



Dr.  
V. Dudarev

# RESEARCH DATA MANAGEMENT SYSTEM IN MATERIALS SCIENCE COLLABORATIVE RESEARCH CENTER 1625

Victor Dudarev



Christian-Albrechts-Universität zu Kiel



MAX-PLANCK-INSTITUT FÜR  
CHEMISCHE ENERGIEKONVERSION



FRITZ-HABER-INSTITUT  
MAX-PLANCK-GESELLSCHAFT



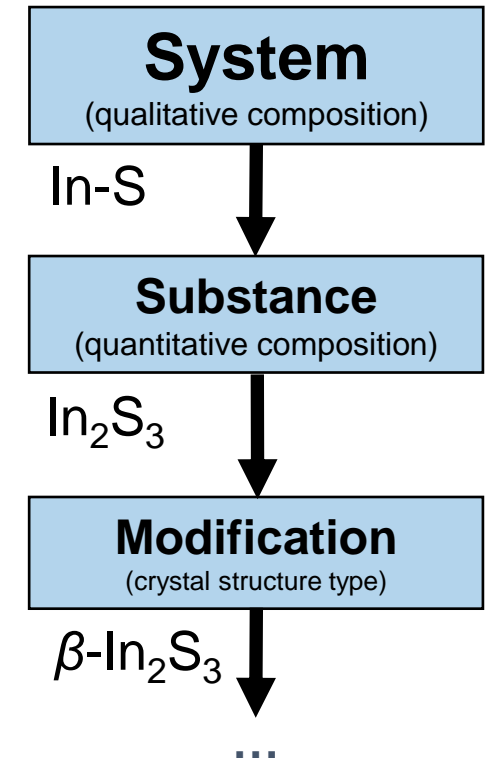
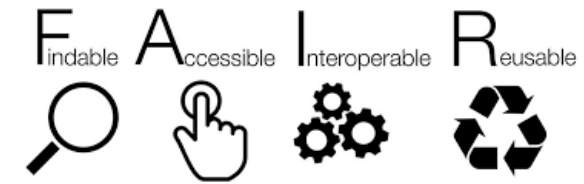
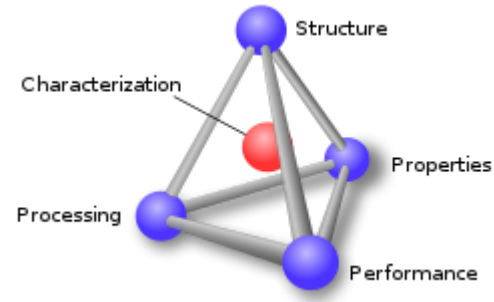
Deutsche  
Forschungsgemeinschaft  
German Research Foundation

# Agenda

- RDMS Core Feature at a glance
- Some New Features:
  - Data Transfer between Tenants
  - Object Links for Projects
  - Advanced Objects Interlinking (graph)
- Sample Type and Lifecycle Workflow (states & splitting)
- Handover Event (Sample Tracking)
- Platform Customisation:
  - User Defined Object Types & Templates: no-code approach
  - External Web Services: Validation & Data Extraction API
  - External Visualisation (plotting outsource)
- Reports & API

# RDMS Core Features

- Support for multiple **Tenants**;
- Flexible **Project Tree**;
- **User-defined Object Type** support:
  - **Properties Templates**
  - **Table Templates**
  - Document **Validation**
  - Data **Extraction**
  - Data **Visualization**
- **Batch Import** of Documents
- **API**
- **Rich Customizations** (with HTML/JS/CSS injection & external Web Services)



**Priorities: Security, Scalability, Flexibility**

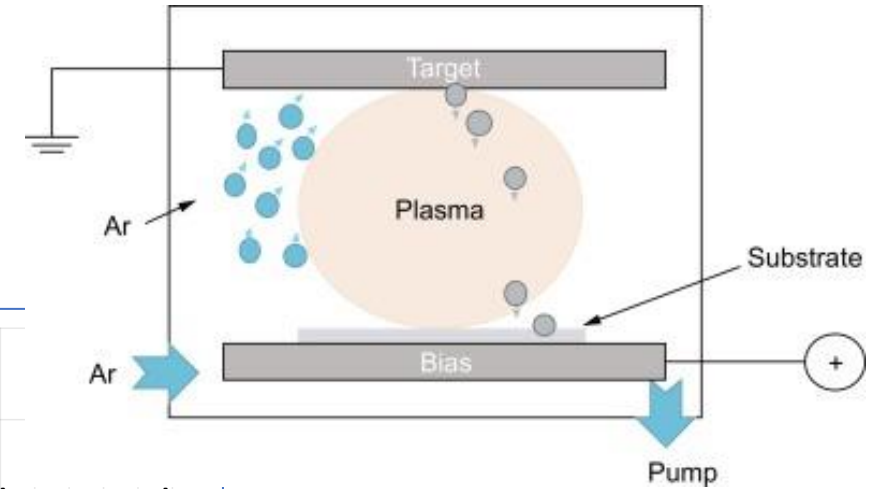
# RDMS in Production: mdi.matinf.pro (since 13.11.2023)

## Objects statistics (top 10+ types):

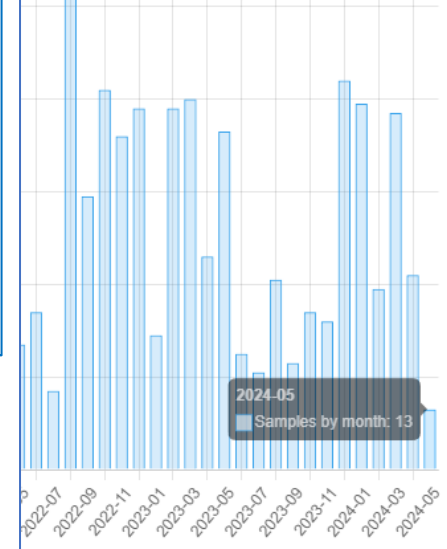
Object Count	TypeId	Type Name	Type Comment
28937	!= 0	All objects	All types (TypeId != 0)
14648	8	Composition	Composition information (compound)
4217	6	Sample	Sample information (sputter deposition)
3287	18	Synthesis	Synthesis Document (parameters from Compact)
1687	7	Raw Document	Data file on disk in any format (file extension required to identify MIME-type)
1119	12	Photo	Photo (bitmap in jpg, png, etc...)
816	13	EDX CSV	CSV of EDX with Header row (e.g. "Index,V,Mn,Co,Ni,Ho")
766	24	SEM (image)	from Compact
684	26	TEM image	from Compact
291	15	EDX Image	Image (bitmap) that refers to EDX
215	33	HTTS Resistance TXT	TXT of HTTS Resistance. Via HTTPs by Azadeh
202	40	Thickness Other (zip, opj, opju, pdf)	Thickness Other files (from Compact)

## Projects and object interlinks statistics:

969	Project	<b>Projects</b>	<b>All projects</b>
42028	Link	<b>Object Links</b>	<b>All links between objects</b>



Samples by month  
4200+ samples in total



### Typical Materials Library Workflow:

#### Sputtering:

- Add **Sample** object (chemical system + substrate)
- Add **Synthesis** object (sputtering parameters)

#### Characterisation:

- Add processed **EDX** (342 MA)
- Add Other Characterisation Documents (**Photo, XRD, Resistance, Thickness, Bandgap**, etc...)

Storage, Curation, **Search** (on properties retrieved from devices / file formats), Reporting

# RDMS: Data transfer between tenants (MDI -> CRC 1625)

MDI Database Search Tree Edit List Edit Handover Reports find SampleID Hello vic.dudarev@gmail.com! Logout

Industry Nonindustry / CRC 1625

CRC 1625

Please, fill here the data for **CRC 1625** according to the [project structure from the root](#). All data are to be transferred to [crc1625.mdi.ruhr-uni-bochum.de](http://crc1625.mdi.ruhr-uni-bochum.de) tenant.

Take a look at subprojects:

- Area A

Add Subproject Hide Non-samples

Add Sample

Area A

DFG

ERC DEMI

Partner projects

Werkstoffpraktikum

Data from **MDI tenant**  
(Nonindustry / CRC 1625 project)  
transferred to  
**CRC 1625 tenant**

Automatic Data Transfer  
(on a daily basis)

CRC 1625 Database Search Tree Edit List Edit Handover Reports find SampleID Hello vic.dudarev@gmail.com! Logout

Area A Area A / A01

A01

Please, be aware that in this data in this project are updated from [mdi.matinf.pro](#) tenant.

Take a look at subprojects:

- Ag-Au-Pd-Pt-X
- Au-Pd-Pt-Rh-X
- Preliminary work for proposal and evaluation Ag-Au-Cu-Pd-Pt
- Preliminary work for proposal and evaluation Ir-Pd-Pt-Rh-Ru
- Preliminary work for proposal and evaluation - other

Add Subproject Hide Non-samples

Add Sample

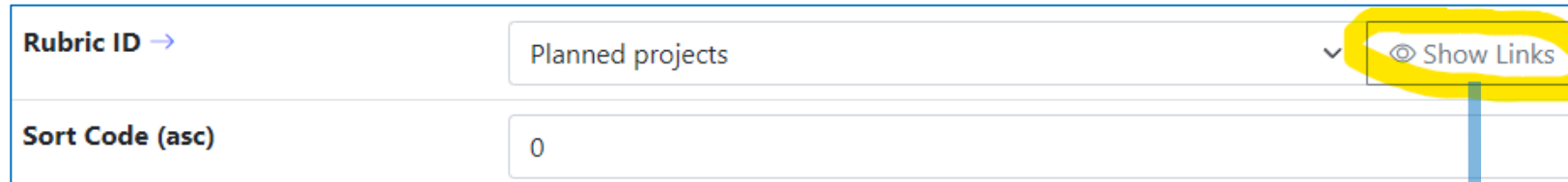
Description of partial project from proposal Protected Publication

**Tenant** – separate (distinct) instance of a software application that is used by an organization / group of users / consortia.

# Object Links for Projects: object modification form

**Motivation:** Object is displayed in a project to which it belongs. Sometimes it's required to associate object with multiple projects.

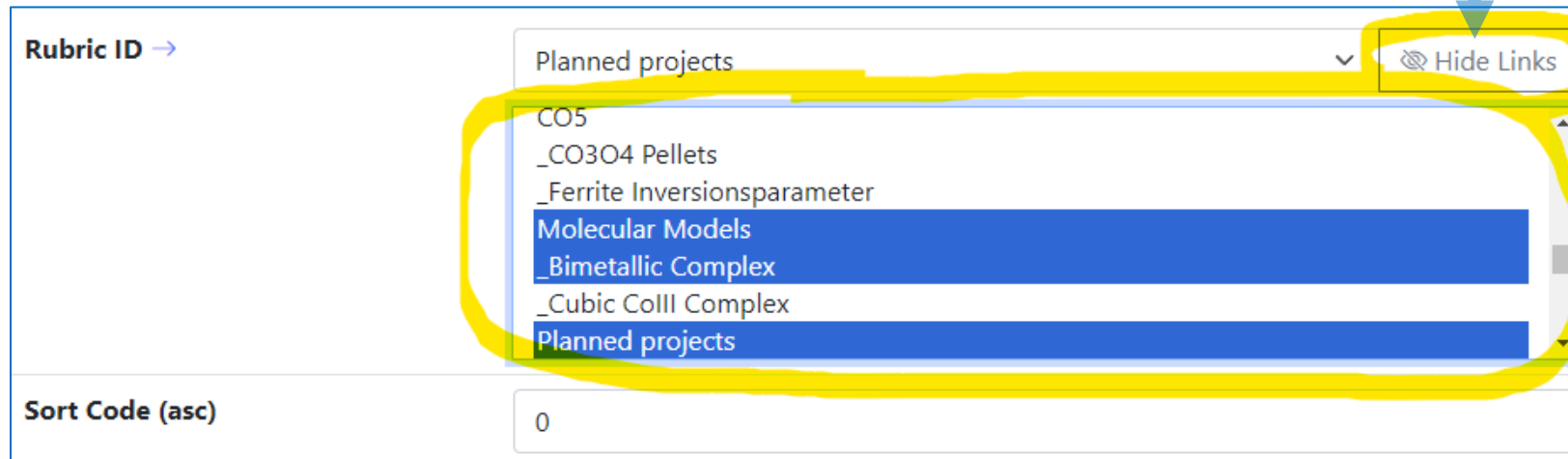
## Object modification interface update



Rubric ID → Planned projects ✓ **Show Links**

Sort Code (asc) 0

Select projects that should additionally display the object:

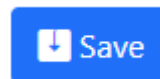


Rubric ID → Planned projects ✓ **Hide Links**

- CO5
- \_CO3O4 Pellets
- \_Ferrite Inversionsparameter
- Molecular Models**
- \_Bimetallic Complex**
- \_Cubic Coll Complex**
- Planned projects**

Sort Code (asc) 0

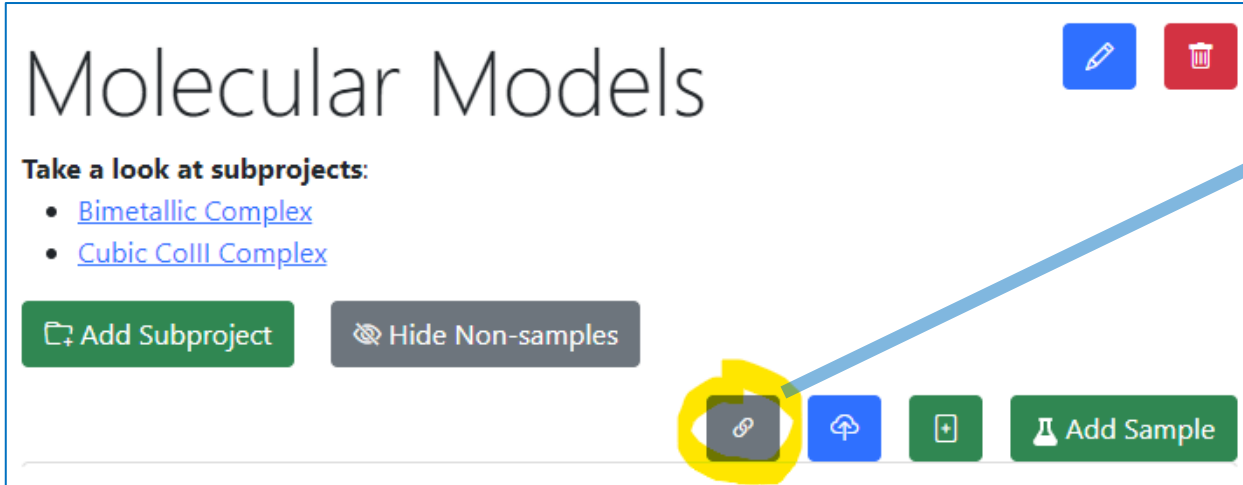
Press "Save" button



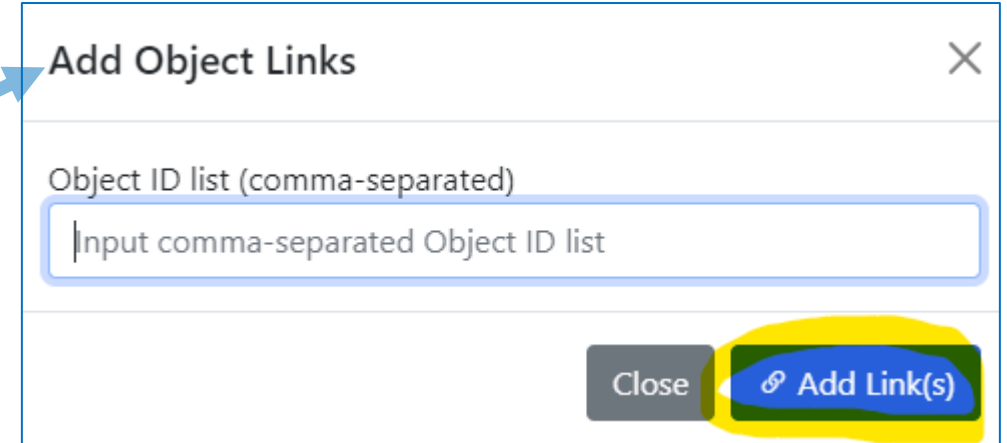
# Use-Case: add object links to a project

**Motivation:** allow everybody (PowerUser+) to add object links to a project.

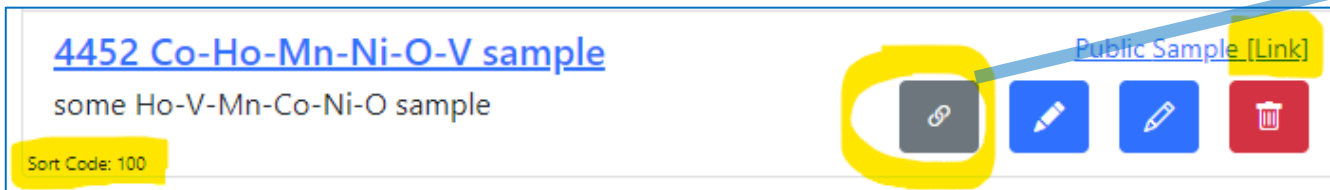
1) In a project click gray “Add Link” button



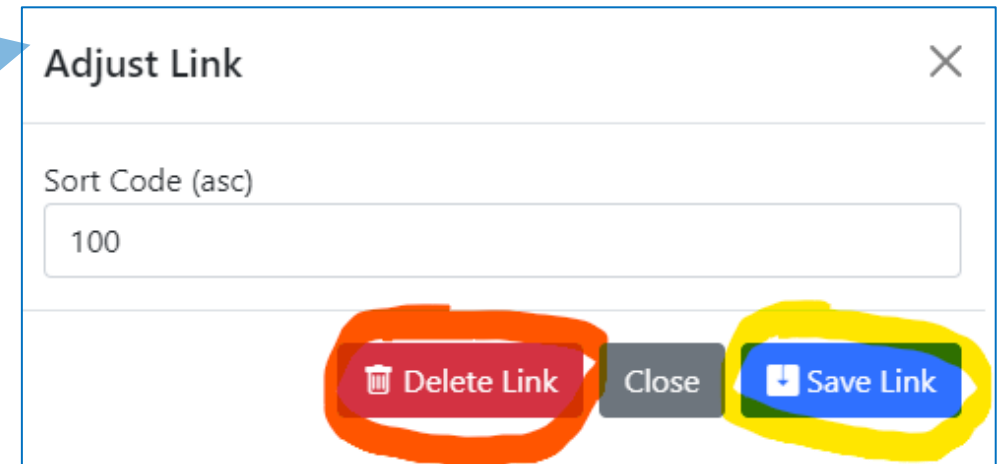
2) Popup window: add links for object(s)



3) Linked object layout in the project (with link button):

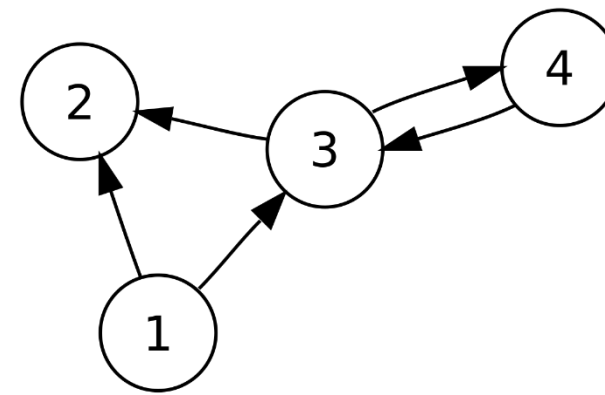


4) Popup: adjust link order or delete it





# RDMS: Advanced Interlink Objects



## Example: establishing graph above objects

Associated Objects

Sample [1] Composition [1]

[Measurement Area 003 from 0008081\\_EDX\\_ZGH\\_Ho L default\\_Ok](#) [Public Composition](#)

Automatically created via import from file (MA 003)

[Manual](#)

EDX Image [1]

Associated Objects

- 4421 Co3O4\_anisotropic [Sample] [Manual](#)
- Measurement Area 003 from 0008081\_EDX\_ZGH\_Ho L default\_Ok [Composition] [Manual](#)
- Probe 1\_SpZn\_x10k\_14\_verbessert [EDX Image] [Auto](#)

[Save](#)

Drag&Drop objects here

Search EDX Image objects: Search

Link type Auto

[Probe 1\\_SpZn\\_x10k\\_14\\_verbessert \[EDX Image\]](#)

## To add associated objects:

- 0) Go to “List edit” or select object to edit
- 1) Select associated object type (optional filter)
- 2) Input search phrase, contained in the Name of the desired object to be associated
- 3) Select (if required) **link/association type** object.
- 4) Drag & Drop desired object(s) from search result list to the area / adjust list
- 5) Save changes

**Important:** Reverse associations allow to browse reverse-directional associations.



# Sample Customised Form

 - Customised sample input

## Creating a new sample

**Substrate\***

**Type\***

**Name\***

**Description**

**Chemical System\***

**Access Control (accessibility)\***

List from Service / Substrates subproject

-- select the substrate --

-- select the substrate --

- Alumina
- Fused Silica
- other
- Sapphire**
- Si
- Si + SiO2

List from JS-template

Materials Library (342-grid)

unknown

**Materials Library (342-grid)**

- Stripe
- No Gradient
- Stress Chip
- Piece

Standard chemical system

Select elements

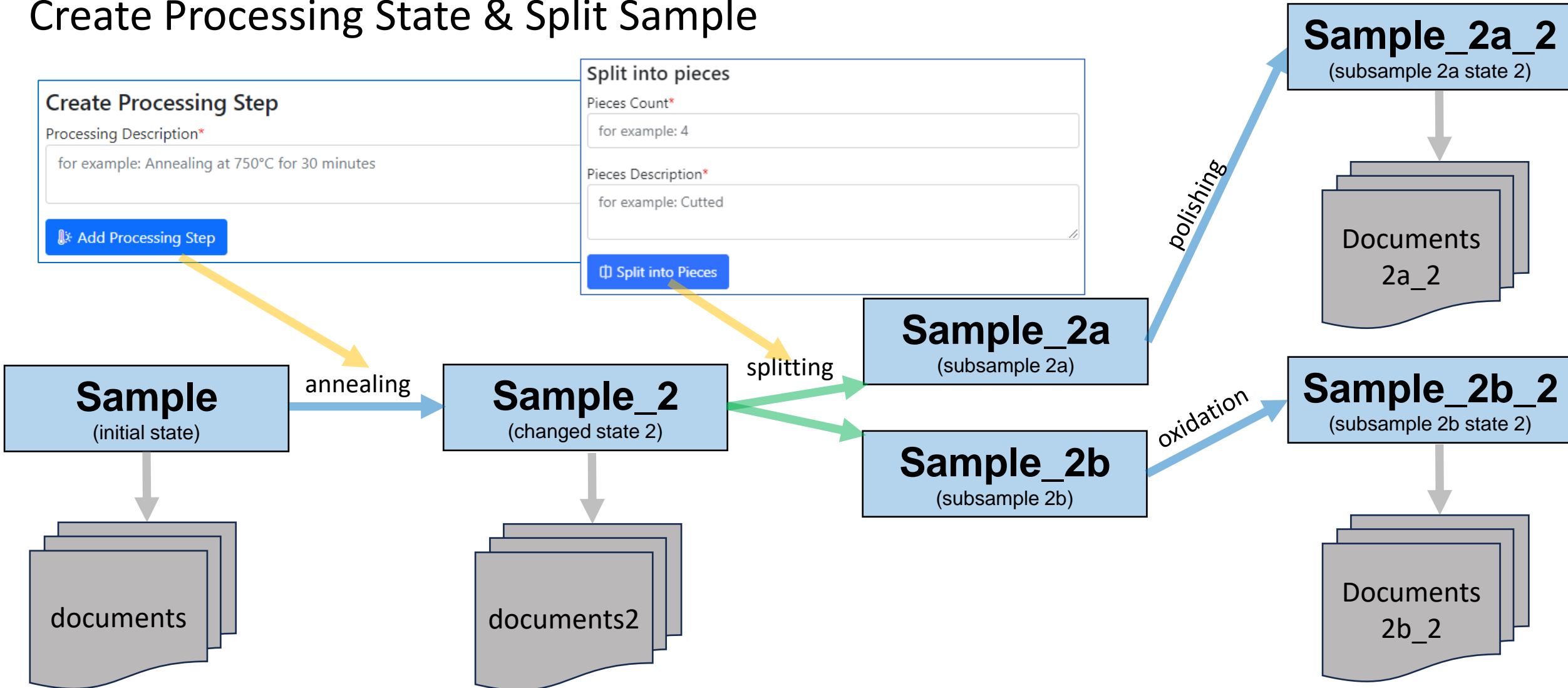
Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	H																	He
2	Li	Be											B	C	N	O	F	Ne
3	Na	Mg											Al	Si	P	S	Cl	Ar
4	K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
5	Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
6	Cs	Ba	*Lu	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
7	Fr	Ra	**Lr	Rf	Db	Sg	Bh	Hn	Mt	Ds	Rg	Cn	Nh	Fl	Mc	Lv	Ts	Og
			*La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb		
			**Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No		

Chosen elements: Al-Ni

## Customisation with UI form / JS

# Use Case: Sample Transformation Workflow (Sample edit form)

## Create Processing State & Split Sample



# Tracking Samples or Handover event: initialisation

1) In “physically transferrable” objects (samples), find a green handover button

## Handover

No handover events found.



2) Select the recipient and press a blue button.

The form is titled "Handover event" and contains the following sections:

- Recipient (destination)**: A dropdown menu with the placeholder text "-- select recipient --". An arrow points to this field with the label "Select recipient".
- Amount**: Two input fields labeled "Amount" and "Measurement Unit". An arrow points to the "Amount" field with the label "Specify Amount (optional)".
- Comments**: A text area with the placeholder text "Comments for the recipient". An arrow points to this area with the label "Specify Comments".
- Buttons**: At the bottom, there are two buttons: a grey "Close" button and a blue "Handover" button with a hand icon. An arrow points to the "Handover" button with the label "Click to initialise".

A list of recipients for the handover event, with "A01) Akhil Hareendran [akhil.hareendran@rub.de]" selected. The list includes:

- select recipient --
- Carina Marek [carina.marek@rub.de]
- furong.yan@uni-due.de [furong.yan@uni-due.de]
- A01) Akhil Hareendran [akhil.hareendran@rub.de]**
- A01) Catalina Leiva Leroy [catalina.leivaleroy@rub.de]
- A02) Adarsh Koul [adarsh.koul@ruhr-uni-bochum.de]
- A02) Alejandro Esteban Perez Mendoza [alejandro.perez-mendoza@uni-due.de]
- A02) Raphael Otto [raphael.otto@uni-due.de]
- A02) Shivam Shukla [shivamandshukla160999@gmail.com]
- A02) Shubhadeep Chandra [Shubhadeep.Chandra@ruhr-uni-bochum.de]
- A02) Thomas Quast [Thomas.quast@rub.de]
- A04) Arno Bergmann [abergmann@fhi-berlin.mpg.de]
- A04) Timon Wagner [wagner@fhi-berlin.mpg.de]
- A05) Niklas Sülzner [niklas.suelzner@rub.de]
- A06) Amir Hossein Omranpoor [ahomranpoor@gmail.com]
- A06) Eckhard Spohr [eckhard.spohr@uni-due.de]
- A06) Falonne Bertholde Sharone Nkou [falonnenkou@gmail.com]
- A07) Hauke Scheele [hscheele@ac.uni-kiel.de]
- A07) Jan Smyczek [jansmyczek.work@gmail.com]
- A07) Jessica Wulfes [Wulfes@pctc.uni-kiel.de]

# Tracking Samples or Handover event: delivering / delivered


1) Handover event is created

Handover						
Sender	Sent	Amount	Sender Comments	Recipient	Received	Recipient Comments
<a href="#">Victor Dudarev</a> <a href="#">INF</a>	2/13/2024 6:53:05 PM		Please measure something	<a href="#">Victor Dudarev</a> <a href="#">INF</a>		Comments for sender (required)  <input checked="" type="button" value="Confirm handover"/>

3) Recipient should add comments and confirm the handover.

2) Recipient is notified by e-mail

Handover event registered: 8389 sample handover to Victor Dudarev [vic\_dudarev@mail.ru] on 2/13/2024 6:53:05 PM

 MDI robot

Dear Victor Dudarev,

[Victor Dudarev](#) [INF](#) registered a handover event for you: 8389 sample handover to Victor Dudarev [vic\_dudarev@mail.ru] on 2/13/2024 6:53:05 PM  
Comment: *Please measure something.*  
When you receive the sample, please [confirm this](#) in the system by specifying your comments for sender!

[crc247.mdi.ruhr-uni-bochum.de](mailto:crc247.mdi.ruhr-uni-bochum.de) automatic notification service

4) Confirmed

Handover						
Sender	Sent	Amount	Sender Comments	Recipient	Received	Recipient Comments
<a href="#">Victor Dudarev</a> <a href="#">INF</a>	2/13/2024 6:53:05 PM		Please measure something	<a href="#">Victor Dudarev</a> <a href="#">INF</a>	2/13/2024 7:12:36 PM	Got it!

+ e-mail notification for sender

Creation by sender -> Recipient notification -> Recipient confirmation -> Sender notification.

# Tracking Samples or Handover event: user overview

CRC/TRR247 Database Search Tree Edit List Edit **Handover** Reports find ObjectId Hello vic.dudarev@gmail.com! Logout

## Incoming Handovers for you

Name	Amount	Sender	Sent	Sender Comments	Received	Your Comments
<a href="#">8220 sample 2 g handover to Victor Dudarev [vic.dudarev@gmail.com] on 2/5/2024 10:03:28 PM</a>	2 g	<a href="#">Victor Dudarev (INF)</a>	2/5/2024 10:03:28 PM	This is first 2 grams	2/5/2024 10:07:03 PM	Thank you!
<a href="#">8220 sample handover to Victor Dudarev [vic.dudarev@gmail.com] on 12/20/2023 8:23:15 PM</a>		<a href="#">Victor Dudarev (INF)</a>	12/20/2023 8:23:15 PM	Take it	12/20/2023 8:31:21 PM	I've got it!

## Outgoing Handovers from you

Name	Amount	Sent	Your Comments	Recipient	Received	Recipient Comments
<a href="#">8389 sample handover to Victor Dudarev [vic_dudarev@mail.ru] on 2/13/2024 6:53:05 PM</a>		2/13/2024 6:53:05 PM	Please measure something	<a href="#">Victor Dudarev (INF)</a>	2/13/2024 7:12:36 PM	Got it!
<a href="#">8220 sample 2 g handover to Victor Dudarev [vic.dudarev@gmail.com] on 2/5/2024 10:03:28 PM</a>	2 g	2/5/2024 10:03:28 PM	This is first 2 grams	<a href="#">Victor Dudarev (INF)</a>	2/5/2024 10:07:03 PM	Thank you!
<a href="#">8220 sample handover to Victor Dudarev [vic.dudarev@gmail.com] on 12/20/2023 8:23:15 PM</a>		12/20/2023 8:23:15 PM	Take it	<a href="#">Victor Dudarev (INF)</a>	12/20/2023 8:31:21 PM	I've got it!

To add a handover, please go to the sample page and create a handover associated with the sample.

Incoming

Outcoming

# Tracking Samples or Handover event: project overview

Area C / C03

## C03

[Show Handover Report for Project](#)

### Incoming Handovers for project C03

No incoming handovers.

### Outgoing Handovers from project C03

Name	Amount	Sender	Sent	Comments	Recipient	Received	Recipient Comments
<a href="#">8387 sample 200 mg handover to Furong Yan [furong.yan2023@gmail.com] on 2/7/2024 1:30:26 PM</a>	200 mg	<a href="#">Carsten Placke-Yan (C03)</a>	2/7/2024 1:30:26 PM	Handed over in Person	<a href="#">Furong Yan (B03)</a>	Comments for sender (required) <input type="checkbox"/> Confirm handover	
<a href="#">8386 sample 120 mg handover to Furong Yan [furong.yan2023@gmail.com] on 2/7/2024 1:29:53 PM</a>	120 mg	<a href="#">Carsten Placke-Yan (C03)</a>	2/7/2024 1:29:53 PM	Handed over in Person	<a href="#">Furong Yan (B03)</a>	Comments for sender (required) <input type="checkbox"/> Confirm handover	
<a href="#">cpy126Cr01_calc sample handover to Dana Krenz [dana.krenz@uni-due.de] on 1/31/2024 9:54:06 AM</a>		<a href="#">Carsten Placke-Yan (C03)</a>	1/31/2024 9:54:07 AM	ca. 150mg for Lasertreatment (do we see dispersion of Cr in the bulk?)	<a href="#">Dana Krenz (C05)</a>	2/7/2024 10:19:52 AM got it thanks	

To add a handover, please go to the sample page and create a handover associated with the sample.

Handover overview is displayed in project that matches by name with one of **Project** claims for users.

Contains overview for a group of users having the same **Project** claim.

Pending handovers (unconfirmed)

Successful handover

# Ways of customization



- **Basic** (no-code)

Introducing User-Defined Data Types:

- List templates
- Table templates

- **Simple** (client-side code only)

Page-code injection:

- HTML
- Javascript
- CSS

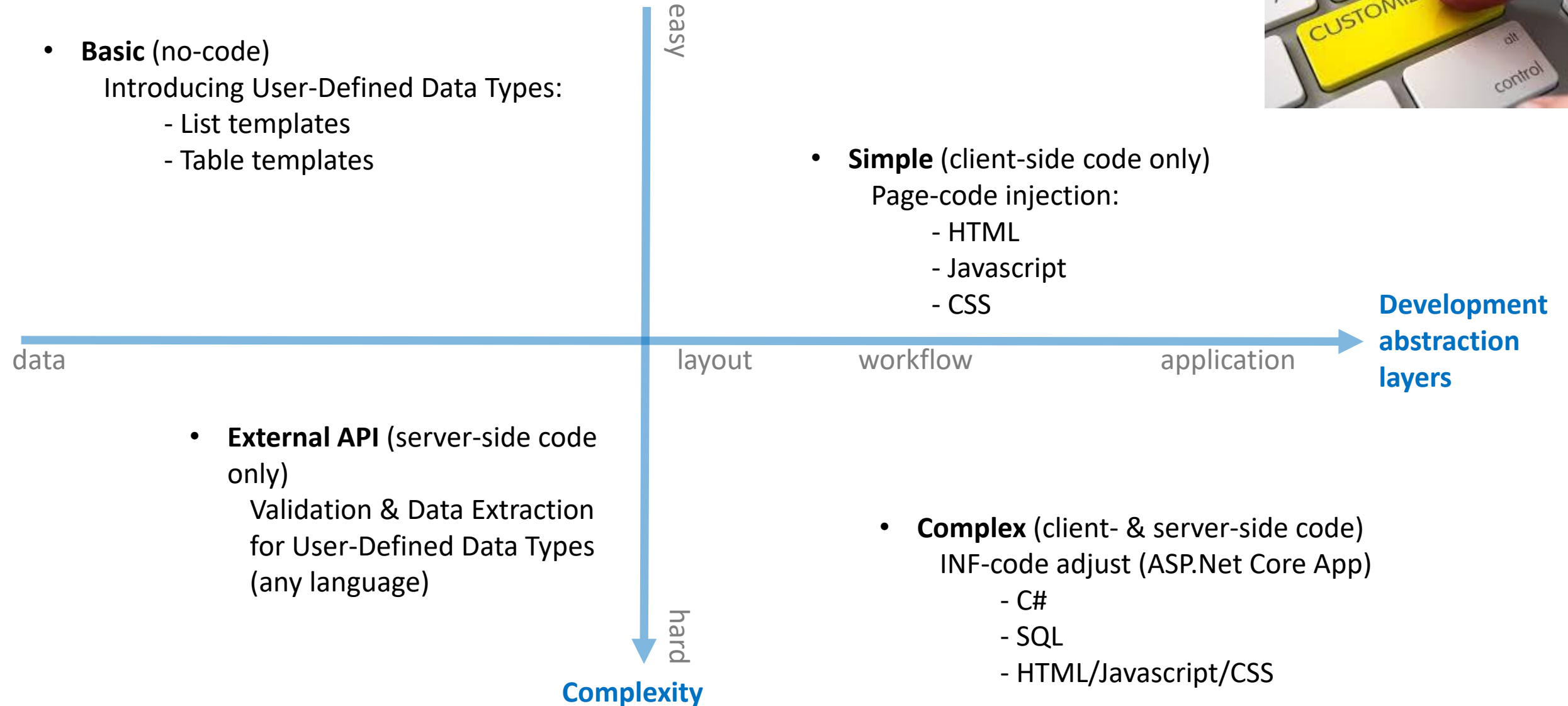
- **External API** (server-side code only)

Validation & Data Extraction for User-Defined Data Types (any language)

- **Complex** (client- & server-side code)

INF-code adjust (ASP.Net Core App)

- C#
- SQL
- HTML/Javascript/CSS






# User-Defined Type Support

**Motivation:** to support new user-defined object type (UDT) **without modifying core RDMS source code**

**Given:** New object type with/without a new arbitrary file format (could new measurement type result).

**Solution** (or 5 keys to success):

- **Templates** – define object model (data structure) in terms of key-value pairs / table (properties)
- **Search** – enable search for UDT objects based on defined data model 
- **Validation** – provide check for data correctness (valid/invalid)
- **Data Extraction** – if valid, extract data in terms of properties / related object creation
- **Visualization** – display data, including drawing charts / diagrams, etc...



# UI to create a new user-defined type

## Creating a new type

**Type Name**

**Table Name (Data Structure)**

**Hierarchical classifier**

**Validation Schema**

**Data Schema**

**File Required**

**Settings (JSON)**

**Description**

**Type Name** – user-defined type name

### Core Object Types:

- **ObjectInfo** – the simplest (bare) object
- **Sample** – chemical system (set of chemical elements), for example **Co-O**
- **Composition** – chemical composition (set of elements and quantities), for example **Co<sub>3</sub>O<sub>4</sub>**
- **Reference** – literature reference (Authors, Title, Year, DOI, etc.)

**Validation & Data Schema** – only if you have external Web Services ready to validate & extract data from files

**File Required** – sets the file mandatory (boolean)

**Settings (JSON)** – type customization in RDMS

```
{ "CustomEditPath": "/custom/editsample",
  "UrlPostVisualizer": "https://abc.de/VisualizeRT" }
```

**Description** – comments for type designation

Defining data & behavior

# RDMS no-code customisation: Templates for object type

## Template

### List

(specify sorted properties list with separators)

- Create an object with “\_Template” name
- Specify properties set for a type

### Table

(specify table structure: columns & types)

Type	Name	Value	Comment
▼ general process parameters			
Float	start vacuum	1.8e-007	Torr
String	process duration	55:36	s
Float	substrate temperature	527	°C
Float	substrate rotation	0	RPM
String	process gas		
String	reactive gas 1		
Float	flow RG 1		
Float	table height		
Float	process pressure		
Float	flow PG		
▼ target source configuration			
▼ Cathode 1			
▼ Mag1-RF			
Float	ignition power cathode 1		
Float	pre clean power 1		
Float	pre clean time 1		
String	pre clean pressure cathode 1		

Key-value pairs

Type	Name	Value	Comment
▶ chamber configuration			
▼ general process parameters			
Float	start vacuum	1.8e-007	Torr
String	process duration	55:36	s
Float	process duration		s
Float	substrate temperature	527	°C
Float	substrate rotation	0	RPM
String	process gas	Ar	
String	reactive gas 1	O2	
Float	flow RG 1	40	scm
Float	flow RG 2		scm
String	reactive gas 2		
Float	table height	25	mm
Float	substrate position		*
Float	process pressure	3	mTorr

Modification	CrystalSystem	SpaceGroup	Temperature	E <sub>g</sub>	IsCalculated	ReferenceId	Comment
α	Monoclinic	P2 <sub>1</sub> /c	300	0.87	0	581	Optical absorption spectroscopy. Direct transition. Thin film, single crystal
β	Cubic	Im3(-)m	296	1.03	0	581	Photoconduction. Thin film, polycrystalline
β	Cubic	Im3(-)m	77	1.23	0	581	Photoconduction. Thin film, polycrystalline

Data in a table with predefined structure



# RDMS (no-code effort): Extended Properties Search

## All values are to be stored using appropriate data types!

**DEFINITION:** Data type is a collection or grouping of data values, usually specified by a **set of possible values**, a **set of allowed operations** on these values, and/or a representation of these values as machine types.

Supported data types (and their visualization in the search form):

- **Int** – integer values (64-bit signed integer) in range [-9 223 372 036 854 775 808; 9 223 372 036 854 775 807].

Property type:  Name:  Property Value:  ≤ value ≤

- **Float** – double-precision floating-point (8 bytes, IEEE-754),  $\pm 5.0 \times 10^{-324}$  to  $\pm 1.7 \times 10^{308}$ .

Property type:  Name:  Property Value:  ≤ value ≤

- **String** – any characters sequence no longer than 4096 symbols.
- **BigString** – any characters sequence no longer than 2 147 483 647 symbols (2 Gb).

Property type:  Name:  Property Value:

- equal
- starts with**
- ends with
- contains

# Materials Library: integrating heterogeneous measurement data

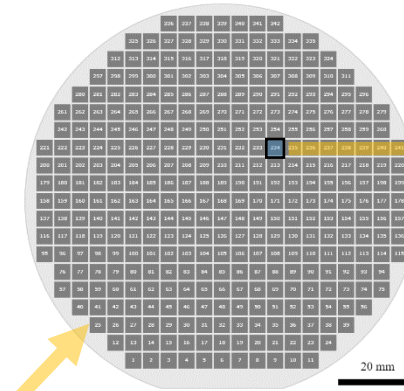
## Typical Materials Library Data Lifecycle:

### Sputtering:

- Add **Sample** object (chemical system + substrate)
- Add **Synthesis** object (sputtering parameters)

### Characterisation:

- Add processed **EDX** (342 Composition objects)  
Composition == Measurement Area
- Add Other Characterisation Documents (**Photo**, **XRD**, **Resistance**, **Thickness**, **Bandgap**, etc...)



### Measurement Area (composition)

- Phases: [ Phase<sub>1</sub>,  
...,  
Phase<sub>N</sub> ]
- Resistance: \_\_ Ohm
- Thickness: \_\_ nm
- Bandgap: \_\_ eV
- ...

### Storage, Curation, **Search** (on properties retrieved from devices / file formats), Reporting:

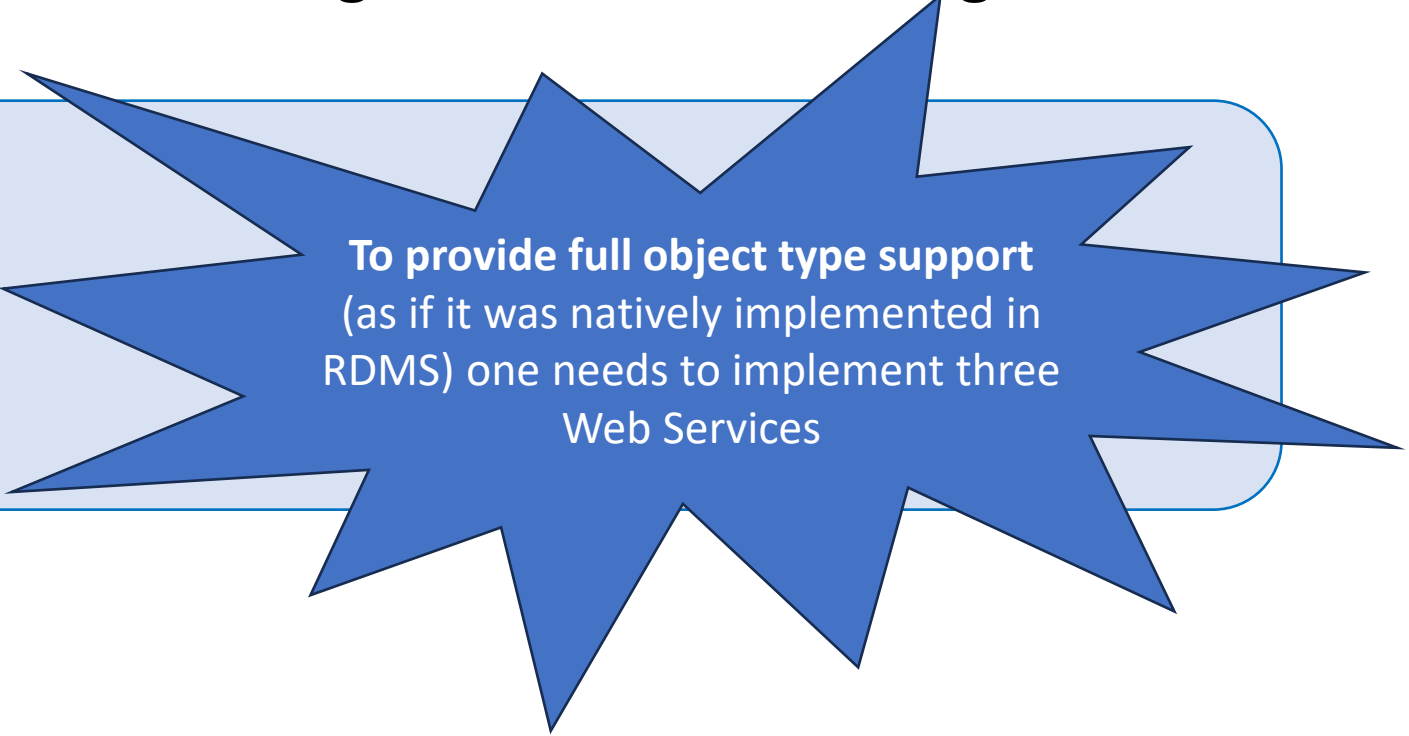
- Find compositions with given properties in a certain range (across all known materials libraries)
- Find by synthesis parameters
- Find by chemical system / creation date / author / text / properties, etc...
- Reports

# Customisation via External Services

**Motivation:** support new data formats (models) without RDMS code change

**Solution:** RDMS can call external services to get assistance dealing with new data formats with respect to:

- **Validation**
- **Data Extraction**
- **Visualisation**



To provide full object type support (as if it was natively implemented in RDMS) one needs to implement three Web Services

# Type configuration: Data Validation types

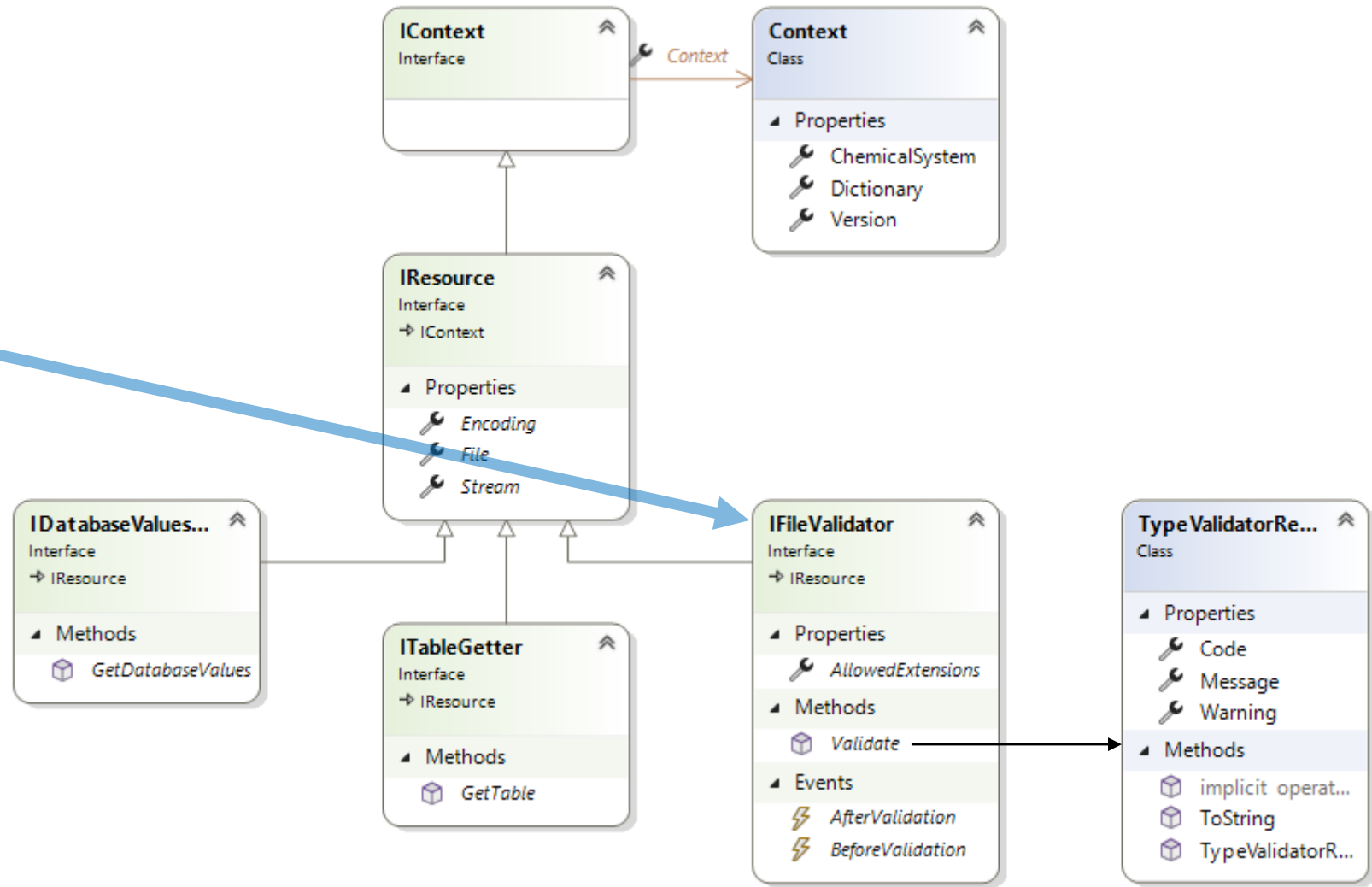
Provide a validator for every object type. Implemented:

- **type:**

any .Net-compatible type, that implements **IFileValidator** interface and is available in Application Domain (object is instantiated by type name using Reflection);

- **https:**

URL to a REST Service (OpenAPI specification).





# Type configuration: Data Validation via external Web Service

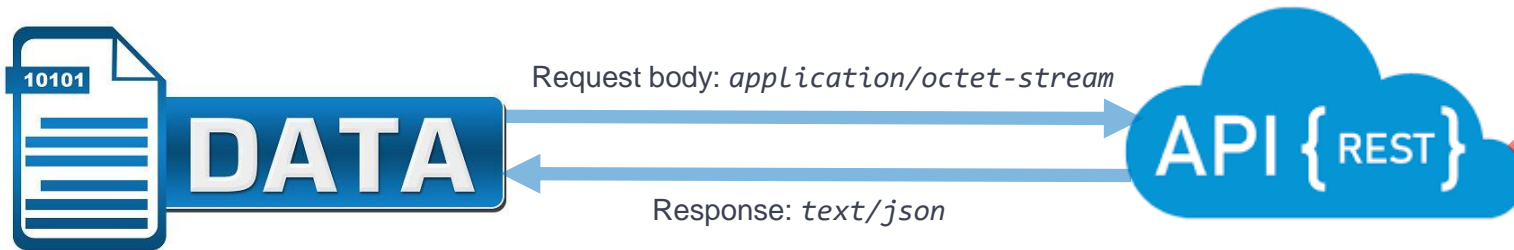
Validation task is delegated to an external service

**Request URL:** `HTTPS_ValidationSchema + "/body"`

**Example:**

Validation Schema (for EDX): <https://validation.matinf.pro/edx/validation>

Request URL: <https://validation.matinf.pro/edx/validation/body>



Validation result  
via external API call

```
{  
  "Code": 0,  
  "Message": null,  
  "Warning": null  
}
```

External service decision

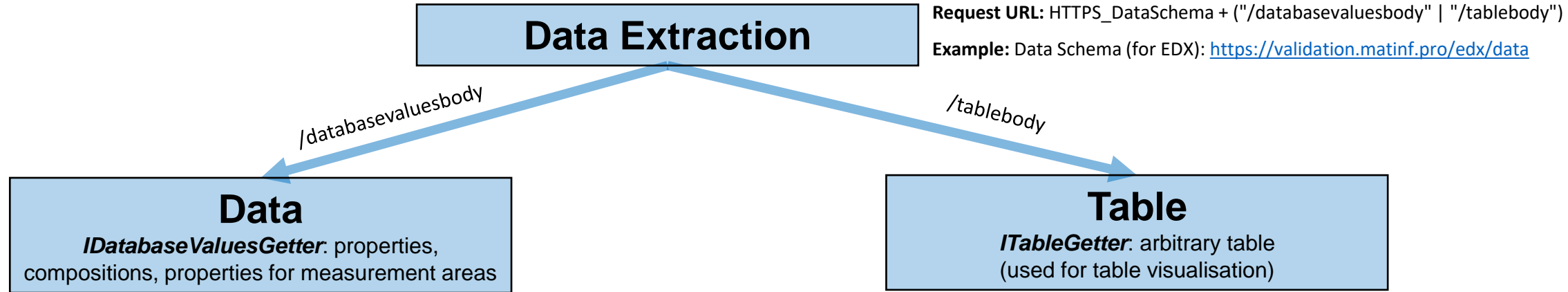
```
{  
  "Code": 500,  
  "Message": "error text",  
  "Warning": null  
}
```

Validation schema values (current status):

Validation Schema*	Comments	Author
<code>type:TypeValidationLibrary.TypeValidator_Ok</code>	default validator: checks extensions only	Victor
<code>type:TypeValidationLibrary.TypeValidator_EDX_CSV</code>	EDX (Victor)	Victor
<a href="https://validation.matinf.pro/edx/validation">https://validation.matinf.pro/edx/validation</a>	EDX (Victor)	Victor
<a href="https://htts.matinf.pro/resistance/csv/validation">https://htts.matinf.pro/resistance/csv/validation</a>	HTTS Resistance CSV	Azadeh
<a href="https://htts.matinf.pro/resistance/txt/validation">https://htts.matinf.pro/resistance/txt/validation</a>	HTTS Resistance TXT	Azadeh
<a href="https://thickness.matinf.pro/thickness/xlsx/validation">https://thickness.matinf.pro/thickness/xlsx/validation</a>	Thickness XLSX	Azadeh
<a href="https://thickness.matinf.pro/thickness/txt/validation">https://thickness.matinf.pro/thickness/txt/validation</a>	Thickness TXT	Azadeh

# Type configuration: Data Extraction (via external Web Service)

Data extraction & conversion to table tasks are delegated to an external service



**Request URL:** HTTPS\_DataSchema + ("/databasevaluesbody" | "/tablebody")

**Example:** Data Schema (for EDX): <https://validation.matinf.pro/edx/data>

Request URL (data): <https://validation.matinf.pro/edx/data/databasevaluesbody>

Request URL (table): <https://validation.matinf.pro/edx/data/table>

see next slides...

```

{
  "DataTable": [
    {
      "Index": "1",
      "V": 31.3,
      "Mn": 2.7,
      "Co": 22.6,
      "Ni": 6,
      "Ho": 37.4
    },
    ...
  ],
  "Coordinates": {
    "CoordinateNames": [],
    "HasCoordinates": false
  },
  "Dependencies": {
    "DependencyNames": [
      "Index"
    ],
    "HasDependencies": true
  }
}

```

Data schema values (current status):

Data Schema	Comments	Author
type:TypeValidationLibrary.TypeValidator_EDX_CSV	EDX (Victor)	Victor
<a href="https://validation.matinf.pro/edx/data">https://validation.matinf.pro/edx/data</a>	EDX (Victor)	Victor
<a href="https://htts.matinf.pro/resistance/csv/data">https://htts.matinf.pro/resistance/csv/data</a>	HTTS Resistance CSV	Azadeh
<a href="https://htts.matinf.pro/resistance/txt/data">https://htts.matinf.pro/resistance/txt/data</a>	HTTS Resistance TXT	Azadeh
<a href="https://thickness.matinf.pro/thickness/xlsx/data">https://thickness.matinf.pro/thickness/xlsx/data</a>	Thickness XLSX	Azadeh
<a href="https://thickness.matinf.pro/thickness/txt/data">https://thickness.matinf.pro/thickness/txt/data</a>	Thickness TXT	Azadeh

<https://validation.matinf.pro/>

# Data Extraction: creating properties

**Task:** create properties for object that contains file

```
{ "DeletePreviousProperties": true,
  "Properties": [
    { "PropertyId": 0,
      "Type": 1,
      "Name": "V",
      "Value": 0,
      "ValueEpsilon": null,
      "SortCode": 10,
      "Row": 1,
      "Comment": "Minimal V content"
    },
    { "PropertyId": 0,
      "Type": 1,
      "Name": "V",
      "Value": 36.400001525878906,
      "ValueEpsilon": null,
      "SortCode": 10,
      "Row": 2,
      "Comment": "Maximal V content"
    }
  ]
}
```

Metadata for EDX (range values):

V	Mn	Co	Ni	Ho
0	0	0	0	20.8
36.4	16.1	46.5	19.1	100

Administrative interface:

0008081 EDX ZGH Ho L default.csv  remove file  
3B4B03C48001257C0609D88D978A862636E15633019D5C7C1D5F4D94BB373125

Upload replacement:  
 No file chosen

Compositions

# Data Extraction: creating compositions

**Task:** create compositions for every measurement area

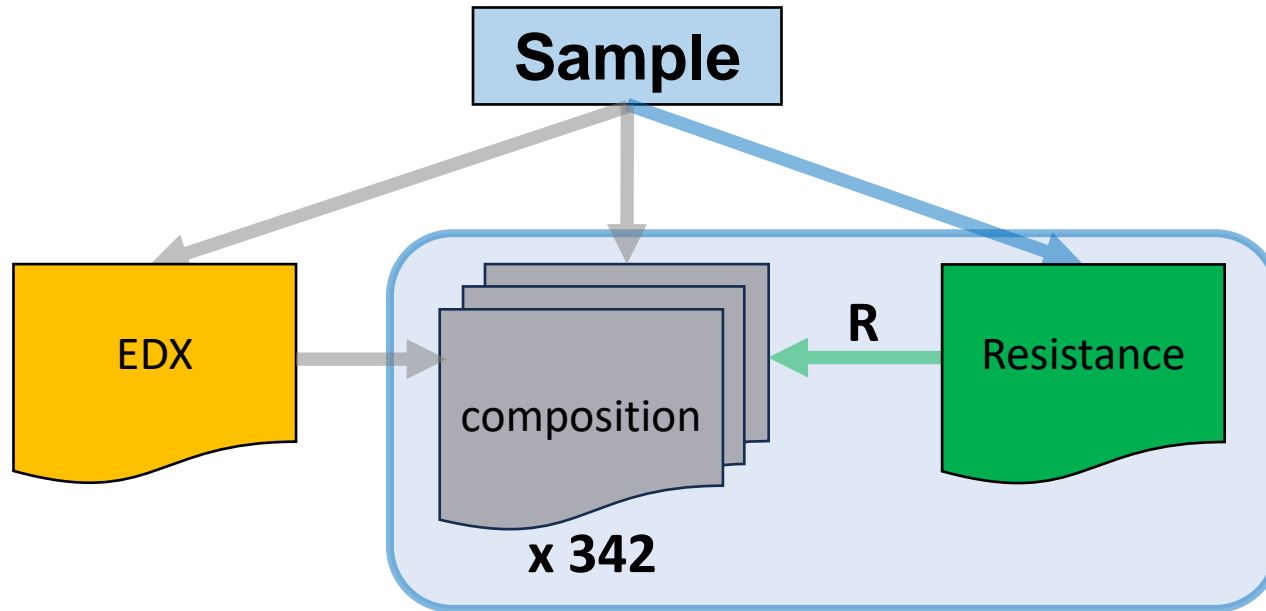
**Data from EDX (all values == 342 compositions):**

```
{
  "CompositionElements": [
    { "CompoundIndex": 0, "ElementName": "V", "ValueAbsolute": null, "ValuePercent": 31.299999237060547 },
    { "CompoundIndex": 10, "ElementName": "Mn", "ValueAbsolute": null, "ValuePercent": 2.700000047683716 },
    { "CompoundIndex": 20, "ElementName": "Co", "ValueAbsolute": null, "ValuePercent": 22.600000381469727 },
    { "CompoundIndex": 30, "ElementName": "Ni", "ValueAbsolute": null, "ValuePercent": 6 },
    { "CompoundIndex": 40, "ElementName": "Ho", "ValueAbsolute": null, "ValuePercent": 37.400001525878906 }
  ],
  "DeletePreviousProperties": true,
  "Properties": [
    {
      "PropertyId": 0, "Type": 2, "Name": "Measurement Area", "Value": 1, "ValueEpsilon": null,
      "SortCode": 0, "Row": null, "Comment": null
    }
  ]
}
```

Index	V	Mn	Co	Ni	Ho
1	31.3	2.7	22.6	6	37.4

# Data extraction: adding data to Measurement Area (Composition)

Use case: adding characterization data after EDX is done.



## Data import:

- 1) Locate parent sample object.
- 2) Find measurement area collection.
- 3) Find a certain Measurement Area (corresponding to a predicate, e.g. MA index).
- 4) Add properties values to a composition corresponding to a measurement area index.

```
{
  "CompositionsForSampleUpdate": [
    {
      "Predicate": {
        "Properties": [
          {
            "Type": 2,
            "Name": "Measurement Area",
            "Value": 1
          }
        ]
      },
      "DeletePreviousProperties": false,
      "Properties": [
        {
          "PropertyId": 0,
          "Type": 1,
          "Name": "R",
          "Value": 34143346,
          "ValueEpsilon": null,
          "SortCode": 10,
          "Row": null,
          "Comment": "Resistance (Room Temperature) for Measurement Area 1"
        }
      ]
    },
    ...
  ],
  "DeletePreviousProperties": true,
  "Properties": [
    {
      "PropertyId": 0,
      "Type": 1,
      "Name": "R",
      "Value": 97692.43229,
      "ValueEpsilon": null,
      "SortCode": 10,
      "Row": 1,
      "Comment": "Minimal Resistance (Room Temperature) of Materials Library (of all 342 MAs)"
    },
    {
      "PropertyId": 0,
      "Type": 1,
      "Name": "R",
      "Value": 34143346,
      "ValueEpsilon": null,
      "SortCode": 10,
      "Row": 2,
      "Comment": "Maximal Resistance (Room Temperature) of Materials Library (of all 342 MAs)"
    }
  ]
}
```

# Type configuration via JSON (extendable)

**Task:** configure type behavior (stored in TypeInfo.SettingsJson)

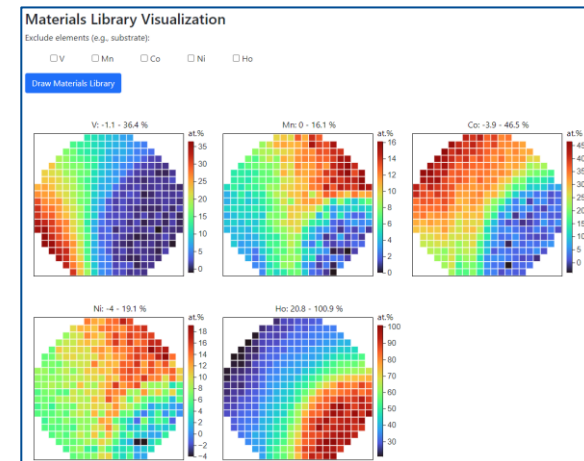
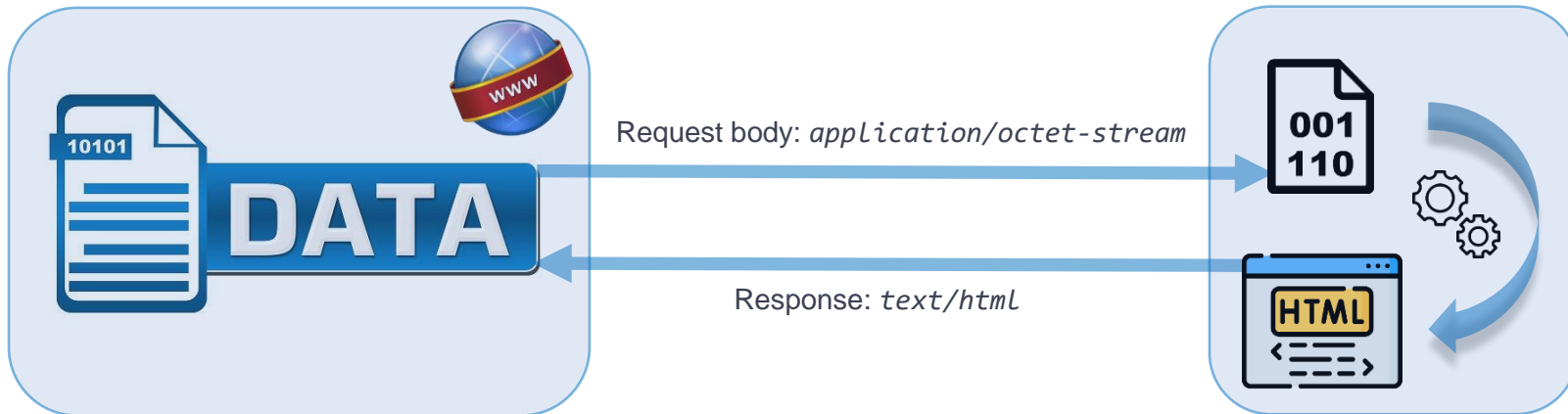
```
{  
  "ApplyForTypeIds": [ 6 ],  
  "CustomEditPath": "/custom/editsample",  
  "AllowedExtensions": [ ".txt", ".csv" ],  
  "UrlPostVisualizer": "https://..."  
}
```

Type is applicable to enlisted types (For Handover Event)

For Sample Type (custom edit form implementation)


Array of allowed extensions for data files

External Visualization (with Web Application)



# Reports & Statistics: Objects

Reports ▾

MDI Database Search Tree Edit List Edit Handover Reports ▾ find SampleID  Hello vic.dudarev@gmail.com! Logout

## Objects Statistics

Person:

Object Count	TypeId	Type Name	Type Comment
18315	!= 0	All objects	All types (TypeId != 0)
4592	8	Composition	Composition information (compound)
4100	6	Sample	Sample information (sputter deposition)
3143	18	Synthesis	Synthesis Document (parameters from Compact)
1670	7	Raw Document	Data file on disk in any format (file extension required to identify MIME-type)
1022	12	Photo	Photo (bitmap in jpg, png, etc...)
783	13	EDX CSV	CSV of EDX with Header row (e.g. "Index,V,Mn,Co,Ni,Ho")
733	24	SEM (image)	from Compact
660	26	TEM image	from Compact
256	15	EDX Image	Image (bitmap) that refers to EDX
214	33	HTTS Resistance TXT	TXT of HTTS Resistance. Via HTTPs by Azadeh
196	40	Thickness Other (zip, opj, opju, pdf)	Thickness Other files (from Compact)
146	28	Topography	from Compact

### Objects Statistics:

- Objects (w.r.t. types)
- Projects
- Object Links

### Feature:

- Filter by person



# Reports & Statistics: Users

Reports ▾

## Users Statistics

Rank	Name	Project	Entities Count
1	Tobias Piotrowiak	C04	1225
2	Carsten Placke-Yan	C03	46
3	Timo Fockenberg	C03	24
4	Leon Müller	C02	16
5	Victor Dudarev	INF	13
6	Carina Marek		6
7	Dana Krenz	C05	2
8	Sven Reichenberger	C05	1
8	Akhil Hareendran	A01	1
8	Alfred Ludwig	C04, INF	1
8	Adarsh Koul	A02	1
8	Wulyu Jiang	B06	1

### Users Statistics:

• Objects



• Projects

• Links

# Project-wide Statistics / Person-wide Statistics

Even more is to be implemented...

## Questions:

- How productive is project/person in terms of adding data (objects creation)?
- Is project subtree clearly structured (for manual subprojects traversing)?
- What is the most referenced / liked object?



?



## “Data Creator” awards (the winner is...):

- The most productive data creator award project/person
- The people’s choice award: project subtree clarity
- The citation / popularity award



# Use Case: data upload Sample -> EDX -> Resistance & Search

- 1) Create **Project**
- 2) Add **Sample** to Project
- 3) Add **EDX** document (342 Measurement Areas (MA), i.e. 342 Compositions)
- 4) Add **Resistance** document (associate property values to every MA)

## Search for Composition with a given Resistance

- Composition – object representing one of 342 measurement area
- Container for other properties, associated with a certain measurement area

Chosen elements: Co-Ho-Mn-Ni-V [\[clear\]](#)

	Absolute			Percentage				
V	min	≤ N ≤	max	or	min	≤ N % ≤	max	
Co	min	≤ N ≤	max	or	min	≤ N % ≤	max	
Mn	min	≤ N ≤	max	or	min	≤ N % ≤	max	
Ho	min	≤ N ≤	max	or	20	≤ N % ≤	80	
Ni	min	≤ N ≤	max	or	min	≤ N % ≤	max	

**Additional filters** | Specific search

Object type: Composition | Search phrase: text in Name or Description

Property type: Float | Name: R | Property Value: 1 ≤ value ≤ 8

Created by: | from: mm/dd/yyyy | till: mm/dd/yyyy |

# Flexible data queries for advanced users

Show samples having Au-Pd-Rh-Pt and how many characterization documents for them we have (EDX, XRD, 4PP, SDC, SECCM)

SQL execution tool [show schema](#)

```

SQL Query
DECLARE @dt as TABLE([Value] [varchar](2) NOT NULL);
INSERT INTO @dt VALUES('Au'),('Pd'),('Rh'),('Pt');
select ExternalId as SampleId, ObjectName, ElemNumber, RTRIM(LTRIM([Elements], '-')) as [Elements],
(
select TOP 1 OI.ObjectName from vro.vroObjectLinkObject as OLO
INNER JOIN vro.vroObjectInfo as OI ON OLO.LinkedObjectId=OI.ObjectId AND OI.TenantId=S.TenantId
WHERE OLO.ObjectId=S.ObjectId AND OI.TypeId=5) as [SubstrateMaterial]
, (select count(ObjectLinkId) FROM [vro].[vroObjectLinkObject] WHERE ParentObjectId=S.ObjectId AND ChildTypeId IN (13, 19)) as EDX
, (select count(ObjectLinkId) FROM [vro].[vroObjectLinkObject] WHERE ParentObjectId=S.ObjectId AND ChildTypeId IN (17, 44, 97)) as XRD
, (select count(ObjectLinkId) FROM [vro].[vroObjectLinkObject] WHERE ParentObjectId=S.ObjectId AND ChildTypeId IN (14, 33)) as [4PP]
, (select count(ObjectLinkId) FROM [vro].[vroObjectLinkObject] WHERE ParentObjectId=S.ObjectId AND ChildTypeId IN (85)) as [SDC]
, (select count(ObjectLinkId) FROM [vro].[vroObjectLinkObject] WHERE ParentObjectId=S.ObjectId AND ChildTypeId IN (86, 87)) as [SECCM]
from vro.vroSample as S
where TypeId=6/'Sample/' and NOT EXISTS (select top 1 [Value] from @dt WHERE CHARINDEX ('-'+[Value]+'-', S.[Elements])=0)
    
```

**Execute** **Save CSV data**

**Query Execution result**

SampleId	ObjectName	ElemNumber	Elements	SubstrateMaterial	EDX	XRD	4PP	SDC	SECCM
10304	10304 Au-Pd-Pt-Rh on 15nm Ta Library 1	4	Au-Pd-Pt-Rh	Sapphire	0	0	0	0	0
10269	10269 Ag-Au-Pd-Pt-Rh on 15nm Ta Library 1	5	Ag-Au-Pd-Pt-Rh	Sapphire	1	0	0	0	0
10275	10275 Ag-Au-Pd-Pt-Rh on 15nm Ta Library 2	5	Ag-Au-Pd-Pt-Rh	Sapphire	1	0	0	0	0
10311	10311 Au-Pd-Pt-Rh-Ru on 15nm Ta Library 1	5	Au-Pd-Pt-Rh-Ru	Sapphire	0	0	0	0	0

Showing 1 to 4 of 4 entries

Show samples (identifiers) that have all enlisted characterisations: EDX, XRD, Resistance, Thickness, Photo

SQL execution tool [show schema](#)

```

SQL Query
SELECT ParentObjectId as ObjectId, ParentExternalId as SampleId FROM [vroObjectLinkObject] where ParentTypeId=6 and ParentTenantId=7 and ChildTypeId IN (13, 19) -- 921: EDX
INTERSECT
SELECT ParentObjectId, ParentExternalId FROM [vroObjectLinkObject] where ParentTypeId=6 and ParentTenantId=7 and ChildTypeId IN (17, 31) -- 88: XRD
INTERSECT
SELECT ParentObjectId, ParentExternalId FROM [vroObjectLinkObject] where ParentTypeId=6 and ParentTenantId=7 and ChildTypeId IN (14, 33, 16) -- 346: HTTPS Resistance CSV=14 // HTTPS Resistance TXT=33 // 16=HTTPS Resistance Image
INTERSECT
SELECT ParentObjectId, ParentExternalId FROM [vroObjectLinkObject] where ParentTypeId=6 and ParentTenantId=7 and ChildTypeId IN (27, 38, 39, 40) -- 456: Thickness Excel=27 // Thickness TXT=38 // 39=Thickness Image, 40=Thickness Other (zip, opj, opju, pdf)
INTERSECT
SELECT ParentObjectId, ParentExternalId FROM [vroObjectLinkObject] where ParentTypeId=6 and ParentTenantId=7 and ChildTypeId IN (12, 15, 16, 24, 26, 31, 39) -- 2780: PHOTO
    
```

**Execute** **Save CSV data**

**Query Execution result**

ObjectId	SampleId
51614	5210
51753	5255

Showing 1 to 2 of 2 entries

# API for running arbitrary read-only queries

https://mdi.matinf.pro/vroapi/v1/execute

POST https://mdi.matinf.pro/vroapi/v1/execute

Params Authorization Headers (9) Body Scripts Settings Cookies

none form-data x-www-form-urlencoded raw binary GraphQL

Key	Value	Description
sql	select * from vroTypeInfo	

Body Cookies Headers (6) Test Results 200 OK 407 ms 21.03 KB Save as example

Pretty Raw Preview Visualize JSON

```
1 [
2   {
3     "TypeId": -2,
4     "IsHierarchical": false,
5     "TypeIdForRubric": 2,
6     "TypeName": "Link Type",
7     "TableName": "ObjectInfo",
8     "UrlPrefix": "linktype",
9     "TypeComment": "Link Type objects are used in objects that are linked together to express
10      the meaning of the link (and probably quantify it using SortCode).",
11     "ValidationSchema": null,
12     "DataSchema": null,
13     "SettingsJson": null,
14     "FileRequired": false,
15     "_date": "2024-04-11T15:56:29.137"
16   },
17 ]
```

## POST /vroapi/v1/execute

Executes SQL statement and returns result in JSON

## Python wrapper to use API:

# 1. Specify tenant URL (since numerous tenants exist)

```
service_url = "https://mdi.matinf.pro" # tenant URL
```

# 2. Specify API Key ("VroApi" user claim)

```
api_key = "<api-key>" # your api_key here
```

# 3. Create a client object

```
client = MatInfWebApiClient(service_url, api_key)
```

# 4. Execute SQL and get result

```
result = client.execute("select * from vroTypeInfo")
```

# result - contains JSON

# client.dataframe - contains pandas dataframe

```
In [8]: # see response as dataframe
client.dataframe
```

	TypeId	IsHierarchical	TypeIdForRubric	TypeName	TableName	UrlPrefix	TypeComment
0	-2	False	2.0	Link Type	ObjectInfo	linktype	Link Type objects are used in objects that are...
1	-1	False	2.0	Handover	Handover	handover	Handover event
2	1	True	NAN	Organisation Structure	RubricInfo	organisation-structure	Hierarchy of organisational units within the te...
3	2	True	NAN	Project	RubricInfo	project	Project is a container to include all relevant information about publications that are con...
4	3	False	2.0	Literature Reference	Reference	literature-reference	Information about publications that are con...
56	95	False	2.0	Stress Measurement (DSM)	ObjectInfo	stress-measurement-dsm	Asked by Elsieh
57	95	False	2.0	Report	ObjectInfo	report	Asked by Elsieh
58	97	False	2.0	XRD Raw ZIP (xy)	ObjectInfo	xrd-raw-zip-xy	Jill
59	98	False	2.0	PowerPoint Presentation Slides (ppt, pptx)	ObjectInfo	powerpoint-presentation_slides_ppt_pptx	PowerPoint Presentation Slides (ppt, pptx)
60	99	False	2.0	Calculation/Computational Sample	Sample	calculation-computational_sample	Calculation/Computational Sample. Request from...

61 rows x 12 columns

# API for downloading documents

https://mdi.matinf.pro/vroapi/v1/download?id=43511

GET https://mdi.matinf.pro/vroapi/v1/download?id=43511

Params Auth Headers (7) Body Scripts Settings Cookies

Query Params

<input checked="" type="checkbox"/>	Key	Value	Description	...	Bulk Edit
<input checked="" type="checkbox"/>	id	43511			
	Key	Value	Description		

Body 200 OK 240 ms 23.68 KB Save as example

Pretty Raw Preview Visualize Text

```
1
2 Project: 160322-K1-2 (1)
3 Owner:
4 Site: Site of Interest 1
5
6 Sample: Sample 1
7 Type: Default
8 ID:
9
10 Processing option : All elements analysed (Normalised)
11
12 All results in atomic%
13
14 X and Y are stage positions
15
16 Spectrum In stats. X (mm) Y (mm) O Si Ti Ta
17
18 Spectrum 1 {1} Yes -43.811 -40.975 47.60 51.82 2.33 -1.75
19 Spectrum 1 {2} Yes -39.308 -40.975 51.66 46.76 1.63 -0.05
20 Spectrum 1 {3} Yes -34.808 -40.975 44.69 54.14 2.06 -0.89
21 Spectrum 1 {4} Yes 30.310 40.975 36.30 61.80 0.76 1.06
```

**GET /vroapi/v1/download**

Downloads document by specified ObjectId

**Python wrapper to use API:**

```
# 1. Specify tenant URL (since numerous tenants exist)
service_url = "https://mdi.matinf.pro" # tenant URL
# 2. Specify API Key ("VroApi" user claim)
api_key = "<api-key>" # your api_key here

# 3. Create a client object
client = MatInfWebApiClient(service_url, api_key)

# 4. Download object file by ObjectId
result = client.download(43511) # ObjectId
# result - contains HTTP response
# result.content - contains file data
# client.file_name - contains file name
```



# Source Control (GitLab)

Everybody's welcome!

Feel free to contribute to open source!



Core INF project:

<https://gitlab.ruhr-uni-bochum.de/vic/infproject>

Shared projects:

1) Administration User Interface (Identity Manager UI):

<https://gitlab.ruhr-uni-bochum.de/vic/identitymanagerui>

2) Web Application General Library (WebUtilsLib):

<https://gitlab.ruhr-uni-bochum.de/vic/webutilslib>



<https://gitlab.ruhr-uni-bochum.de>

Name	Last commit	Last update
InfProject	Search Done	58 minutes ago
.gitignore	InitialCommit (tree editing ready)	2 months ago
InfProject.sln	InitialCommit (tree editing ready)	2 months ago
global.json	InitialCommit (tree editing ready)	2 months ago

Commit	Author	Date	Commit hash
849afc7	Victor Dudarev	1 day ago	849afc7
83cbdfc	Victor Dudarev	15 days ago	83cbdfc
5eea0a7	Victor Dudarev	25 days ago	5eea0a7
787fa0c	Victor Dudarev	1 month ago	787fa0c
c57dacf	Victor Dudarev	2 months ago	c57dacf
7075154	Victor Dudarev	2 months ago	7075154
db276bb	Victor Dudarev	2 months ago	db276bb



**Thanks for your kind attention**

# Questions & Answers

Victor Dudarev

[Victor.Dudarev@rub.de](mailto:Victor.Dudarev@rub.de)

# RDMS: Setting the Object Access Level

## What is an Object?

Object (=document) is a data entry within RDMS that has user-defined access level and reflects an object of the real world (e.g. sample) or its model. Ultimately, object has its unique Web page (with unique URL address).

**To establish data access policy Data Access Levels are introduced in RDMS.**

## Four Access Levels:

- **Public** (default, we are doing open science, FAIR, aren't we?):
  - objects are available to everybody regardless of authorization (visible to internet search engines);
- **Protected** (visible to the community only):
  - objects are visible to the community (available to authorized users with at least **User** role assigned);
- **ProtectedNDA** (restricted visibility):
  - objects are visible to the authorized users with **NDA claim** set OR to Administrators;
- **Private** (person's secret, but open for Administrators):
  - objects are visible to the user-creator (at least **PowerUser** role assigned);
  - objects are visible to all users of **Administrator** role.



# RDMS: Providing Project Description

**Motivation:** allow to add to a project arbitrary text (HTML).

Workshop2023 - INF

Here you can find presentations from the CRC247 INF Workshop (7-9 November 2023, Bochum), devoted to **Research Data Management**. Related materials can also be found in the [INF project](#).

Add Subproject

Hide Non-samples

please in edit project click “show more”:

Editing Workshop2023 - INF [Id=329]

**Name**

Workshop2023 - INF

name of the node

[show more parameters](#)

Close Save

Editing Workshop2023 - INF [Id=329]

**Name**

Workshop2023 - INF

name of the node [hide more parameters](#)

**Parent Id**

307

Don't change it if you are not sure what you are doing

**Sort Code**

0

within a parent all children are sorted by this number (ascending)

**Access Control (accessibility)**

public

public - to all; protected - to authorized users; protectedNDA - to authorized users with NDA claim; private - to you only

**Text**

Here you can find presentations from the CRC247 INF Workshop (7-9 November 2023, Bochum), devoted to

description to show (HTML is allowed)

Close Save