



28

April
2022

Natural Language Processing

NLP Study

DIGIT.D2 - Interoperability.

interoperable
europe

Webinar Practicalities



Click on <<Connect audio>> but please mute your microphones



You can share your questions and comments via the chat



The webinar will be recorded



Welcome

Conductor:

- Miguel Alvarez Rodriguez – Policy Officer, expert on European Interoperability Framework and supervisor of the NLP study

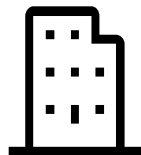
Facilitator:

- Zsófia Dudàs – Independent advisor

Speakers:

1. Louis Matha, Nathan Ghesquiere and Emidio Stani – Authors of the NLP study
2. Luca TANGI - Project Officer at Joint Research Centre
3. Cecile GUASCH - Consultant for the European Commission- Information Systems Architect & Vidas Daudaravicius - Scientific Project Officer at European Commission Joint Research Centre
4. Emmanouil MARAGKOUidakis - Project owner at DG DIGIT
5. Georgi GITCHEV - Lawyer at DG CNECT

The catalogue of Service Action



The Catalogue of Services Action supports **public administrations** in building their **digital catalogues of public services**, to allow citizens, businesses and public administrations across Europe to access and understand the information they need.

Test and validate your solutions



Make your service description fit for exchange

Help you create your catalogue of services



Advise you in adopting standards and technologies



Exchange knowledge



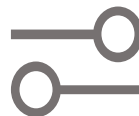
Assignments part of this project



CPSV-AP



Controlled
vocabularies



Tools for
mapping



Studies and
reports



Supporting
activities to
SDG policies



Knowledge
sharing and
community
building



Courses

Natural Language Processing (NLP)

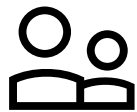


“the machine’s ability to identify, process, understand and/or generate information in written and spoken human communications”

Goals of this webinar



Provide knowledge on NLP applied to the Public Sector



Give hands-on experience to implement NLP projects



Present the necessary regulatory compliance about NLP

Agenda

01

Presentation of the NLP study

02

Benefits - How can the public sector benefit from the use of NLP?

03

Demo - How can the public sector use NLP?

04

What can we learn from the DEAP project?
- A successful NLP initiative

05

Regulatory Compliance - How can public administrations comply with regulatory requirements?



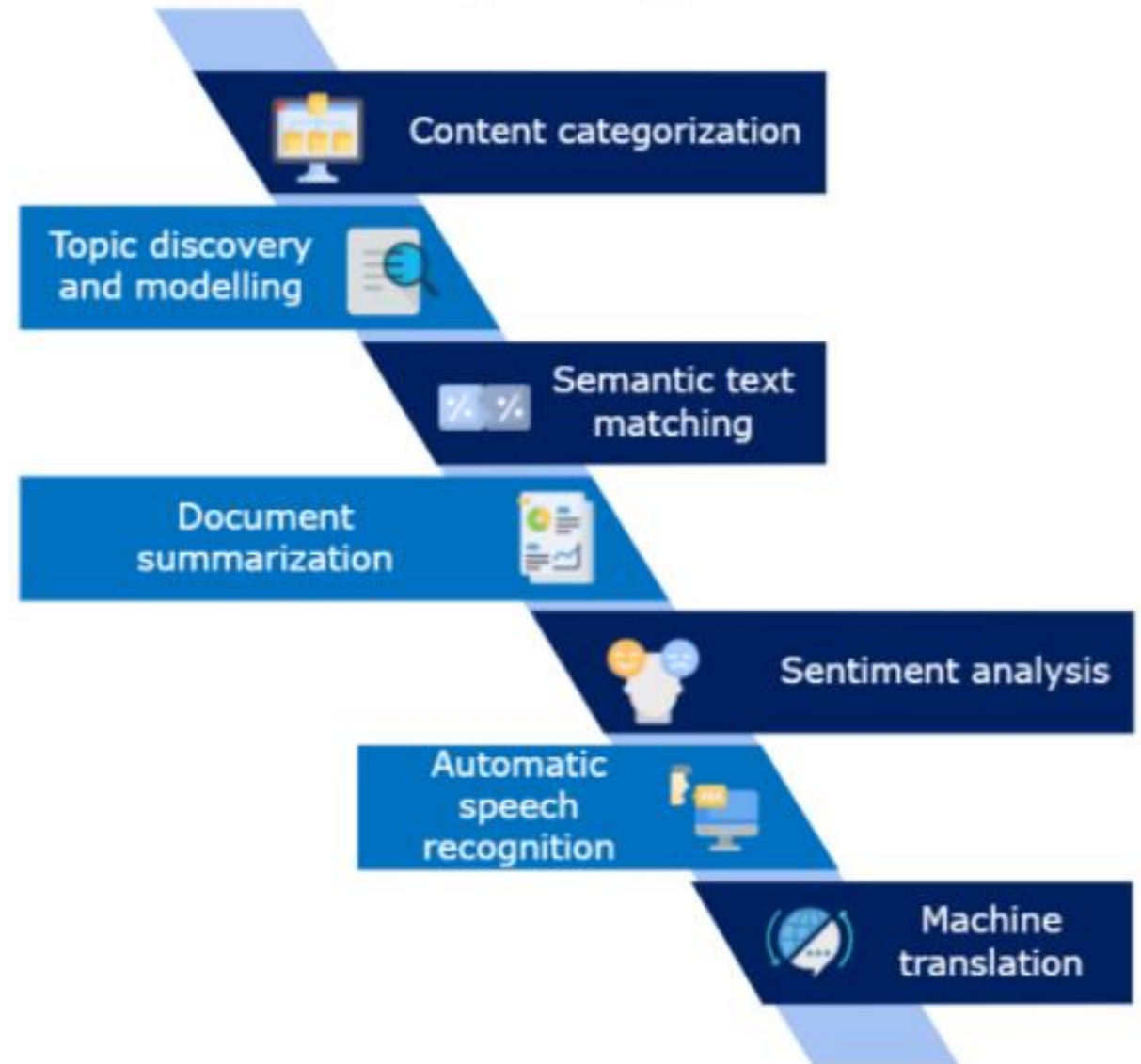
NLP Study

**Presentation
of the NLP
Study**

*Catalogue of
Services*

Application areas for NLP

- In the public sector, NLP can be used for various different purposes



Application areas for NLP

01 Content categorization

Content categorization is a fundamental application area of text classification algorithms.



Example: Case classification to remove guesswork in populating case fields



Content categorization

02 Topic discovery and modelling

Topic modelling is the process of automatically identifying topics present in a corpus of texts and to derive hidden patterns present in this corpus.



Example: Tracking geographical locations using a geo-aware topic model for analysing social media data

Topic discovery and modelling

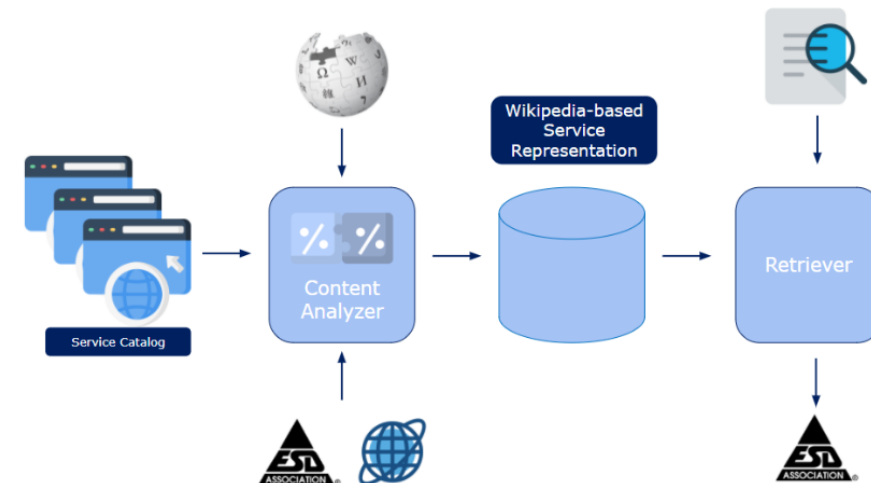


03 Semantic text matching

Semantic text matching is the task of estimating the similarity between the source and the target text pieces

Example: Link e-government services between each other

Semantic text matching



Application areas for NLP

04 | Sentiment analysis

Comparable to content categorization models, sentiment analysis (SA) applications are also based on text classification.



Example: Tracking of internal social network to interpret employee sentiment



Sentiment analysis

05 | Automatic speech recognition

Automatic speech recognition (ASR) is the application of NLP which converts spoken words into computer texts.



Example: Speech recognition to improve productivity of the police

Automatic speech recognition



06 | Document summarisation

Document summarisation is the NLP method relying on keyword extraction enabling machines to condense a piece of text to a shorter version without losing essential information



Example: Enhancement of Public Services in Meadville using text summarisation

Document summarization



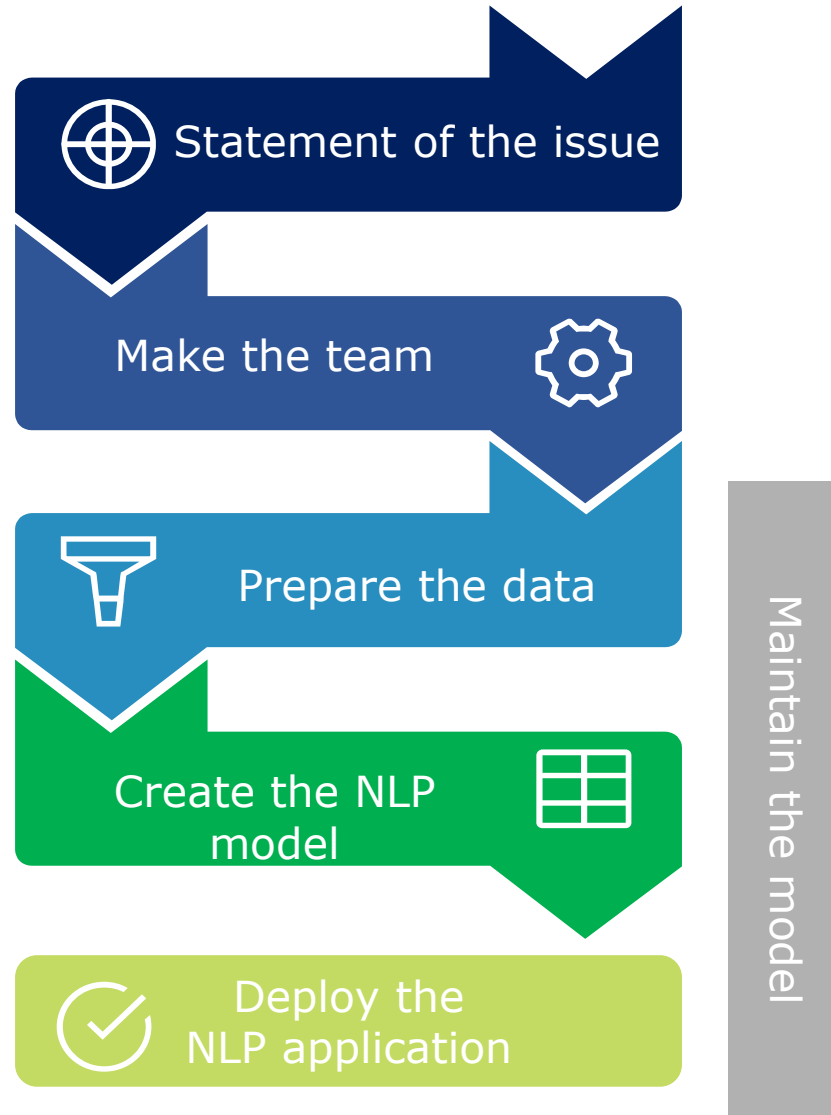
07 | Machine translation

Machine translation (MT) is an application area that uses NLP to translate text from one language to another



Machine translation

Structure of the NLP Project - Overview



Structure of NLP project

01 | Statement of the issue



Define the scope



Set the technical requirements



Identify the data needed



Draw the IT architecture



02 | Create the team



All aspects of the NLP project needs to be tackled



Roles vary depending on the scale of the project



No strict roles defined for NLP project yet

Chief Data Officer

Data Analyst

Data Scientist

Data Engineer

Machine Learning Engineer

Data Architect

Data Steward

Legal Advisory

Structure of NLP project

03 | Prepare the data

The objective of the transformation step is to shape raw data into a format that can be used in different applications

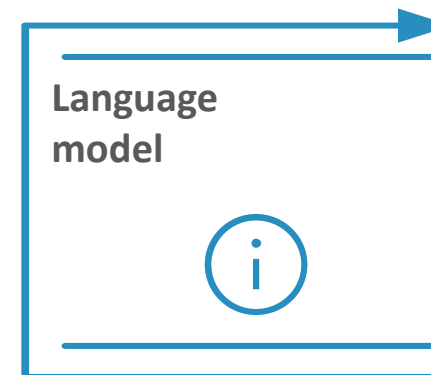


Once the data sources has been identified, it can be collected to a target database

The cleaning process remove all the insignificant words and characters not needed when training an NLP model



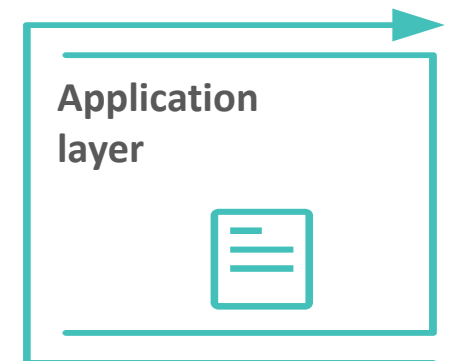
04 | Create the language model



The language model understands and process textual data.

There are two types of language models:

- Pre-trained
- To-be-trained



The next step is to parse the data according to the NLP application of the project.

Structure of NLP project

05 | Deploy the NLP application



Serverless

The serverless architecture is the simplest deployment option. However, it can lead to a higher latency.



Container platforms

Container platforms need expertise to work with them but they offer the most flexibility: they can be deployed either on-premise or in the cloud.



06 | Maintain the model



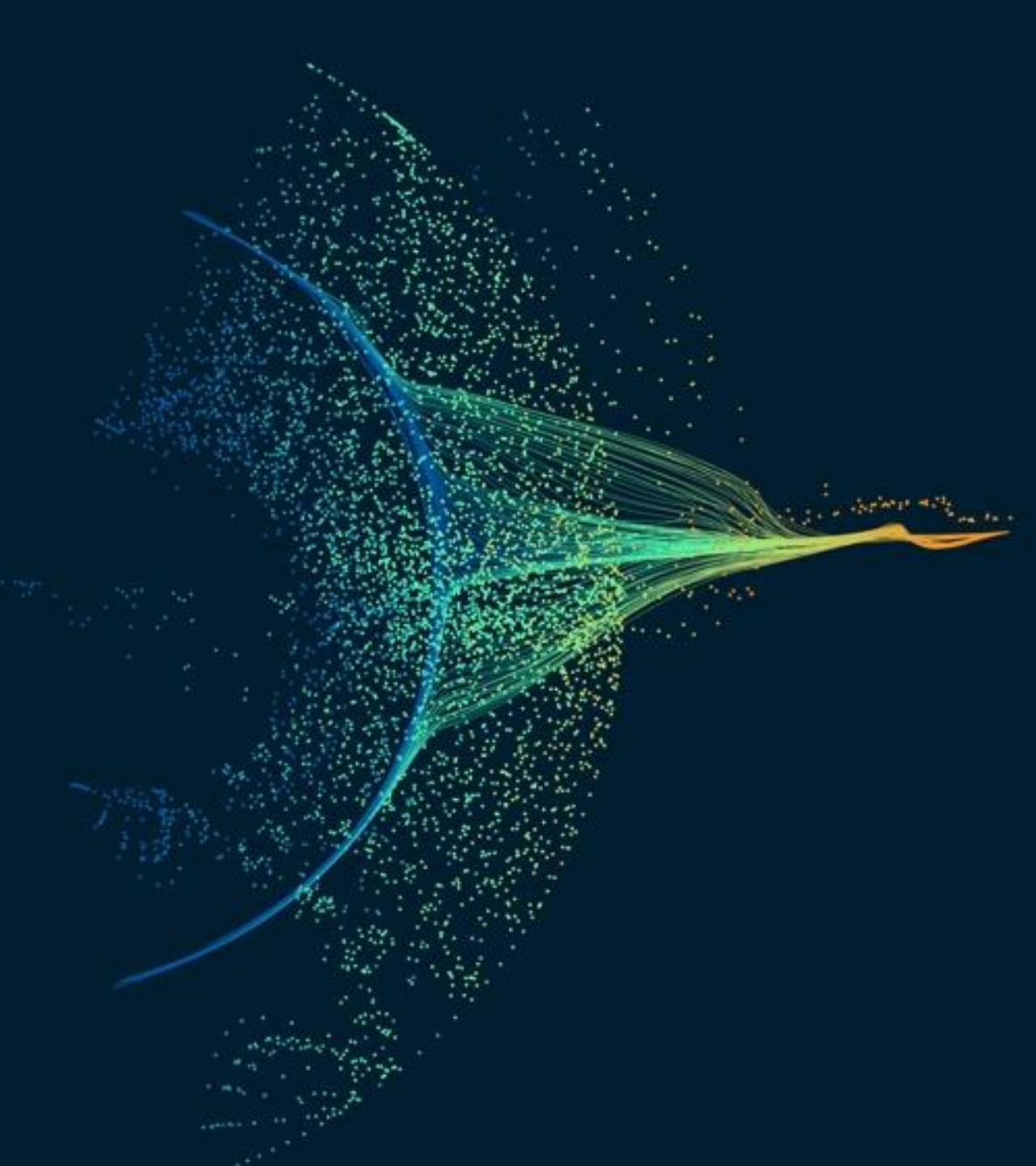
Time-based

This approach aims at retraining the model at regular intervals. It requires a good understanding of how the data changes over time.



Threshold-based

This approach prefers to retrain the model when it is needed based on key performance indicator (accuracy, bias ..).



NLP study

Q&A



Benefits of NLP



How can the public sector benefit from the use of NLP?

Agenda

- Brief JRC presentation
- Our projects on AI in the public sector
- Results: landscaping overview
- NLP overview
- Some NLP examples

Brief JRC presentation



JRC Mission

As the science and knowledge service of the European Commission our mission is to support EU policies with independent evidence throughout the whole policy cycle.

JRC sites

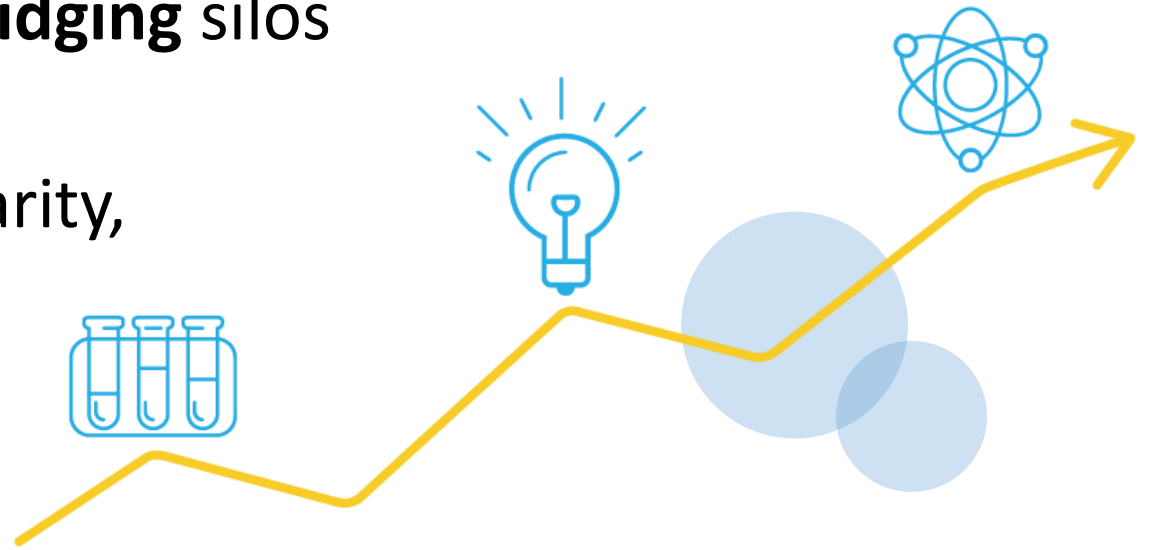
Headquarters in **Brussels**
and research facilities located
in **5 Member States:**

- Belgium (Geel)
- Germany (Karlsruhe)
- Italy (Ispra)
- The Netherlands (Petten)
- Spain (Seville)



JRC research

- Fully **policy-relevant** and world class **knowledge production**
- Priorities driven **knowledge and competence management**
- One JRC – **anticipating** emerging issues, **understanding** complexities and **bridging** silos
- **Addressing challenges of research**
(information deluge, multidisciplinary, integrity, reproducibility)



Our projects on AI in the public sector

AI Watch – the Knowledge Service to monitor the Development, Uptake and Impact of AI for Europe



[AI for the public sector](#)



[AI Landscape and Dashboard](#)



[Strategic Actions and Coordination](#)



[AI History Timeline](#)



[European Policy on AI](#)



[A Storymap on AI in Europe](#)

Innovative Public Services



Innovative Public Services

 **24**
Events

 **24**
News

 **7**
Documents

2020

Scientific methods, data and information

Information society

Innovative Public Services Observatory

Abstract: The Innovative Public Services Observatory (IPSO) is a platform jointly created by DIGIT and JRC in the framework of the IPS Action of the EU ISA² Programme, with the purpose of monitoring the adoption and use across Europe of emerging and disruptive technologies - as AI, DLT, IoT, APIs - for the provision of public services. This collection includes the data produced by the observatory.

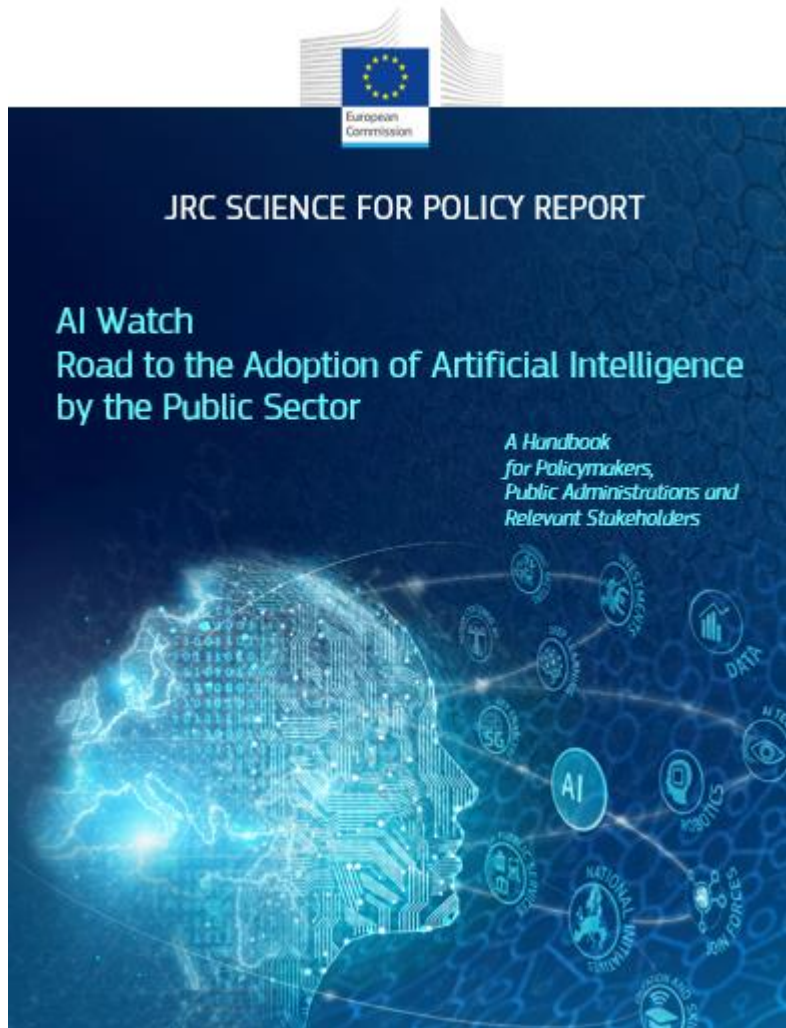
Authors: [PEREGO Andrea](#); [ULRICH Peter](#); [DALLA BENETTA Alessandro](#)

Citation: Perego, A., Ulrich, P. and Dalla Benetta, A., Innovative Public Services Observatory, European Commission, 2020, JRC120247.

Publisher: European Commission

Results: landscaping overview

Road to the adoption of AI by the public sector



4 Areas of interventions

16 Recommendations

Road to the adoption of AI by the public sector

Area 1

Promote an EU-value oriented, inclusive and human-centric AI in the public sector

- 1.1** Harmonise and complement EU regulations to promote fair, non-discriminatory and transparent AI-enabled public services for all citizens.
- 1.2** Promote the adoption of ethical principles the development of guidelines and the development of mitigating measures to minimize the risks of deployment of AI by the public sector.
- 1.3** Develop and promote dedicated AI-enabled solutions based on co-creation approaches to increase citizens' and businesses' relevance and confidence in the use of AI by the public sector.

Area 2

Enhance coordinated governance, convergence of regulations and capacity building

- 2.1** Create an EU-wide network of governance bodies for AI in the public sector.
- 2.2** Design national and European, capacity-building programs for public sector innovators aiming to adopt AI in support to the digital transformation of public services.
- 2.3** Build upon and promote the use of regulatory sandboxes for public administrations, allowing experimentation of AI-enabled solutions in controlled environments.
- 2.4** Optimise funding in support to AI in the public sector to promote the spreading and scaling of reusable solutions.
- 2.5** Promote the development of multilingual guidelines, criteria, and tools for public procurement of AI solutions in the public sector throughout Europe.

Area 3

Build a shared and interactive AI digital ecosystem

- 3.1** Support multidisciplinary research and knowledge creation amongst European universities and R&D institutions around AI for the Public Sector.
- 3.2** Build a common European Data Space for public sector bodies and their operators, drawing from the compilation of relevant AI datasets throughout Europe.
- 3.3** Reinforce and advance existing initiatives on open data and interoperability.
- 3.4** Share reusable and interoperable AI components at all operational levels of European public administrations.
- 3.5** Create a European marketplace for GovTech solutions in support to public sector digital transformation.

Area 4

Applying value oriented AI impact assessment frameworks

- 4.1** Set up and observatory on AI, built on a pan-European network of AI national observatories to gather, share and collectively manage best practices and experiences learned from different stakeholders in the Public Sector throughout Europe.
- 4.2** Develop and apply umbrella impact assessment frameworks based on key influencing factors to measure the impact and related use of AI in the public sector.
- 4.3** Support Green AI in the public sector in compliance with environmental sustainability principles, and promote civic engagement to that end.

Mapping AI use in public services in the EU

■ Science for policy report

Misuraca, G., and van Noordt, C., Overview of the use and impact of AI in public services in the EU, EUR 30255 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-19540-5, doi:10.2760/039619, JRC120399



230 Cases

■ Open Data repository

Available at the [JRC Data Catalogue, AI Watch collection, "selected AI cases in the public sector"](#)

The information box contains the following details:

- Organisation: European Commission, Joint Research Centre
- Point of contact: ✉ ec-ai-watch@ec.europa.eu
- Title: Selected AI cases in the public sector
- Data access:
 - CSV - Comma Separated Value format
 - ODS - Open office format
 - XLS - Microsoft Excel format

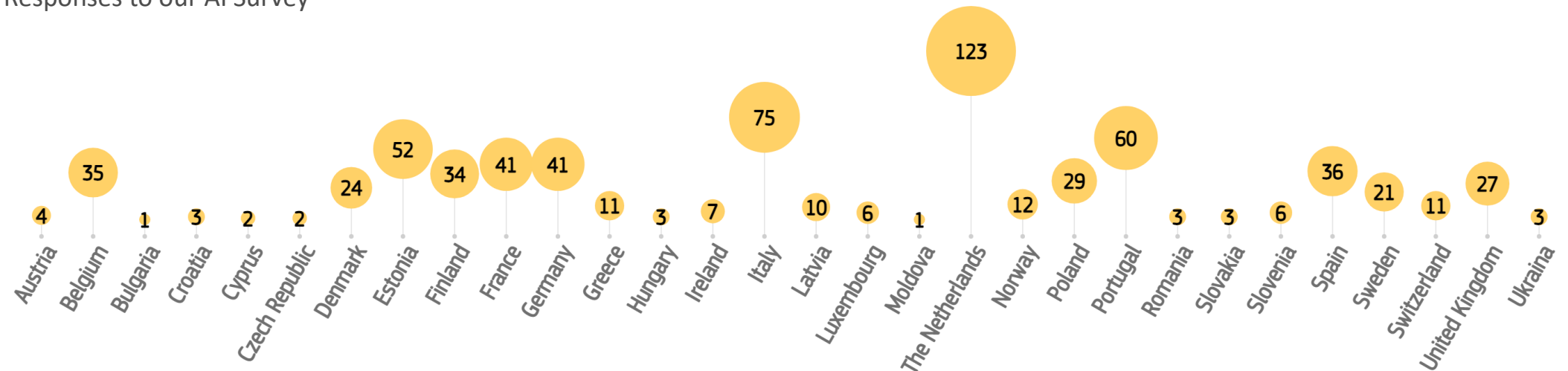
142 Cases

Ongoing activity - AI Case Collection

686 AI Cases collected and validated

How we collect cases?

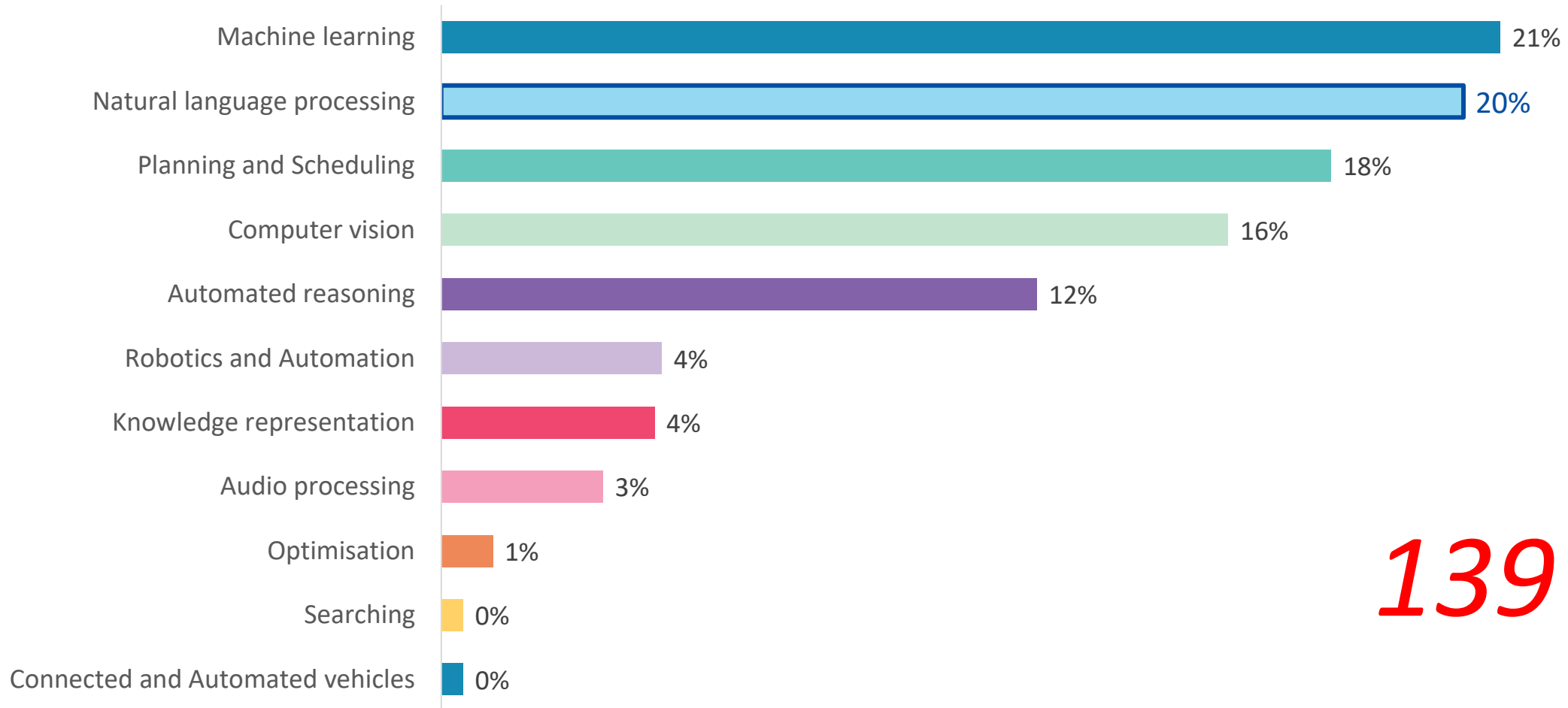
- Country repositories or research studies
- News articles
- Responses to our AI Survey



n.b. the cases are not statistically representative. No comparison can be done among the different countries

Published cases: [Joint Research Centre Data Catalogue - Selected AI cases in the public sector - European Commission \(europa.eu\)](https://ec.europa.eu/jrc/data-catalogue/)

AI Cases by Technology

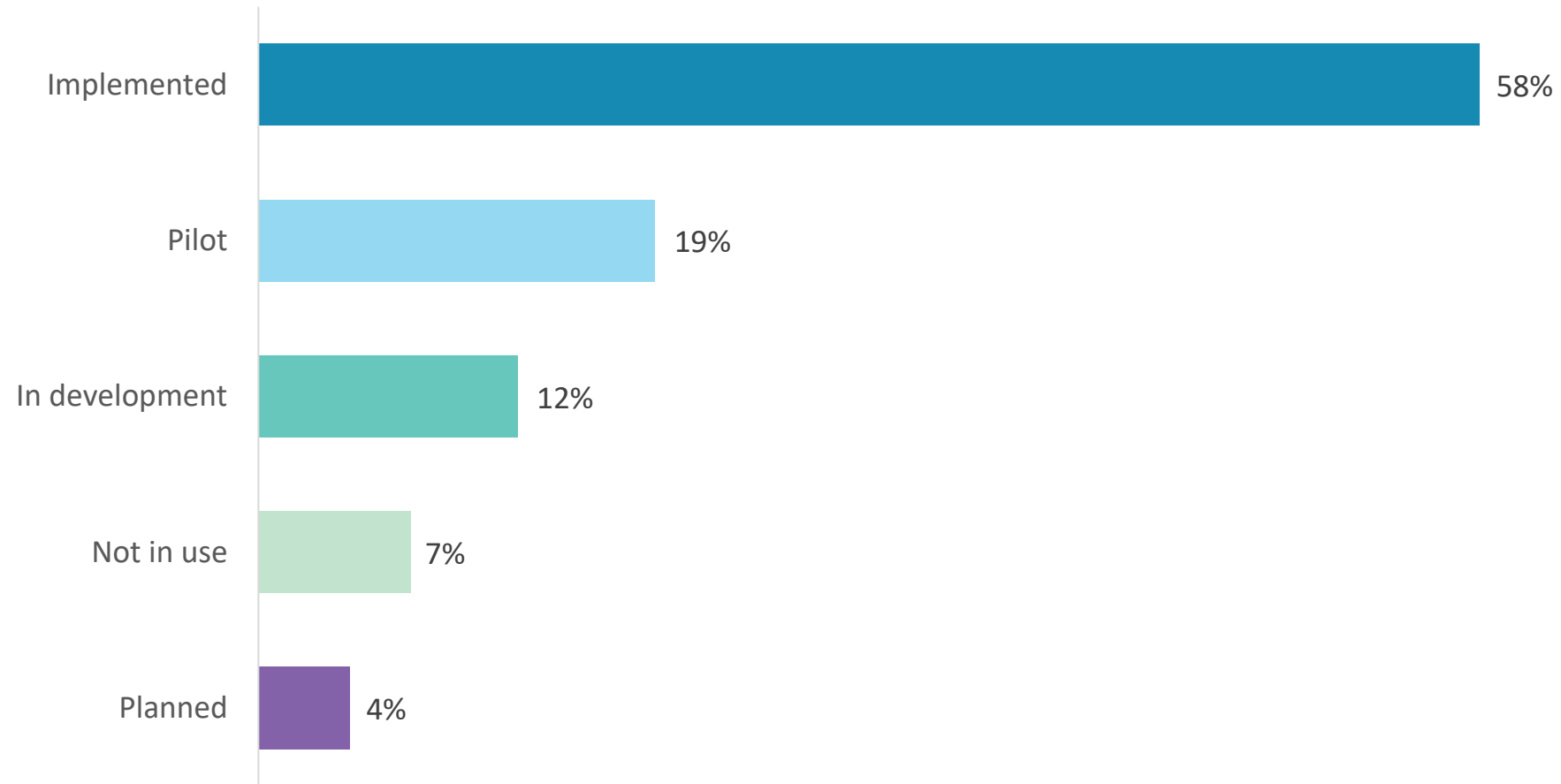


139 NLP Cases

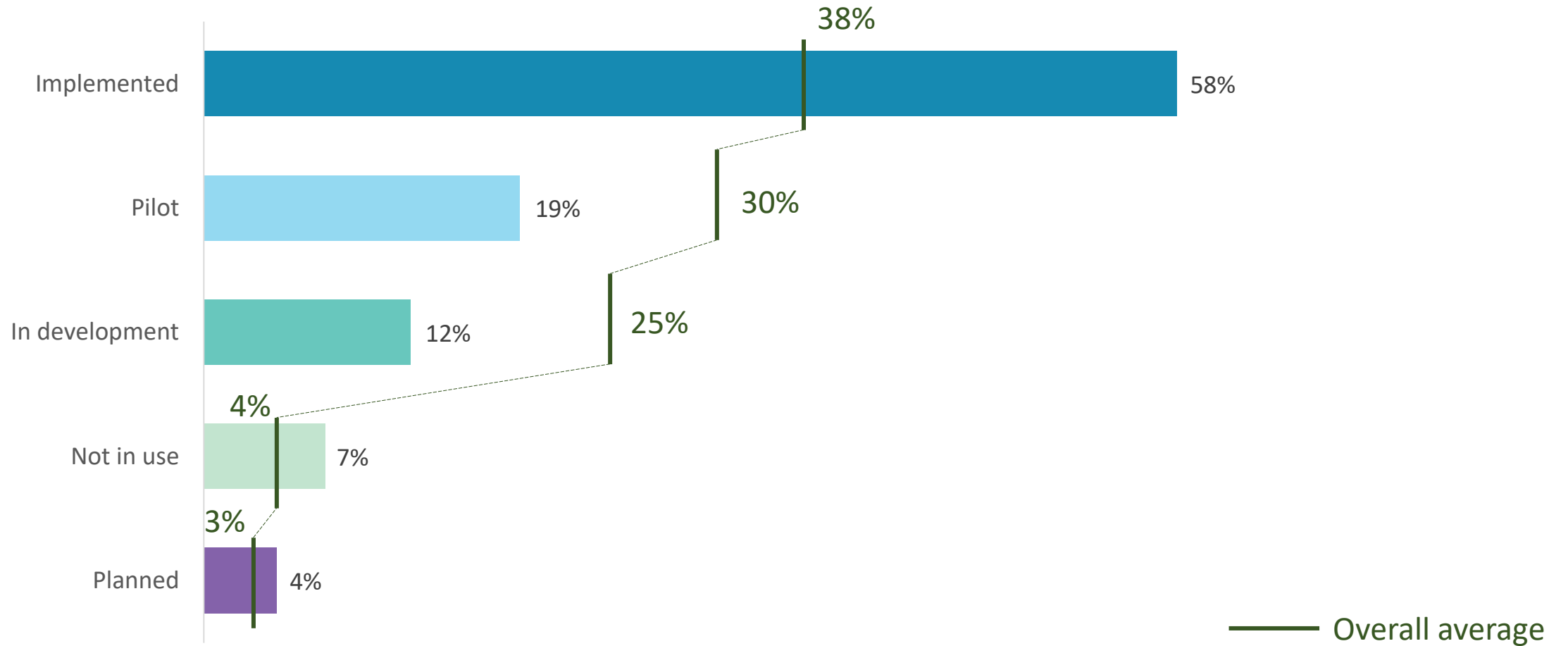
NLP cases in PS

Some data

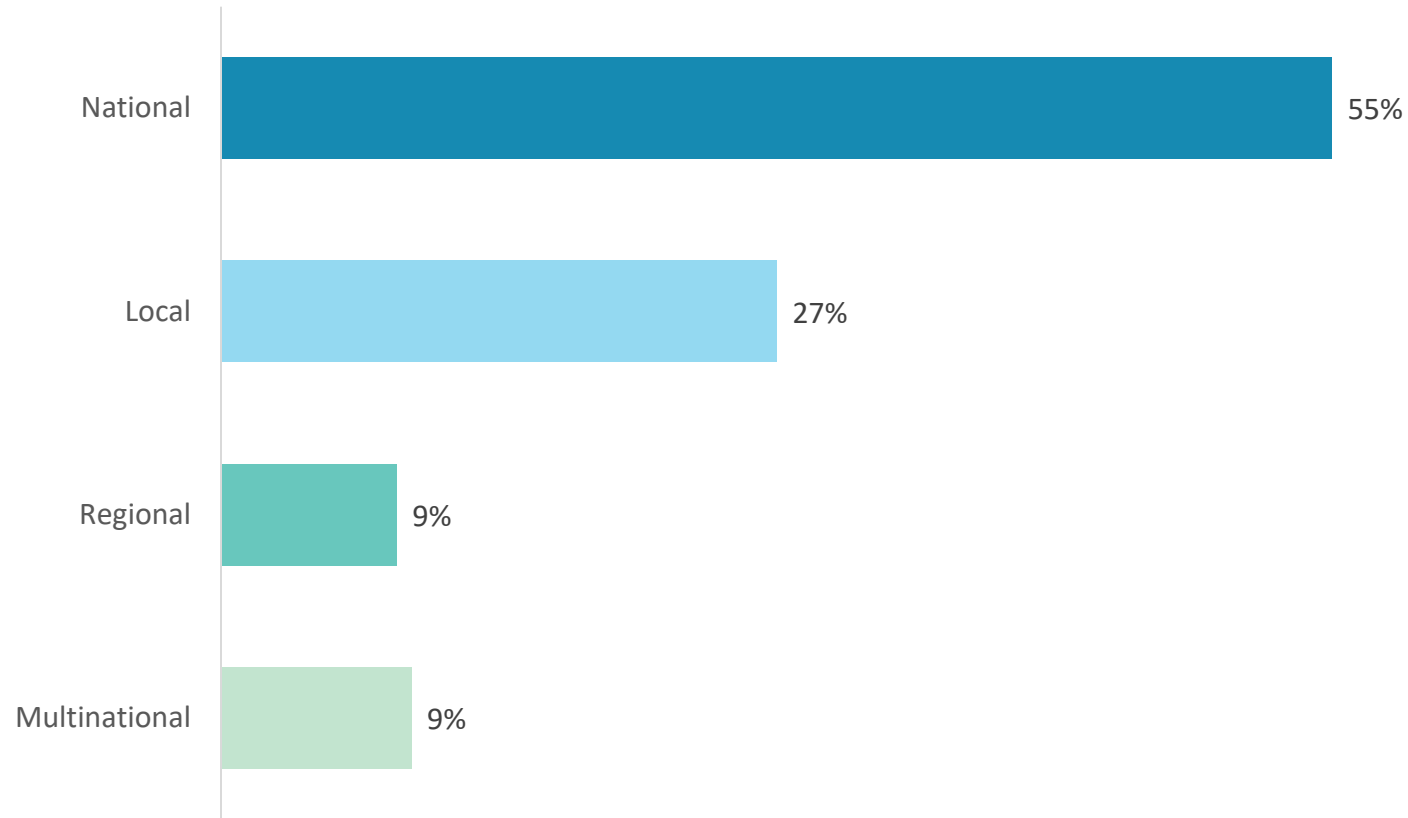
NLP Cases by Status



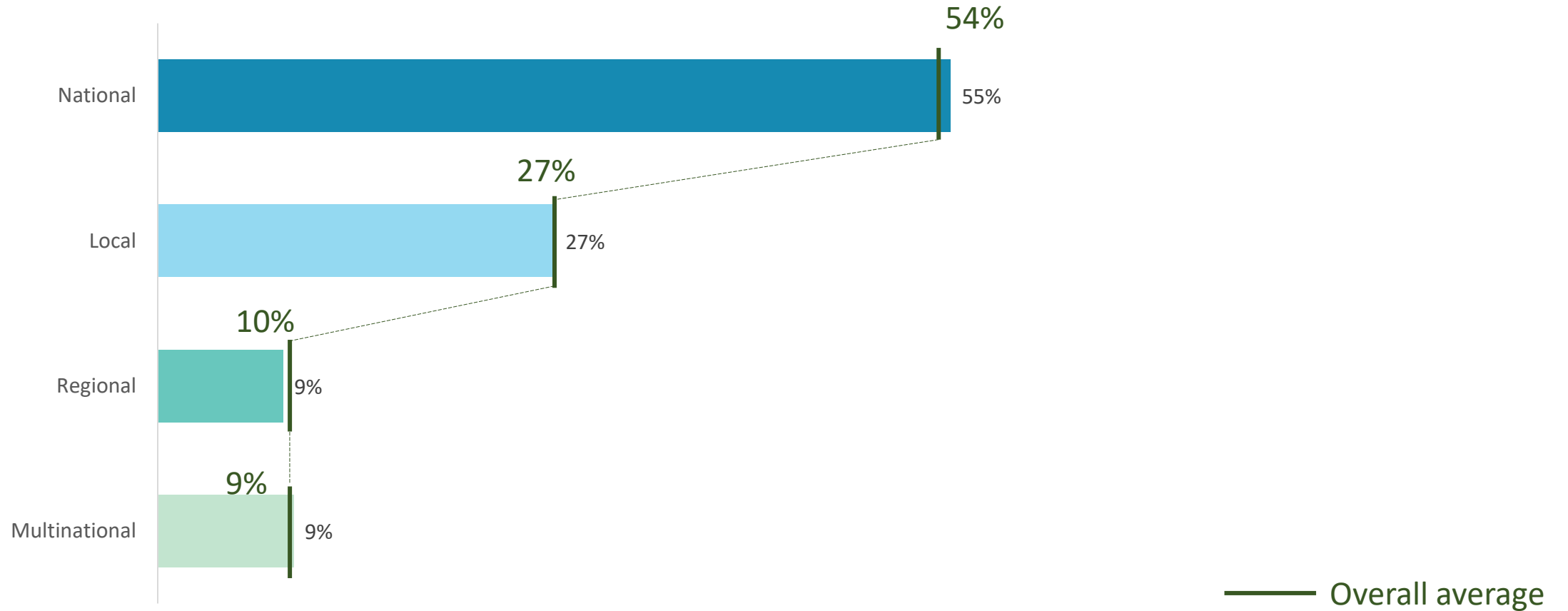
NLP Cases by Status



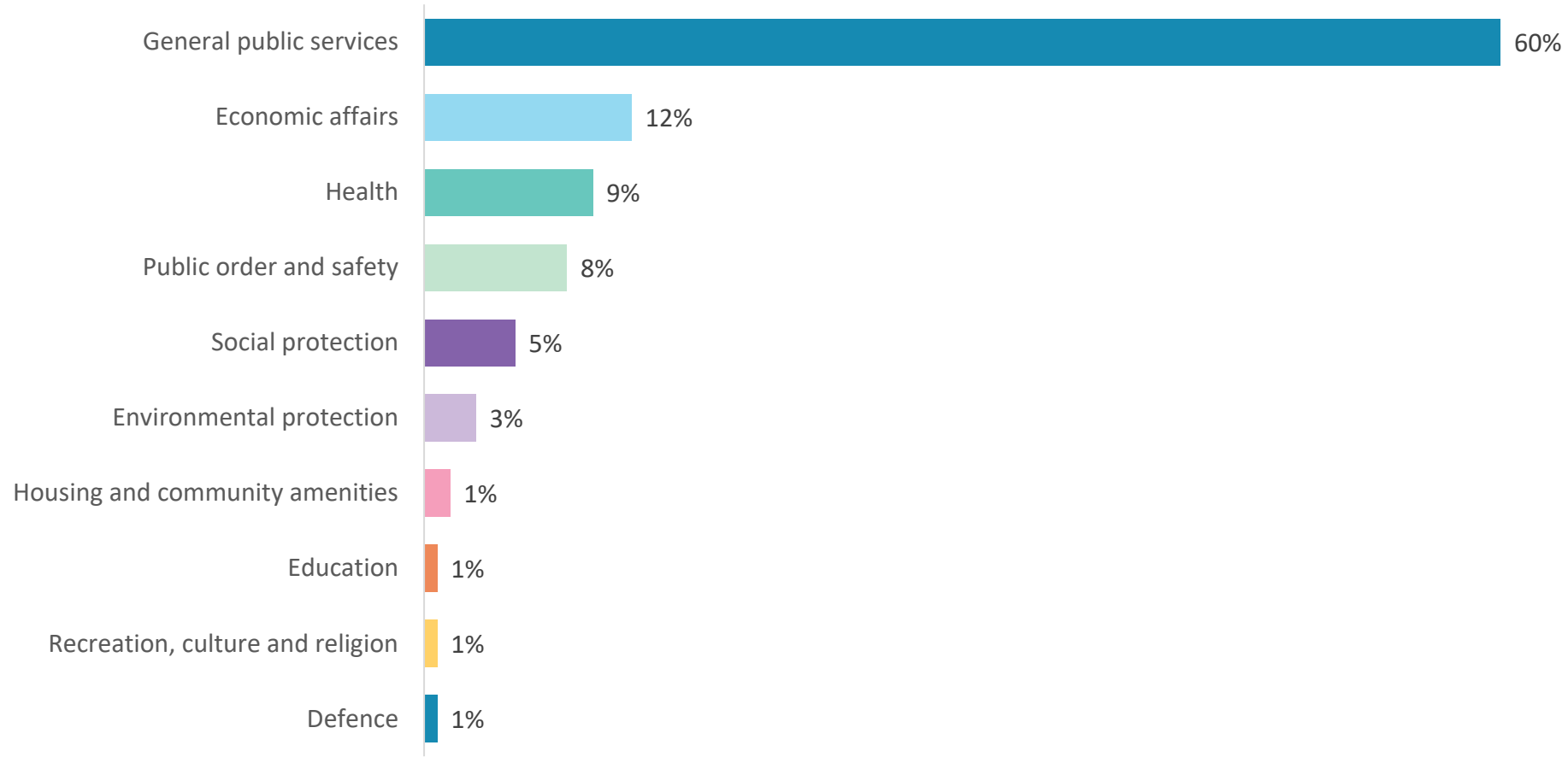
NLP Cases by Geographical Extent



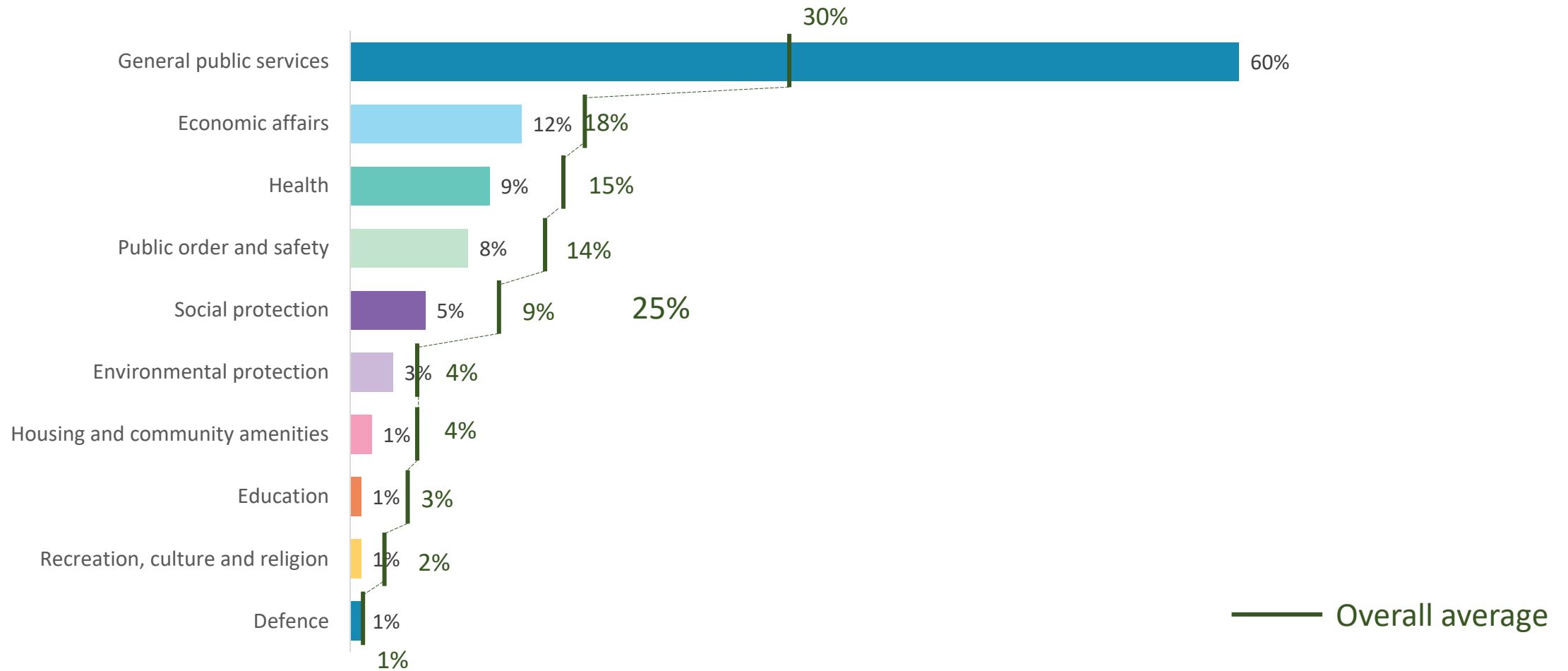
NLP Cases by Geographical Extent



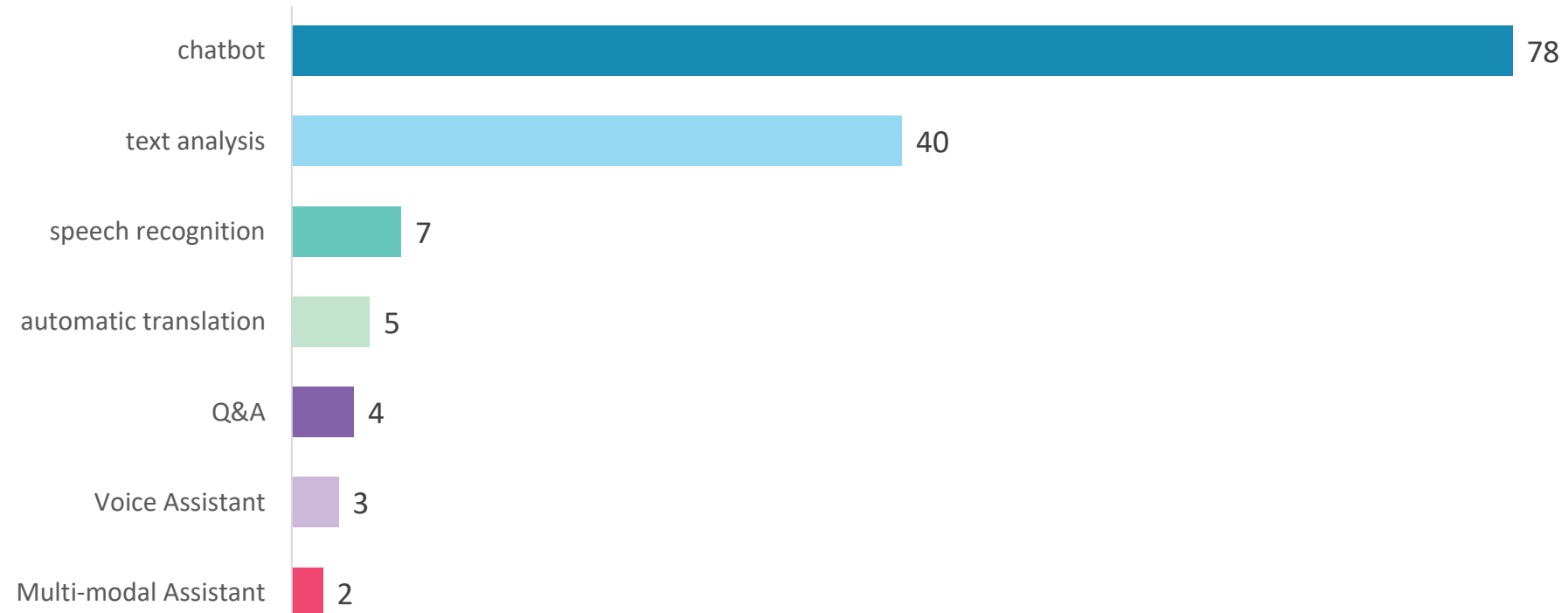
NLP cases by COFOG I



NLP cases by COFOG I



NLP cases by AI Keywords



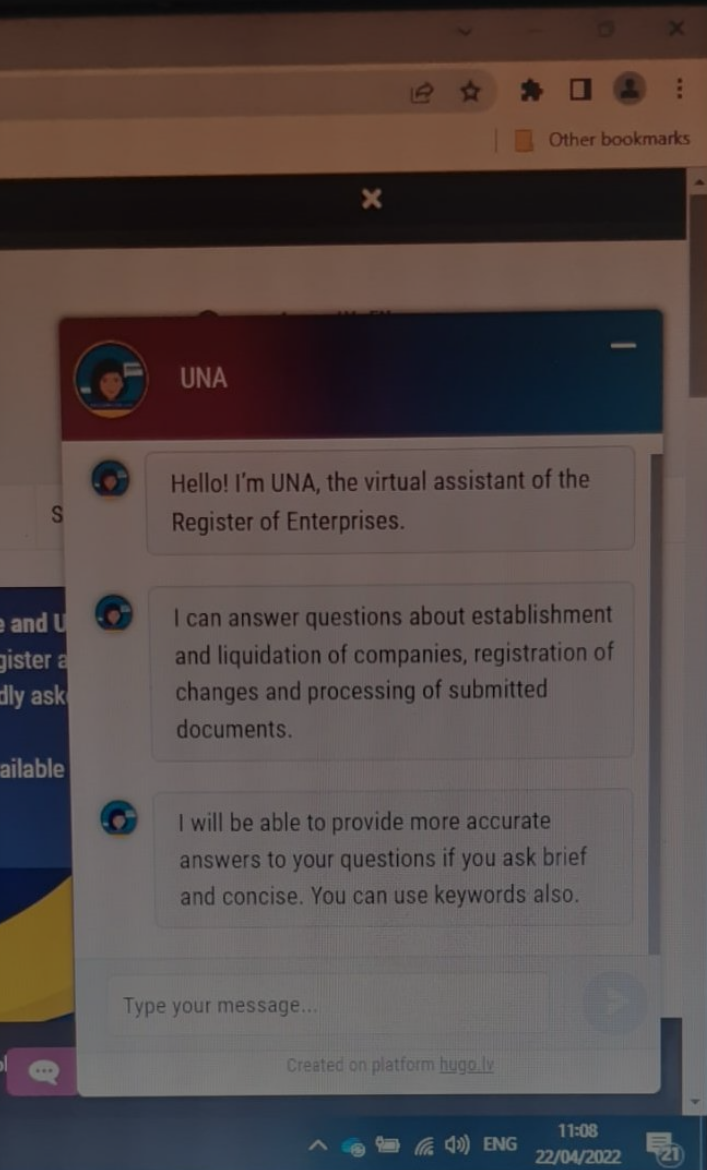
NLP cases in PS

Some examples



UNA Chatbot

- _ done by the Register of Enterprises of Latvia
- _ is able to answer about information on businesses (registration, liquidation, etc.)
- _ if citizens have an application in progress, it is able to answer on the progress of the documents
- _ according to the first performance indicators, 44% of the questions asked on UNA are easily answered by the Chatbot





Classification of phone calls

- _ NLP is used to automatically classify incoming phone calls at the [Flemish Infoline](#)
- _ it detects and categorizes the incoming questions faster and more 'finely-meshed', it integrates them into the editorial management
- _ it also suggests standardised answer to the operator



HAL - classification of documents

- _ the Ministry of Foreign Affairs (MFA) receives up to 6,000 reports from Norwegian embassies, delegations, etc.
- _ it analyses and classifies reports the content of these documents and to find almost all relevant information on a given subject matter
- _ the [AI solution](#) is also used to extract key information in reports and prepare summaries.



VeriPol - Detect false police reports

- _ [VeriPol](#) use a combination of NLP and ML classification algorithms, capable of estimating the probability of false police reports with significant accuracy
- _ it also enables insights into the differences between false and true police reports
- _ the system is integrated into the existing Spanish National Police information system

Thank you

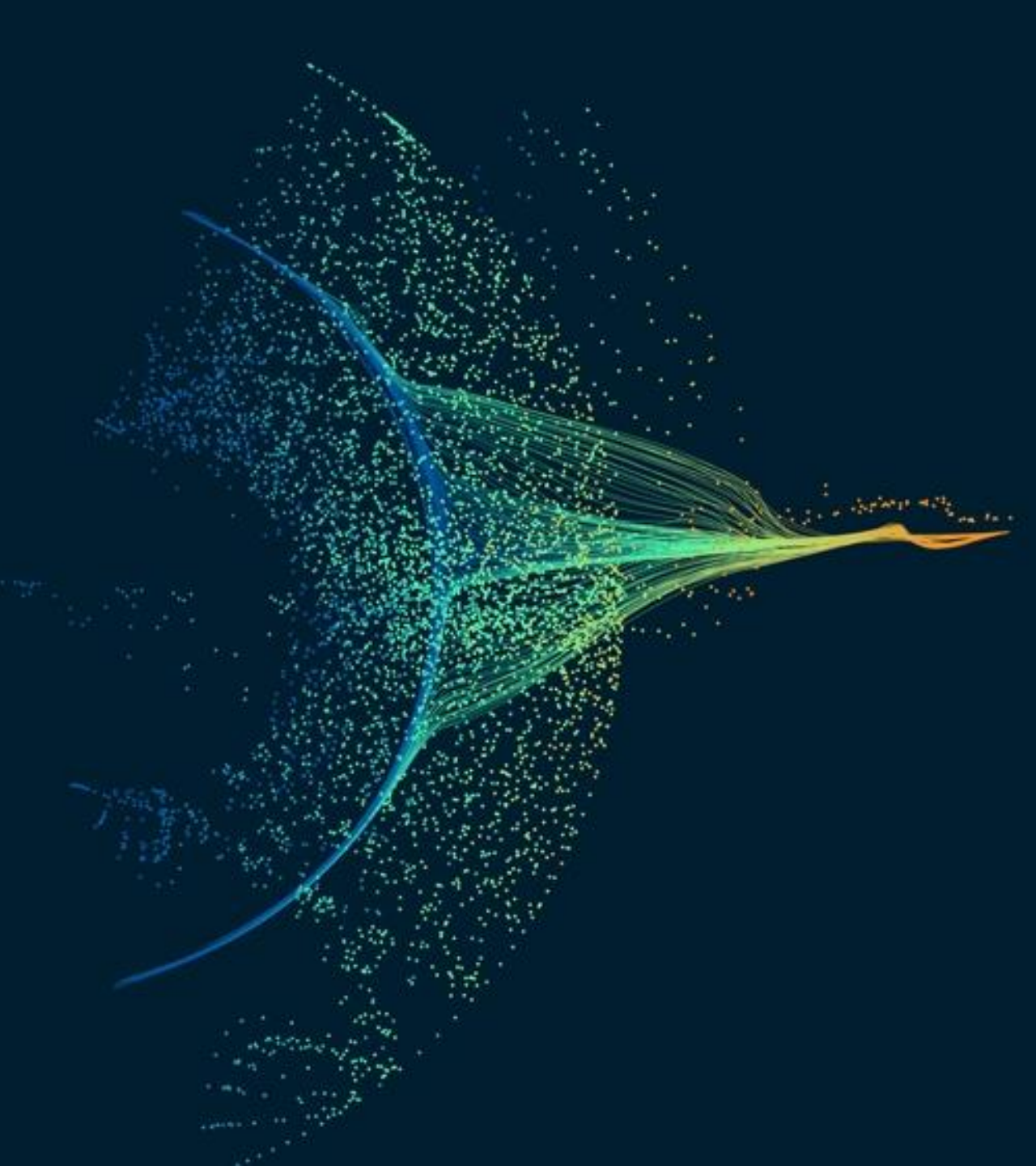
luca.tangi@ec.europa.eu
EC-AI-WATCH@ec.europa.eu



© European Union 2021

Unless otherwise noted the reuse of this presentation is authorised under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license. For any use or reproduction of elements that are not owned by the EU, permission may need to be sought directly from the respective right holders.

Images: <https://unsplash.com/>



Benefits of NLP

Q&A



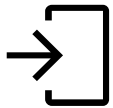
DEMO - NLP



**How can the
public sector
use NLP?**

Context

How to improve public services accessibility while saving time using Natural Language Processing?



The Single Digital Gateway

A repository of links has been created to foster the accessibility of public services

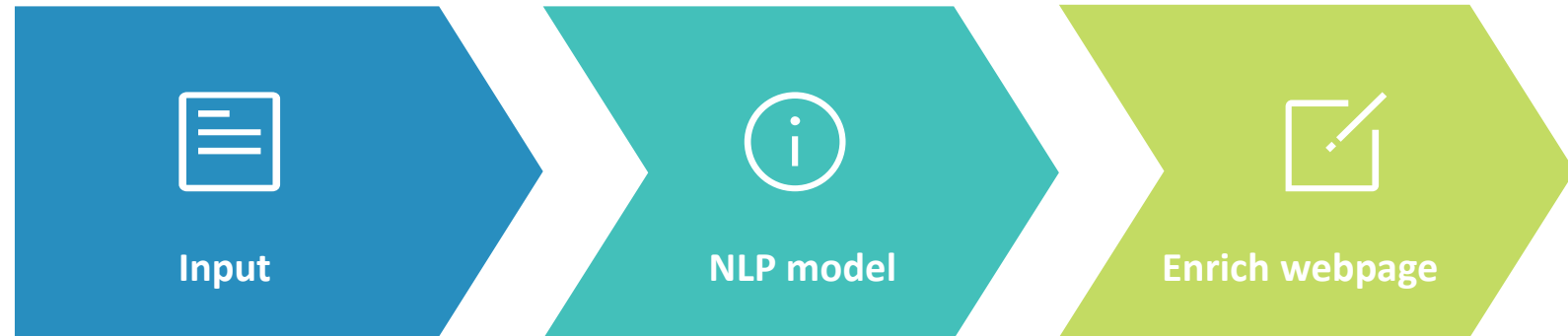


The Goals

- Improve the ranking of the web pages describing public services
- Reduce the time needed for sharing required information

NLP Prototype for the Single Digital Gateway

Multiple text classification techniques allow automatic retrieval of metatags based on the URL or the content of the page



Webpages that need to be tagged according to the SDG Search Service model.

Automatically enrich webpages in the HTML code

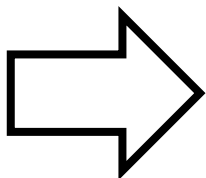
Benefits for users



Automation of previously manual work



The prototype can be scaled up easily



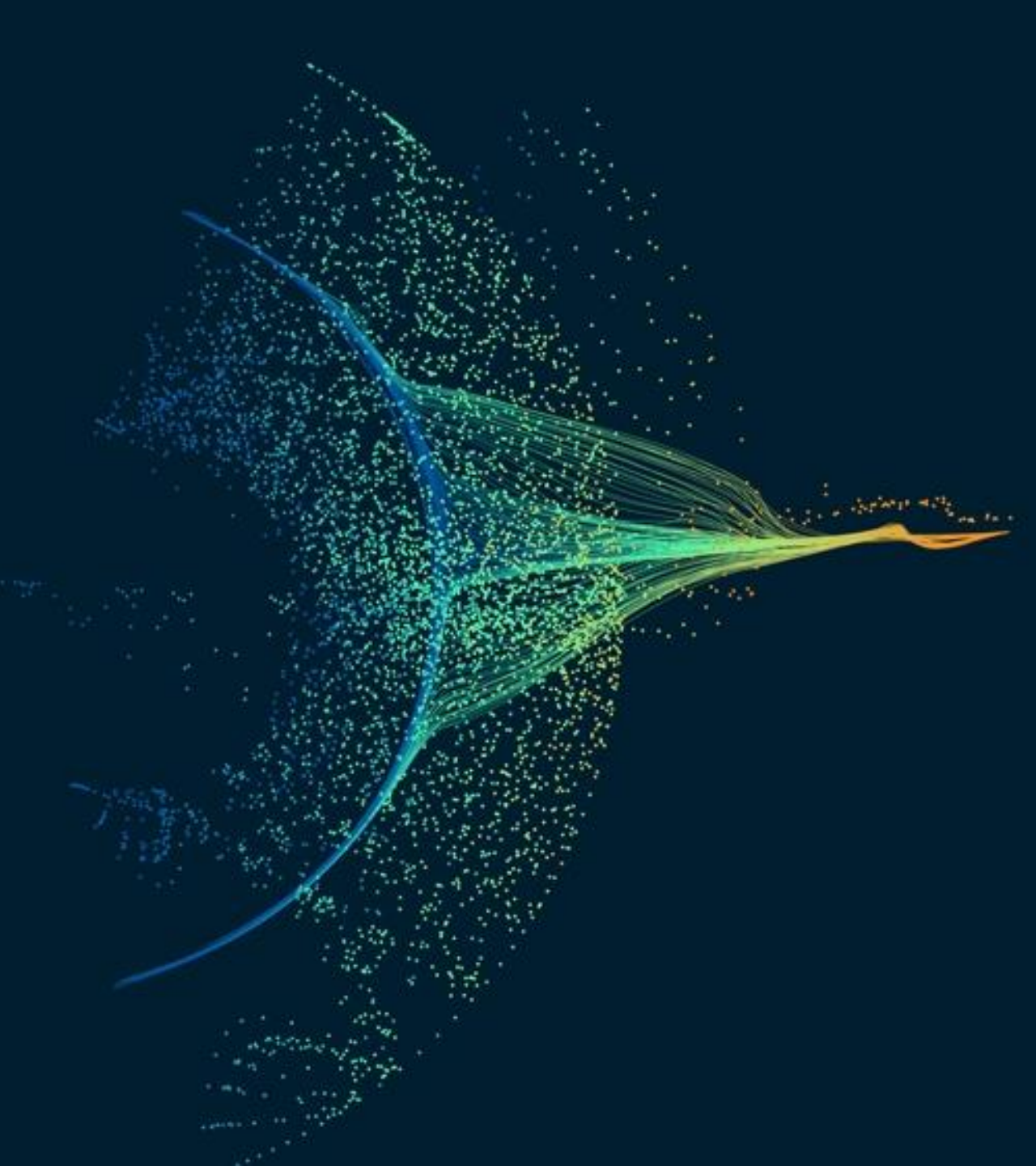
In the context of the SDG requirement for local authorities to provide SDG information by the end of 2022, this prototype could relief a great amount of manual work

The image features a complex network graph visualization on a dark blue background. The graph consists of numerous nodes, represented by small dots, and edges, represented by thin lines. The nodes and edges are color-coded, with a gradient from blue to green to yellow. The overall structure is symmetrical and resembles a butterfly or a pair of wings, with a central hub and many smaller clusters branching out. The text "Demonstration of POC" is centered in the middle of the image in a white, sans-serif font.

**Demonstration
of POC**

The image features a complex network graph visualization on a dark blue background. The graph consists of numerous nodes and edges, forming a dense, multi-lobed structure. The nodes are represented by small, glowing particles in shades of blue, green, and yellow. The edges are thin, curved lines connecting these nodes, also in similar colors. The overall shape is roughly symmetrical, with a central core and several radiating, tapering sections. In the center of the image, there is a white text overlay that reads "Demonstration of SeTA tool (Demo)".

**Demonstration
of SeTA tool
(Demo)**



NLP Demos

Q&A



DEAP Project

**What can we
learn from the
DEAP project? –
A successful NLP
Initiative**

Who we are

- Directorate General for Education, Youth, Sports and Culture (DG EAC)
- Directorate C. Innovation, Digital Education and International Cooperation
- Unit C4. Digital Education
- Portfolio of Commissioner Mariya Gabriel

- Directorate General for Informatics (DIGIT)
- Directorate D. Digital Services
- Unit D1. Data Services
- Portfolio of Commissioner Johannes Hahn



Digital Education Action Plan 2021-2027



Strategic priority 1

Fostering the development of a high-performing digital education ecosystem

Strategic priority 2

Enhancing digital skills and competences for the digital transformation



Stronger coordination and cooperation through the European Digital Education Hub

High quality and inclusive digital education

Digital Education Action Plan 2018-2020

Priority 1

Making better use of digital technology for teaching and learning

Priority 2

Developing relevant digital competences and skills for the digital transformation

Priority 3

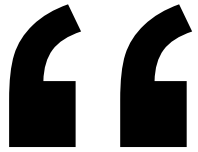
Improving education through better data analysis and foresight



Action 10:

Artificial Intelligence and Analytics Pilots

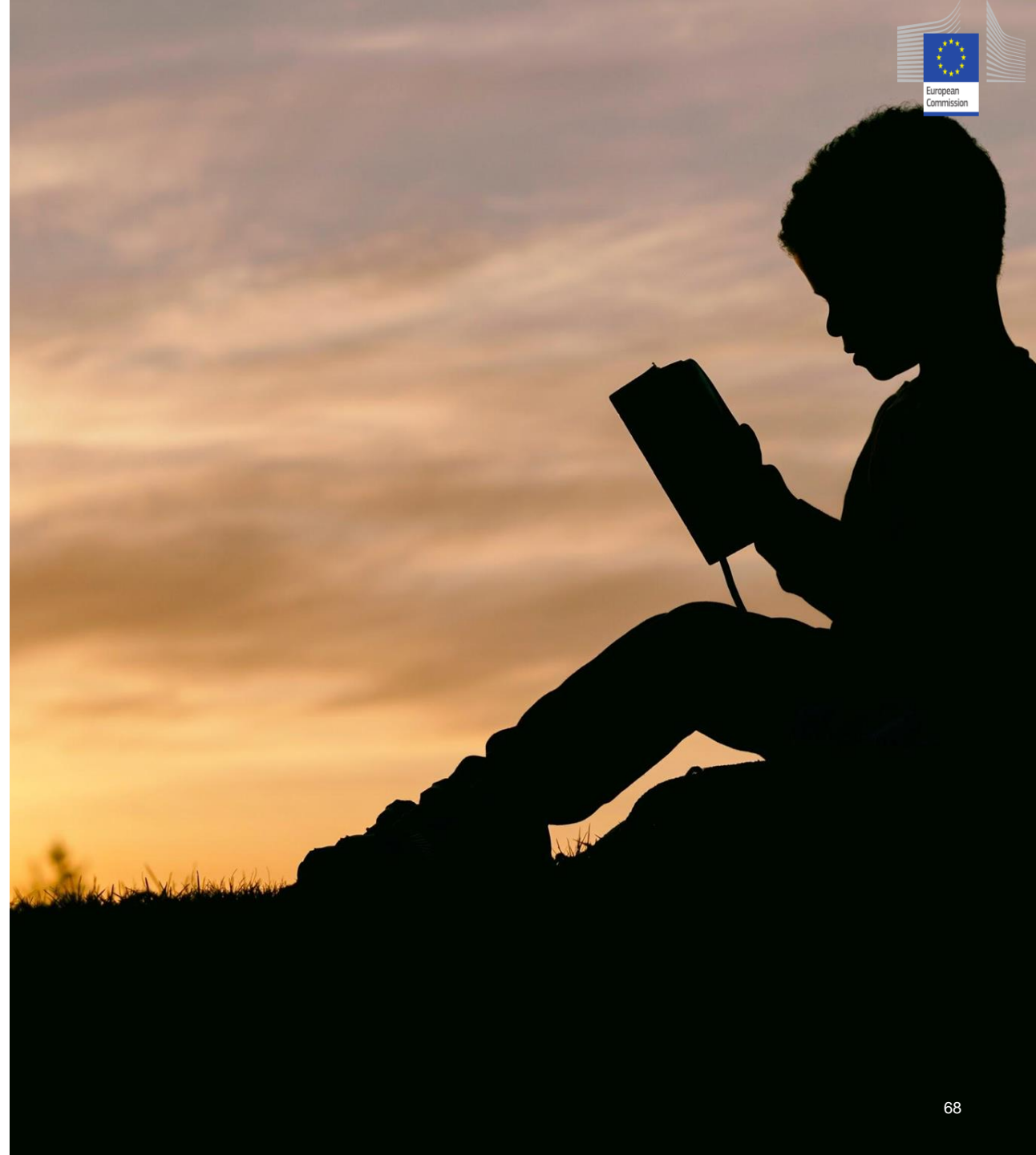
Digital Education Action Plan



The Commission has adopted a Digital Education Action Plan ([DEAP](#)) which includes 11 actions to support technology-use and digital competence development in education.

Action 10 is about AI and educational analytics.

In simpler words, "how can we make the best use of the data available online to capture insights and improve decision-making in education"?.



Digital Education Action Plan



Use data and AI to predict future skills and skills shortages to demonstrate the added value of analytics for the definition of education policies.



The Challenge

- Data is vital for education and training. Using technology, we have created data that can be **exploited**. But how?
- Data also helps to identify and address needs for evidence-based required policy measures, but **comparative** data especially is rarely used.
- Education data and trends are generally collected in a top-down way, led by international organisations and governments. The **user's perspective** is often not sufficiently considered.
- Foresight: from lagging behind to anticipating change. Education and training institutions are trying to **catch up** with technological developments.

Early developments

DG EAC
+
DIGIT

Pilots to study possible applications of AI in education. Two prototypes developed:

- PoC 1: Skills Map Panorama
- PoC 2: Impact of AI on Sectors – Curriculum generator

EAC-DIGIT AI pilots

Context: DG EAC started cooperating with DG DIGIT in 2019 to identify possible pilot projects using AI and analytics in education. Within six months, **two Proofs-of-Concept (PoCs)** have been identified and developed.

From PoC to Project: The two PoCs were implemented as prototypes and represent a good input for reflections on the impact of AI on educational analytics and on how a smart use of the data can support education design and policymaking.

DG EAC is carrying out a project of development of the two PoCs to transform them in mature services, potentially supporting education institutions interested in improving their educational programmes (***Go-Live by June 2022***).

The services will be made available online, use the infrastructure of the European Commission Data Platform and be open for use through the internet without any restriction.

EAC-DIGIT AI pilots

Skills Map Panorama

- **Connecting educational programs to jobs and vice versa**

Summary:

- How AI and Natural Language Processing help policy makers or DG EAC to visually inspect the levels of similarities of various MSc programs across EU member states to occupations

Why is it important?

- DEAP Action 10 is about AI and educational analytics. In simpler words, "how can we make the best use of the data available online to capture insights and improve decision-making in education"?

EAC-DIGIT AI pilots

Impact of AI on Sectors – Curriculum generator

- **Finding the most relevant AI topics in professional sectors**

Summary:

- How AI can help identifying which topics of IT are emerging in various professional sectors. Educational stakeholders such as University managers, policy makers, etc. can select a set of professional sectors, provided by the [NACE](#) classification, and visualize which AI-topics are emerging, based on the top-ranked research communities for these sectors.

Why is it important?

- Education is a concept of high volatility, new technologies, methods and ideas are emerging. Educational domain experts can use AI to modernize traditional sectors by introducing new courses that can help them boost their educational impact.

EAC-DIGIT: the AI4Educat Project

The *Skills Map Panorama* PoC has been transformed to:

Service 1: The skills & education matchmaking tool

User Story
(overall activity)

“As an employee from HEI, I would like to access data and analytics that help me understand the relation between education programmes and real-world skills and occupations, so that I can compare what we are teaching to students to what is required on the job market, and therefore model my curriculum accordingly”.

The *Impact of AI on Sectors* PoC has been transformed to:

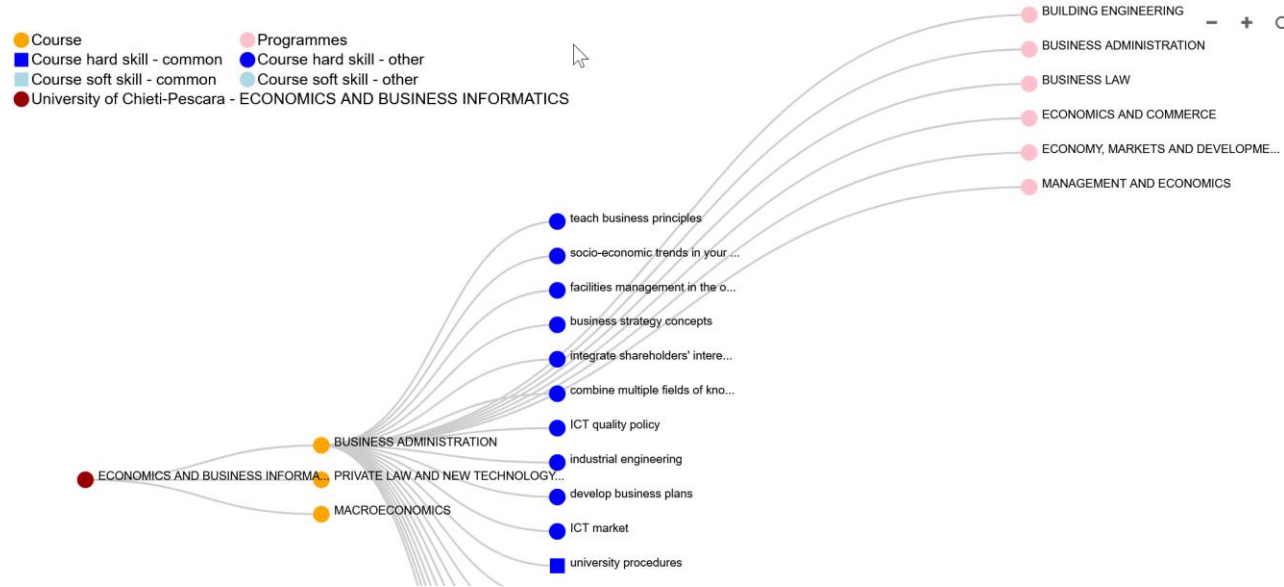
Service 2: Technology watch for education

User Story
(overall activity)

“As an employee from HEI, I would like to access to macro level data, in order to understand what AI is, how AI technologies are evolving and to compare this data to my own curriculum, so that I can review and adjust the content of it accordingly”.

EAC-DIGIT: the AI4Educat Project

Service 1: The skills & education matchmaking tool



Service 2: Technology watch for education

Explore AI courses

Search criteria

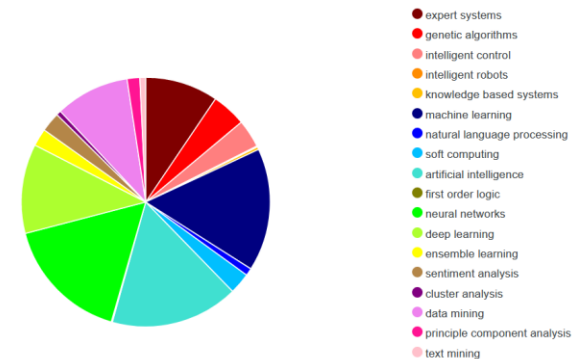
Country: Lithuania

HEI: Kaunas University of Technology

Programme: Intelligent Robotics Systems

DISCLAIMER: The quality of results depends on the volume of the data collected.

AI technologies for the AI courses



EAC-DIGIT: the AI4Educat Project

Data Sources

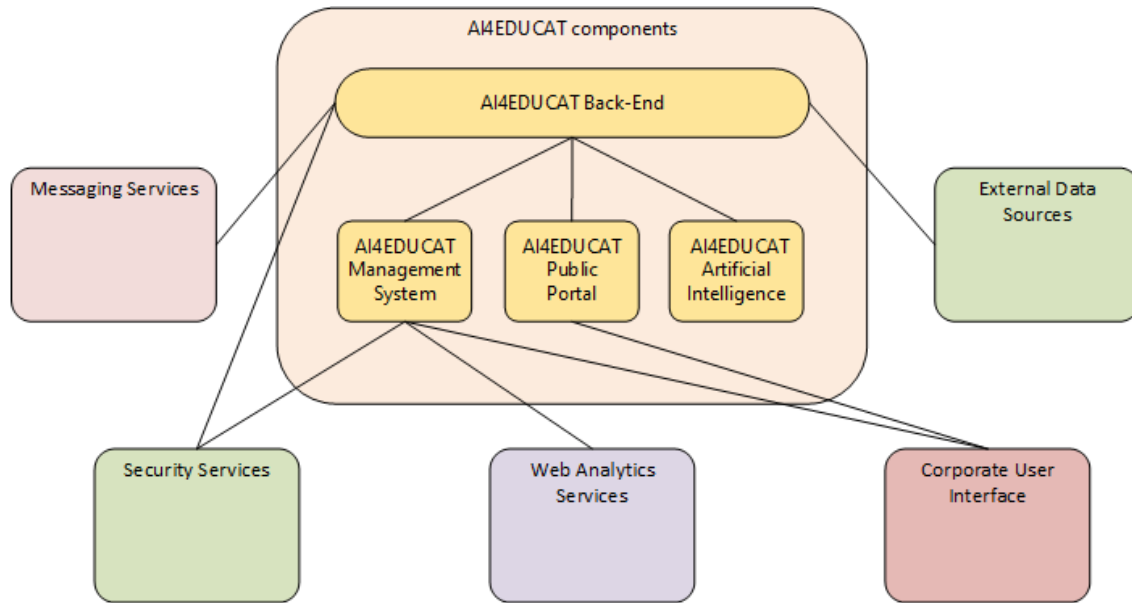
Category	Amazon RDS	Amazon S3
Education Data	45GB ²	4GB
Publications	200MB (for 5 years)	N/A
ESCO Dataset	100MB	700MB
Trained ML Models (Architectural Decision AD-001)	N/A	Up to 2GB
Total	45.3GB	6.7GB

Architectural NLP Decisions

- Use Sentence-BERT for producing Embeddings for textual content
- Use Spark NLP for the pre-processing task of cleaning text and lemmatization
- Use ESCO Dataset as basis to construct a Training Dataset for text classification
- Use CSO Classifier to classify Courses and Publications to AI Technologies
- Use LightGBM as Classifier to identify the relevant pairs of Courses and Skills
- Host AI4EDUCAT in AWS Cloud Environments of DIGIT
- Use Amazon EMR as big data platform
- Use GitLab as CI/CD tool
- Use Amazon ECS and Docker for deploying the Spring Boot application

EAC-DIGIT: the AI4Educat Project

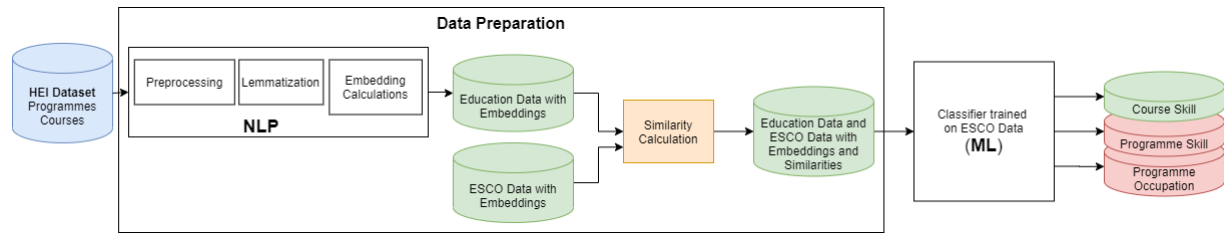
High-Level Context and Component Diagram



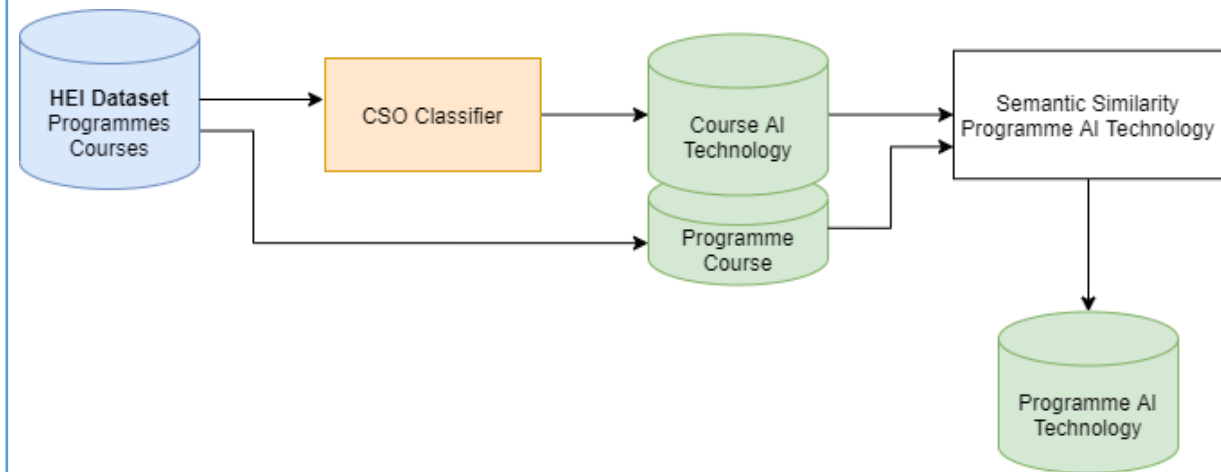
Name	Description
AI4EDUCAT Management System (AI4EDUCAT-MS)	Web Application front-end based on eUI that allows HEI Employees to provide curriculum data and System Administrators to manage the system in general
AI4EDUCAT Public Portal (AI4EDUCAT-PP)	Web Application front-end based on eUI that allows Public Users to explore and analyse curriculum data in various ways
AI4EDUCAT Back-End (AI4EDUCAT-BE)	Web Application back-end that allow both Web Application front-ends to integrate with corporate building blocks and to interact with AI4EDUCAT specific data; it allows also to HEI systems to provide curriculum data, as well as, retrieves data from external sources, like Publications
AI4EDUCAT Artificial Intelligence (AI4EDUCAT-AI)	Set of NLP/ML algorithms for all needed AI Processing, like calculating Embeddings for Courses and Programmes, calculating similarities between Courses and Skills, computing AI Technologies for Courses and Publications, etc.

EAC-DIGIT: the AI4Educat Project

Data Pipeline for Classification of Educational Data for Service 1

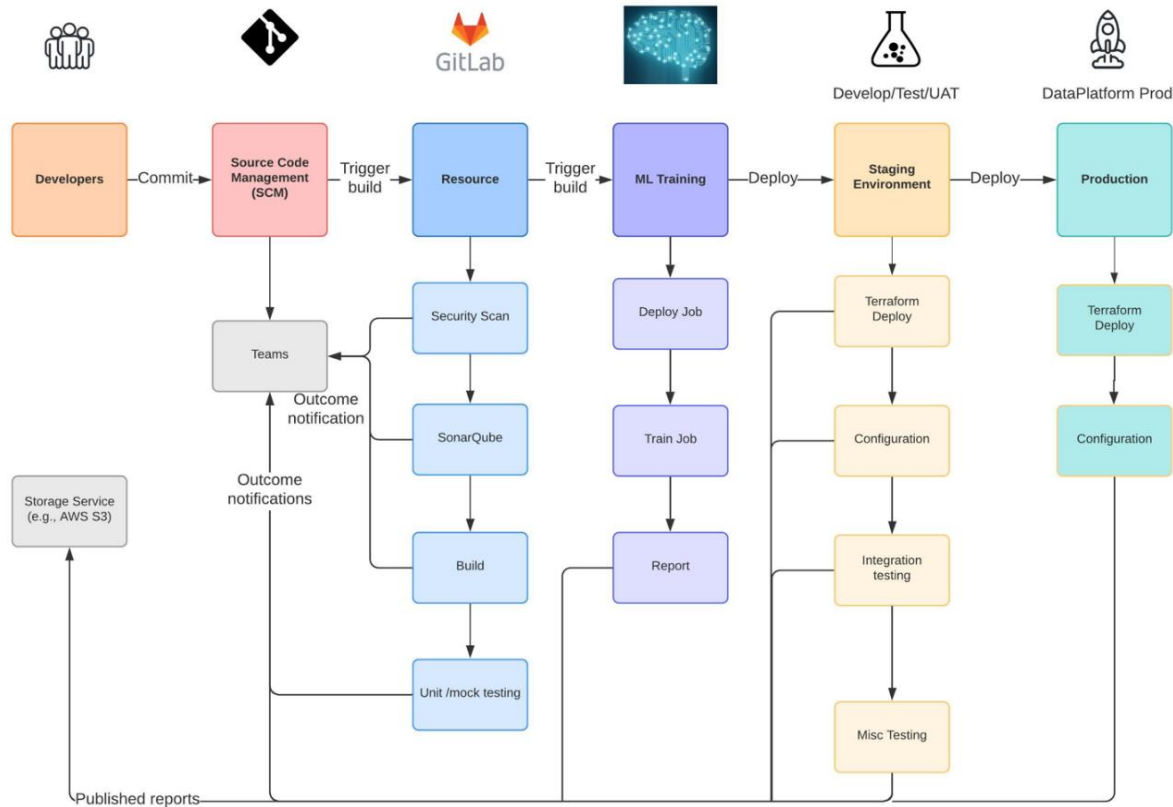


Data Pipeline for Classification of the Publication Dataset for the Online Service 2

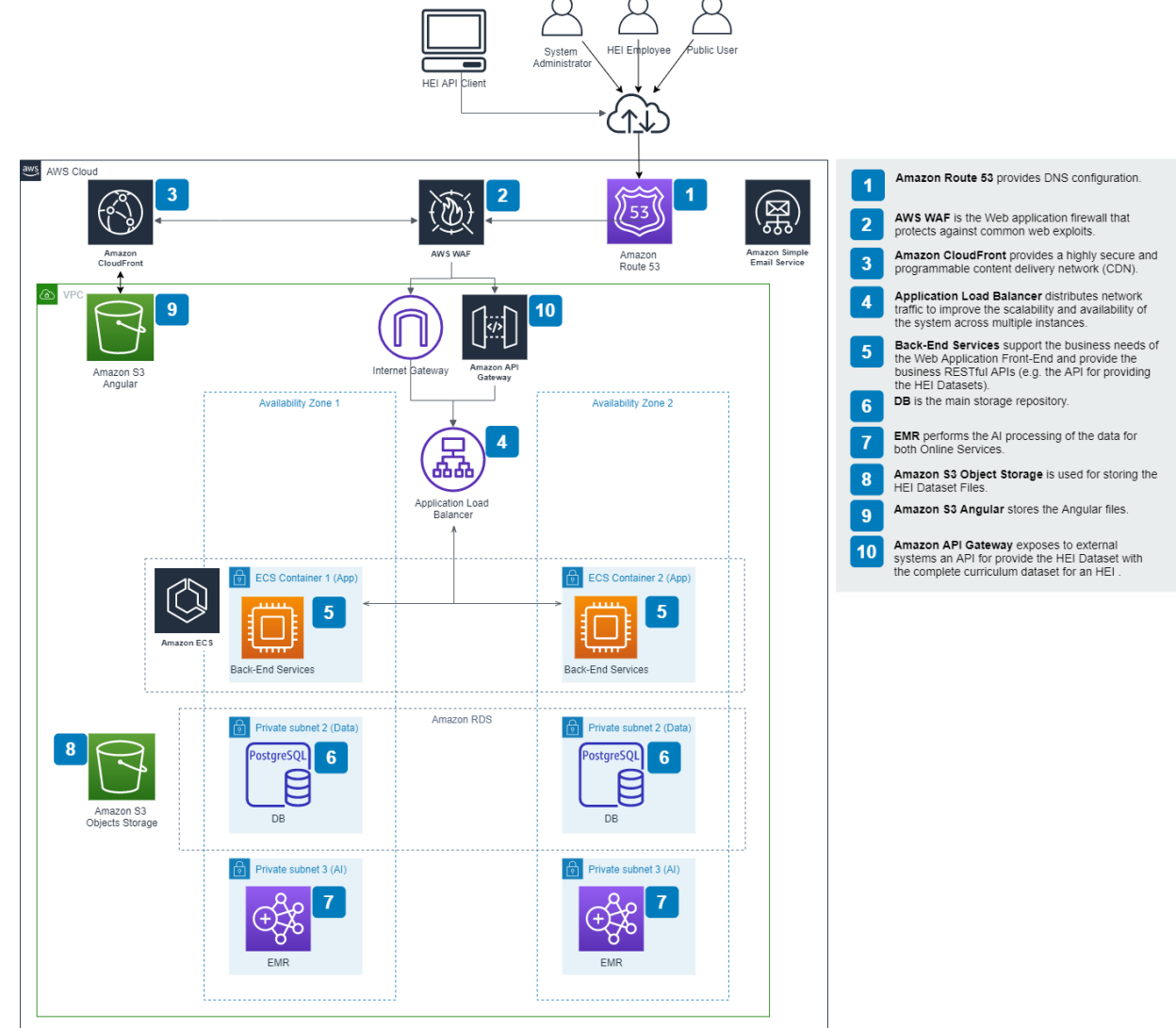


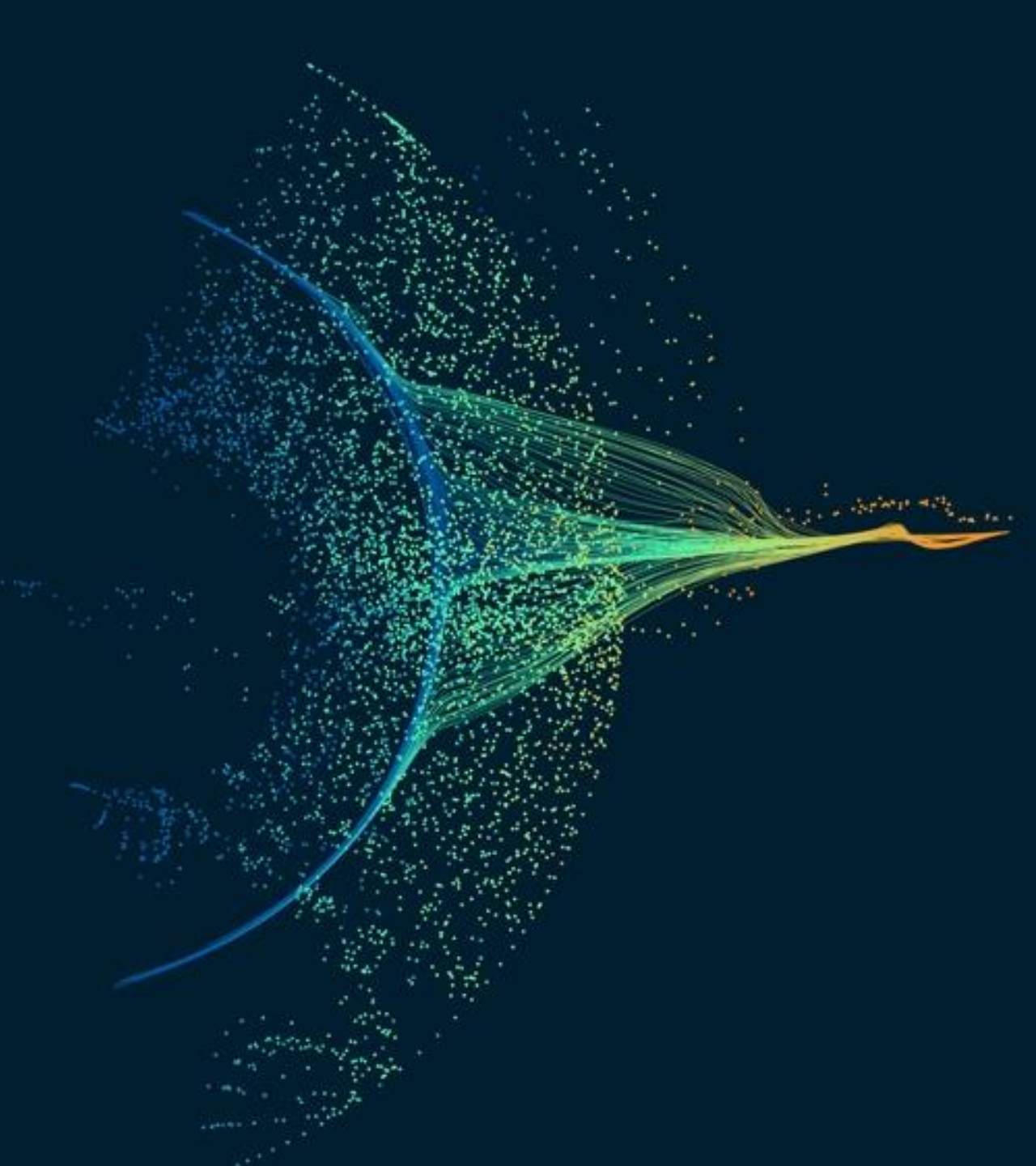
EAC-DIGIT: the AI4Educat Project

Generic Overall Pipeline Structure



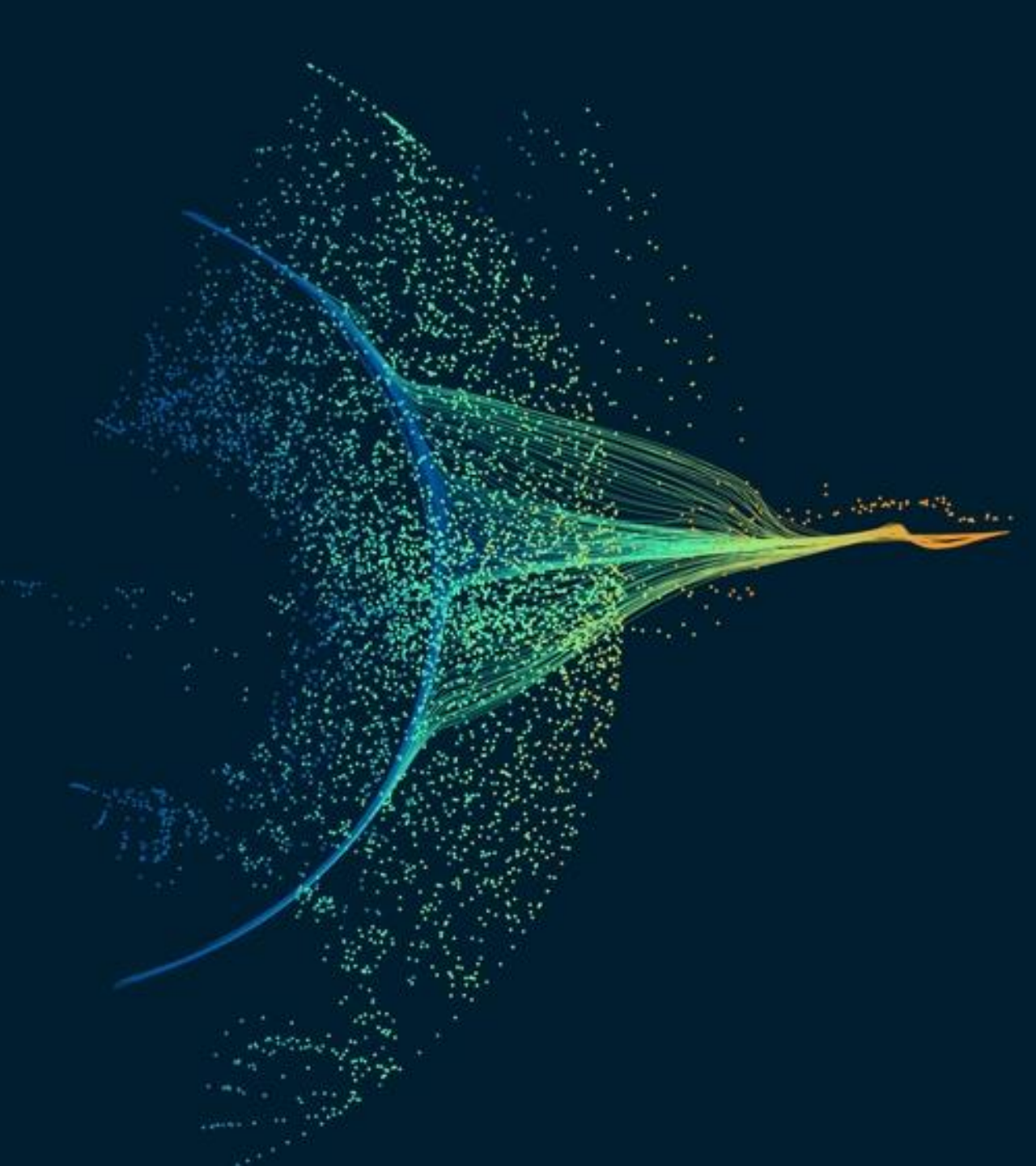
AWS Architecture Diagram





DEAP project

Q&A



Mural:
Time for some interactions!



Regulatory Compliance

How can public administrations comply with regulatory requirements



EXCELLENCE
& TRUST

The graphic features a stylized brain with circuitry patterns in shades of purple and blue. Below the brain, two hands are shown holding a glowing, spherical object. The background is a dark purple circle with a bokeh effect of light spots.

2021 Artificial Intelligence Package

**Georgi Gitchev, Legal and Policy Assistant, Unit A2, DG
CONNECT, European Commission**

**Knowledge Sharing Session:
Natural Language Processing for Public Services**

28 April 2022

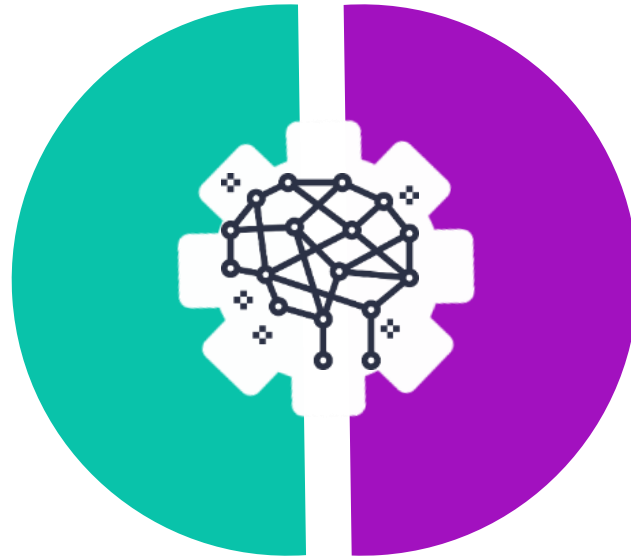
2021 European Commission AI Package: Proposal for an Artificial Intelligence Act

- **Introduction: AI Act Logic and Rational**
- **Main elements and principles:**
 - scope of application
 - horizontal risk-based approach
 - main requirements
 - implications for NLP systems
- **State of play and next steps**
- **Other relevant legislation**



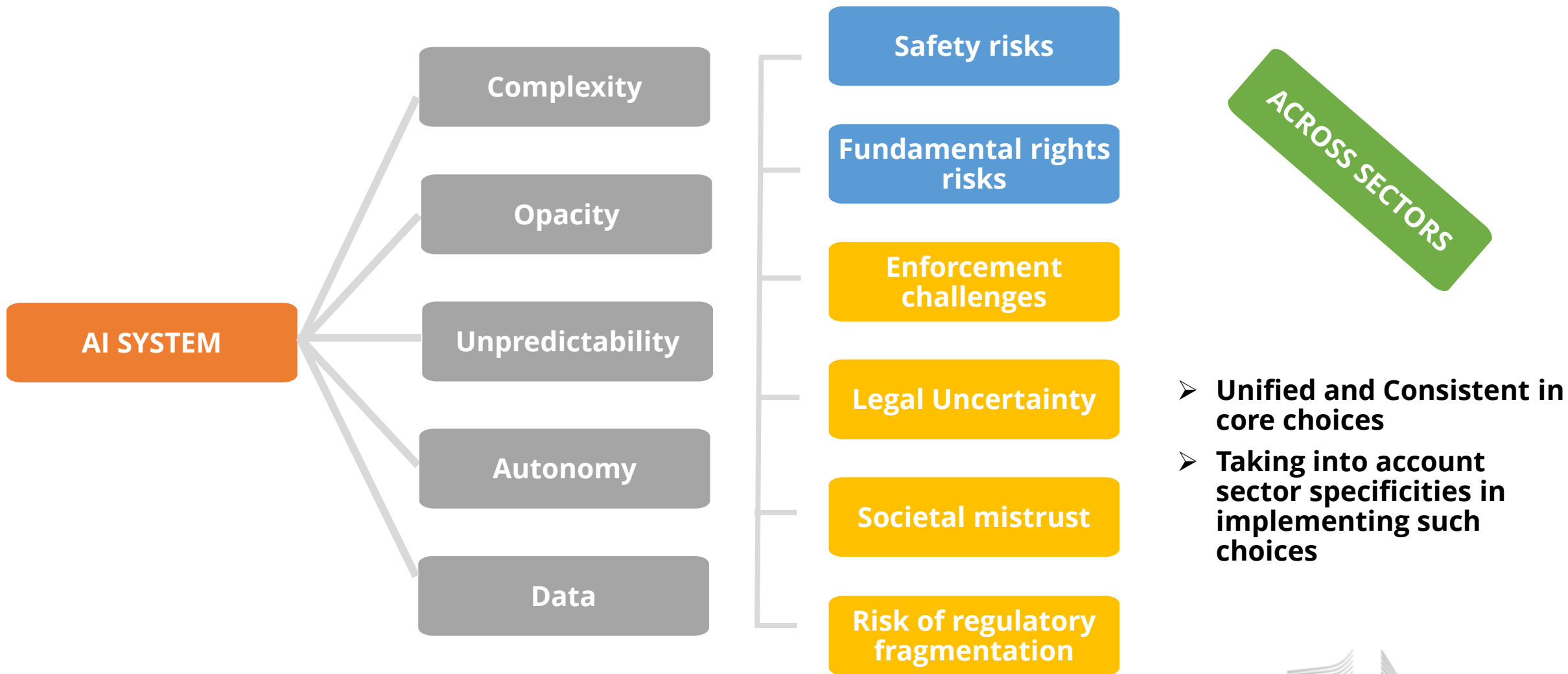
AI is good ...

- For citizens
- For business
- For the public interest



... but creates some risks

- For the safety of consumers and users
- For fundamental rights



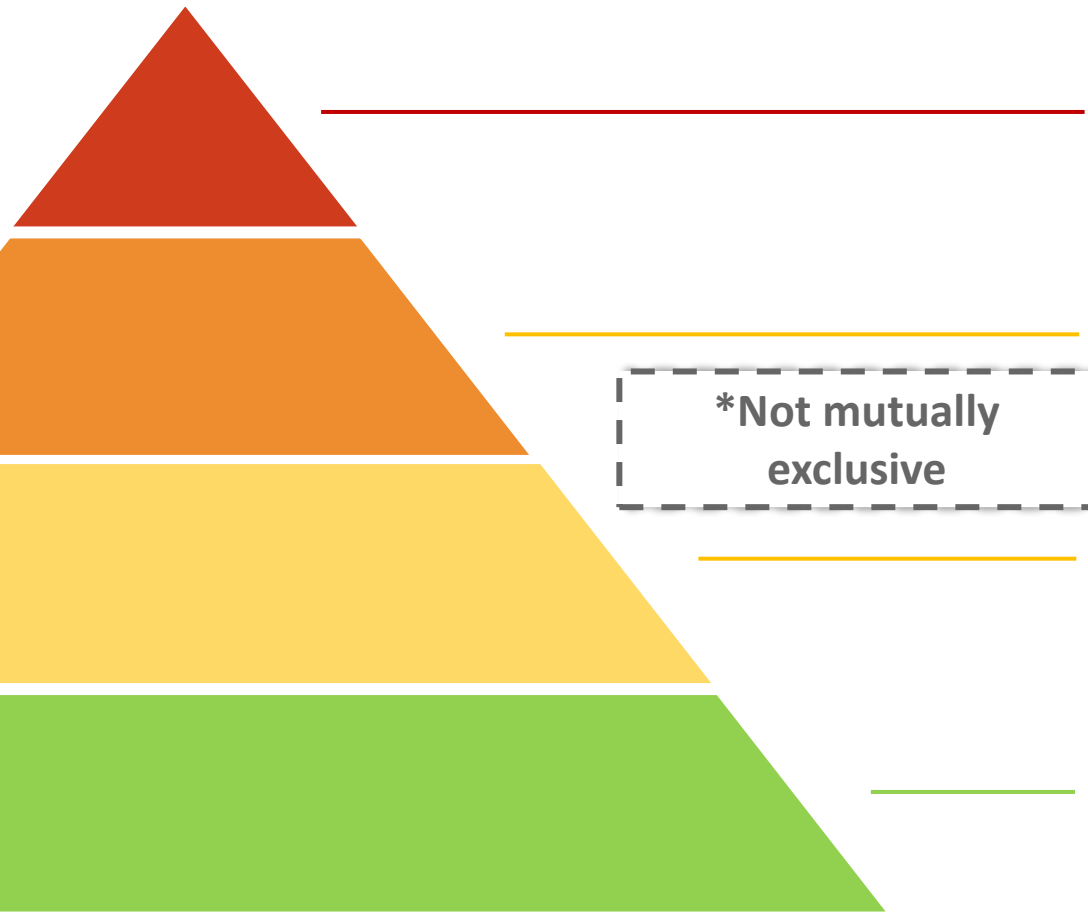
Regulation applicable to:

- ▶ **Providers (public or private)** placing on the market or putting into service AI systems in the Union independent from their origin
- ▶ **Users (public or private)** located within the Union
- ▶ **Providers and users** located in a third country, where the output produced by the system is used in the Union

Excluded from the scope:

- ▶ Public authorities in a third country or international organisations who use AI systems in the framework of international agreements for law enforcement and judicial cooperation with the Union or with one or more Member States
- ▶ AI developed or used exclusively for military purposes
- ▶ National security also out of scope

▶ **“AI system” definition: Article 3(1) + Annex I**



Unacceptable risk

e.g. social scoring, real time RBI for law enforcement purposes

Prohibited

High risk

e.g. recruitment, social benefits, law enforcement

Permitted subject to **compliance with AI requirements and ex-ante conformity assessment**

*Not mutually exclusive

'Transparency' risk

'e.g. chatbots, emotion recognition, biometric categorisation)

Permitted but subject to information/transparency obligations

Minimal or no risk

Permitted with no restrictions
Possible voluntary codes of conduct

High-risk AI Systems

1 SAFETY COMPONENTS OF REGULATED PRODUCTS

(e.g. medical devices, machinery) which are subject to third-party assessment under the relevant sectorial legislation

2 CERTAIN (STAND-ALONE) AI SYSTEMS IN THE FOLLOWING FIELDS

- ✓ Biometric identification and categorisation of natural persons
- ✓ Management and operation of critical infrastructure
- ✓ Education and vocational training
- ✓ Employment and workers management, access to self-employment
- ✓ Access to and enjoyment of essential private services and public services and benefits
- ✓ Law enforcement
- ✓ Migration, asylum and border control management
- ✓ Administration of justice and democratic processes



- NLP is a form of AI that gives systems the ability to identify, process, understand and/or generate information in written and spoken human communications (*Commission study on Natural Language Processing for Public Services*).
- General purpose AI → is not by itself considered a high-risk AI system unless intended to be used in one of the high-risk areas.

BUT

- ▶ Consider the purpose of the NLP application
 - ▶ If intended to be used in a high-risk use case (e.g. Annex III, point 6 (d)), it will be classified as a high-risk AI system
- ▶ The user who uses or further develops the NLP system for a high-risk intended purpose will be considered as a “provider” of a high-risk AI system
 - ▶ Responsible for compliance with the AI Act requirements

NB: This is according to the proposal as it currently stands and subject to changes!

Requirements for high-risk AI systems

Establish and implement risk management processes

Use high-quality **training, validation and testing data**

Establish **documentation** and design logging features

Ensure **transparency** and provide users with **information**

Ensure **human oversight**

Ensure **robustness, accuracy and cybersecurity**

Obligations for high risks AI Systems

Provider obligations
(incl. public bodies developing in-house)

- ▶ Undergo **conformity assessment** (self assessment for Annex III except for RBI systems) and potentially re-assessment of the system (in case of significant modifications) to demonstrate compliance with AI requirements
- ▶ Establish and Implement **quality management** system in its organisation
- ▶ Draw-up and keep up-to-date **technical documentation**
- ▶ **Keep logs** to monitor the operation of the high-risk AI system (when empowered by law or the user)
- ▶ **Register** stand-alone AI system in public EU database
- ▶ **Affix CE** marking and sign declaration of conformity
- ▶ Conduct **post-market monitoring** and take corrective action
- ▶ **Report serious incidents and malfunction** that can pose risks to fundamental rights
- ▶ **Collaborate** with market surveillance authorities

User obligations
(AI developed in-house or bought off the shelf)

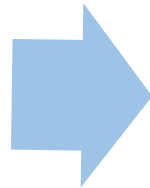
- ▶ Operate AI system in accordance with **instructions of use**
- ▶ Ensure **human oversight** when using AI system (for public authorities essential)
- ▶ **Monitor** operation for possible risks
- ▶ **Inform the provider or distributor about any serious incident** or any malfunctioning
- ▶ Use the information given by the provider for the **data protection impact assessment** (where applicable)
- ▶ **Existing legal obligations** for users continue to apply (e.g. under GDPR, public administrative law)



CE marking is an indication that a product complies with the requirements of a relevant Union legislation regulating the product in question. In order to affix a CE marking to a high-risk AI system, a provider shall undertake **the following steps**:

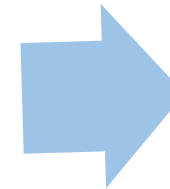
1

Determine whether its AI system is **classified as high-risk** under the new AI Regulation



2

Ensure design and development and quality management system are **in compliance with the AI Regulation**



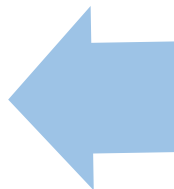
3

Conformity assessment procedure, aimed at assessing and documenting compliance



4

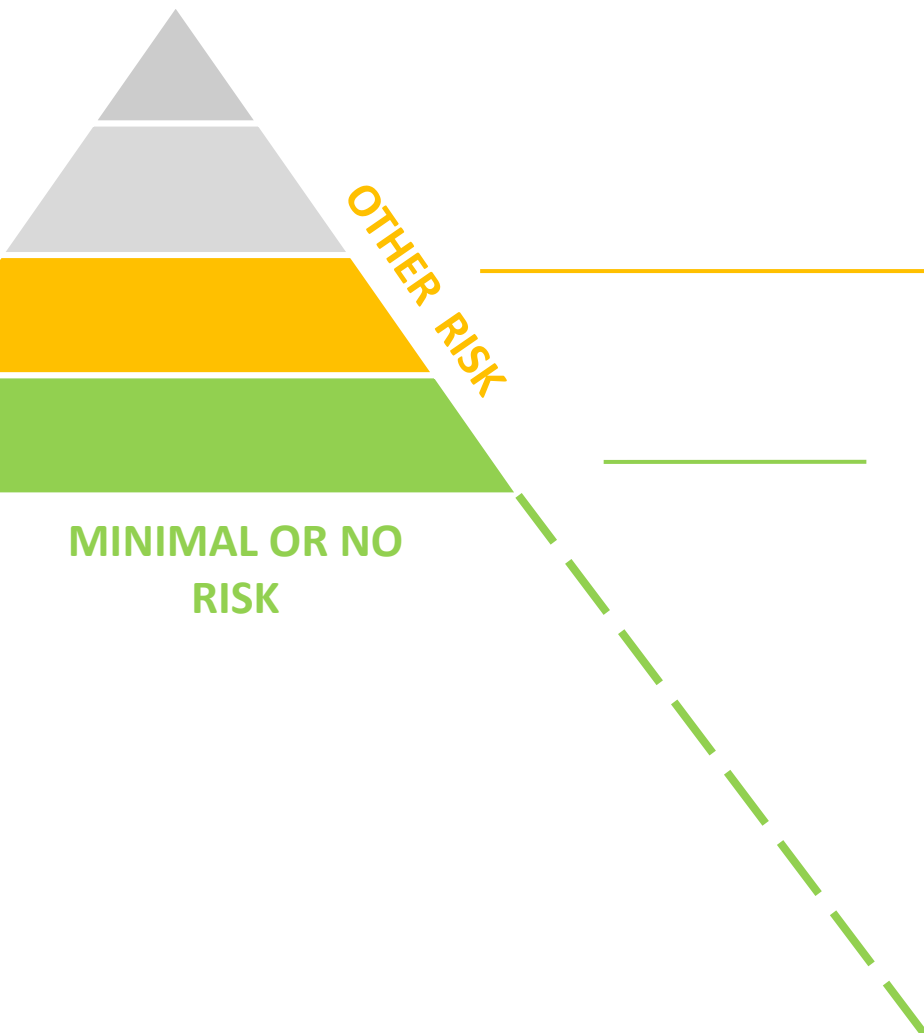
Affix the CE marking to the system and sign a declaration of conformity



5

PLACING ON THE MARKET or PUTTING INTO SERVICE





New transparency obligations for certain AI systems (Art. 52)

- ▶ **Notify humans** that they are **interacting with an AI system** unless this is evident.
- ▶ Applicable to chat bots, among others.

Possible voluntary codes of conduct for non-high-risk AI (Art. 69)

- ▶ Commission and Board to encourage drawing up of codes of conduct intended to foster the **voluntary application of requirements to non-high-risk AI systems**.

National level

Key role for enforcement

National Market
Surveillance Authorities



Rules for Penalties



**EDPS acts as a supervisory
authority for EU institutions,
bodies and agencies**

European level

Coordination of implementation and
exchange

Artificial Intelligence
Board



Commission to act as
Secretariat



Expert Group*



*Not foreseen in the regulation but the Commission intends to introduce it in the implementation process.

Council

- Completed full detailed presentation of the proposal by the Commission with Q&A by Member States and proposed a first set of amendments.
- The FR Presidency aims to complete a full reading of the proposal, and possibly a partial general approach.
- **Timeline TBC**

Parliament

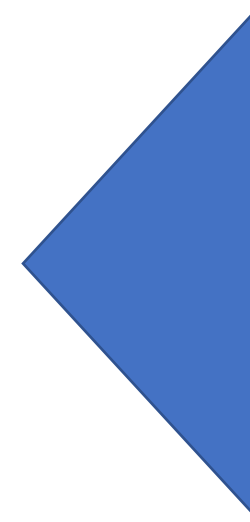
- **Joint competence of IMCO & LIBE committees:** draft report from 20 April.
- In addition, five other committees have partial competence on different parts of the file: JURI, ITRE, CULT, ENVI & TRAN.
- IMCO & LIBE vote scheduled for October, **plenary vote in November.**

-
- Once adopted: **2 years of transitional period** before the **AI Act becomes directly applicable** across the EU.

Some relevant issues under discussion:

- General-purpose AI systems and implications for NLP systems
- List of high-risk AI systems extended?
- More substantial obligations for public authorities?
- Enforcement mechanism: role of the Commission?

- **General Data Protection Regulation (EU) 2016/679:**
 - To the extent that personal data is being processed for training and testing the NLP system, or if the NLP system processes personal data while being used.
- **Sectorial legislation** applicable to the concrete activity, when used within regulated sectors.
- **Intellectual Property law :**
 - E.g. Directive (EU) 2016/943 on trade secrets, Directive 2009/24/EC on the legal protection of computer programs
- **Digital Services Act** (when adopted):
 - If NLP systems are used for content moderation

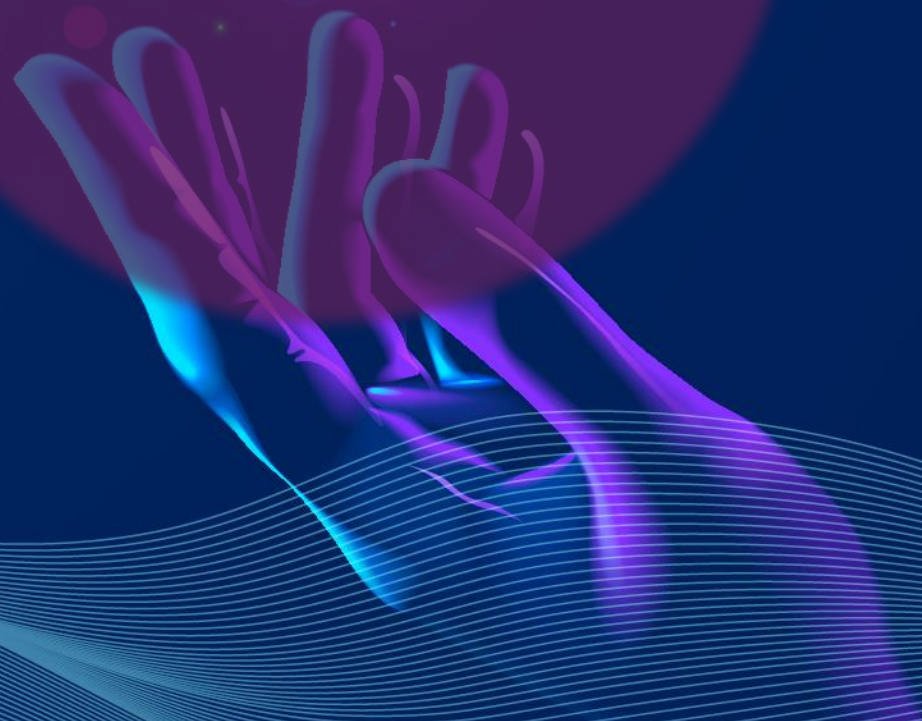
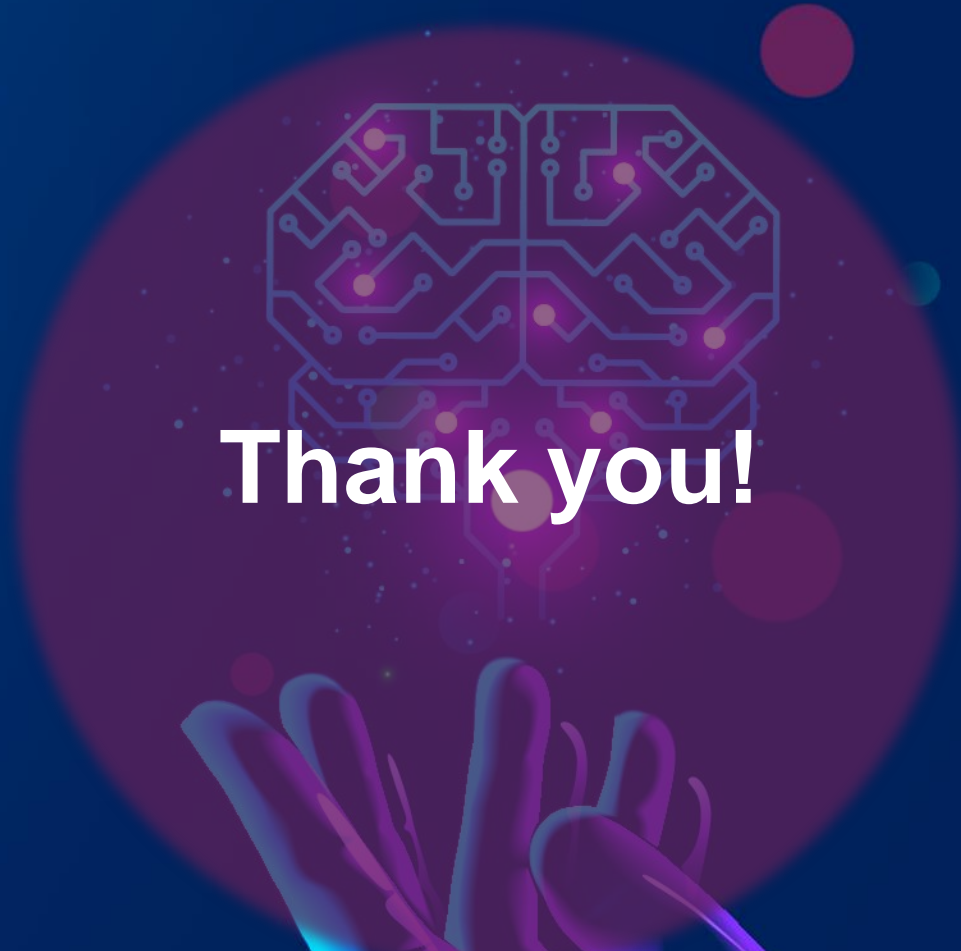


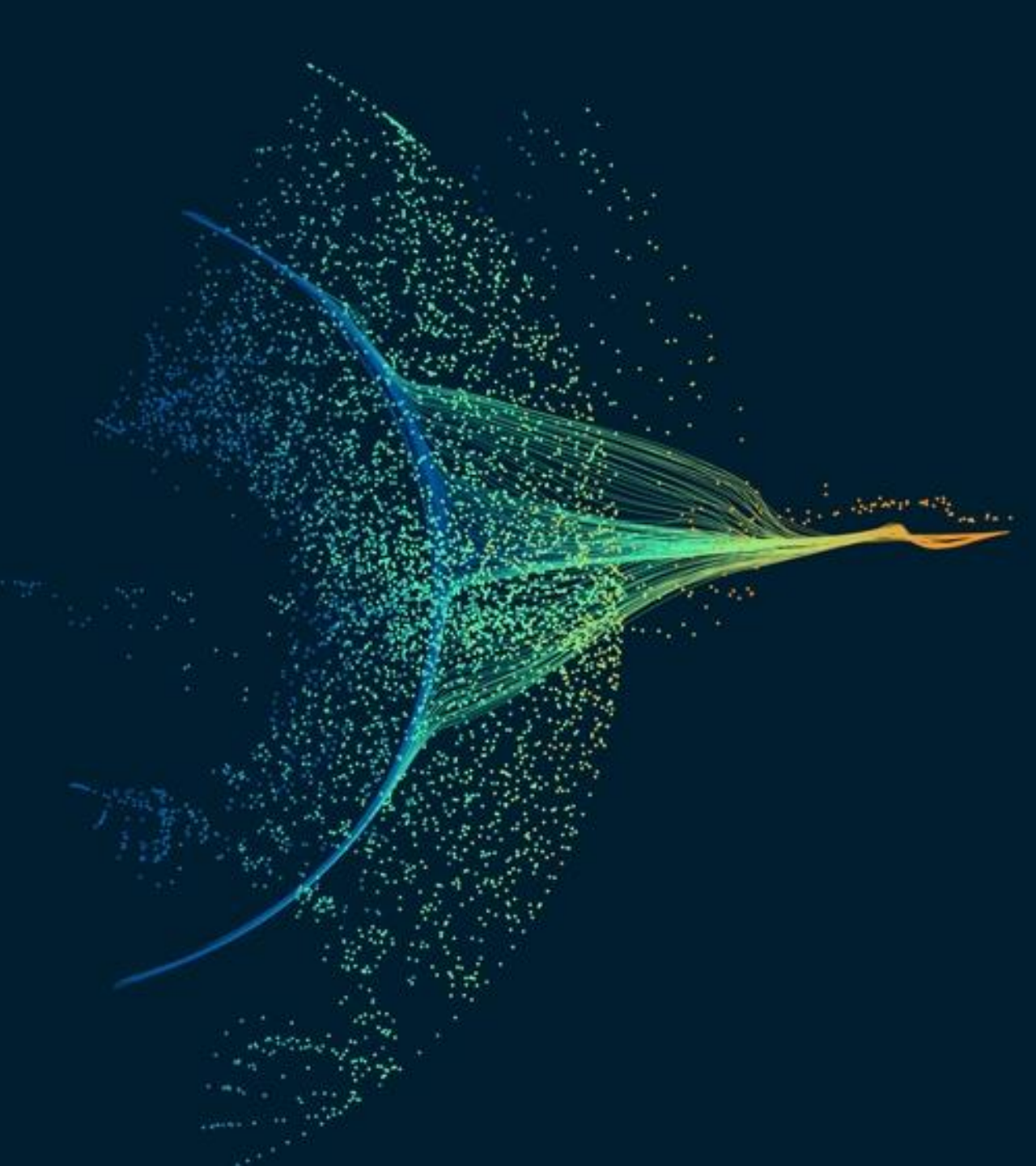
**Depends on the
context of the
application**

AI
ARTIFICIAL
INTELLIGENCE

EXCELLENCE
& TRUST

Thank you!





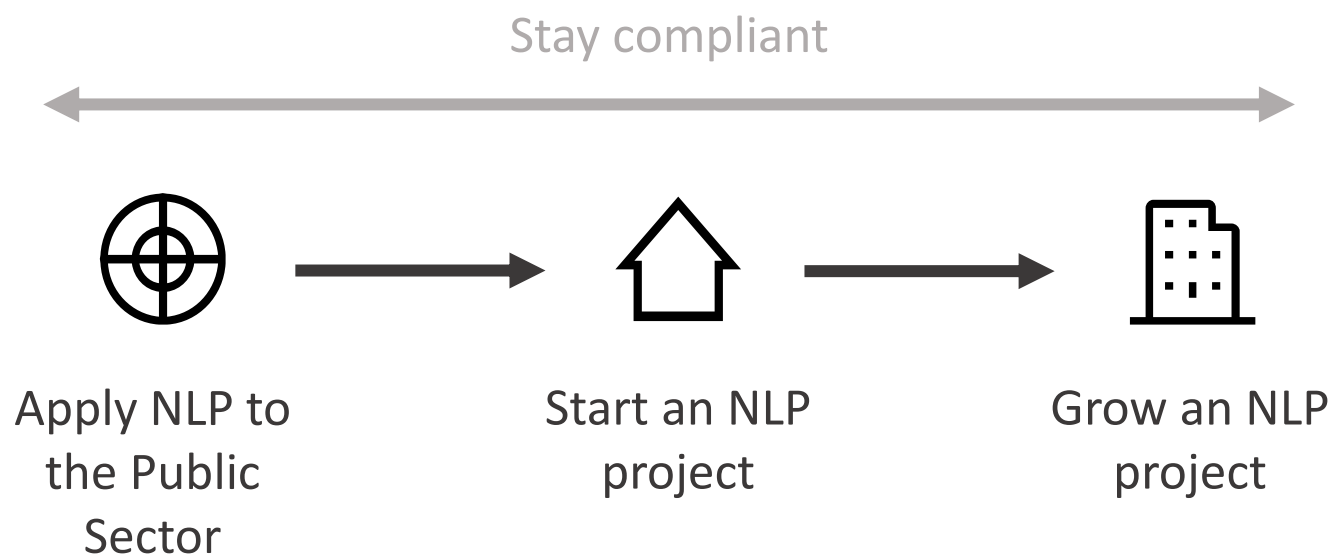
Regulatory compliance

Q&A

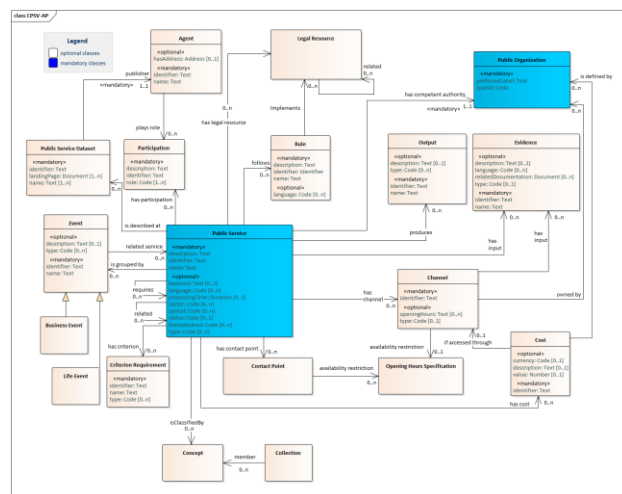


Conclusion

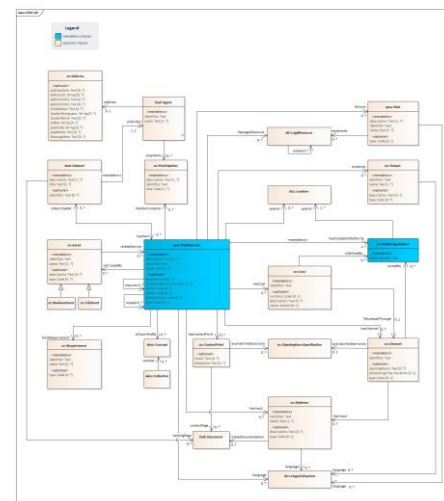
What have you learnt?



Upcoming activities



New course on CPSV-AP



Release of CPSV v3.0



Follow up on the SDG



interoperable europe

innovation ∞ govtech ∞ community

Stay in touch



[\(@InteroperableEU\) / Twitter](#)



[Interoperable Europe - YouTube](#)



[Interoperable Europe | LinkedIn](#)



DIGIT-INTEROPERABILITY@ec.europa.eu



<https://joinup.ec.europa.eu/collection/interoperable-europe/interoperable-europe>