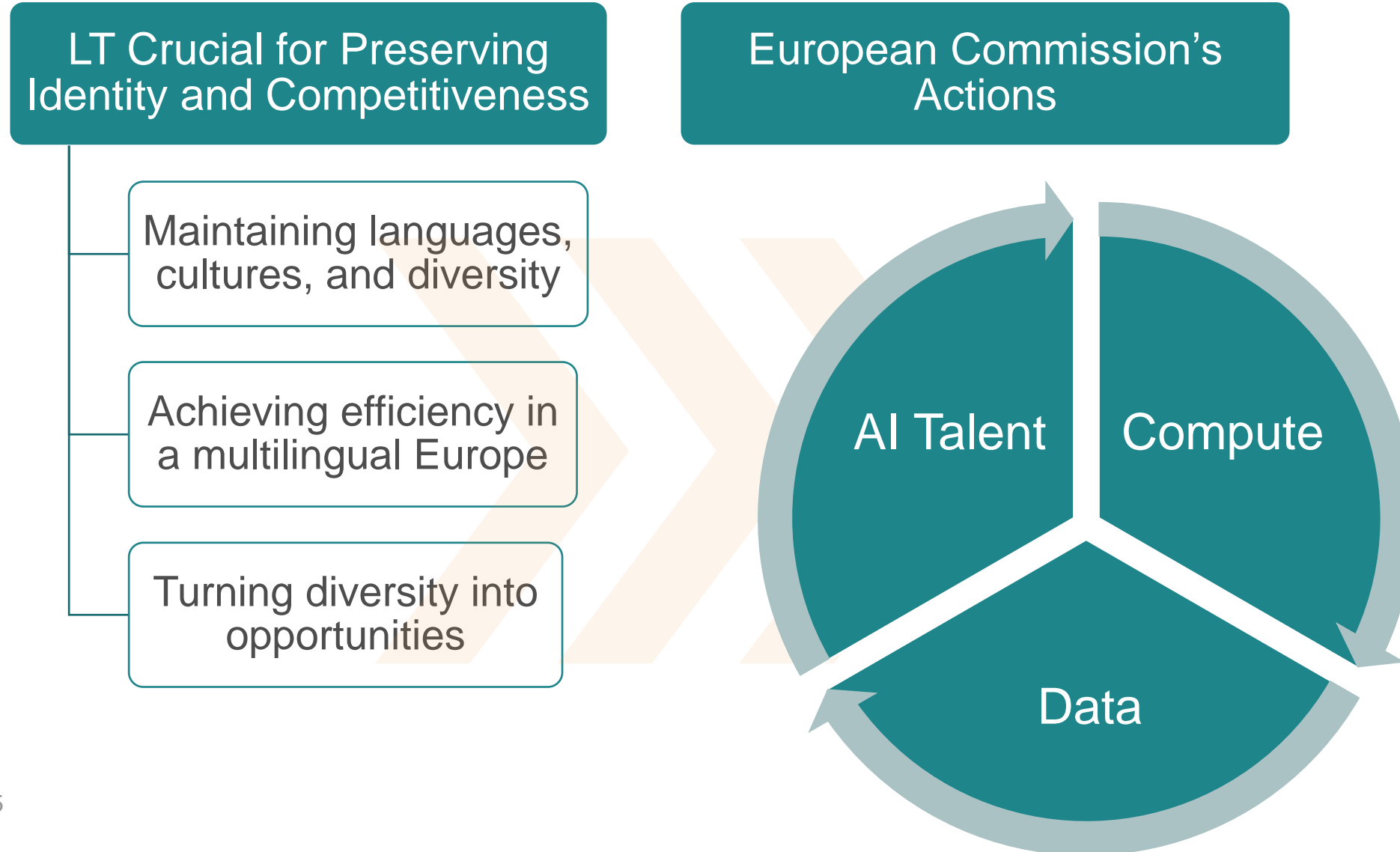Microelectronics
Blockchain
5G

Landscape

LT Crucial for Preserving Identity and Competitiveness
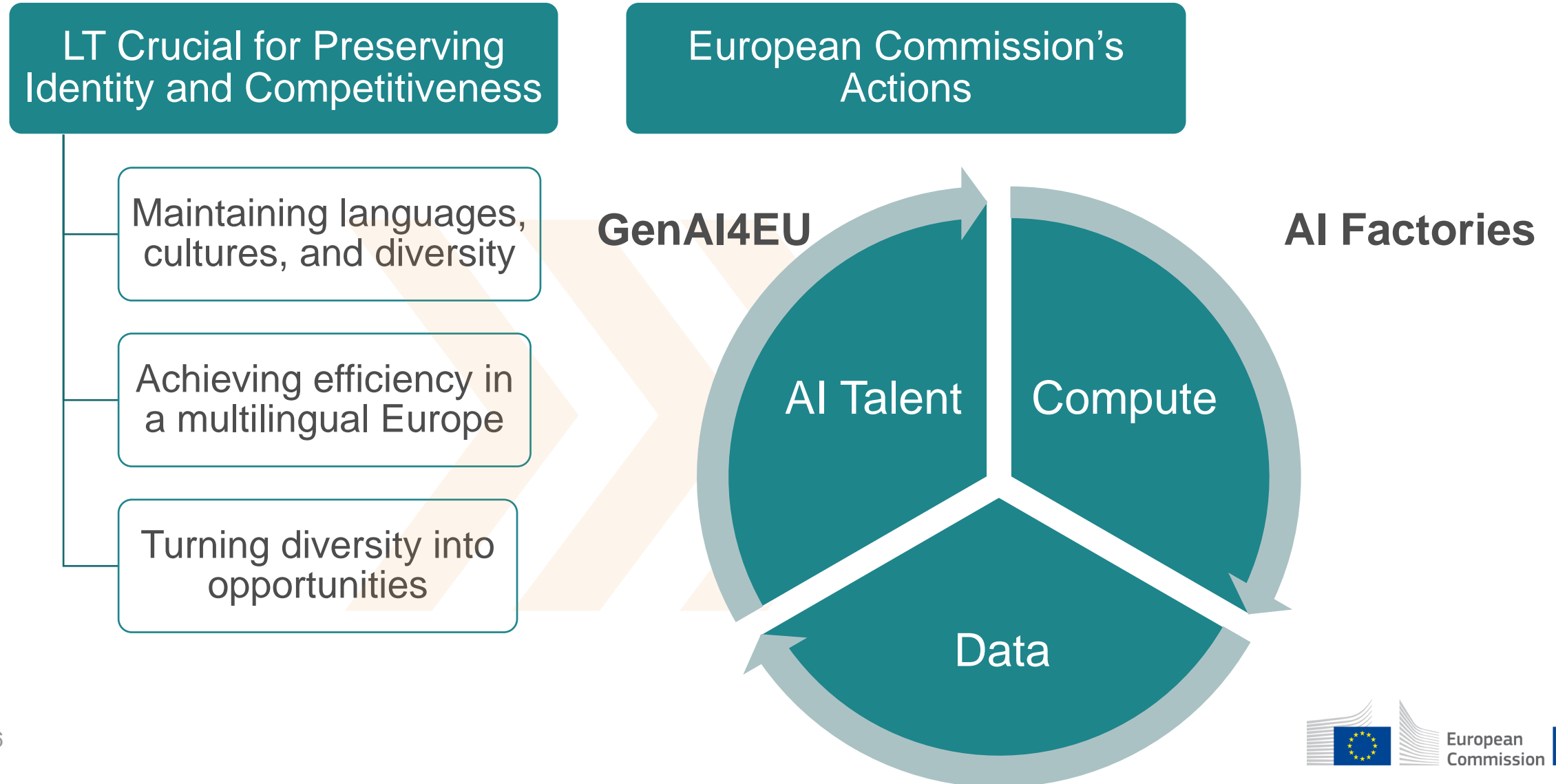
- Maintaining languages, cultures, and diversity

- Achieving efficiency in a multilingual Europe

- Turning diversity into opportunities

European Commission

Landscape

LT Crucial for Preserving Identity and Competitiveness

- Maintaining languages, cultures, and diversity
- Achieving efficiency in a multilingual Europe
- Turning diversity into opportunities

European Commission's Actions

- AI Talent
- Compute
- Data

European Commission

Landscape

LT Crucial for Preserving Identity and Competitiveness

Maintaining languages, cultures, and diversity

Achieving efficiency in a multilingual Europe
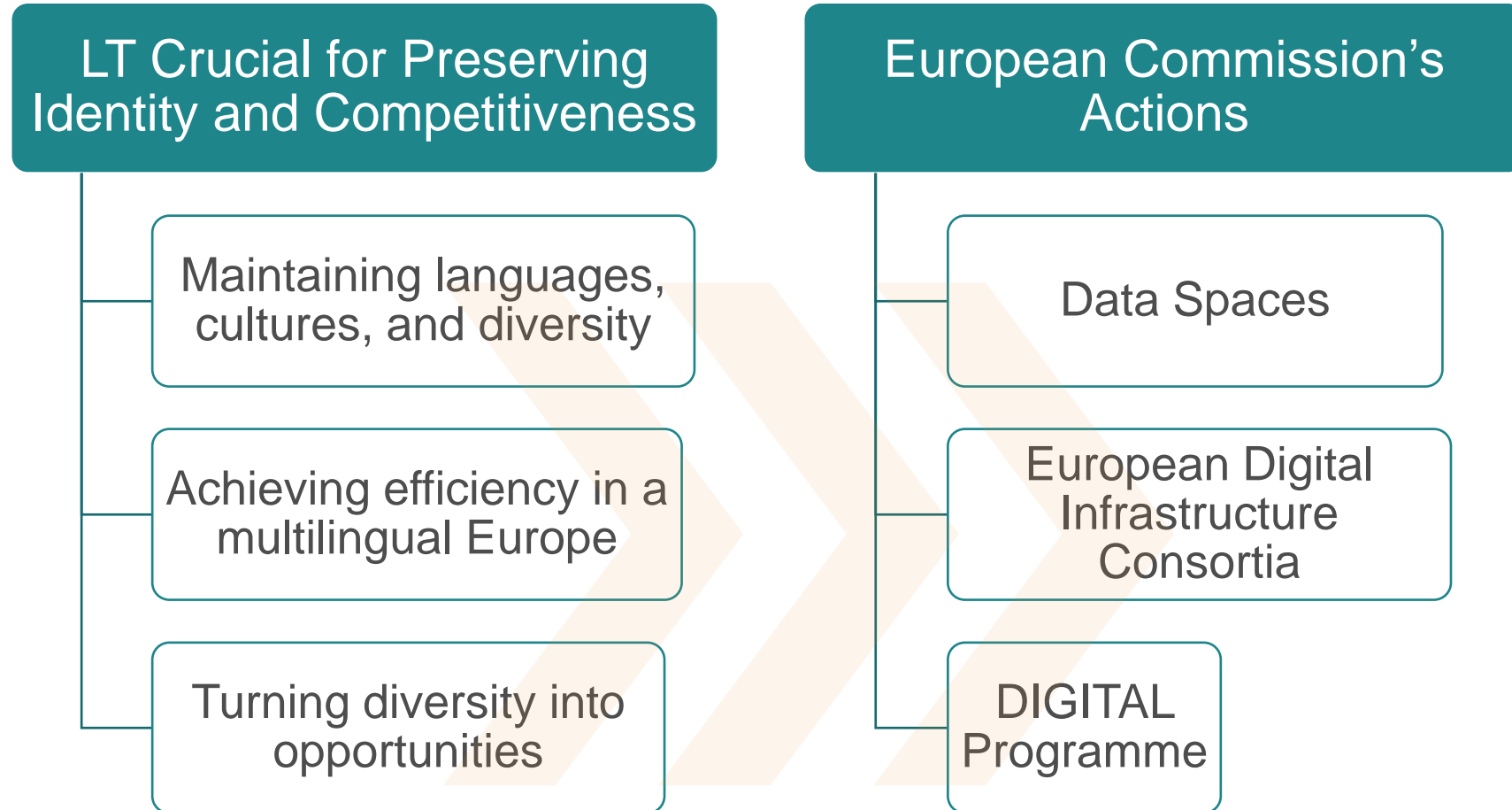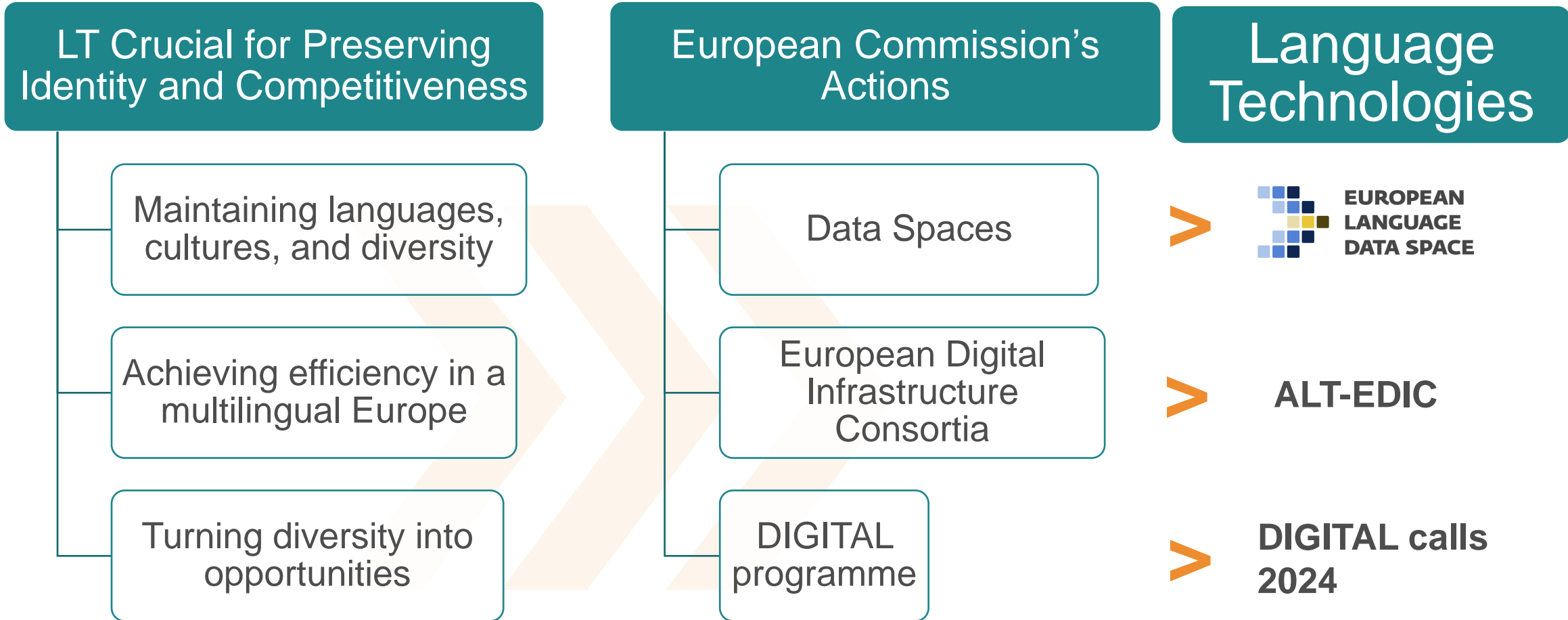
Turning diversity into opportunities

European Commission's Actions

GenAI4EU

AI Factories

AI Talent

Compute

Data

European Commission

Landscape

LT Crucial for Preserving Identity and Competitiveness

- Maintaining languages, cultures, and diversity
- Achieving efficiency in a multilingual Europe
- Turning diversity into opportunities

European Commission's Actions

- Data Spaces
- European Digital Infrastructure Consortia
- DIGITAL Programme

European Commission

Landscape

LT Crucial for Preserving Identity and Competitiveness	European Commission's Actions	Language Technologies
Maintaining languages, cultures, and diversity	Data Spaces	> EUROPEAN LANGUAGE DATA SPACE
Achieving efficiency in a multilingual Europe	European Digital Infrastructure Consortia	> **ALT-EDIC**
Turning diversity into opportunities	DIGITAL programme	> **DIGITAL calls 2024**

European Commission

Landscape

Public Sector

> EUROPEAN LANGUAGE DATA SPACE

> **ALT-EDIC**

> **DIGITAL calls 2024**

Private Sector

European Commission

Landscape

Public Sector

> EUROPEAN LANGUAGE DATA SPACE

> ALT-EDIC

> DIGITAL calls 2024

Private Sector

European Commission

The Common European Language Data Space (LDS)

- Overview

- Architecture

European Commission

Language Data Space – Overview

What	**Platform** and **marketplace** for the collection, creation, sharing and re-use of multilingual and multimodal language data
	6 Mio € **Procurement** Contract from the **DIGITAL** Work Programme 2021-2022
Why	Increase **availability** of clean, high-quality, compliant language data
	Support the development of state-of-the-art European language technologies
Who	**Consortium**: DFKI, ELDA, TILDE, Athena Institute
	Governance: Member States representatives and Industry User group
	Stakeholders: Industry, Public Administration, Research, etc.

How

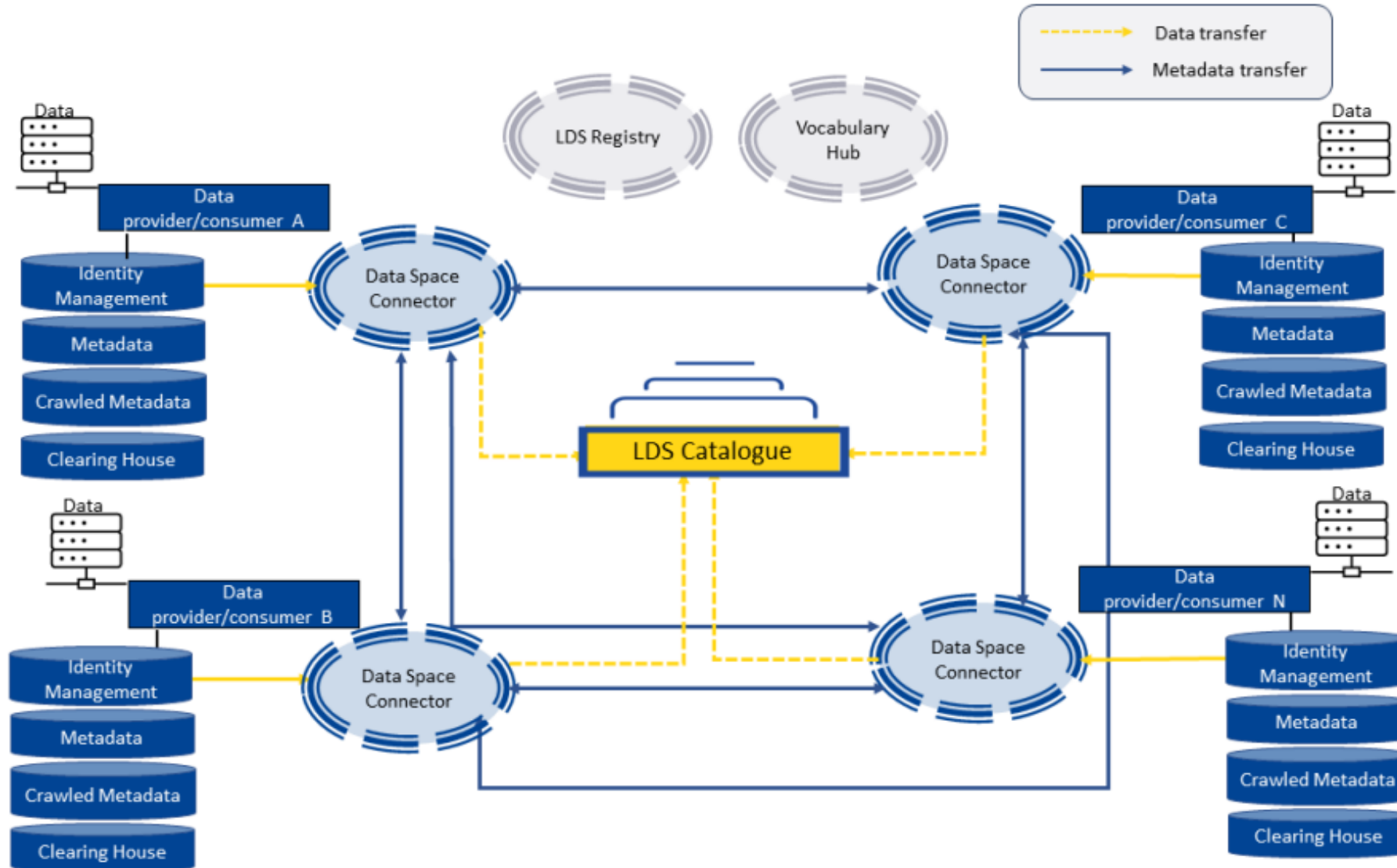|	
- Technical Architecture and Infrastructure	- Openess
- Governance Framework	- Respect of EU Rules and Values
- Promotion	- Overcoming Legal Barriers (IPR, GDPR, etc.)

When

Started in **January 2023** for Three years – until end 2025 (Possible contract renewal for another year).

First version expected in **Q4 2024**

Language Data Space – Architecture

ALT-EDIC – Alliance for Language Technologies EDIC

What	DDPP Multi-Country Project in the area of "European Common Data Infrastructure and Services"
	Build a European ecosystem around **Language Data and Technologies**
Why	Federate **data** and resources from EU and Member States to boost Data and AI competitiveness
	Promote technology and service deployment for European leadership and strategic autonomy
Who	**Coordinated by: France** (Ministry of Finance and Digitalization & Ministry of Culture)
	Members: BG, CZ, DK, ES, FR, GR, HR, HU, IE, IT, LT, LU, LV, NL, PL, SI
	Observers: AT, BE, EE, FI, MT, PT, RO, SK, CY – **Interested**: DE, SE
How	**~25 Mio €** Member States contribution
	~25 Mio € EU funding (DIGITAL Programme)
When	**7 February 2024**: Set up by Commission Implementing Decision (EU) 2024/458
	Q1 2025: Fully operational

European Commission

Slovakia has currently status of an observer within ALT-EDIC

European Commission

Digital Europe Programme

Work Programme 2024

DIGITAL Work Programme 2024 (3 calls - **Closed**)

DIGITAL-2024-AI-06-LANGUAGE-02

Alliance for Language Technologies: **Ecosystem**

€ 4 millions (CSA)

Support the Coordination of a European Language Technologies Ecosystem.

DIGITAL-2024-AI-06-FINETUNE

Open-source European foundation model

€ 25 millions (SME Grant)

Develop and make available one open-source large language foundation model as an infrastructure designed to be largely used by public or private users.

DIGITAL-2024-AI-B-06-LANGUAGE

Alliance for Language Technologies: **Data and Finetunig**

€ 20 millions (Simple Grant)

- Support the Collection of Language Data
- Support Fine-tuning of Large Language Models

European Commission

Thank you

European Commission