The Information Platform for chemical monitoring - IPCheM

Objectives of the meeting

Open the discussion for:

- Possibility to access to chemical data and metadata under EMODnet
- Possibility to publish in IPCheM data and metadata on chemical concentrations
- How can we start the collaboration plan next steps
**Problem statement**
The lack of information on the chemical exposure and burden on the humans and the environment is a major gap in knowledge base for the European chemical policies.

- It is not possible to assess the real impact of chemicals and their mixtures
- It is difficult to assess effectiveness of policies
- Ad-hoc collections are very time demanding and inefficient

**Key Policy Question**
support improved understanding of the chemical mixtures to which human populations and the natural environment are actually exposed

- EC Communication: The combination effects of chemicals – Chemical mixtures COM/2012/0252 final
- 7th Environment Action Programme. Priority objective 5: to improve the knowledge and evidence base for Union environment policy

---

**Definition and Objectives**

*2012: JRC started the design and development of IPCheM “Information Platform for chemical monitoring”, under the project coordination of DG ENV (founding the project)*

**IPCheM is**
a single access point for discovering and retrieving chemical monitoring data in Europe

- implementing searching facilities to discovery, access and retrieve chemical monitoring data
- Offering hosting facilities for “orphan” data or data currently not easily accessible
- Proposing and working on chemical monitoring documentation (Metadata and data templates) of defined quality
And...

- **It covers only chemical monitoring data**: other data (e.g. hazard, emissions, physicochemical properties) are not integrated to IPChem but can be linked to it
- **Data providers**: Commission Services, EU Agencies, EU Member States, international and national organisations, researchers
- **Users**: Commission Services, EU Agencies, MSs, international and national organisations, researchers and public, respecting existing access and use conditions

![DIOXIN DB (DE) AIRBASE BIOSOIL WFD priority substances](image)

**Structure**

- **Environmental monitoring data**
  - EEA
  - EEA
- **Human Biomonitoring**
  - EEA
- **Food and feed**
  - EFSA
- **Indoor air and products**
  - IHCP JRC

**4 thematic modules**

- **PROJECT GOVERNANCE**
  - Key Masters: SMRT, ESDL, OCLCT...
  - Chef de fil DG ENV
  - Module coordinators
  - Technical Coordinator JRC
- **Data providers**
  - DGs, EU Agencies, MS Agencies
- **Working groups**
  - Terms and data policy
  - Interoperability

IPChem allows to search and access data related to...
**Challenges**

- Handle heterogeneity of data (formats, web applications, providers, thematic areas)
- Promote data quality and data quality documentation
- Improve the data comparability (inter and across thematic modules and media)
- Support the response to (already identified) policy questions

WITHOUT REQUIRING (OR MINIMASING) THE SUPPORT (TECHNICAL/PHYSICAL/ECONOMIC) OF THE DATA PROVIDERS

**Actions**

- Design and implementation of a decentralised system
- IPCheM Data Policy Definition and implementation guideline
- Metadata and data content proposal for the 4 thematic modules
- Working group activities to identify policy questions and interoperability solutions and data to be integrated
- Data quality: auto-evaluation questionnaire
- Data quality: Quality checking rules and tools
Design solution

IPCheM is designed as a decentralised system establishing remote access to existing data collections

- No data duplication and information systems replication
- Condition/restriction of data access and use defined by Data Providers are always respected

All the transformations to publish the data are carried out by IPCheM team
Through the “Extract, Transform and Load” (ETL) layer (Kattle)

Design solution

1ST level: metadata page
Information about data collection, e.g.:
- Spatial and temporal coverage
- Media
- Chemicals and related UoM
- Conditions of data access and use
- Responsible Body and PoC
- Links (home page, db app.)

2ND level: DB Console
- Visualisation of data values (concentrations) into the Master Table
- Additional specific attributes and filters based on single data structure
- Sample site/station info
- Temporal trends
- Spatial search and query refinement
- Possibility to save the results into the Basket

Unique search interface

- Chemical name
- CAS number
- Module
- Media
- Spatial coverage

Information provided by DP or automatically extracted from the DBs
Software components

IPCheM implementation relies on the following OS Software components

### Front End
- Apache
- JQuery
- Backbone.js
- OpenLayers

### Back End
- Apache Solr
- Apache Tomcat
- ECAS client (integration in Tomcat)
- Virtuoso Open-Source Triple Store

### Data Hosting
- PostgreSQL DBMS

### Data Transformation
- DPentaho Kettle

### Features

- **search tools**: to discover metadata (FIRST LEVEL) and data (SECOND LEVEL) of the data collections integrated in IPCheM, starting by the name of the substances (IOE toxic identifiers) and spatial coverages, refining by molecule and ready.

- **view metadata tool**: to read about the general information about the data collection, to open the data collection web site, if available.

- **DB Console**: to view or take the results of the search, to filter the value using specific filter based on data structure, to select data to save into the basket, to show the location site of the sample and discover info about them, to refine the search by geographical area.

- **Basket tool**: to save the result of the search, to compare results of the same data collection or different data collection, to download result as csv, excel, the number and type of columns, to save/export results as pdf file.

- **news feed (EMEA-Medisys)**: to discover and react new on chemical, update every ST and based on Medisys, the Medical Information System developed by IPCS-PSI, which displays those articles with Interest to Public Health, grouped by disease or disease type.

- **Feedback tool**: to report bugs, ideas, questions or problem related with the platform and to track solutions.

- **Map viewer/creator**: to access and overlay chemical concentration measurements of 3-n substances in different media, starting by a geographical location, integration of base map layers and thematic map layers to facilitate the interpretation of the results.

- **User authentication and access management service**: to control access to data with particular restriction, access management.
Status of data integration

- Number of current Data Collections: **16**
- Number of Data Collections for September: **21**
- Number of concentration measurements: **21,592,958 across Europe and beyond**

Status of information system linkages

- Links with other information systems and services
  - Based on CAS nr or the Chemical name:
  - ChemAgora service (XML) customized for IPCheM
  - Providing a selection of links on Third-party repositories
  - Created by IHCP Systems Toxicology

- News on Media, searching by name,
- Based on EMM-MEDISYS
- Updated every 15'
- Developed in Collaboration with IPSC

- Links to BUMA and BUMAC, providing emissions and health relevant thresholds
- Set up in collaboration with IHCP Chemical Assessment and Testing Unit