



The SOLUTIONS project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 603437

Solutions for present and future emerging pollutants in land and water resources management

THEME

ENV.2013.6.2-2

Toxicants, environmental pollutants and land and water resources management

Start date of project: 1st October 2013

Duration: 5 years

Deliverable

D5.1 Integrated Data Portal for SOLUTIONS (IDPS) and web-based knowledge base integrated with IPCheM and RiBaTox

Due date of deliverable: 31/03/2018

Actual submission date: 07/06/2018

Author(s)

Sara Comero, Joint Research Centre – European Commission, JRC, Ispra, Italy

Bernd Manfred Gawlik, Joint Research Centre – European Commission, JRC, Ispra, Italy

in cooperation with
all SOLUTIONS partners

| Dissemination Level | | |
|----------------------------|--|----------|
| PU | Public | X |
| PP | Restricted to other programme participants (including EC) | |
| RE | Restricted to a group specified by the consortium (including EC) | |
| CO | Confidential, only for members of the consortium (including EC) | |

Summary

The Integrated Data Portal for SOLUTIONS has been ideates for providing solutions for monitoring, assessment, prioritisation, abatement and regulatory problems in water resources management.

Managing the information through the portal it is possible to overcome the need for a more coherent and combined data management by the integration of all information in one knowledge base as interlinked database architecture in order to enables information from multiple data sources to be available to each member of the project, thus facilitating a multidisciplinary. IDPS allows data to be available and searchable through one end-user friendly portal, which also allow an optimal interfacing with the models.

Its main goals are to exchange compound- and structure associated data within the consortium, hold and link to additional survey data, provided also by external databases and providers, and present SOLUTIONS products to stakeholders. Moreover, IDPS will help identify links between exposure and epidemiological data, to explore potential biological effects and to lead to improved health outcomes.

Infographic

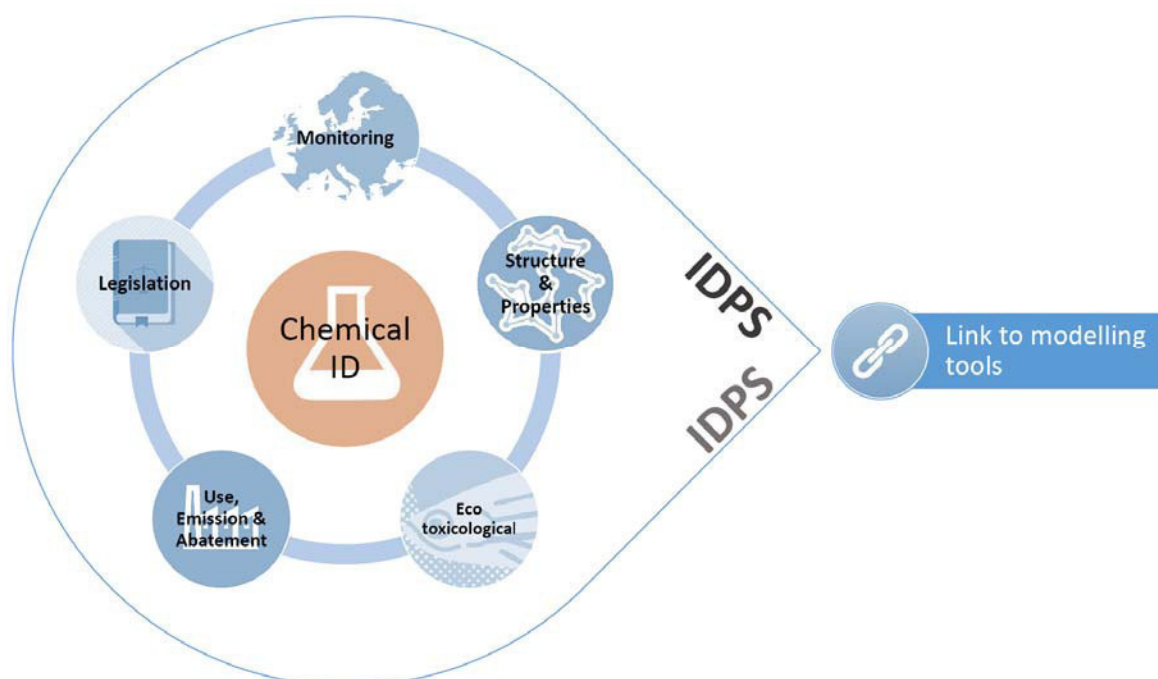


Table of contents

| | | |
|-------|---|----|
| 1. | IDPS information | 4 |
| 2. | Homepage..... | 4 |
| 2.2 | Main information..... | 4 |
| 2.3 | Search for a Chemical..... | 4 |
| 2.4 | Information on available databases | 6 |
| 3 | Module page..... | 6 |
| 3.1 | Environmental Monitoring module | 6 |
| 3.2 | Eco-Toxicology module | 8 |
| 3.3 | Structure & Properties module | 10 |
| 3.4 | Legislation module | 11 |
| 4 | Result page | 12 |
| 4.1 | Environmental Monitoring result page..... | 12 |
| 4.2 | Eco-Toxicology result page | 14 |
| 4.3 | Structure & Properties result page..... | 15 |
| 4.4 | Legislation module | 16 |
| 5 | Basket page | 16 |
| Annex | | 18 |
| | Development of the Solr database..... | 19 |
| | Database schema | 19 |
| | Environmental Monitoring Schema | 21 |
| | Toxicology Schema | 33 |
| | Ecology Schema..... | 54 |
| | Structure & Properties Schema | 56 |
| | Legislation Scheme | 60 |
| | Create command..... | 61 |
| | Populate command | 62 |
| | Substructure tool | 63 |
| | Configuration | 63 |
| | Entry script | 65 |
| | Table creation | 66 |
| | Entry file..... | 67 |

List of Abbreviations

| | |
|---------|--|
| CAS | Chemical Abstracts Service |
| IDPS | Integrated Data Portal for SOLUTIONS |
| InChI | International Chemical Identifier |
| IPCHEM | Integrated Platform for Chemical Monitoring |
| RiBaTox | Decision support tool for River Basin specific Toxicants |
| SMILES | Simplified molecular-input line-entry system |

1. IDPS information

The Integrated Data Portal for SOLUTIONS (IDPS), has been ideates for providing solutions for monitoring, assessment, prioritisation, abatement and regulatory problems in water resources management.

Its main goals are to exchange compound- and structure associated data within the consortium, hold and link to additional survey data, provided also by external databases and providers, and present SOLUTIONS products to stakeholders.

Managing the information through the portal it is possible to overcome the need for a more coherent and combined data management by the integration of all information in one knowledge base as interlinked database architecture in order to enables information from multiple data sources to be available to each member of the project, thus facilitating a multidisciplinary. IDPS allows data to be available and searchable through one end-user friendly portal, which also allow an optimal interfacing with the models.

The following Sections provide a user manual of the portal with screenshot of the main functionalities. In addition, the Annex contains the complete software developer technical documentation and programming of IDPS.

2. Homepage

2.2 Main information

At the bottom of the main screen, information on the SOLUTIONS project and the IDPS portal are given. Links to RiBaTox and IPCHEM are also provided.



Figure 1: Bottom of the main page with main information

2.3 Search for a Chemical

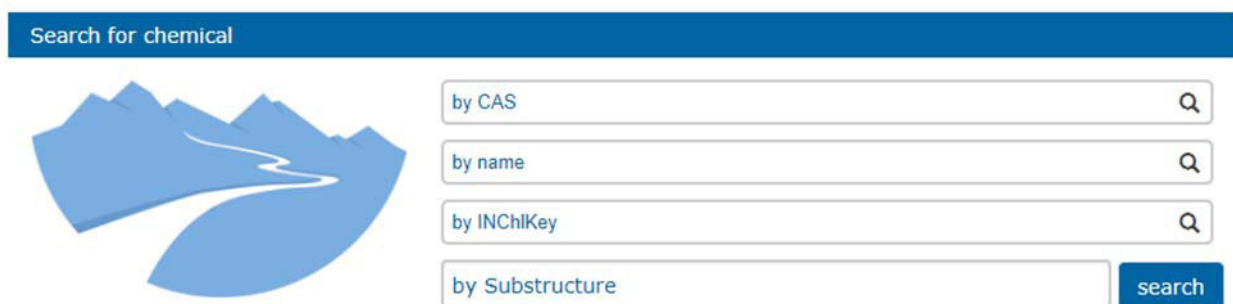
To access the information stored in the available databases implemented in IDPS, it is mandatory to search by a chemical and then select one of the five IDPS modules.

The IDPS homepage provides a search interface that allow users to perform a single chemical search using one of the following three identifiers (Figure 2):

- by CAS number
- by name (common English name)
- by INChIKey

To start the search, edit one of the chemical identifier in the corresponding free-text fields or select a value from the drop-down lists. When start editing in the CAS and INChIKey boxes, the drop-down list includes all the “start with” entries; in the case of the search by name, the drop-down list includes all chemicals which “contains” the edited text. An examples of search by name is given in Figure 3.

After a chemical is selected, the IDPS modules which do NOT contain the chemical appear in grey. In the example of Figure 4, following the selection of “Diclofenac”, the last three modules appear in grey, while the Environmental Monitoring and the Eco-Toxicology modules are clickable, indicating the availability of the selected chemical in their databases.



Search for chemical

by CAS

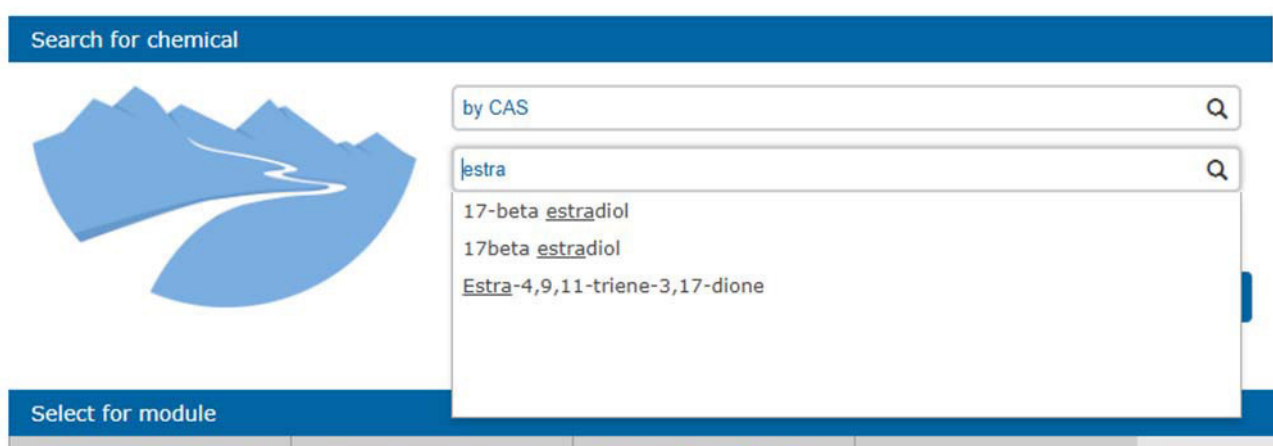
by name

by INChIKey

by Substructure

search

Figure 2: Homepage: search interface



Search for chemical

by CAS

by name

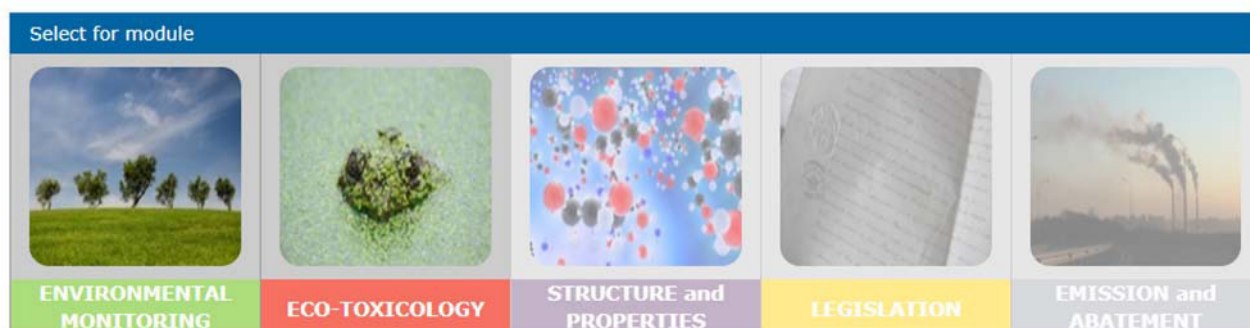
17-beta estradiol

17beta estradiol

Estra-4,9,11-triene-3,17-dione

Select for module

Figure 3: Screenshot of the main page: selection of a chemical



Select for module

ENVIRONMENTAL MONITORING

ECO-TOXICOLOGY

STRUCTURE and PROPERTIES

LEGISLATION

EMISSION and ABATEMENT

Figure 4: Screenshot of the main mage: module highlighted after the selection of a chemical

IDPS also provide to search a chemicals by a substructure (the last box in Figure 2). The

substructure search has been implemented only using the SMILES codification.

To search for a substructure, insert the SMILES in the respective free-text field and then click the Search button (Figure 5). The search result(s) will be displayed in a drop-down list, where only a single value can be selected at a time.

Figure 5: Substructure search tools

2.4 Information on available databases

Prior to select a chemical, it is possible to access the list of the available datasets for each IDPS module.

Clicking on one of the modules, the list of implemented databases for the selected module is displayed in a new page. For each database, the metadata can be accessed by clicking on the info icon.

3 Module page

After selecting a chemical, one of the modules must be chosen to access the list of available databases.

A common layout for each module page has been adopted, but some minor changes in the structure have been applied to each module separately to take into account the specificity of the information to be displayed. Moreover, to facilitate the readability and the navigation through the IDPS pages, a different colour has been adopted and maintained for each module.

3.1 Environmental Monitoring module

The Environmental Monitoring module page can be accessed by selecting a chemical (for example: Diclofenac) and then the corresponding module icon from the homepage. The search result is displayed in a new window (Figure 6).

A geographical map of Europe visualises all the sampling points collected in the available databases, listed below the map. Using the map, it is possible to refine the search by selecting one or more countries: only databases containing sampling points in the selected country/ies are displayed. The user may further define the appropriate search by including an additional filter based on the matrix type.

Next to the map, the IPCHEM logo provides a link to the IPCHEM search page. IPCHEM is the European Commission's reference access point for searching, accessing and retrieving chemical occurrence data collected and managed in Europe.

Under the SOLUTIONS logo, the map button which is red-circled in Figure 6, allows to select one or more countries from the map and freeze the selection for a further search both in the same module, by changing the substance, or in other modules. To activate the freezing option, click on the map button and then click on the chosen country/ies (they turn into a dark grey colour). Then, to confirm the selection, click again on the map button: selected countries will be red coloured (Figure 7). With this "Freeze" option, starting a new substance search or changing the module will maintain the country selection. To deactivate the country-freezing option, click again on the map button.



Figure 6: Environmental Monitoring module page

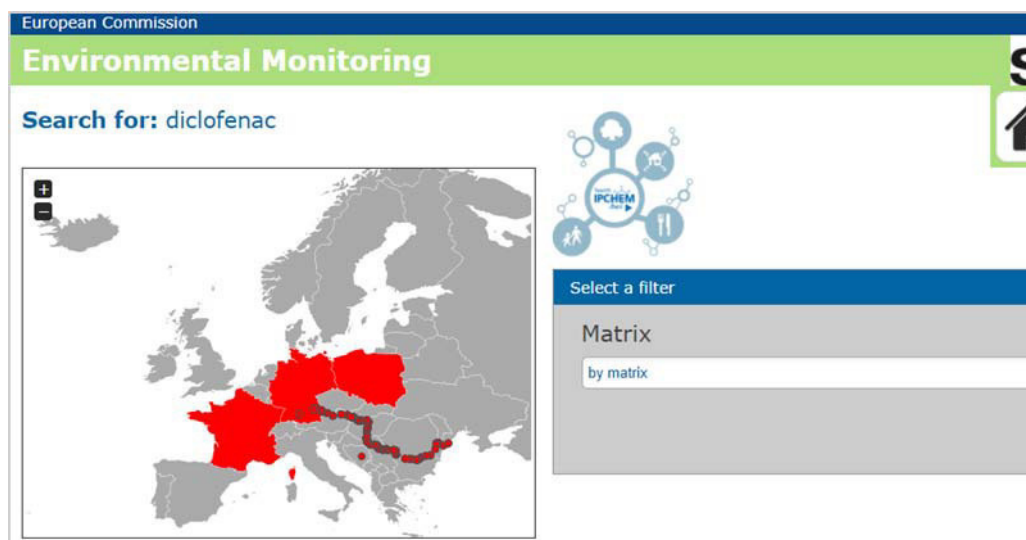


Figure 7: Environmental Monitoring module page: freeze option

The bottom part of the screen, lists all of the available databases. If filtering options (country or matrix-based) are included, the list is updated removing those databases that do not match with the selected filters. Clicking on the info button will open the corresponding metadata page. To enter the result page of one of the available databases, click the arrow at the right of the database name; the corresponding result page will be displayed in a new window (See Section 4.1).

3.2 Eco-Toxicology module

The Eco-Toxicology module page can be accessed by selecting a chemical (for example: Chlorofen) and then the corresponding module icon from the homepage. The search result is displayed in a new window (Figure 8).

Specifically for this module, two sub-categories have been created to best fit the information provided by the datasets: Toxicology and Ecology.

The **Toxicology** category (it is the default category) includes all databases related to the toxicity of chemicals. Geographical specific information can be visualised in the map, when available. Using the map, it is possible to refine the search by selecting one or more countries: only databases containing sampling points in the selected country/ies are displayed. Similar to the Environmental Monitoring module, the “Freeze” option is available to memorise selected country/ies.

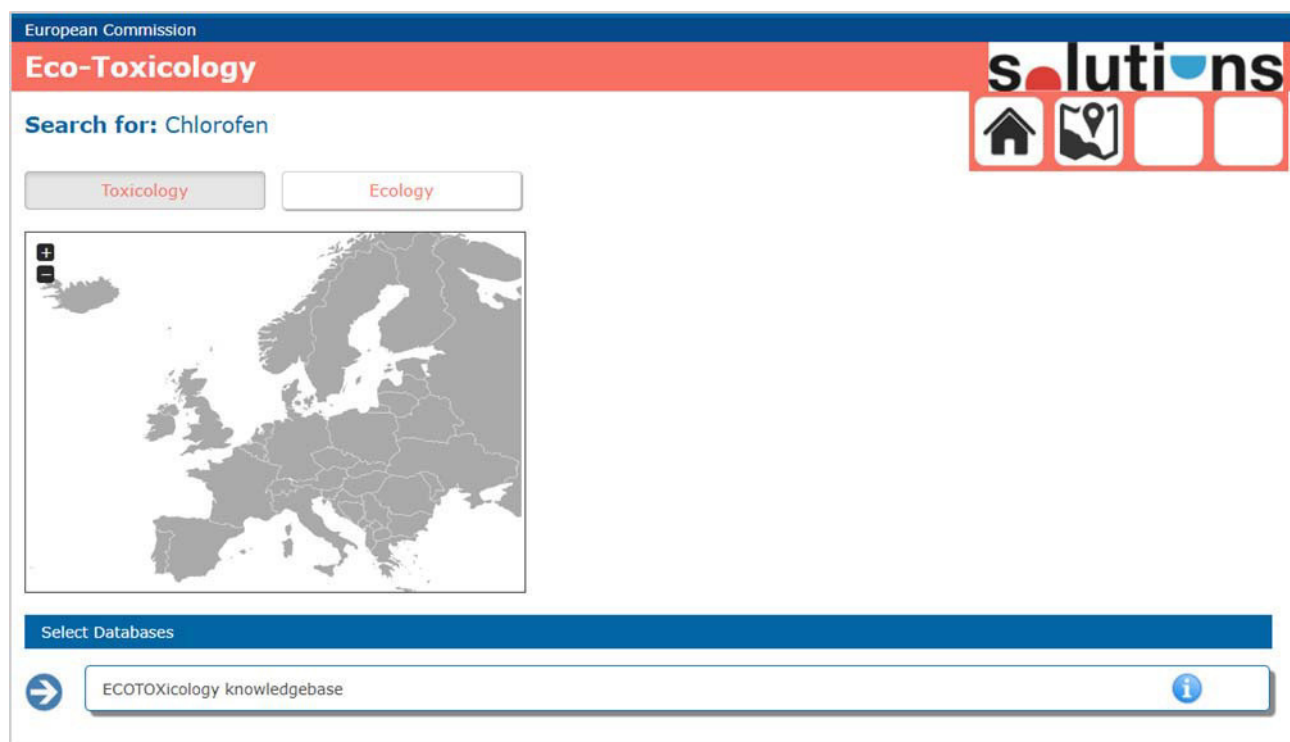


Figure 8: Eco-Toxicology module, Toxicology category

The **Ecology** category (Figure 9) display all traits-based databases which are available in the ISPD. Since this category is not chemical-specific (no information on chemical is given in the datasets), it can be displayed for every selected chemical. If the information in the stored databases is georeferenced, samples will be displayed on the map and a country-base filter can be applied to refine the search.

The user may further define the appropriate search by including an additional filter based on the specie taxonomic group, family or taxon/scientific name.

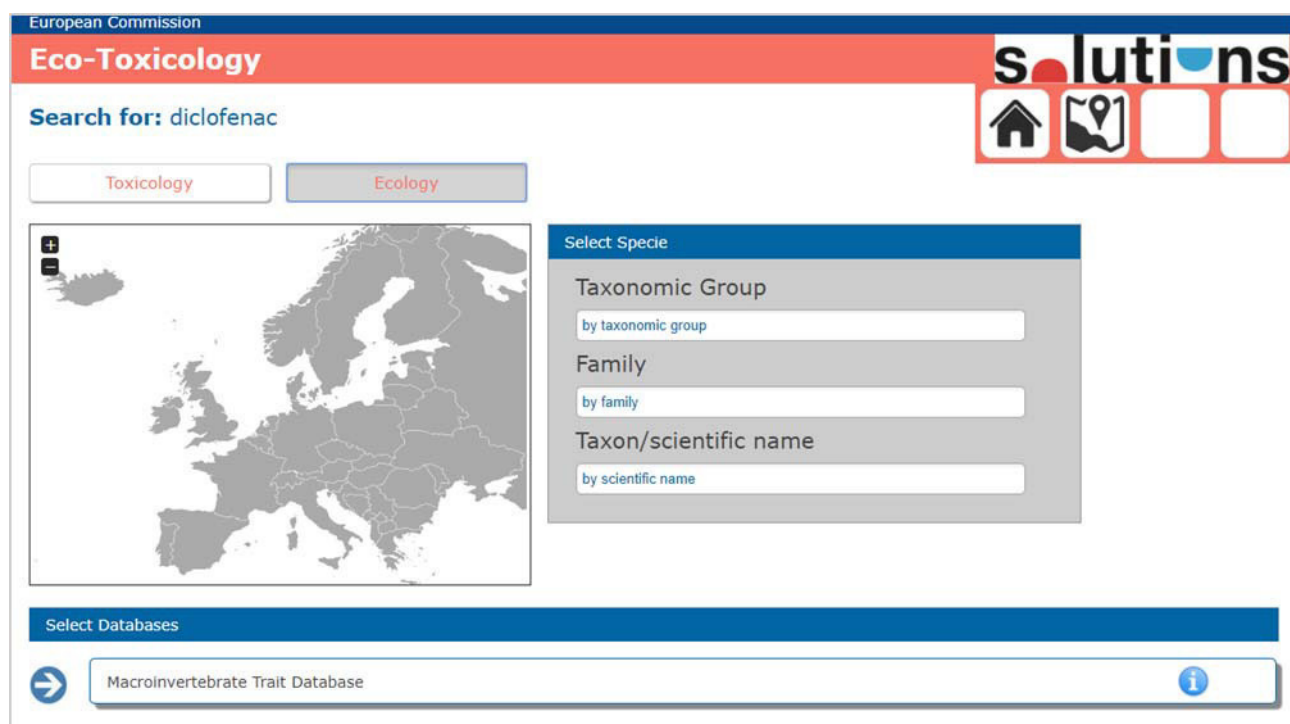


Figure 9: Eco-toxicology module, Ecology category

For both categories, the bottom part of the screen, lists all of the available databases. If filtering options (country or species-based) are included, the list is updated removing those databases that do not match with the selected filters. Clicking on the info button will open the corresponding metadata page.

To enter the result page of one of the available databases, click the arrow at the right of the database name; the corresponding result page will be displayed in a new window (See Section 4.24.1).

3.3 Structure & Properties module

The Structure & Properties module page can be accessed by selecting a chemical (for example: Diclofenac) and then the corresponding module icon from the homepage. The search result is displayed in a new window (Figure 10).

In this module, the map has been replaced by a summary box showing the main information and the chemical structure of the selected substance. This information is retrieved by the MassBank database (<https://massbank.eu/MassBank/>) provided by the UFZ SOLUTIONS partner. No filtering options are available in this module.

The bottom part of the screen, lists all of the available databases. Clicking on the info button will open the corresponding metadata page. To enter the result page of one of the available databases, click the arrow at the right of the database name; the corresponding result page will be displayed in a new window (4.14.3).

European Commission

Structure and Properties

Search for: Diclofenac

STRUCTURAL INFORMATION AND PROPERTIES

Chemical Name: Diclofenac
CAS Number: 15307-86-5
InChIKey: DCOPUUMXTXDBNB-UHFFFAOYSA-N

Select Databases

MassBank


Figure 10: Structure & Properties module

3.4 Legislation module

The Legislation module page can be accessed by selecting a chemical (for example: Mercury) and then the corresponding module icon from the homepage. The search result is displayed in a new window (Figure 11).

In this particular case, the only available and accessible information is a set of documents, which can be downloaded as text files (downloading format is the same of the provided format). For this reason, the Legislation module page corresponds to the Legislation result page.

A map is available to refine the search by selecting one or more countries, but this works only in the case of databases which provided for georeferenced data.

The bottom part of the screen, lists all of the available databases. Clicking on the info button will open the corresponding metadata page. The document icon () is a static link which gives access to the on-live version of the database (this feature only work when the on-line version of the database is available). Click on the arrow icon, will display a drop-down list of the legislation documents associated to the selected chemical (see example of Figure 12). If needed, the user may directly download one or more documents from the drop-down list, by clicking on them.

The basket icon at the left of each database name (red circle in Figure 12), can be used to save in the basket the list of documents of the corresponding database.



Figure 11: Legislation Module

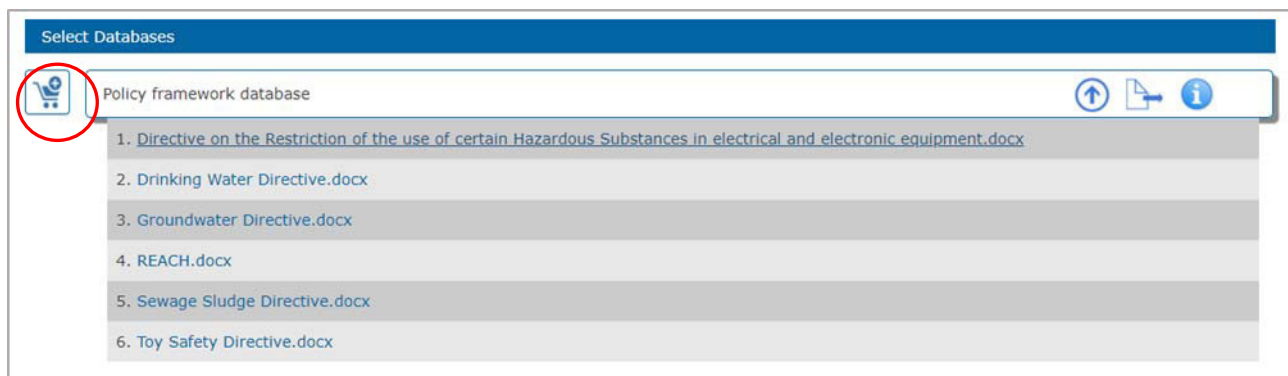


Figure 12: Legislation Module –list of available documents for the selected database

4 Result page


Following the selection of a chemical in a specific module in the homepage, and then the selection of one of the available databases, the result page will be displayed in a new windows. If filtering options have been applied in the module page (for example: selection of one or more countries, selection of a media, selection of a specific specie name) the result page will display filtered data.

4.1 Environmental Monitoring result page





















At the bottom of the page, the table with the available samples is given (Figure 13). Only the first ten entries are displayed; to view all the entries the user can select different data rows from the navigation bar under the table. If a country-based or media-based filter was applied in the module page, only the filtered sampled are shown.

From the table, it is possible to select all the data entries (option “Select the full-table” at the top of the table) or single data elements by selecting the samples one-by-one, for further save them in the basket. All data elements in the table can be ordered in ascending or descending order by clicking on the arrow next to the column's header.

When available, further information of samples can be accessed by clicking the info button under the interactive column. An example of the info box is given in Figure 14.

Under the interactive column it is also possible to visualise each sample on the map, by clicking on the location icon 















☐ Select the full-table

| Basket | ID sample | Country | Concentration | UoM | LoD | LoQ | Interactive |
|-------------------------------------|-----------|----------|---------------|-----|-----|-------|---|
| <input checked="" type="checkbox"/> | JDS14 | Austria | 0,0082 | N/A | 0 | 0,002 |   |
| <input type="checkbox"/> | JDS17 | Slovakia | 0,0057 | N/A | 0 | 0,002 |   |
| <input type="checkbox"/> | JDS25 | Slovakia | 0,0073 | N/A | 0 | 0,002 |   |
| <input checked="" type="checkbox"/> | JDS37 | Hungary | 0,0084 | N/A | 0 | 0,002 |   |
| <input checked="" type="checkbox"/> | JDS44 | Hungary | 0,009 | N/A | 0 | 0,002 |   |
| <input type="checkbox"/> | JDS48 | Croatia | 0,0064 | N/A | 0 | 0,002 |   |
| <input type="checkbox"/> | JDS57 | Serbia | 0,0041 | N/A | 0 | 0,002 |   |
| <input type="checkbox"/> | JDS65 | Serbia | 0,0034 | N/A | 0 | 0,002 |   |
| <input type="checkbox"/> | JDS70 | Romania | 0,0022 | N/A | 0 | 0,002 |   |
| <input type="checkbox"/> | JDS74 | Bulgaria | 0,0095 | N/A | 0 | 0,002 |   |

Showing 1 to 10 of 457 entries

Previous 1 2 3 4 5 ... 46 Next

Figure 13: Table with available samples and main information

| | | | | | | | |
|-------------------------------------|---------|---------|--------|-----|---|-------|---|
| <input type="checkbox"/> | JDS20 M | Germany | 0,0055 | N/A | 0 | 0,001 |   |
| <input checked="" type="checkbox"/> | JDS21 M | | | | | 0,001 |   |
| <input type="checkbox"/> | JDS22 M | | | | | 0,001 |   |
| <input type="checkbox"/> | JDS24 M | | | | | 0,001 |   |
| <input type="checkbox"/> | JDS25 M | | | | | 0,001 |   |
| <input type="checkbox"/> | JDS26 M | | | | | 0,001 |   |
| <input type="checkbox"/> | JDS28 M | Hungary | 0,001 | N/A | 0 | 0,001 |   |

Other INFO

TITLE OF THE PROJECT: JDS2 (Joint Danube Survey 2)
 LOCATION NAME: Vah (Rkm 0.8), Danube/Vah (rkm 1766), Danube (SK)
 NATIONAL CODE: SK
 FRACTION: Dissolved fraction
 SAMPLING DATE: 24/8/2007
 ORGANISATION: N/A

Figure 14: Box with further information available for each sample

A summary of the selected option is available at the top of the result page (Figure 15), which list the selected module, the searched chemicals and the name of the database. To save the selected samples in the basket page, click on the basket icon in the summary box.

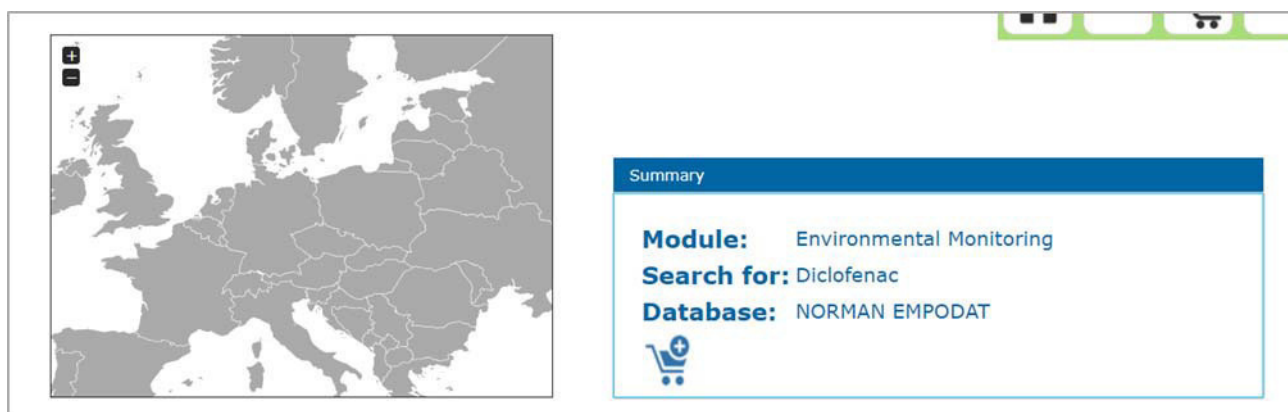


Figure 15: Summary box with basket button

4.2 Eco-Toxicology result page

Under the Eco-toxicology module, two sub-categories have been created: Toxicology and Ecology. For the first category, the selection of a chemical is mandatory, as well as for the other IDPS modules. An exception is due for the Ecology module which is always accessible, since it is not directly linked with chemicals.

The layout of the **Toxicology** result page, is similar to the Environmental Monitoring Result page. A summary box lists all the main information related to the specific search. The data table provides the list of samples, which can be sorted by clicking on the arrow next to the column's header. Samples can be selected and then stored in the basket. Under the interactive column, the info box provides for additional information (See example in Figure 16) and the location icon can visualises each sample in the map. When the location icon is deselected, no geographical information was provide for the sample.

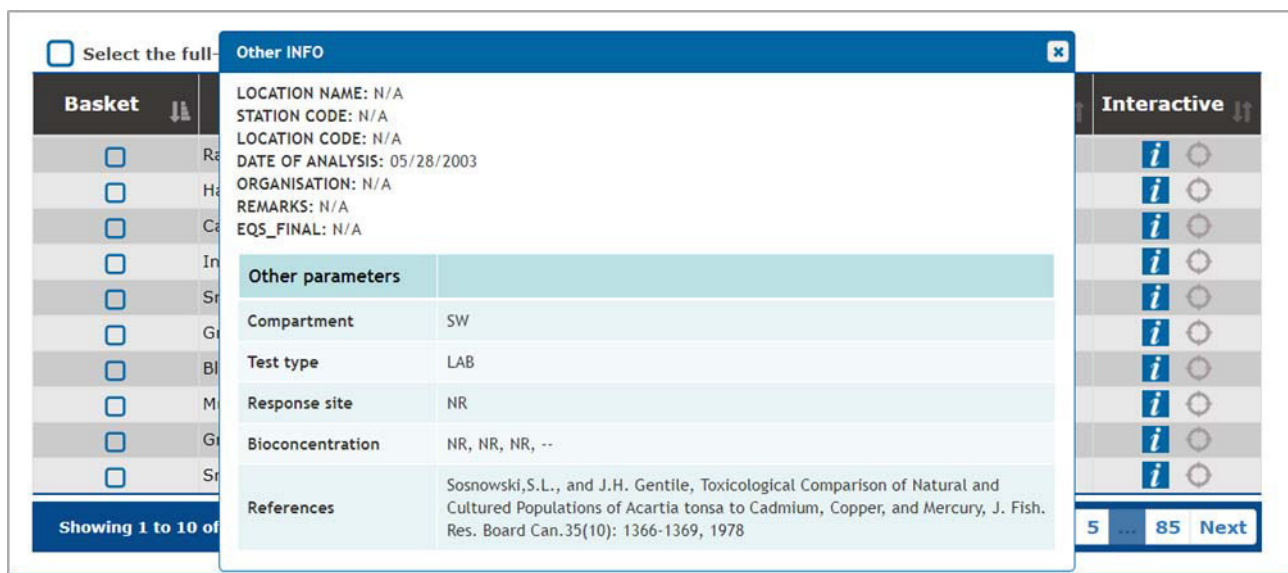


Figure 16: Additional info box

In the **Ecology** module, the summary list the main information used in the search. The data table (Figure 17) lists all the samples, with the possibility for the user to further select specific

measured parameters in drop-down lists (Traits and Category columns). The “Values” column will display the result of the selected parameters. Additional information available for each samples can be visualised by clicking on the info icon under the Interactive column.

☐ Select the full-table

| Basket | Taxon Name | Trait | Category | Value | Interactive |
|--------------------------|----------------------------|-------------------------------------|----------|-------|-------------|
| <input type="checkbox"/> | Chelicorophium curvispinum | select trait | | | <i>i</i> |
| <input type="checkbox"/> | Chelicorophium robustum | select trait | | | <i>i</i> |
| <input type="checkbox"/> | Chelicorophium sowinskyi | Maximal potential size | | | <i>i</i> |
| <input type="checkbox"/> | Corophium sp. | Life cycle duration | | | <i>i</i> |
| <input type="checkbox"/> | Dikerogammarus bispinosus | Potential number of cycles per year | | | <i>i</i> |
| <input type="checkbox"/> | Dikerogammarus haemobaphes | Aquatic stages | | | <i>i</i> |
| <input type="checkbox"/> | Dikerogammarus sp. juv. | Reproduction | | | <i>i</i> |
| | | Dispersal | | | <i>i</i> |
| | | Resistance forms | | | <i>i</i> |
| | | Respiration | | | <i>i</i> |
| | | Locomotion and substrate relation | | | <i>i</i> |
| | | Food | | | <i>i</i> |
| | | Feeding habits | | | <i>i</i> |
| | | Transversal distribution | | | <i>i</i> |
| | | Longitudinal distribution | | | <i>i</i> |

Figure 17: Filtering option under the result data table

☐ Select the full-table

| Basket | Taxon Name | Trait | Category | Value | Interactive |
|--------------------------|----------------------------|----------------|----------|-------|-------------|
| <input type="checkbox"/> | Chelicorophium curvispinum | Aquatic stages | | 0,33 | <i>i</i> |
| <input type="checkbox"/> | Chelicorophium robustum | select trait | | | <i>i</i> |
| <input type="checkbox"/> | Chelicorophium sowinskyi | select trait | | | <i>i</i> |

Other INFO

WUR CODE: 84

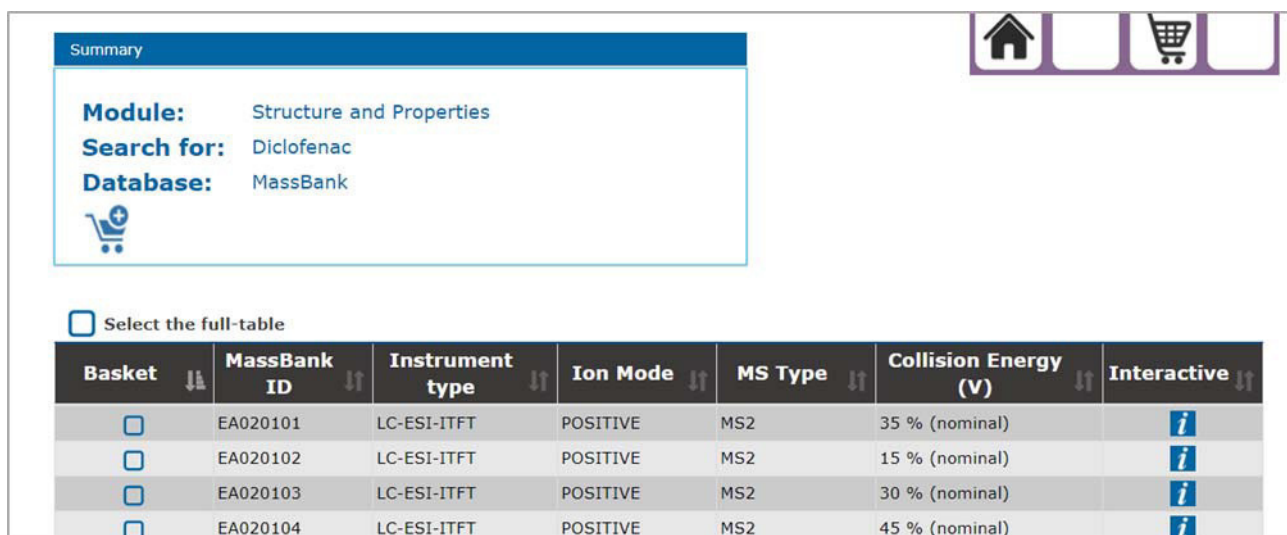
SAMPLING METHOD: MHS

Figure 18: Other info box

For both sub-category result pages, selected samples can be saved in the basket for further download by clicking the basket icon in the summary info box.

4.3 Structure & Properties result page

In the Structure & Properties result page, the summary box summarized the selected searching options. The data table list all the entries corresponding the selected substance in the selected database (no filtering options are available in the respective module page). All data elements in the table can be ordered in ascending or descending order by clicking on the arrow next to the column's header. Samples can be selected one-by-one or as a full-table and stored in the basket page by clicking on the basket icon. Further information on samples can be visualised by clicking the info icon (Figure 20).



| Basket | MassBank ID | Instrument type | Ion Mode | MS Type | Collision Energy (V) | Interactive |
|--------------------------|-------------|-----------------|----------|---------|----------------------|-------------------|
| <input type="checkbox"/> | EA020101 | LC-ESI-ITFT | POSITIVE | MS2 | 35 % (nominal) | i |
| <input type="checkbox"/> | EA020102 | LC-ESI-ITFT | POSITIVE | MS2 | 15 % (nominal) | i |
| <input type="checkbox"/> | EA020103 | LC-ESI-ITFT | POSITIVE | MS2 | 30 % (nominal) | i |
| <input type="checkbox"/> | EA020104 | LC-ESI-ITFT | POSITIVE | MS2 | 45 % (nominal) | i |

Figure 19: Summary box and data Table, Structure & Properties module



| Other INFO |
|--|
| DATE: 2017.01.05 |
| AUTHORS: Schulze T, Krauss M, Department of Effect-Directed Analysis, Helmholtz Centre for Environmental Research - UFZ GmbH, Leipzig, Germany |
| COPYRIGHT: Copyright (C) 2017 |
| PUBLICATION: N/A |
| EU LINK: massbank.eu |

Figure 20: Other info box, Structure & Properties module

4.4 Legislation module

Only for the Legislation module, the result page correspond to the result page, described in Section 3.4.

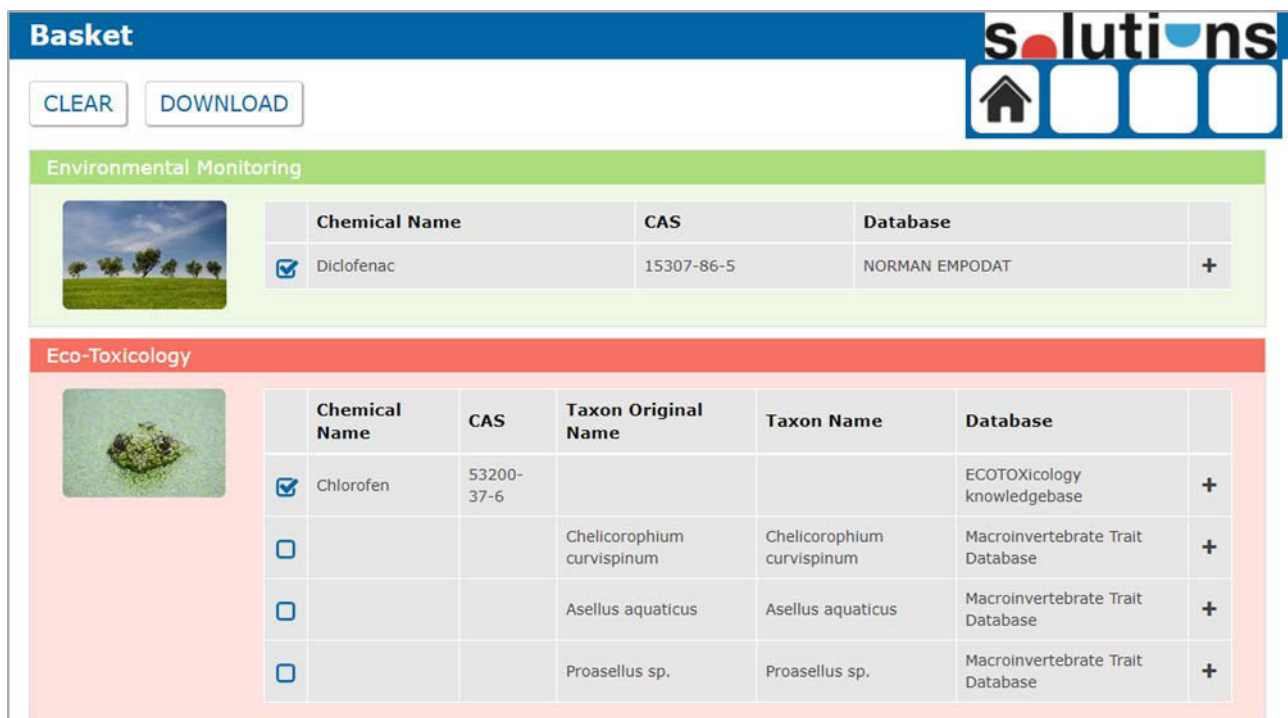
5 Basket page

Selected samples in each of the available IDPS modules can be saved in the basket from the respective result page. The basket page can be accessed from each result page by clicking on the basket icon on the top of the page (Figure 21).



Figure 21: example of the basket icon in the Environmental Monitoring result page


The basket page displays all the entries saved in each of the module result pages, both as single samples and full-table. For each module, a dedicated box with the list of saved result is give. An example of basket page is given in Figure 22 and Figure 23.



Basket


CLEAR DOWNLOAD

Environmental Monitoring



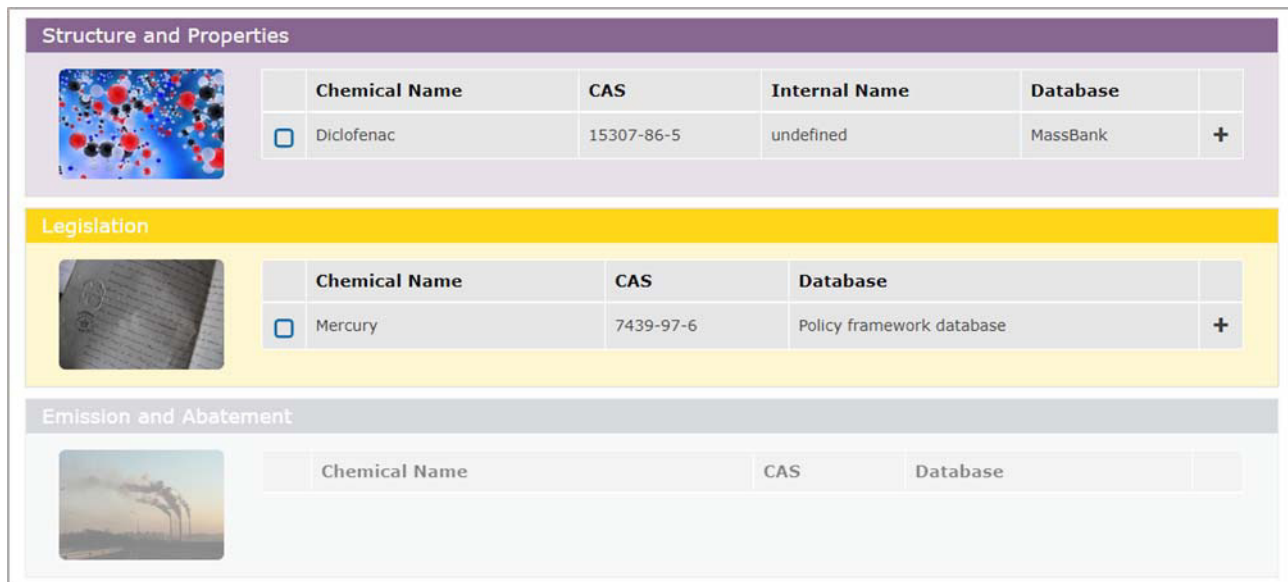
| Chemical Name | CAS | Database | |
|--|------------|----------------|---|
| <input checked="" type="checkbox"/> Diclofenac | 15307-86-5 | NORMAN EMPODAT | + |

Eco-Toxicology

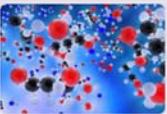


| Chemical Name | CAS | Taxon Original Name | Taxon Name | Database | |
|---|------------|----------------------------|----------------------------|----------------------------------|---|
| <input checked="" type="checkbox"/> Chlorofen | 53200-37-6 | | | ECOTOXicology knowledgebase | + |
| <input type="checkbox"/> | | Chelicorophium curvispinum | Chelicorophium curvispinum | Macroinvertebrate Trait Database | + |
| <input type="checkbox"/> | | Asellus aquaticus | Asellus aquaticus | Macroinvertebrate Trait Database | + |
| <input type="checkbox"/> | | Proasellus sp. | Proasellus sp. | Macroinvertebrate Trait Database | + |

Figure 22: example of the basket page




Structure and Properties




| Chemical Name | CAS | Internal Name | Database | |
|-------------------------------------|------------|---------------|----------|---|
| <input type="checkbox"/> Diclofenac | 15307-86-5 | undefined | MassBank | + |

Legislation



| Chemical Name | CAS | Database | |
|----------------------------------|-----------|---------------------------|---|
| <input type="checkbox"/> Mercury | 7439-97-6 | Policy framework database | + |

Emission and Abatement



| Chemical Name | CAS | Database | |
|---------------|-----|----------|--|
|---------------|-----|----------|--|

Figure 23: example of the basket page

One or more saved entries can be selected to be downloaded. All downloads are saved as .csv files. It is also possible to clear one or more basket entries using the “Clear” button. The “+” (plus) symbol at the end of each entry will display the metadata of the corresponding database.

Annex

IDPS technical documentation provided by the software developer

Development of the Solr database

Following the procedure for the development of the Solr database.

Database schema

```
{
  "add-field": {
    "name": "moduleId",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "databaseName",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "institution",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "role",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "PoCName",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "PoCEmail",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
```

```
        "name": "accessAndUse",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "dbVersion",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "generalInfo",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "QAAQC",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "module-specific-notes",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    }
}
```

Environmental Monitoring Schema

```
{
  "add-field": [
    {
      "name": "countryName",
      "type": "string",
      "indexed": true,
      "stored": true,
      "multiValued": false
    },
    {
      "name": "locationName",
      "type": "string",
      "indexed": true,
      "stored": true,
      "multiValued": false
    },
    {
      "name": "nationalCode",
      "type": "string",
      "indexed": true,
      "stored": true,
      "multiValued": false
    },
    {
      "name": "countryCode",
      "type": "string",
      "indexed": true,
      "stored": true,
      "multiValued": false
    },
    {
      "name": "sampleID",
      "type": "string",
      "indexed": true,
      "stored": true,
      "multiValued": false
    },
    {
      "name": "eastWest",
      "type": "string",
      "indexed": true,
      "stored": true,
      "multiValued": false
    },
    {
      "name": "longitudeDegrees",
      "type": "string",
```

```
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "longitudeMinutes",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "longitudeSeconds",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "longitudeDec",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "latitudeNorthSouth",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "latitudeDegrees",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "latitudeMinutes",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "latitudeSeconds",
        "type": "string",
        "indexed": true,
```



```
        "stored": true,
        "multiValued": false
    },
    {
        "name": "latitudeDec",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "precisionOfCoordinates",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "samplingMatrix",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "other1",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "CASNumber",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "smiles",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "chemicalName",
        "type": "string",
        "indexed": true,
        "stored": true,
```

```
        "multiValued": false
    },
    {
        "name": "individualCompound",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "concentrationData",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "concentration",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "concentrationValue",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "value",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "unitOfMeasures",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "samplingDay",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    }
```

```
    },
    {
      "name": "samplingMonth",
      "type": "string",
      "indexed": true,
      "stored": true,
      "multiValued": false
    },
    {
      "name": "samplingYear",
      "type": "string",
      "indexed": true,
      "stored": true,
      "multiValued": false
    },
    {
      "name": "hour",
      "type": "string",
      "indexed": true,
      "stored": true,
      "multiValued": false
    },
    {
      "name": "minute",
      "type": "string",
      "indexed": true,
      "stored": true,
      "multiValued": false
    },
    {
      "name": "analysedFraction",
      "type": "string",
      "indexed": true,
      "stored": true,
      "multiValued": false
    },
    {
      "name": "nameOfRiverEstuaryLakeReservoirSea",
      "type": "string",
      "indexed": true,
      "stored": true,
      "multiValued": false
    },
    {
      "name": "nameOfTheWiderAreaOfSampling",
      "type": "string",
      "indexed": true,
      "stored": true,
      "multiValued": false
    },
  },
```

```
{
  "name": "riverBasinName",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
{
  "name": "riverKm",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
{
  "name": "pH",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
{
  "name": "temperatureDegree",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
{
  "name": "hardness",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
{
  "name": "remark",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
{
  "name": "soilTexture",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
{
```

```
        "name": "concentrationNormalisedForTheParticleSize",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "grainSizeDistribution",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "percentageOfTotalDryWeight",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "other2",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "additionalInformation",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "typeOfDataSource",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "typeOfMonitoring",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "otherInformation",
```

```
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "projectTitle",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "organisation",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "mail",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "laboratory",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "laboratoryID",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "references",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "LOD",
        "type": "string",
```

```
        "indexed": true,  
        "stored": true,  
        "multiValued": false  
    },  
    {  
        "name": "LOQ",  
        "type": "string",  
        "indexed": true,  
        "stored": true,  
        "multiValued": false  
    },  
    {  
        "name": "percentage",  
        "type": "string",  
        "indexed": true,  
        "stored": true,  
        "multiValued": false  
    },  
    {  
        "name": "samplePreparationMethod",  
        "type": "string",  
        "indexed": true,  
        "stored": true,  
        "multiValued": false  
    },  
    {  
        "name": "other3",  
        "type": "string",  
        "indexed": true,  
        "stored": true,  
        "multiValued": false  
    },  
    {  
        "name": "analyticalMethod",  
        "type": "string",  
        "indexed": true,  
        "stored": true,  
        "multiValued": false  
    },  
    {  
        "name": "other4",  
        "type": "string",  
        "indexed": true,  
        "stored": true,  
        "multiValued": false  
    },  
    {  
        "name": "code",  
        "type": "string",  
        "indexed": true,
```

```
        "stored": true,
        "multiValued": false
    },
    {
        "name": "other5",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "number",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name":
"hasTheUsedMethodBeenValidatedAccordingToOneOfTheBelowProtocols",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "haveTheResultsBeenCorrectedForExtractionRecovery",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "wasAFieldBlankChecked",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "isTheLaboratoryAccreditedAccordingToISO17025",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "isTheLaboratoryAccreditedForTheGivenAnalyte",
        "type": "string",
        "indexed": true,
```



```

        "stored": true,
        "multiValued": false
    },
    {
        "name":
"doesTheLaboratoryParticipateInInterlaboratoryStudiesForTheGivenDeterminand",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name":
"summaryOfPerformanceOfTheLaboratoryInInterlaboratoryStudyForTheGivenDeterminand",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "areControlChartsRecordedForTheGivenDeterminand",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "areTheDataControlledByCompetentAuthority",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "remark2",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "databaseId",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    {
        "name": "moduleId",
        "type": "string",

```

```

    "indexed": true,
    "stored": true,
    "multiValued": false
  }
}

```

Toxicology Schema

```
{
  "add-field": {
    "name": "sample_size_mean_op",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "sample_size_mean",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "sample_size_min_op",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "sample_size_min",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "sample_size_max_op",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "sample_size_max",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "sample_size_unit",
    "type": "string",
    "indexed": true,
```

```
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "sample_size_comments",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "obs_duration_mean_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "obs_duration_mean",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "obs_duration_min_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "obs_duration_min",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "obs_duration_max_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "obs_duration_max",
        "type": "string",
        "indexed": true,
        "stored": true,
```

```
        "multiValued": false
    },
    "add-field": {
        "name": "obs_duration_unit",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "obs_duration_comments",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "endpoint",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "endpoint_comments",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "trend",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "effect",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "effect_comments",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    }
```

```
},
"add-field": {
  "name": "measurement",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "measurement_comments",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "response_site",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "response_site_comments",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "effect_pct_mean_op",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "effect_pct_mean",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "effect_pct_min_op",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
}
```

```
"add-field": {
  "name": "effect_pct_min",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "effect_pct_max_op",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "effect_pct_max",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "effect_pct_comments",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "concl_type",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "ion1",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "concl_mean_op",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
```

```
    "name": "concl_mean",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "concl_min_op",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "concl_min",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "concl_max_op",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "concl_max",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "concl_unit",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "concl_comments",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "conc2_type",
```



```
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "ion2",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc2_mean_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc2_mean",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc2_min_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc2_min",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc2_max_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc2_max",
        "type": "string",
```

```
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc2_unit",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc2_comments",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc3_type",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "ion3",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc3_mean_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc3_mean",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc3_min_op",
        "type": "string",
        "indexed": true,
```

```
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc3_min",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc3_max_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc3_max",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc3_unit",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "conc3_comments",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "bcfl_mean_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "bcfl_mean",
        "type": "string",
        "indexed": true,
        "stored": true,
```

```
        "multiValued": false
    },
    "add-field": {
        "name": "bcfl_min_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "bcfl_min",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "bcfl_max_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "bcfl_max",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "bcfl_unit",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "bcfl_comments",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "bcf2_mean_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    }
```

```
},
"add-field": {
  "name": "bcf2_mean",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "bcf2_min_op",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "bcf2_min",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "bcf2_max_op",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "bcf2_max",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "bcf2_unit",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "bcf2_comments",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
}
```

```
"add-field": {
  "name": "bcf3_mean_op",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "bcf3_mean",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "bcf3_min_op",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "bcf3_min",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "bcf3_max_op",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "bcf3_max",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "bcf3_unit",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
```

```
    "name": "bcf3_comments",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "significance_code",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "significance_type",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "significance_level_mean_op",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "significance_level_mean",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "significance_level_min_op",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "significance_level_min",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "significance_level_max_op",
```

```
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "significance_level_max",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "significance_comments",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "chem_analysis_method",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "chem_analysis_method_comments",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "endpoint_assigned",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "organism_final_wt_mean_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "organism_final_wt_mean",
        "type": "string",
```



```
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "organism_final_wt_min_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "organism_final_wt_min",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "organism_final_wt_max_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "organism_final_wt_max",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "organism_final_wt_unit",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "organism_final_wt_comments",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "intake_rate_mean_op",
        "type": "string",
        "indexed": true,
```

```
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "intake_rate_mean",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "intake_rate_min_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "intake_rate_min",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "intake_rate_max_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "intake_rate_max",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "intake_rate_unit",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "intake_rate_comments",
        "type": "string",
        "indexed": true,
        "stored": true,
```

```
        "multiValued": false
    },
    "add-field": {
        "name": "lipid_pct_mean_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "lipid_pct_mean",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "lipid_pct_min_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "lipid_pct_min",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "lipid_pct_max_op",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "lipid_pct_max",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "lipid_pct_comments",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    }
```

```
},
"add-field": {
  "name": "dry_wet",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "dry_wet_pct_mean_op",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "dry_wet_pct_mean",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "dry_wet_pct_min_op",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "dry_wet_pct_min",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "dry_wet_pct_max_op",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "dry_wet_pct_max",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
},
```

```
"add-field": {
  "name": "dry_wet_pct_comments",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "steady_state",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "additional_comments",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "companion_tag",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "created_date",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "modified_date",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
  "name": "old_terretox_result_number",
  "type": "string",
  "indexed": true,
  "stored": true,
  "multiValued": false
},
"add-field": {
```

```
    "name": "latin_name",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "reference_db",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "reference_type",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "author",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "title",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "source",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "publication_year",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "media_type",
```

```

        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "test_location",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    }
}

```

Ecology Schema

```
{
  "add-field": {
    "name": "WUR_code",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "TaxonOriginalName",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "SamplingMethod",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "TaxonName",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "Species",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "Genus",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "Sub_family",
    "type": "string",
    "indexed": true,
```



```
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "Family",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "Order",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "TaxaGroup",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-dynamic-field": {
        "name": "*_eco",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    }
}
```

Structure & Properties Schema

```
{
  "add-field": {
    "name": "accession",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "internalName",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "internalName_B",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "internalName_C",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "exactMass",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "precursorMZ",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": false
  },
  "add-field": {
    "name": "ionMode",
    "type": "string",
    "indexed": true,
```

```
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "formula",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "instr",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "instrType",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "MSType",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "ionization",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "precType",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "license",
        "type": "string",
        "indexed": true,
        "stored": true,
```

```
        "multiValued": false
    },
    "add-field": {
        "name": "Iupac",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "INChIKey",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "Csid",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "PcCid",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "EuLink",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "Authors_SP",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    },
    "add-field": {
        "name": "Copyright_SP",
        "type": "string",
        "indexed": true,
        "stored": true,
        "multiValued": false
    }
```

```
    },  
    "add-field": {  
      "name": "Date_SP",  
      "type": "string",  
      "indexed": true,  
      "stored": true,  
      "multiValued": false  
    },  
    "add-field": {  
      "name": "CollisionEnergy_SP",  
      "type": "string",  
      "indexed": true,  
      "stored": true,  
      "multiValued": false  
    },  
    "add-field": {  
      "name": "Publications_SP",  
      "type": "string",  
      "indexed": true,  
      "stored": true,  
      "multiValued": false  
    }  
  }  
}
```

Legislation Scheme

```
{
  "add-field": {
    "name": "PolicyFrameworkIds",
    "type": "string",
    "indexed": true,
    "stored": true,
    "multiValued": true
  }
}
```

Create command

#Solr 6.6.0 installed as a service from default distribution.

#service solr start (if not started)

```
cd /opt/solr/bin
```

```
sudo -u solr ./solr create -c idps_databases
```

```
sudo -u solr ./solr create -c idps_measures
```

```
curl 'http://mig.tyto.it:8983/solr/idps_databases/schema' -X POST -H 'Content-type:application/json' --data-binary @idps_databases_schema.json
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/schema' -X POST -H 'Content-type:application/json' --data-binary @idps_monitoring_schema.json
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/schema' -X POST -H 'Content-type:application/json' --data-binary @idps_legislation_schema.json
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/schema' -X POST -H 'Content-type:application/json' --data-binary @idps_toxicology_schema.json
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/schema' -X POST -H 'Content-type:application/json' --data-binary @idps_ecology_schema.json
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/schema' -X POST -H 'Content-type:application/json' --data-binary @idps_structure_schema.json
```

Populate command

```
curl 'http://mig.tyto.it:8983/solr/idps_databases/update?commit=true' -H "Content-Type: text/xml" --data-binary @idps_databases.xml
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/update?commit=true' -H "Content-Type: text/csv" --data-binary @idps_monitoring.csv
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/update?commit=true&f.PolicyFrameworkIds.split=true&f.PolicyFrameworkIds.separator=%7C;' -H "Content-Type: text/csv" --data-binary @idps_legislation.csv
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/update?commit=true' -H "Content-Type: text/csv" --data-binary @idps_toxicology_0.csv
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/update?commit=true' -H "Content-Type: text/csv" --data-binary @idps_toxicology_1.csv
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/update?commit=true' -H "Content-Type: text/csv" --data-binary @idps_toxicology_2.csv
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/update?commit=true' -H "Content-Type: text/csv" --data-binary @idps_toxicology_3.csv
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/update?commit=true' -H "Content-Type: text/csv" --data-binary @idps_toxicology_4.csv
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/update?commit=true' -H "Content-Type: text/csv" --data-binary @idps_toxicology_5.csv
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/update?commit=true' -H "Content-Type: text/csv" --data-binary @idps_toxicology_6.csv
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/update?commit=true' -H "Content-Type: text/csv" --data-binary @idps_toxicology_7.csv
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/update?commit=true' -H "Content-Type: text/csv" --data-binary @idps_toxicology_8.csv
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/update?commit=true' -H "Content-Type: text/csv" --data-binary @idps_ecology.csv
```

```
curl 'http://mig.tyto.it:8983/solr/idps_measures/update?commit=true' -H "Content-Type: text/csv" --data-binary @idps_structure.csv
```


Substructure tool

Configuration

Instructions to configure PostgreSQL with bingo, create and populate the database. The bingo version to be used for the installation is the following: bingo-postgres9.2-1.7.9-linux64.zip

```
# postgresQL version
postgresQL 9.2
```

```
# access to postgresQL
sudo -u postgres psql postgres
```

```
# bingo version
bingo-postgres9.2-1.7.9-linux64
```

extract the bingo folder where you want

```
# bingo version and postgresQL version must be the same!!!
```

```
# how to install postgresQL 9.2
```

```
# ubuntu
sudo add-apt-repository "deb https://apt.postgresql.org/pub/repos/apt
YOUR_UBUNTU_VERSION_HERE-pgdg main"
wget --quiet -O - https://postgresql.org/media/keys/ACCC4CF8.asc | sudo apt-key add -
sudo apt-get update
sudo apt-get install postgresql-9.2
```

```
# debian 9.3
sudo add-apt-repository "deb http://apt.postgresql.org/pub/repos/apt/ stretch-pgdg main"
wget --quiet -O - https://postgresql.org/media/keys/ACCC4CF8.asc | sudo apt-key add -
sudo apt-get update
sudo apt-get install postgresql-9.2
```

```
# centos7
https://lintut.com/how_to_install_postgresql_on_fedora_centos_rhel/
```

```
# when postgresQL will install successfully
```

```
# create db
sudo -u postgres createdb idps
```

```
# install bingo
copy file bingo_postgres.so in bingo-postgres9.2-1.7.9-linux64/lib/ to /usr/lib/postgresql/9.0/lib/
```

```
# run bingo-pg-install.sh in bingo-postgres9.2-1.7.9-linux64/
```

Deliverable Report

```

sudo ./bingo-pg-install.sh -pglibdir -schema bingo

# give permission to folder bingo-postgres9.2-1.7.9-linux64/
sudo chmod -R 777 bingo-postgres9.2-1.7.9-linux64/

# install bingo on database idps
sudo -u postgres psql postgres -d idps -f bingo_install.sql

# remove bingo directory
sudo rm -R bingo-postgres9.2-1.7.9-linux64/

# run query
sudo -u postgres psql -f create_table.sql idps
sudo -u postgres psql -f insert.sql idps (insert.sql is created running entry.py)

```

Entry script

The following script creates the insert.sql file with all the entries for the database (to be used unless you need to update the solr data).

```
import requests
import json
import os
import sys

# sys.setdefaultencoding() does not exist, here!
reload(sys) # Reload does the trick!
sys.setdefaultencoding('UTF8')

url =
'http://idps.tyto.it:8983/solr/idps_measures/select?fl=chemicalName,smiles&group.field=chemicalName&group=true&rows=1000000000&q=smiles:*&wt=json'
insert = "INSERT INTO CHEMICALS(NAME, SMILE) VALUES('%s', '%s');"

r = requests.get(url)
data = r.json()

try:
    os.remove("insert.sql")
except OSError:
    pass

file = open("insert.sql", "w+")

for groups in data["grouped"]["chemicalName"]["groups"]:
    for doc in groups["doclist"]["docs"]:
        name = doc["chemicalName"].replace("'", "")
        smile = doc["smiles"]
        file.write(insert % (name, smile) + "\n")

file.close()
```

Table creation

File to be used to create tables

```
DROP TABLE IF EXISTS CHEMICALS;
```

```
CREATE TABLE CHEMICALS (  
    ID SERIAL PRIMARY KEY,  
    NAME varchar(255),  
    SMILE varchar(512)  
);
```

Entry file

The file with all the entries for the db (updated to the last import made in solr)