STAKEHOLDER FORUM 2025

Breakout session 2

Unlocking the Potential of Data

Fabrizio Abbinante
Andrea Gissi

efsa
EUROPEAN FOOD SAFETY AUTHORITY

CONTENT

- Data and AI in EFSA
- Cheminformatics and bioinformatics tools
- An EU platform to collect all data on chemicals

DATA AND AI IN EFSA

Photograph by Kolleen Gladden on Unsplash

DATA AND EVIDENCE AT THE CORE OF EFSA'S RISK ASSESSMENT

Who sends requests to EFSA

- European Commission
- European Parliament
- Member States
- EFSA self mandate

Once the request is received

Problem formulation and protocol development

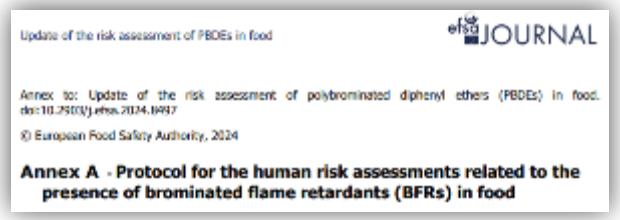
- Data collection
- Evidence appraisal Data validation
- Evidence synthesis/ integration Data analysis

Expert Knowledge Elicitation

EFSA RISK ASSESSMENT

Publication

Communication


Update of the risk assessment of PBDEs in food

EFSA JOURNAL

Annex to: Update of the risk assessment of polybrominated diphenyl ethers (PBDEs) in food.

© European Food Safety Authority, 2024

Annex A - Protocol for the human risk assessments related to the presence of brominated flame retardants (BFRs) in food


SCIENTIFIC OPINION

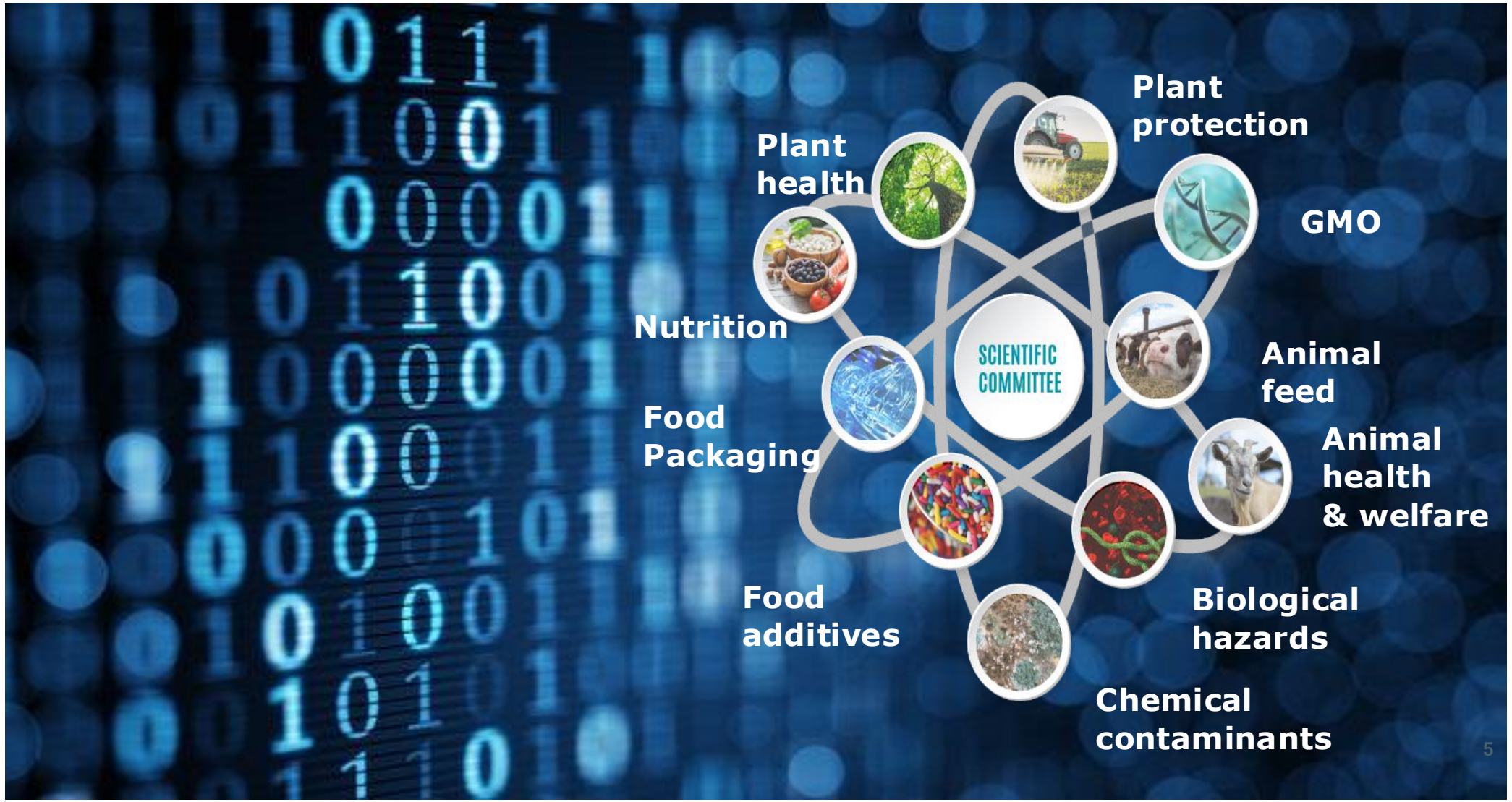
EFSA JOURNAL

Update of the risk assessment of polybrominated diphenyl ethers (PBDEs) in food

EFSA Panel on Contaminants in the Food Chain (CONTAM) | Dieter Schrenk | Margherita Bignami | Laurent Bodin | James Kevin Chipman | Jesús del Mazo | Bettina Grasl-Kraupp | Christer Hogstrand | Laurentius (Roni) Hoogenboom | Jean-Charles Leblanc | Carlo Stefano Nebbia | Elsa Nielsen | Evangelia Ntzani | Annette Petersen | Salomon Sand | Tanja Schwerdtle | Heather Wallace | Diane Benford | Peter Fürst | Andy Hart | Martin Rose | Henri Schroeder | Martine Vrijheid | Sofia Ioannidou | Marina Nikolić | Luisa Ramos Bordajandi | Christiane Vleminckx

Receipt of the Mandate → Allocation to Scientific Panel (10 Scientific Panels + Scientific Committee) → Acceptance → Question n. available in OpenEFSA portal

THE COMPLEXITY OF FOOD SAFETY DATA IS VAST AND DIFFERENT



FULL SPECTRUM OF CENTRALLY MANAGED DATA

Chemical Monitoring data

- Pesticides Residues
- VMPR
- Contaminants, FCM
- Food Additives Food/Feed Flavorings

Biological monitoring data

- Zoonoses AMR FBOs
- TSE
- Avian Influenza/ African Swine Fever
- One Health Surveillance

Food Consumption Composition data

- Food Composition
- Food Consumption (EU MENU)
- FoodEx2 Catalogue
- Food Carbon Footprint (new)

Databases and Endpoints repositories

- EU Legal Limit DB
- Botanical Compendium
- OFT 3.0 NORA FLAVIS
- Knowledge Junction
- Pest in Apple
- Xylella

Regulated products data

- PPP dossiers (IUCLID)
- Non-pest dossiers (ESFC)
- Notification of studies database
- pFCM dossiers (IUCLID integrated with ESFC) — Pilot

DIFFERENT LEVEL OF MATURITY FOR EFSA DATA SERVICES

Monitoring Data Collections Services

Since 2005

NOS Data Services

ESFC Data Services

PPP IUCLID Data Services

Since 2019

Since 2020

Bioinformatic Services

ChemInformatic Services

Since 2023

pFCM IUCLID Data Services

new

ALL THIS IN COLLABORATION WITH MS, INDUSTRY AND EU PARTNERS

Chemical Monitoring Network

Zoonoses Monitoring Network

Food Consumption Network

Pesticide Steering Network
IUCLID sub-group

EU CDPC Governance Body
NEW

Industry Discussion Group for Data Reporting

TSE Subgroup

Advisory Group on Data

OECD IUCLID User Group

Strategic data initiatives are discussed and agreed with AgoD and its subgroups

Advisory Group on Data (AGoD)

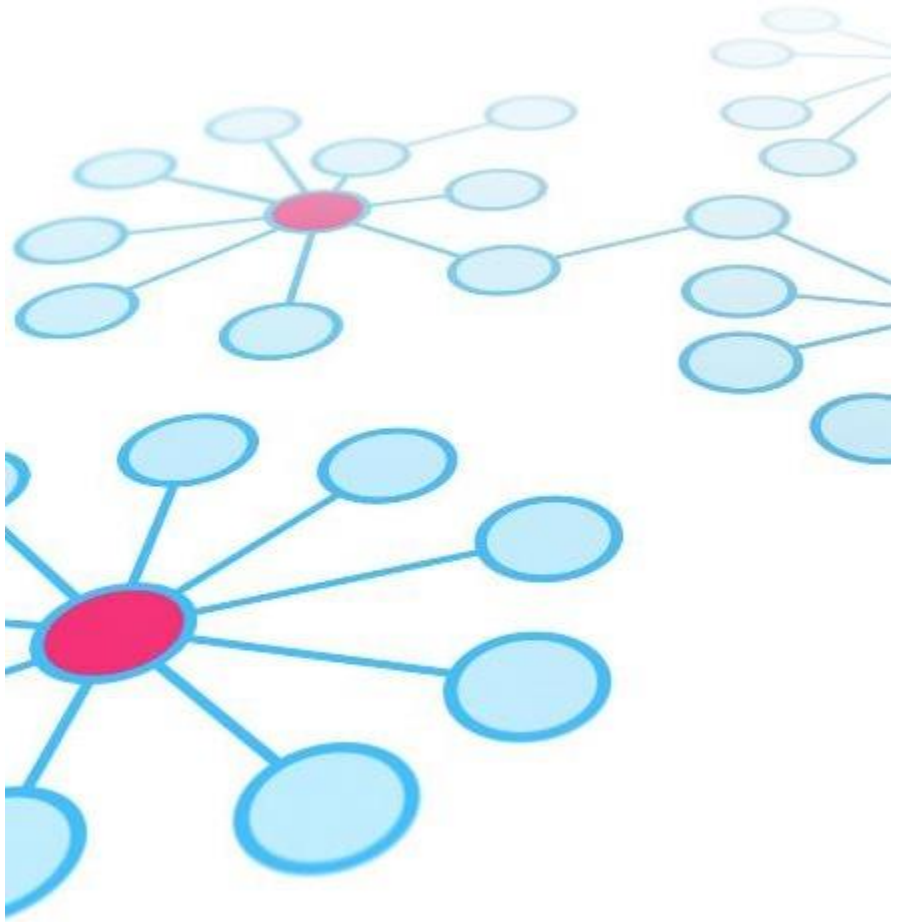
Data Management

Tools & Ecosystems

Innovation

People & Capacity

DATA HAS CHANGED



centralised	***distributed***
collected	***connected/linked***
structured	***unstructured/vary***
owned	***open***
true	***trusted***
slow / asynchronous	***fast / real time***
small	***"Big"***
analysed	***interpreted***
human processed	***A.I./M.L. processed***

NEW CHALLENGES

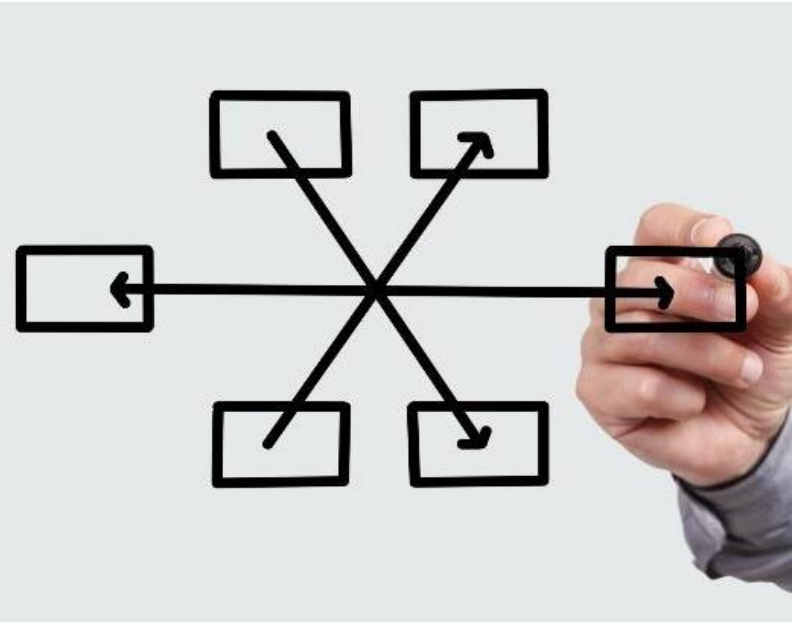
New Data Management needs
Handling large, complex datasets (omics, linked, distributed, sparse) from multiple sources while ensuring interoperability across agencies (e.g., EU CDPC, “Food Data Space”).

New Approach Methodologies (NAMs)
Integrating in vitro and in silico methods into frameworks that predict chemical hazards and risks, reducing animal testing and using advanced analytics like machine learning.

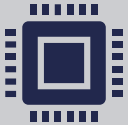
AI-Ready Data
Preparing data so AI systems can process, learn, and act at scale, ensuring reliable and trustworthy outcomes.

Trust and Efficiency
Improving dossier and data quality to enhance and speed up assessment efficiency and build stakeholder trust through transparency.

Data Security
Protecting sensitive personal, scientific, and regulatory data from cyber threats through strict governance, access controls, encryption, and audits.



WHAT WE ARE WORKING ON



Delivering business-domain-led “Data products”.

A Data Product is a governed, curated and self-contained combination of data, metadata, semantics and templates. It must be consumption-ready (trusted by consumers), up-to-date (by engineering teams) and approved for use (by end users). Data products enable various data and analytics use cases, such as data sharing, data integration, domain analytics and application integration.

Delivering a new Data Collection system and scalable Bio-Informatic & Chem-Informatic pipelines

The Rebuild Data Framework (Rebuild DF) is EFSA’s flagship initiative to modernize its data infrastructure, aiming to streamline ingestion, enhance analytics, and improve risk assessment accuracy through agile development, cloud-native tools and robust and scalable Bioinformatic and Cheminformatic Pipelines and tools

Preparing data to embed NAMs in the RA of EFSA and getting ready for the EU CPDC and IUCLID extension

EFSA is actively supporting New Approach Methodologies (NAMs) by integrating them into chemical risk assessment through training, international collaboration, roadmap implementation, and the development of tools and guidance that promote animal-free, mechanistic-based approaches.

AI@EFSA JOURNEY

2017–2022

Exploring

AI for Evidence Management

Roadmap for actions on AI for evidence management in risk assessment

2023–2024

Piloting/Joining Forces

AI4EU Project

First AI Pilots

2024

Preparing

AI Task Force

Agod Symposium on Data Readiness for AI

2025

Setting up

EFSA AI Governance

AI@EFSA Paper

2026

Start Implementing

AI4Science uses cases

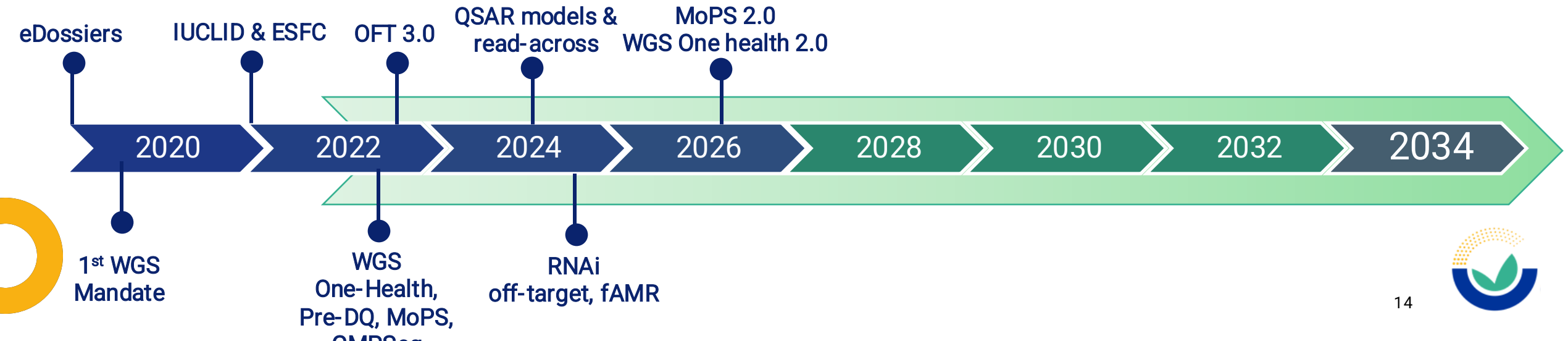
→ Increase speed of EFSA Risk Assessment
→ Ensure preparedness for future 'AI-assisted' RA
→ Empower people in using AI solutions

Cheminformatics and bioinformatics tools

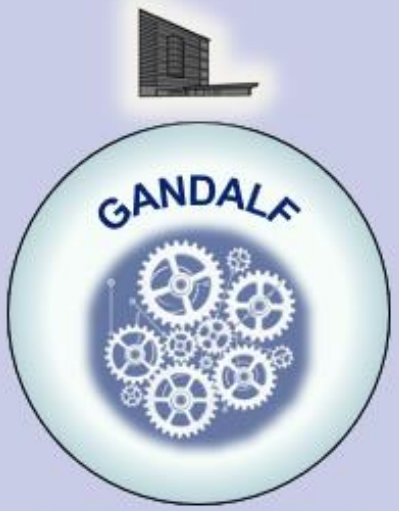
CHEMINFORMATICS, BIOINFORMATICS AND AI

By augmenting *Chemistry* and *Biology* with *Informatics*
we have enabled **ChemInformatics** and **BioInformatics** pipelines, models and tools
with concrete benefits for regulators, industry and academia.

With the advent of AI we can further augment what we have created so far in BioInformatics and ChemInformatics



EFSA BIOINFORMATICS ECOSYSTEM



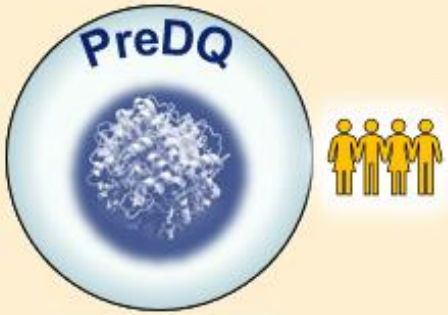
CENTRALIZED PORTAL TO RUN AND MANAGE TOOLS AND PIPELINES

Applications already connected to Gandalf

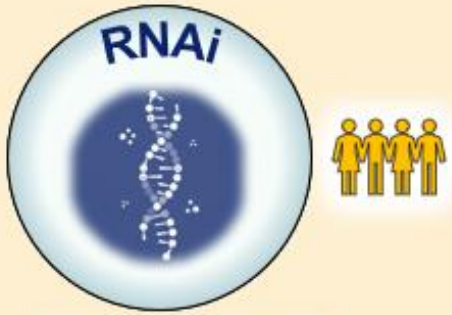
GENOMIC ANALYSIS OF MICROORGANISMS USED IN FOOD AND FEED

GENOMIC ANALYSIS OF FOODBORNE OUTBREAKS

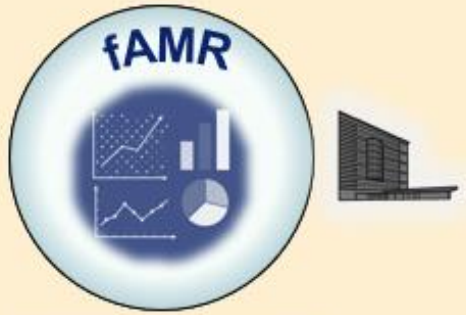
GENETIC ANALYSIS OF GMOs



PREDICTION OF ADVERSE EFFECTS OF PEPTIDES ON CELIACS



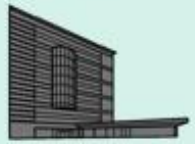
PREDICTION OF RNAi OFF-TARGETS in GMOs



DISTRIBUTION OF AMR GENES

Applications to be connected to Gandalf in 2026

USERS:



INTERNAL

EU AGENCY

MEMBER STATES

APPLICANTS

PUBLIC

EFSA CHEMINFORMATICS ECOSYSTEM

Dossier's intake (partially structured) → Data for RA (structured) → Data integration → Data reuse

PESTICIDES

IUCLID 6

NON-PESTICIDES

ESFC

OFT 3.0

GENOTOX

PROCESSING FACTORS

IUCLID 6 IUCLID 6 IUCLID 6

E-FLAVIS

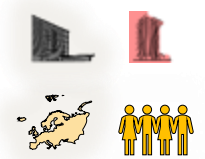
NORA

Risk assessment data deriving from dossiers and EFSA Risk Assessment (end-points databases)

EFSA DATA LAKE

CENTRALIZED AND SECURED REPOSITORY TO RUN AND MANAGE TOOLS AND PIPELINES

EFSA website



Open EFSA, MicroStrategy, IUCLID dissemination instance

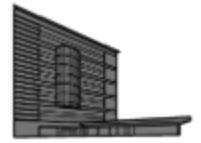
Computational tools

Substance grouping, QSAR predictions

Reporting

Reports and dashboards for risk assessors

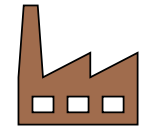
USERS



INTERNAL

EU AGENCY

MEMBER STATES



APPLICANTS

PUBLIC

R4EU: EUROPE'S PLATFORM FOR FOOD SAFETY TOOLS

Cloud-native tools on a secure collaborative platform, developed for industry, with Member States
Leveraging advanced analytics, statistical modelling, machine learning and AI

r4eu.efsa.europa.eu

CHEM- AND BIO- INFORMATICS SPEED UP RISK ASSESSMENT

Benefits of using cheminformatics and bioinformatics in risk assessment:

Speed – Automation

Automation of tools and data analysis, flagging of concerns

- e.g., Use of MoPS to characterise microorganisms used in the food chain and flag potential concerns linked to AMR genes or production of metabolites

Consistency – Standardisation and grouping

Standardisation of approaches, grouping of substances

- e.g., Development and distribution of a workflow in the QSAR Toolbox for the assessment of pesticide residues using read-across and QSAR models

Efficiency – Reduction of testing needs

Reduction of testing needs by predicting (some) hazard properties

- e.g., Use of QSAR to predict simpler toxicological endpoints for data poor substances

An EU platform to collect all data on chemicals

One for all, all for one!

BRINGING DATA TOGETHER REQUIRES HARMONISATION

One substance, one assessment package

1. Consolidating work in the EU agencies and improving cooperation

- Proposal for a regulation on the re-attribution of tasks and improving cooperation among agencies
- Proposal for a directive on the re-attribution of tasks amending RoHS directive

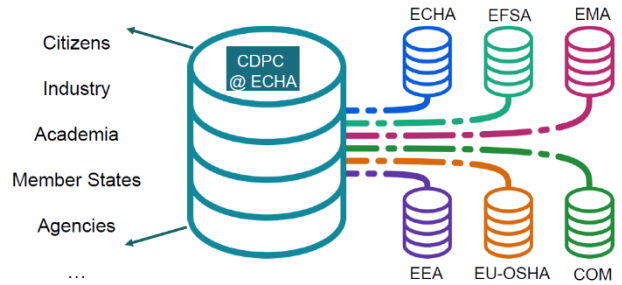
2. Removing barriers to reusing of data and establishing monitoring and outlook framework for chemicals

- Proposal for a regulation establishing a common data platform on chemicals and establishing a monitoring and outlook framework for chemicals

European Commission

Common data platform on chemicals (CDPC)

- **Regulation (EU) 2025/...** of the European Parliament and of the Council establishing a **common data platform** on chemicals, laying down rules to ensure that the data contained in it are findable, accessible, interoperable and reusable, and establishing a **monitoring and outlook framework** for chemicals
- Common data platform on chemicals (CDPC): bring together chemicals data in one database



4

European Commission

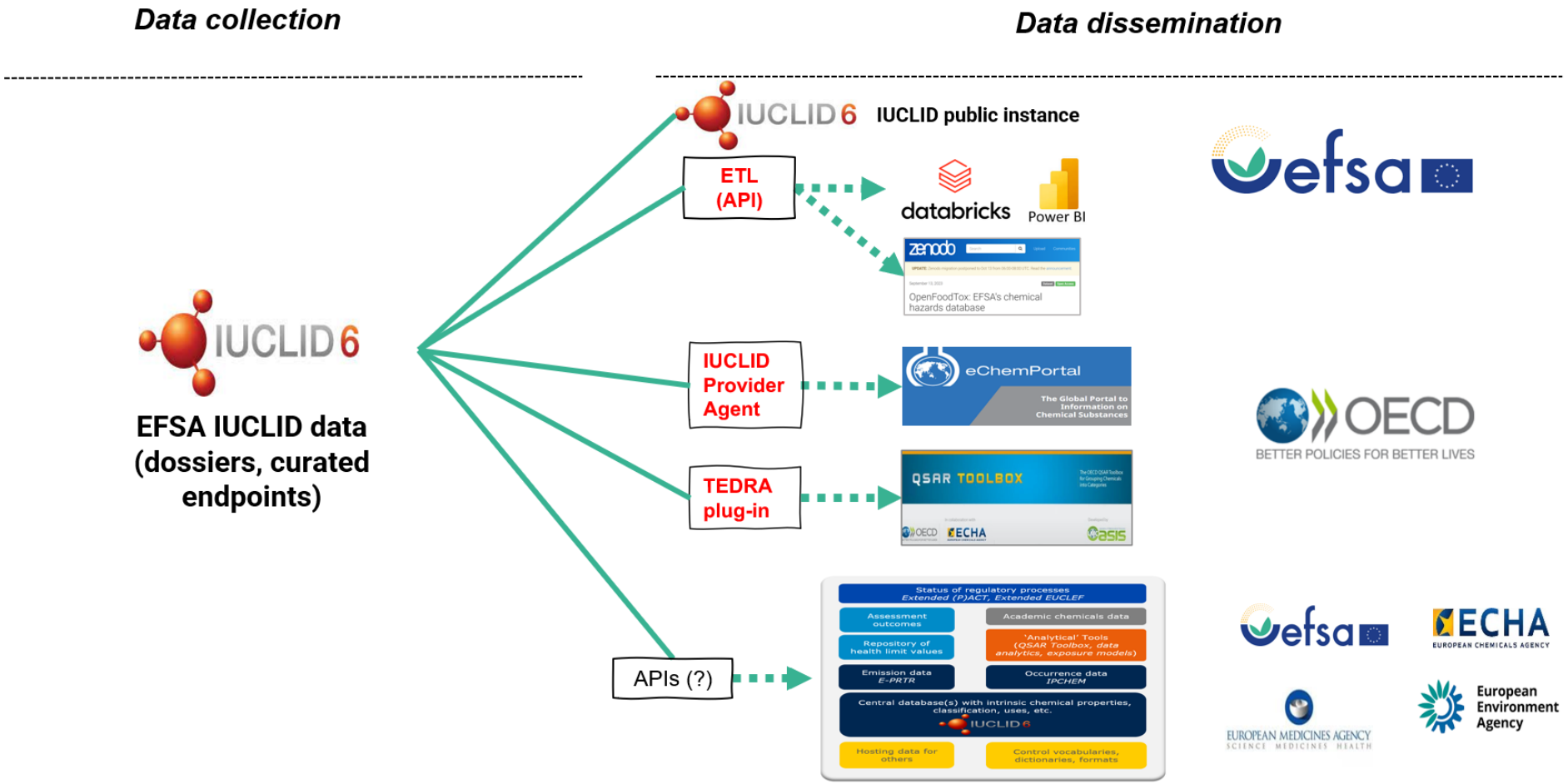
- The regulations in the 1 Substance 1 Assessment legislative package are on track to enter into force within the next year (EU-CDPC already this year)
- An EU Common Data platform (EU-CDPC) on chemicals will be established
- Public and confidential data with different access rights – “originator principle” for confidentiality

TOWARDS STRUCTURED DATA FOR REGULATED PRODUCT

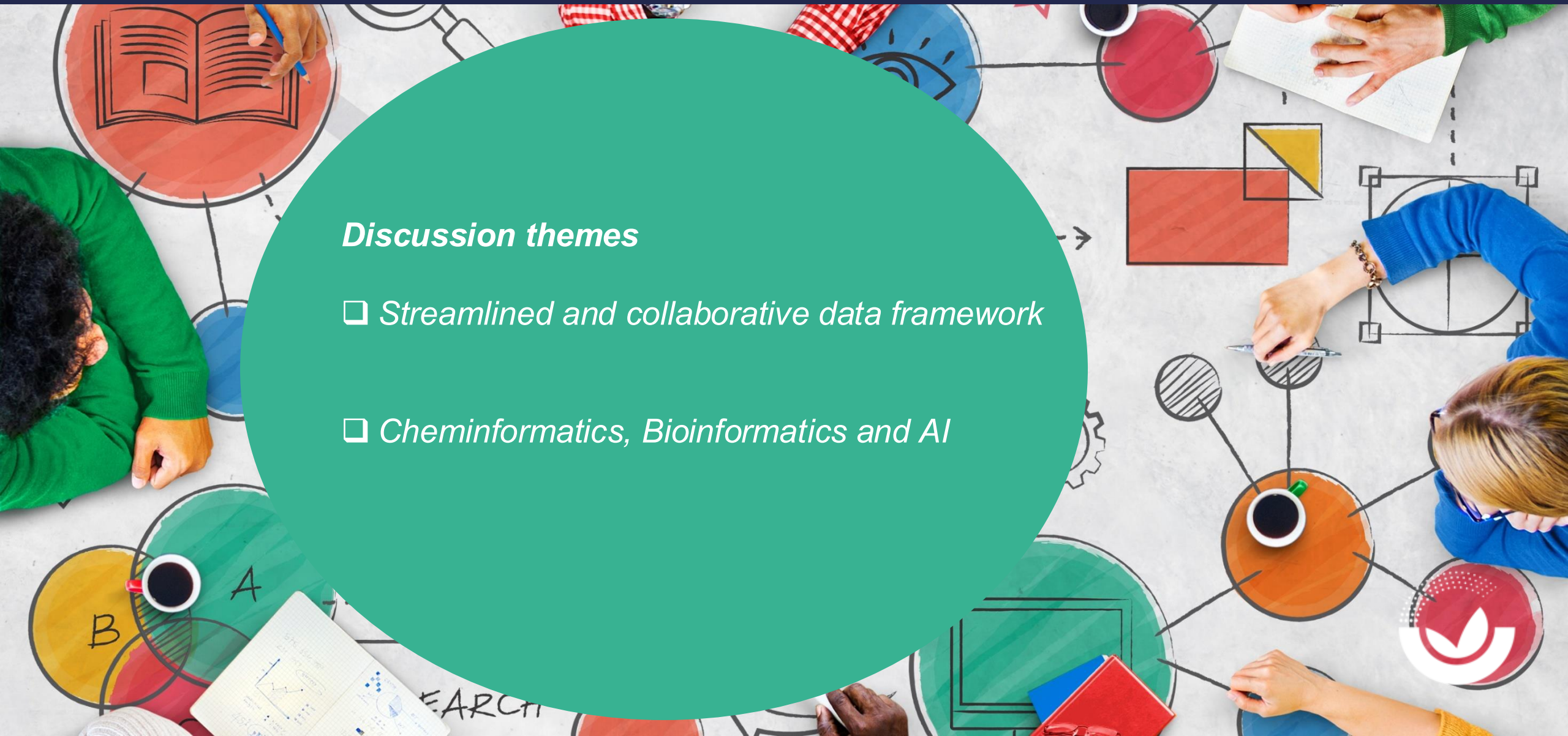
- In the EU-CDPC, EFSA must provide chemical data from applications for chemical food regulations in IUCLID format, which is highly structured
- Today EFSA receives only PPP applications in IUCLID format, while all non-pest applications are submitted via ESFC, where study data are not structured
- 10-year plan to comply with the EU-CDPC requirement to extend the use of IUCLID to structure study data for other domains, while keeping ESFC for the workflow management and for data submission.
- A pilot project to extend IUCLID to plastic Food Contact Material has just begun with DG SANTE and ECHA to initiate the transition. A pilot to integrate ESFC and IUCLID will be ready in early 2027.

DATA IS USEFUL WHEN IT IS (RE)USED

Having dossiers and endpoints data in a structured and harmonized format facilitates data (re)use and integration for scientific and regulatory purposes



OPENING THE FLOOR TO DISCUSSION



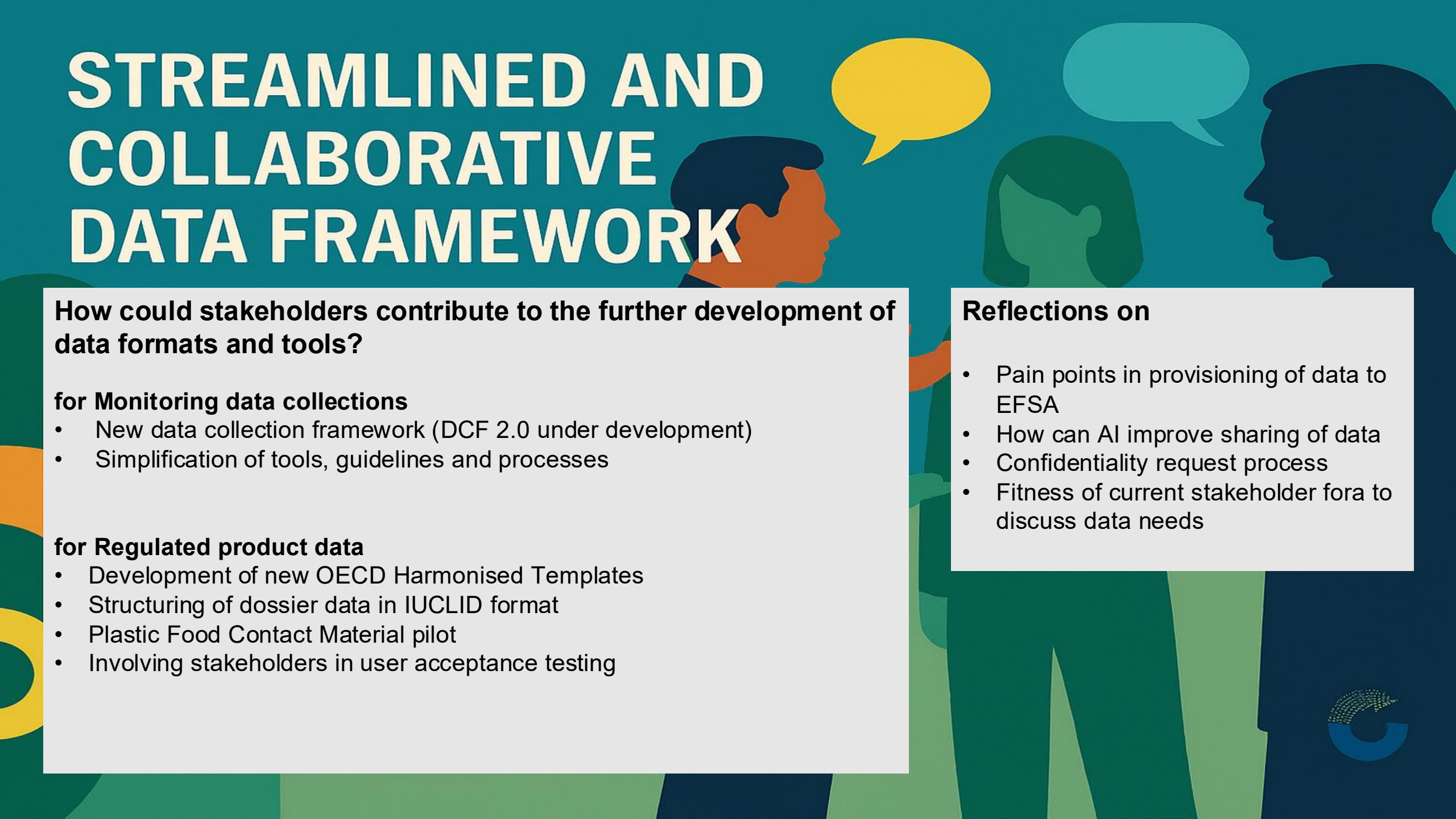
Discussion themes

- ❑ *Streamlined and collaborative data framework*
- ❑ *Cheminformatics, Bioinformatics and AI*

STREAMLINED AND COLLABORATIVE DATA FRAMEWORK

Link to poll

STREAMLINED AND COLLABORATIVE DATA FRAMEWORK



How could stakeholders contribute to the further development of data formats and tools?

for Monitoring data collections

- New data collection framework (DCF 2.0 under development)
- Simplification of tools, guidelines and processes

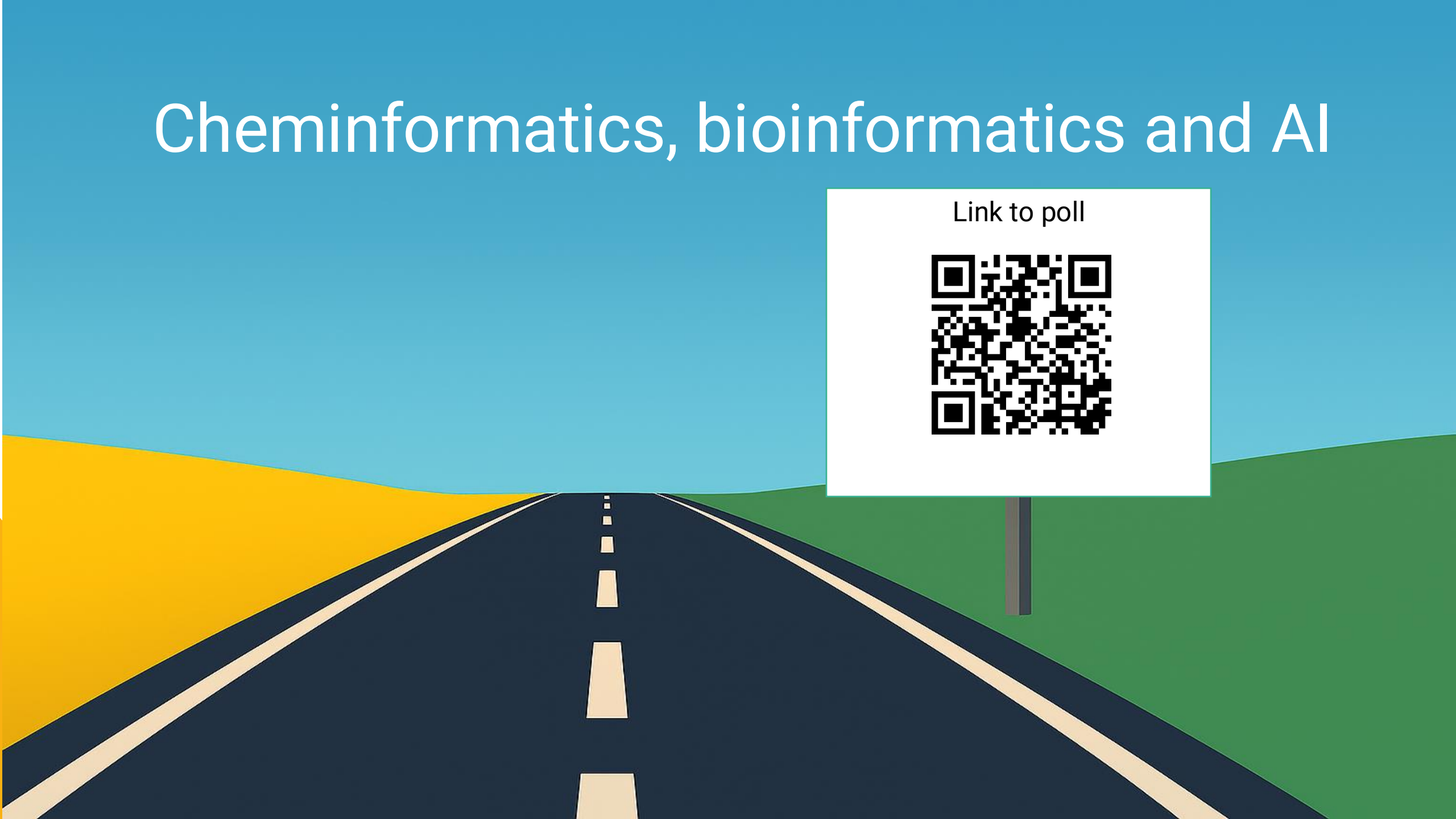
for Regulated product data

- Development of new OECD Harmonised Templates
- Structuring of dossier data in IUCLID format
- Plastic Food Contact Material pilot
- Involving stakeholders in user acceptance testing

Reflections on

- Pain points in provisioning of data to EFSA
- How can AI improve sharing of data
- Confidentiality request process
- Fitness of current stakeholder fora to discuss data needs

Cheminformatics, bioinformatics and AI



Cheminformatics, bioinformatics and AI

How can AI help to:

- Reduce the administrative burden for industry in dossier preparation
- Enhance data quality for regulatory assessments?

How can EFSA facilitate a more extensive use of bioinformatics, cheminformatics, AI?

- Development of tools and methodologies
- Prediction of toxicity for data poor substances
- Regulatory acceptance