Update on Advisory Group on Data

Ákos Jozwiak, Hungary
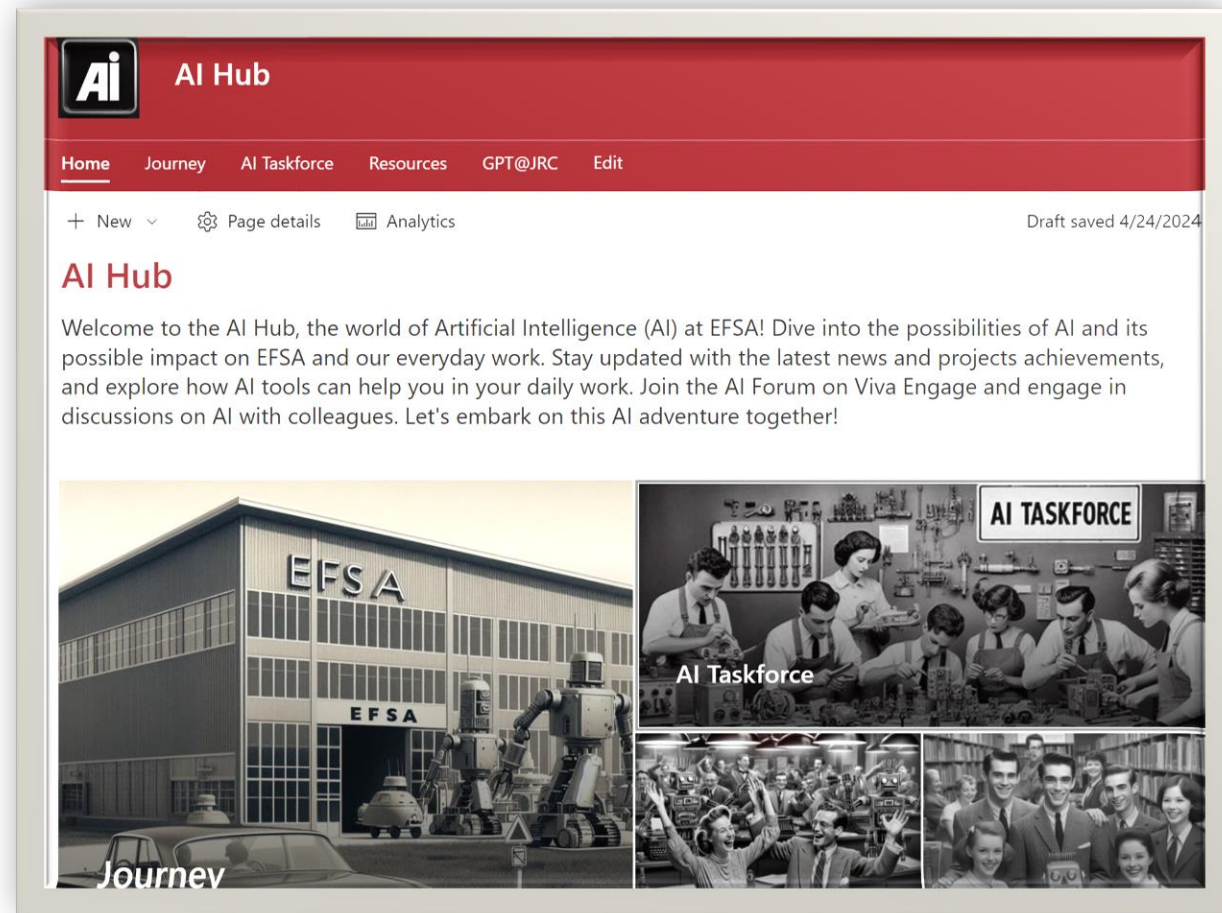
22ND (VIRTUAL) AND 23RD (PHYSICAL) MEETING OF AGOD

- EFSA Data and AI Initiatives
- Subgroups check-in
- Progress of Tailor-made Projects
- Progress of Recommendations
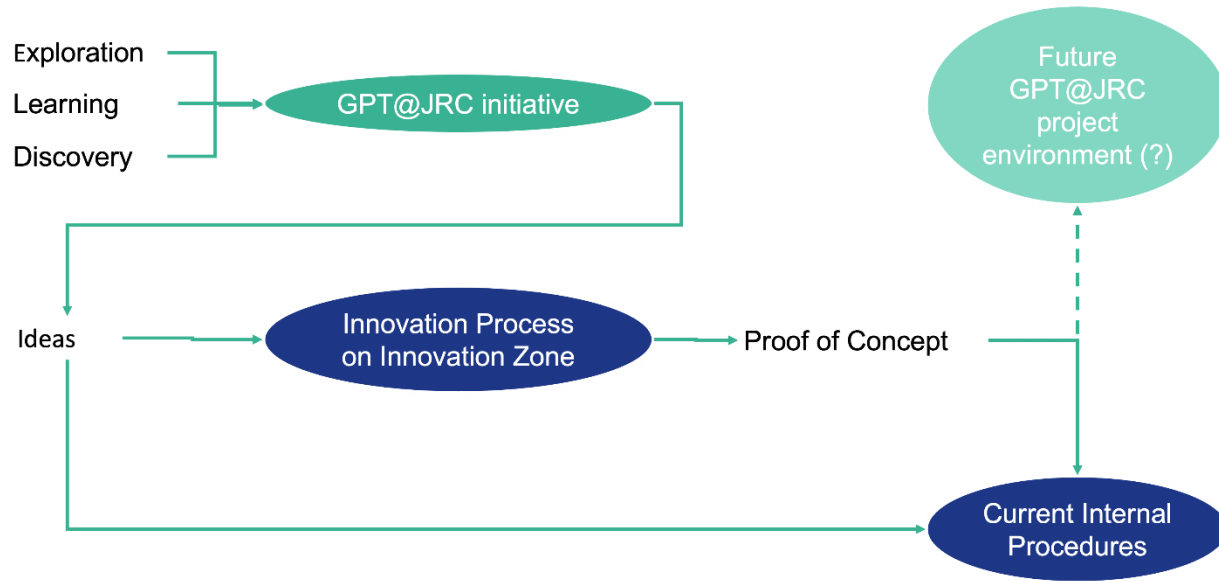- Communicating AGoD activities

UPDATE AI TASKFORCE

- Communication tool "AI HUB"

- Drafting AI@EFSA document

- Smart pilots
  MS Co-Pilot, Journal AI Services, FoodEx SmartCodingApp

- Launched GPT@JRC experiment

- Lots of missions and interactions
  (EMA, EUAN NAPO, JRC, etc.)

GPT@JRC & EFSA

- JRC has developed an initiative called GPT@JRC in which they make available several Large language Models (GPT, Mistral, Llama, etc)
- EFSA (limited number of) users to report back their experience through questionnaires

Objectives of the study

- Allow access to a GPT version endorsed by the commission for exploration
- Learn from user journeys
- Form AI savvy users

EFSA DEEP – DATA AND EVIDENCE PROGRAMME

COLLABORATIVE DATA PLATFORMS

IUCLID
EPA13.02-ED-22 IUCLID project (MVP 2) ART

RASE2
EPA13.02-PLH23-DEEP-W3 Harmonized Dossier Ingestion (RASE2)

DC TRACE
EPA13.02-AMU-28 Data Collection Tracing ORPHAN -->IRASFF

FC TRACE
EPA13.02-AMU-29 Food Classification For Tracing ORPHAN -->IRASFF

IRASFF
EPA13.02-PLH23-DEEP-W5 IRASFF

REBUILD
EPA13.02-TS-02 Rebuild Data Framework IMP

SEED
EPA13.02-ALPHA-11 Integrated spatial analysis for biological RA IMP (SEED)

SIGMA2
EPA13.02-BIOHAW-03 SIGMA 2.0 IMP

SUPPORT FOR RISK ASSESSMENT

BIKE2 – PA
13.02-PLH23-DEEP-W2 - BIKE 2 process analytics

BIKE2 – P2P
13.02-ENV23-FIN-W1 - BIKE 2 - Plan to procure

DOI BUILDER
EPA13.02-LA-16 DOI builder

OPEN FOOD TOX 3
EPA13.02-DATA-31 Open Food Tox 3 IMP

Pesticides genotox → OFT
EPA13.02-PREV-03 Update of the EFSA pesticides genotoxicity database RAMPRO --> Open Food Tox

RUEDIS → OFT
EPA13.02-PRES-05 RUEDIS database --> Open Food Tox

CORSA
EPA13.02-FEEDCO-02 CORSA

HOME
EPA13.02-ENV22-RAL-Z1 Harmonised organization of meetings and events

REM
EPA13.02-EXO-12 Records and Correspondence Management Project

REM 2.0
EPA13.02-PLH23-DEEP-W1 Records and Correspondence Management 2

NEW DATA STREAMS

INNOVATION
EPA13.02-PLH23-DEEP-W4 Enhance and boost innovation

EU MENU
EPA13.02-DATA-01 EU Menu

EU FOOD COMP DB
EPA13.02-DATA-29 Creation of Open Access EU Food Composition Database IMP

FEED SYSTEM & CONS DB
EPA13.02-GMO-16 EFSA Feed classification system and feed consumption database

DATA FOR CHEMICALS
EPA13.02-IDATA-03 Monitoring and surveillance data for chemicals: exploring new opportunities

BEES
EPA13.02-PREV-11 Generate data on food consumption of bees

DATA PRODUCTS & SERVICES

WGS
EPA13.02-BIOCONTAM-19 Interoperating 'One Health' system (WGS)

MOPS
EPA13.02-FEED-08 3 pipelines to analyse WGS data applications IMP (MOPS)

Peptide modelling
EPA13.02-GMO-09.04 PROC_HLADQ peptide modelling 2018 RAMPRO

GMP-SEQ
EPA13.02-GMO-13 GM Plant Sequences IMP (GMP-Seq)

MULTI-OMICS
EPA13.02-IDATA-01 Comparative multi-omics study IMP

RNAi
EPA13.02-NIF-01 Risk assessment methodology for RNAi- applications RAMPRO

PRIMo revision 4
EPA13.02-DATA-25 PRIMo revision 4 (Pesticide Residue Intake model) RAMPRO

Open MCRA SPIDO
EPA13.02-DATA-30 OPEN MCRA SPIDO

OECD Metapath
EPA13.02-PRES-01 OECD MetaPath: Incorporation of pesticide residue data RAMPRO

OPEN INTEROPERABLE DATA

CSS_OSOA
EPA13.02-FIP-27 CSS_OSOA

DATA OUTREACH
EPA13.02-PLH21-DATA-Y1 Open Data discovery, access and visualization IMP (DATA outreach)

IDM
EPA13.02-PLH21-TS-Y1 IDM IMP

EFSA DEEP – PROJECT STATUS

IDEA	INITIATING	PLANNING	EXECUTING				CLOSING	CLOSED
INNOVATION	HOME	DATA OUTREACH	SIGMA2	EU FOOD COMP DB	RNAi	CORSA	OECD METAPATH	RUEDIS → OFT
RASE2/HDI	REM 2.0		SEED	FEED SYSTEM & CONS DB	MULTI-OMICS	DOI BUILDER	EU MENU	IUCLID*
IRASFF	BIKE2 – PA		REBUILD	BEES	OPEN MCRA SPIDO	REM		CSS-OSOA
	IDM		DC TRACE	DATA FOR CHEMICALS	GMP-SEQ	OFT 3		PESTICIDES GENOTOX → OFT
			FC TRACE			BIKE2 – P2P		WGS
								PRIMo REVISION 4
								PEPTIDE MODELLING
								MOPS

EFSA REBUILD PROJECT

The **Rebuild DF project** is EFSA's strategic response to the evolving requirements in data collection and management and it will enhance our capabilities through advanced systems for data ingestion and analysis. The goal of the project is to re-engineer the data collections and data analysis systems in EFSA.
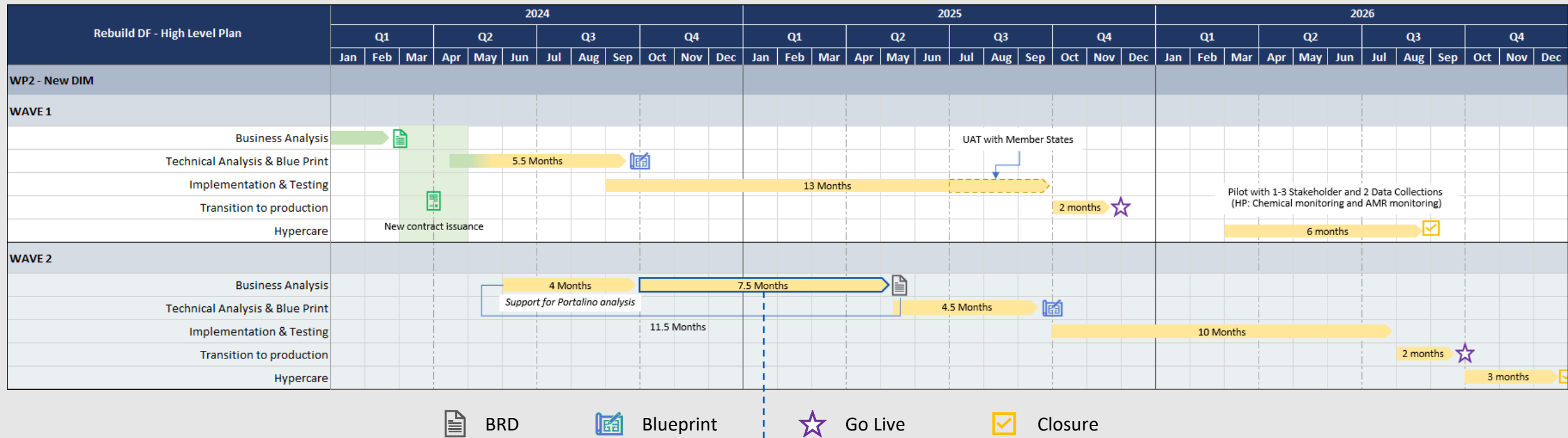
OBJECTIVES

- modernizing the way data is ingested and managed in EFSA;
- increasing the speed of data management systems;
- empowering the autonomy on the way of working and the flexibility of data product users;
- providing users with an environment to prototype an idea or use an existing code which can be modified as needed;
- establishing rules and principles to control proliferation of data products.

MAIN BENEFITS

In the long run, Data Providers of EFSA shall benefit from simplified data collection/connection and EFSA units and its stakeholders shall be empowered to create data products, services, and tools for cloud solutions, whilst maintaining consistency with the overall architectural framework and ensuring that both security and governance are integrated by design.

EFSA REBUILD PROJECT – TIMELINE

EFSA REBUILD PROJECT – MS SUPPORT

OBJECTIVES

- Align with MSs expectations.
- Early engagement in the activities related to the implementation of the data ingestion/data management tools.
- Create a solution in line with the MSs needs.
- Maximize the usage/ adoption of the new solution.

PROPOSED ACTIVITIES (DRAFT)

- Attend EFSA-led training sessions to understand the task, expectations, and the new framework's operation.
- Identify requirements list and TO-BE process for new working methods.
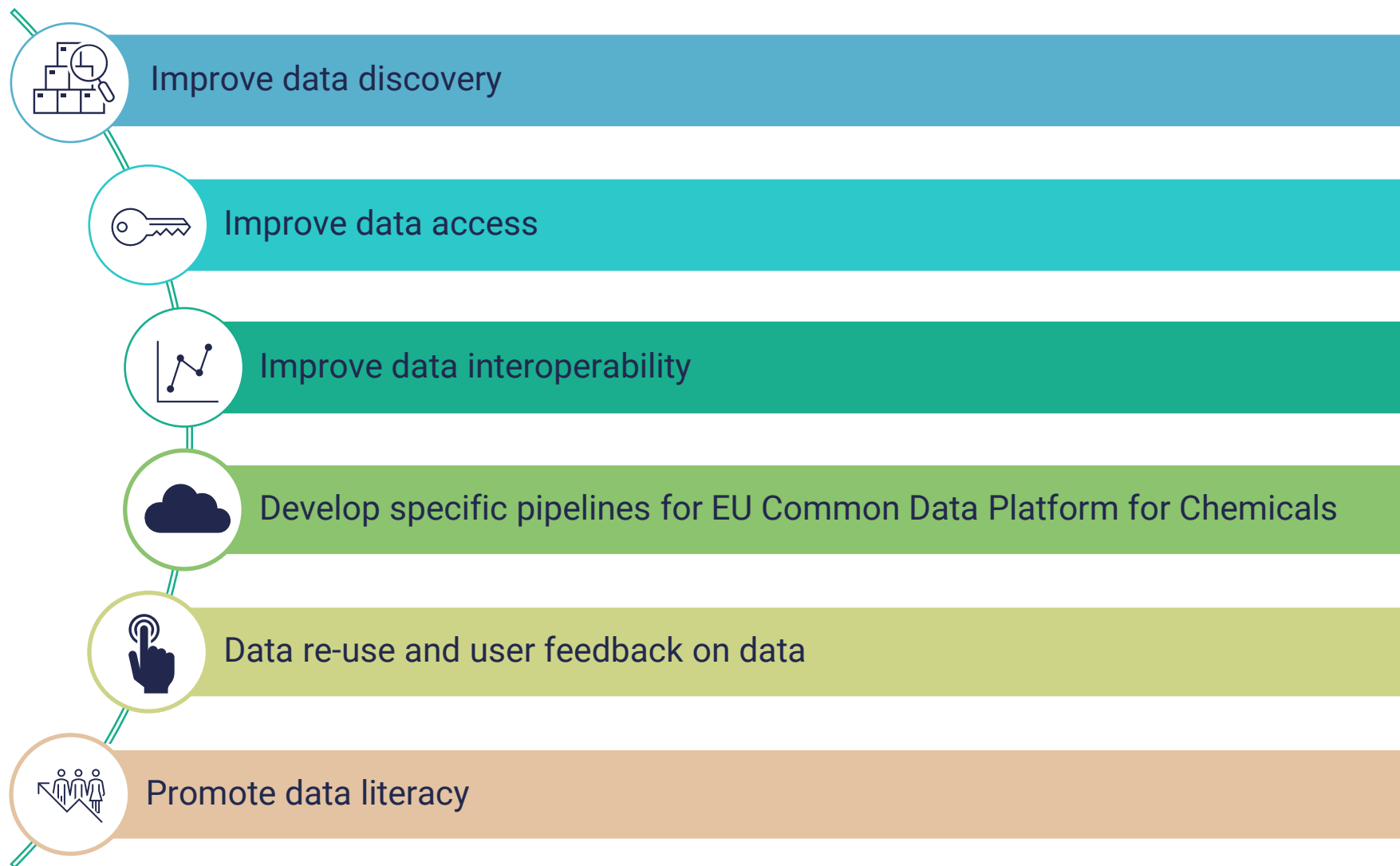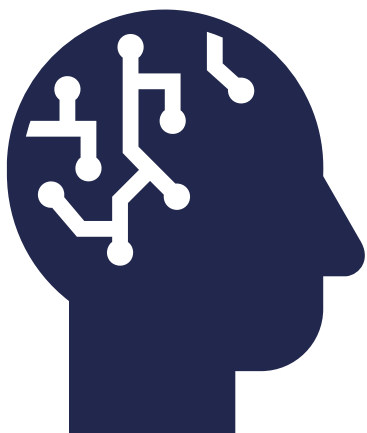- Propose additional actions for the adoption of the solution.

EFFORT EXPECTED

- Effort and activities proposal will soon be finalized and made available to MSs. *(Advanced Draft)*

MS support is submitted as the proposal for the Focal Point tailor-made activity.

EFSA DATA OUTREACH PROJECT

- Improve data discovery
- Improve data access
- Improve data interoperability
- Develop specific pipelines for EU Common Data Platform for Chemicals
- Data re-use and user feedback on data
- Promote data literacy

EFSA DATA OUTREACH PROJECT SCOPE

MACRO ACTIVITIES

Data Visualisation and Storytelling	Data Outreach Platform	API Portal	Data Publication Rules	Data mapping for EU-CDPC Platform	Change Management
Definition and design of data visualisation and storytelling, aiming at creating guidelines for standardised visualisation processes, piloting them with two reports and potential external tools.	Development of the Data Outreach Platform aiming at creating an innovative, user-focused platform for enhancing EFSA's data discovery and visualisation.	Establishment of a secure and scalable API Portal aligned with Data Publication rules by establishing a support community for optimised operations and user satisfaction.	Evaluation of current data sharing protocols to identify improvements, with possible revision of guidelines with stakeholder input prior to implementation.	Analysis of the EU Common Data Platform for Chemicals (EU-CDPC) and designing secure data sharing architecture and access mechanisms to ensure regulatory compliance and confidentiality.	Ensuring a smooth transition by enhancing data interpretation and self-service analytics, sharing guidelines on data visualisation, organising training sessions and workshops to foster continuous improvement within EFSA and its external network.

PROGRESS ON FOCAL POINT TAILOR-MADE PROJECTS

- Data Model Mapping Tool (Sweden)

    Leader: Sweden – Participants: Denmark, Germany, Italy

- Data Model Mapping Tool (Italy)

    Leader: Italy – Participants: Denmark, Germany

- Data Capture @Point of Sampling Tool

    Leader: Portugal – Participants: Denmark

- Inspections and Samples II

    Leader: Netherlands, Austria – Participants: Germany, Portugal, Albania (observer)

- Data Flow Mapping in Member States

DATA MODEL MAPPING TOOL (SWEDEN) – KEY REFLECTIONS

The objectives has been achieved within the allocated budget and timeline

Immense potential of AI technologies, particularly embeddings and large language models

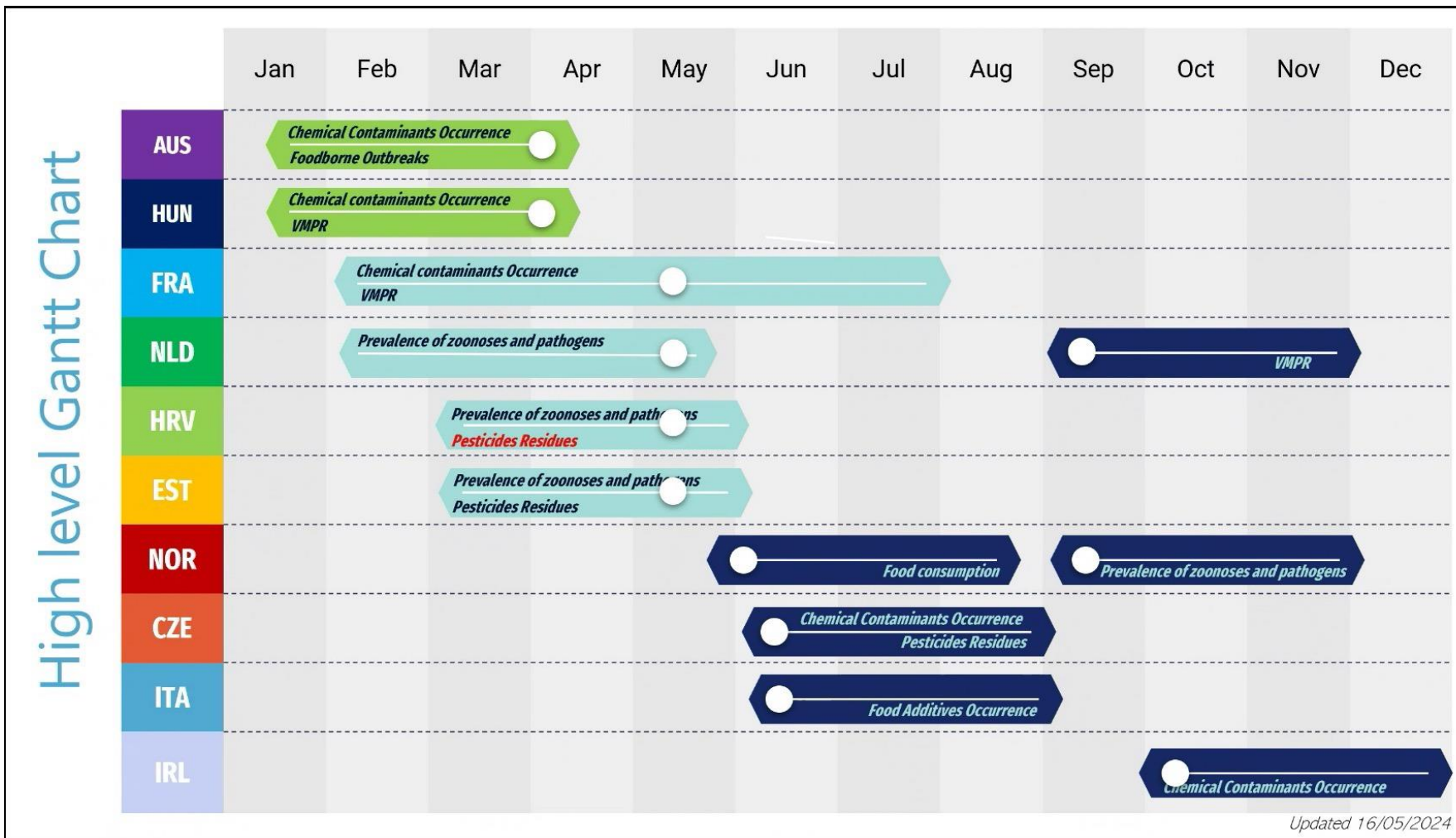The Data Mapping tool will be able to significantly enhance efficiency, accuracy, and scalability

The pre-study has paved the way for the development of a robust, future-proof Data Mapping Tool

We are well-positioned to move forward with the implementation phase

Livsmedelsverket
Swedish Food Agency

PROGRESS ON FOCAL POINT TAILOR-MADE PROJECTS: DATA FLOW MAPPING IN MEMBER STATES
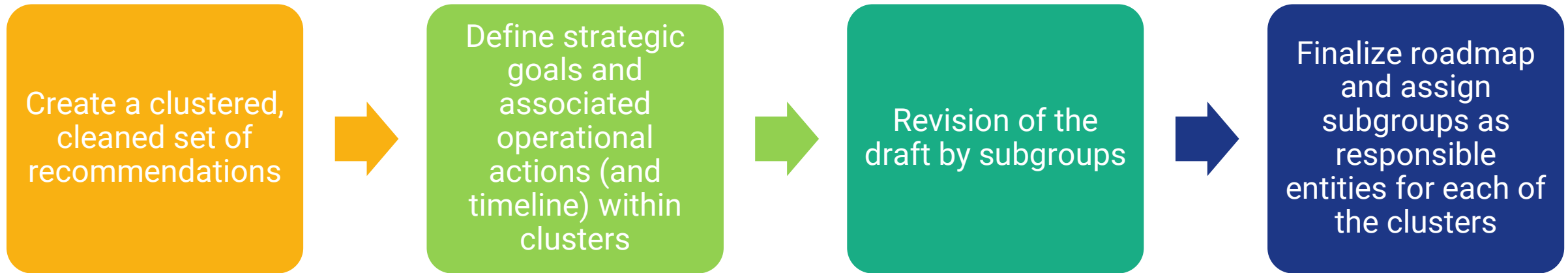

Updated 16/05/2024

14

PROGRESS OF RECOMMENDATIONS

- A new AGoD mandate begins
- Time to revise AGoD multi-annual roadmap
  - o Review systematic implementation of recommendations
  - o Measure achievement and success
  - o Improve communication and outreach
- Update recommendations to reflect impact of emerging technologies

BUILDING AGOD ROADMAP

Create a clustered, cleaned set of recommendations → Define strategic goals and associated operational actions (and timeline) within clusters → Revision of the draft by subgroups → Finalize roadmap and assign subgroups as responsible entities for each of the clusters

WORKSHOP IN PARMA

Activities

1. **Restructure Subgroups**
2. **Discuss ownership**
3. **Prioritize actions**

Restructuring subgroups	Ownership	Prioritisation
New clusters = new subgroups?	Who moves the actions forward?	Which actions are feasible in current mandate?

CLUSTERS AND SUBGROUPS

Data Governance

- Governance bodies and roles
- Collaborative decision-making mechanisms

Data Management

- Data Quality
- Data Modelling and Terminologies
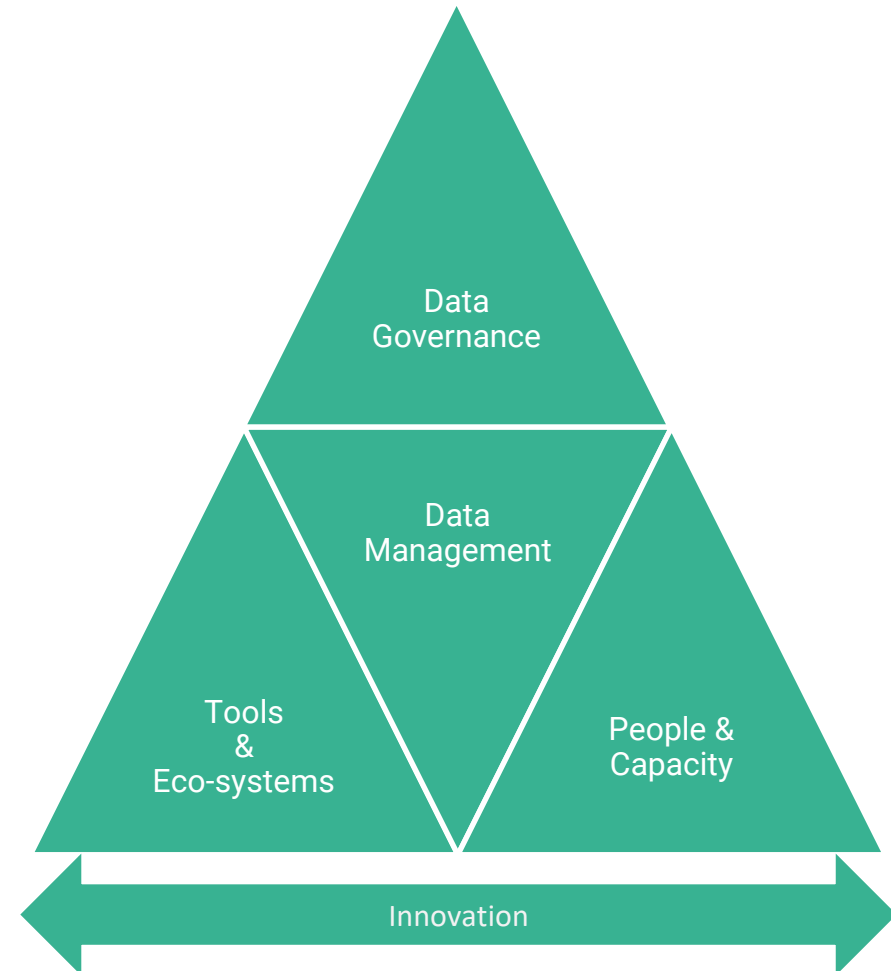- Data Analysis and visualization

Tools & Eco-systems

- Digital Platforms and eco-systems
- Developing and sharing tools

People & Capacity

- Data Literacy
- Data Capacity preparedness

Innovation

- Innovative data analytics
- New data streams
- **A.I. (NEW)**

18

COMMUNICATING AGOD ACTIVITIES

- Annual Report 2023: https://www.efsa.europa.eu/en/supporting/pub/en-8760

- AGoD web page as a part of EFSA website: https://www.efsa.europa.eu/en/advisory-group-on-data

- Data Conference in 2024

  - To bring together members from AGoD, its subgroups Data Collection Networks, and FP TMA participants

  - Planned topic: *Data readiness for AI*

  - Date and venue: *TBD*

WAY FORWARD

- Providing advice for the data-related FP TMT proposals: June-July 2024
- Finalizing AGoD roadmap: summer 2024
- Re-setting Subgroups: summer 2024
- Multi-annual plan: Q3 2024
- Data conference in 2024

AGOD MEETINGS IN 2024

June-July
Ad hoc meeting on FP TMT proposals

31st January online	5th March Belgium	24th April online	21st May Parma	5th Sept online	1st Oct online	3rd Dec Hungary
Subgroup check-in	Subgroup check-in	Subgroup check-in	Subgroup check-in	Subgroup check-in	Subgroup check-in	Subgroup check-in
Progress TMT	Progress TMT	Progress TMT	Progress TMT	Progress TMT	Progress TMT	Progress TMT
Draft new ToR	Progress of recommendations workshop	Revision of FP Tailor Made proposals	New project ideas	Update on AI TF	Progress of recommendations	Revision of FP Tailor Made proposals
AGoD and AI	Update on AI Taskforce	Small conference planning of contents	Session on AI	Outcomes of data flow mapping in MS	Annual Report	Update on AI TF
Annual Report	AGoD website	MS presentations on national initiatives	Rebuild/Data outreach update	Update on EC data initiatives	MS presentations on national initiatives	Outcomes of data flow mapping in MS
			DEEP programme			
			Mapping Tool Sweden			