Dr.
V. Dudarev

Research Data Management Systems in Materials Science based on FAIR Principles

Victor Dudarev

CRC/TRR 247 HETEROGENEOUS OXIDATION CATALYSIS IN THE LIQUID PHASE
RUB
UNIVERSITÄT DUISBURG ESSEN Offen im Denken
C A U Christian-Albrechts-Universität zu Kiel
MAX-PLANCK-INSTITUT FÜR CHEMISCHE ENERGIEKONVERSION
FRITZ-HABER-INSTITUT MAX-PLANCK-GESELLSCHAFT
DFG Deutsche Forschungsgemeinschaft German Research Foundation

Agenda

INF(N)

- INF Project in CRC/TRR 247:
  - Defining requirements
  - Database schema and flexibility
  - Current state: implemented features and future plans
  - Live Demonstration
- Q&A.

RDMS in Chemistry-related Domains: Requirements

INF(N)

Requirements (functional capabilities):

Priorities (SFS): Security + Flexibility + Scalability

- Support for multiple **Tenants**;
- Implement **User Registration** (including e-mail verification) and **Authorization**;
- Implement Administrative Interface to **Control Users and Groups**;
- Establish **Predefined Roles** (Administrator / PowerUser / User) with corresponding permissions;
- Ensure **Data Access Policy** (public / protected / private) for all stored objects based on Object Access Level;
- **Adjust Functionality** of RDMS according to authorized user role (Administrator / PowerUser / User);
- **Upload and Store Objects** (documents) with minimalistic mandatory metadata;
- Provide **Flexible Tree Classification** for stored documents (projects / organizational structure);
- Provide means to **Support Chemical Entities** data types: systems, compositions, crystal structures;
- Enable to **Interlink Objects** manually (associative objects) and show the resulting reverse associations;
- Flexibility: implement **Extended Properties for Objects** and support search on them;
- **Import and Export** object table properties from/to Excel; support table properties templates;
- Provide **Search Interface** on chemical entities and objects with respect to user access level;
- Known object types **Bulk Import from CSV** files (with CSV-schema validation);
- Flexibility: enable to easily **Introduce Additional Object Types** with predefined mandatory fields set*;

To be done:

- Provide **Reports / Charts / Diagrams** on stored data in various aspects;
- **API** to upload data from measurement devices (Bandgap, Resistance, etc…).

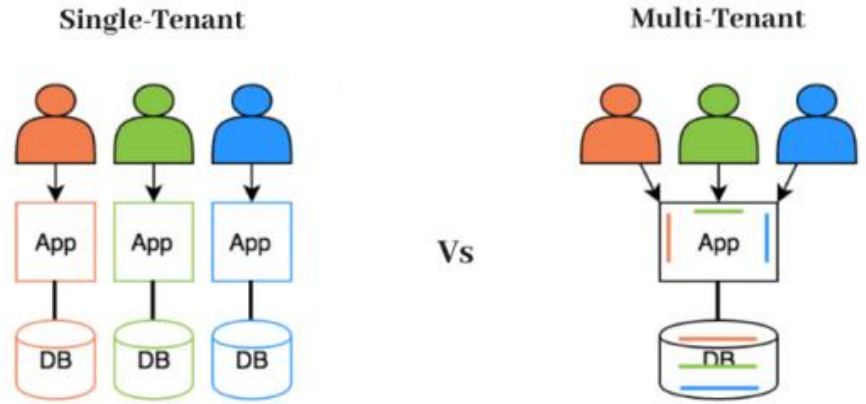
RDMS: Research Data Management System

* - depending on object type complexity may require programming (extra tables introduction in DB)

RDMS: Multiple Tenant Support

INF(N)

What is Tenant?



Short answer:
Tenant – separate (distinct) instance of a software application that is used by a single organization or group of users.

Long answer:
Tenant is a customer account that has its own set of users, data, and configuration settings, and is isolated from other tenants in the same software application. This concept is commonly used in Software as a Service (**SaaS**) applications, where a single software application is hosted in the cloud and is made available to multiple customers as separate tenants. This allows each customer to have their own instance of the application, with their own data, settings, and customizations, while still sharing the underlying infrastructure and resources of the application

Conclusion:
Tenants are useful for separate SFB projects / workgroups / materials data repositories. For example:

- https://**demo**.mdi.ruhr-uni-bochum.de/
- https://**inf**.mdi.ruhr-uni-bochum.de/
- https://**crc247**.mdi.ruhr-uni-bochum.de/

Every tenant must have a unique URL!

(1) https://en.wikipedia.org/wiki/Multitenancy

RDMS Database

Tenants (independent systems):

VicRUBNote.RUB_INF - dbo.Tenant

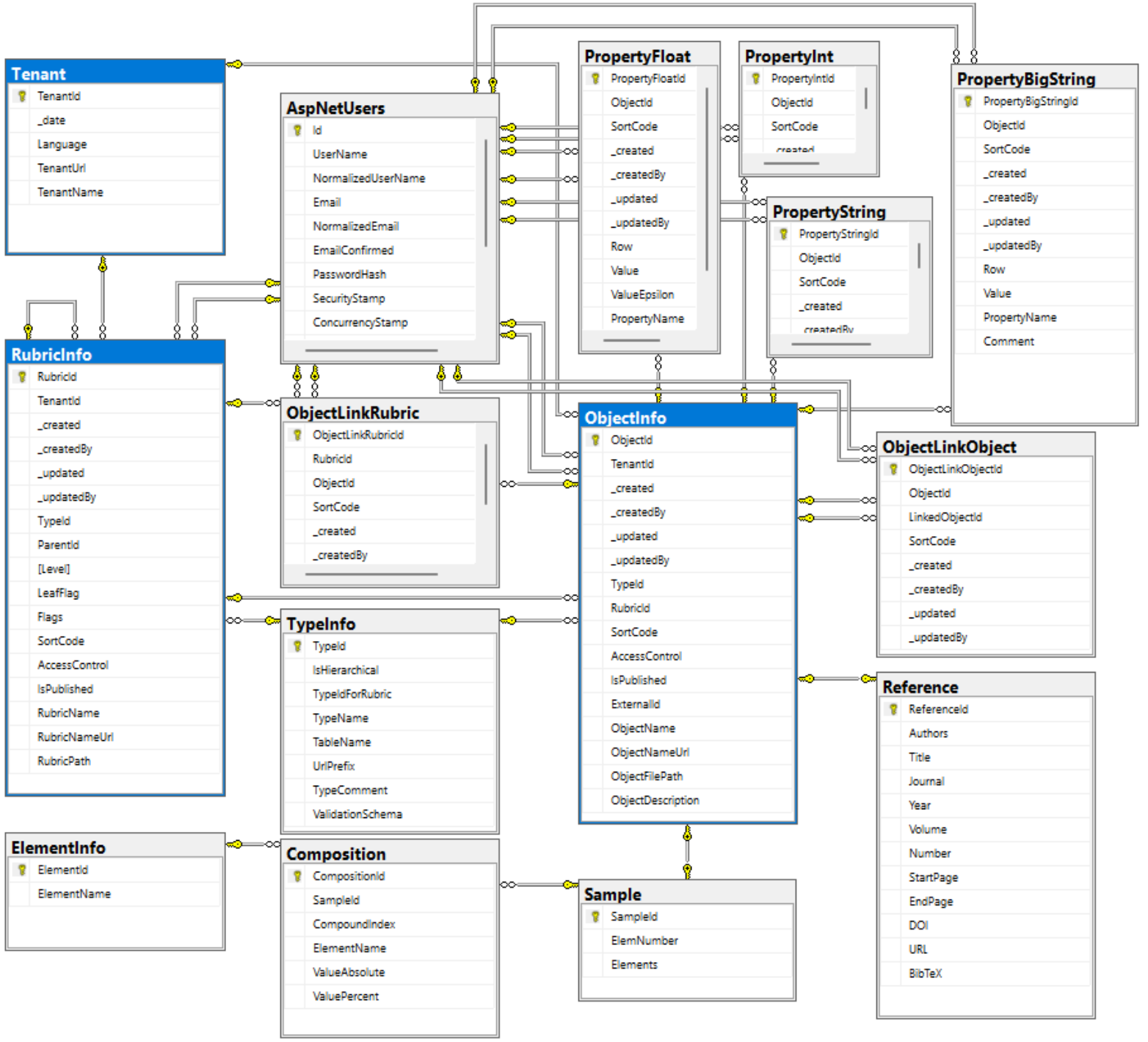
	TenantId	_date	La...	TenantUrl	TenantName
	1	2022-11-09 12:0...	en	inf.mdi.ruhr-uni-bochum.de	CRC/TRR247
	2	2023-01-04 14:4...	en	dim.mdi.ruhr-uni-bochum.de	DIMENSION
	3	2023-01-16 18:3...	en	demo.mdi.ruhr-uni-bochum.de	DEMO

Every tenant must have a unique URL!

Data Types (can be shared among tenants):

VicRUBNote.RUB_INF - dbo.TypeInfo

	TypeId	IsHierarchical	TypeIdForRubric	TypeName	TableName
	1	True	*NULL*	Organisation Structure	RubricInfo
	2	True	*NULL*	Project	RubricInfo
	3	False	2	Literature Reference	Reference
	4	False	2	Publication	Reference
	5	False	2	Substrate	ObjectInfo
	6	False	2	Sample	Sample
	7	False	2	Raw Document	ObjectInfo
	8	False	2	Composition	Composition
	9	False	2	Tag	ObjectInfo
	10	False	2	System	Sample



RDMS: Registration and Authorization

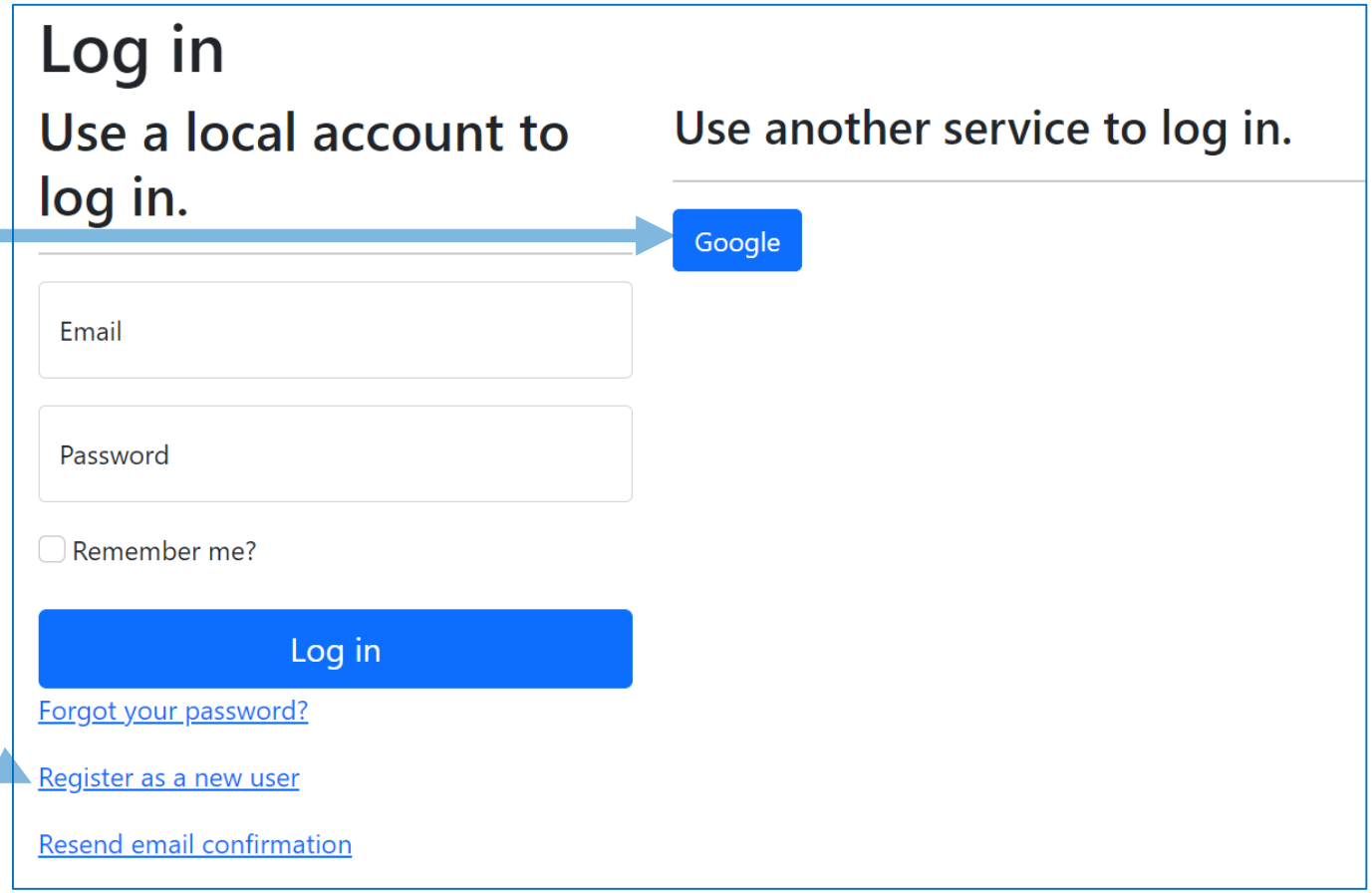
INF(N)

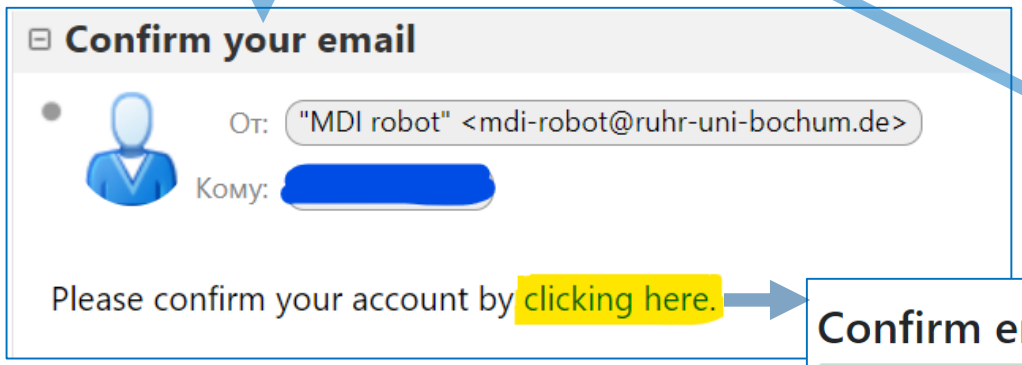
Application access level is determined by the current user context.

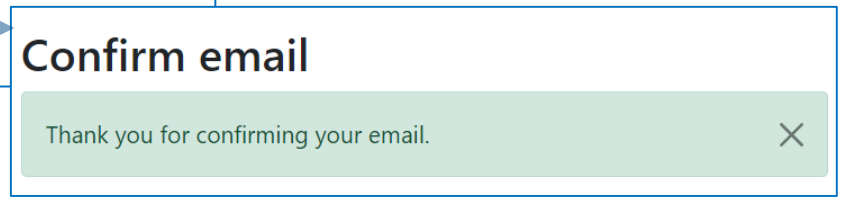
Two-way registration (choose any):

- using **external provider** (OpenID Connect), e.g. Google (external authentication authority is used, no credentials are stored locally).
- using **local account** (credentials are stored internally);

In both cases, an e-mail address is required, and its successful verification is a prerequisite for access.







After e-mail confirmation, the user is considered active, but does not have an assigned user group.

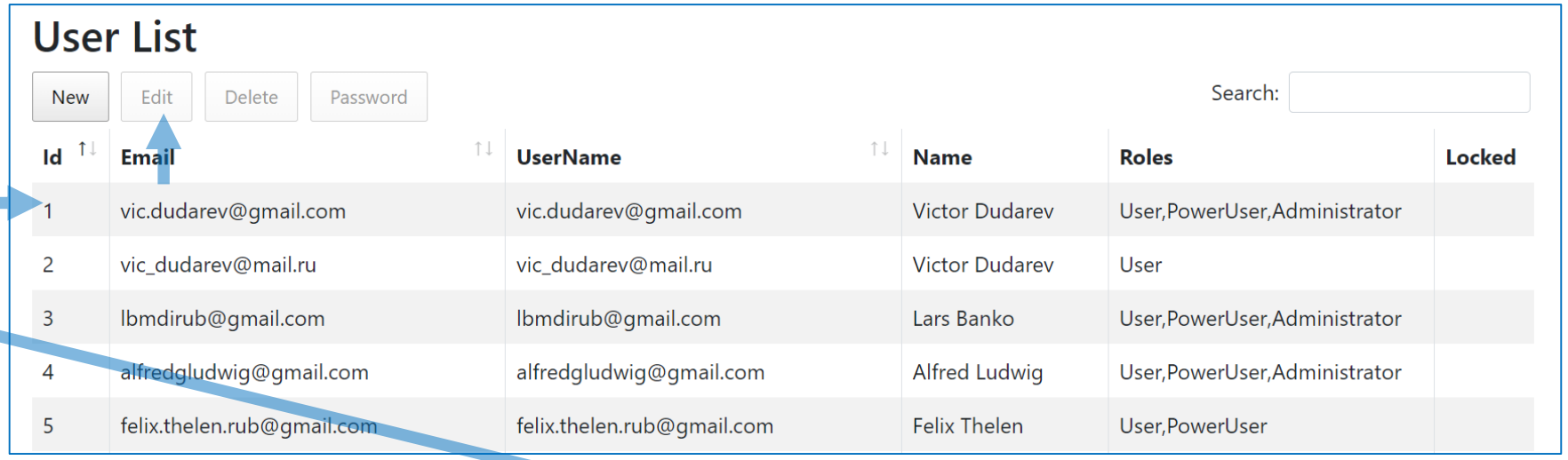
(1) https://openid.net
(2) https://en.wikipedia.org/wiki/Authentication
(3) https://en.wikipedia.org/wiki/Authorization

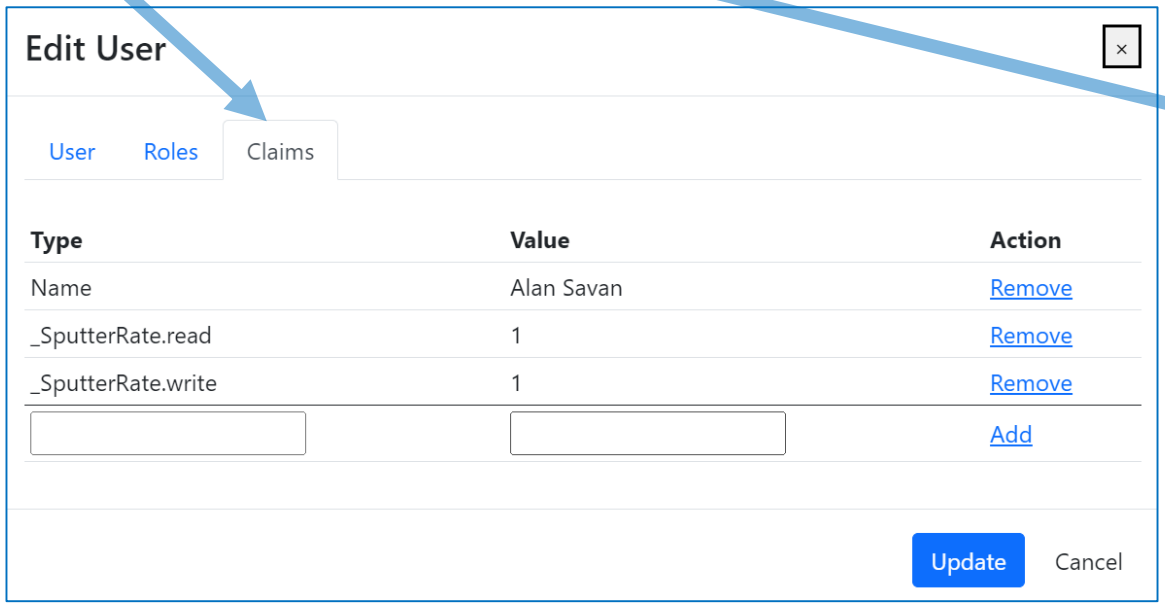
RDMS: Identity User Control Interface

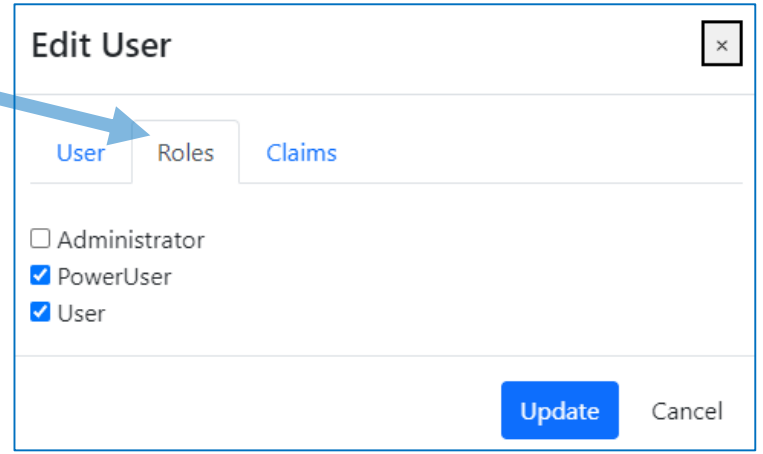
INF(N)

Features:

- Registered users management
- Roles management
- Claims management







Claim is a piece of information (key-value pair) that describes a user. Claims are stored in a user's identity and can be used to determine what actions a user is authorized to perform, such as accessing certain pages or resources.

https://learn.microsoft.com/en-us/aspnet/core/security/authentication/identity

RDMS: Predefined Roles

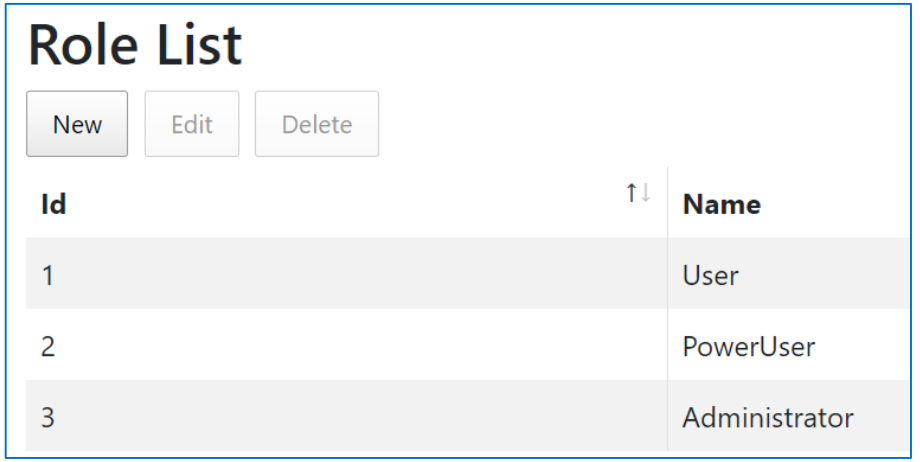
INF(N)

Roles:

- **User:**
  - read-only access to public and protected data

- **PowerUser:**
  - read-only access to public and protected data (+ private data, created by the current user)
  - add data (+ write access to data, created by the current user)

- **Administrator**:
  - full access to all data (CRUD)
  - user management



- **Anonymous (== registered user with no roles assigned)**:
  - read-only access to public data only (be aware: internet search engines will index public data)

https://en.wikipedia.org/wiki/Role-based_access_control

RDMS: Setting the Object Access Level

INF(N)

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 it's unique Web page (with unique URL address).

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

Three 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);

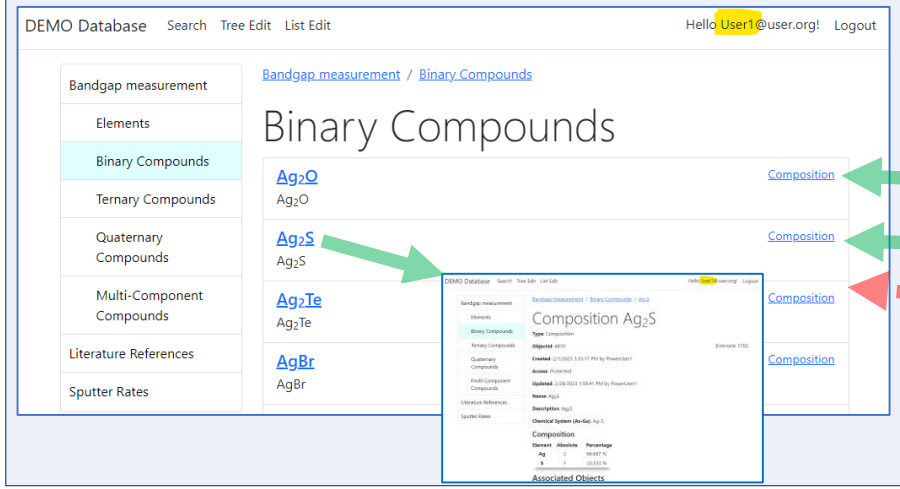
- **Protected** (visible to the community only):
  - objects are visible to the community (available to authorized users with at least **User** role assigned);

- **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.

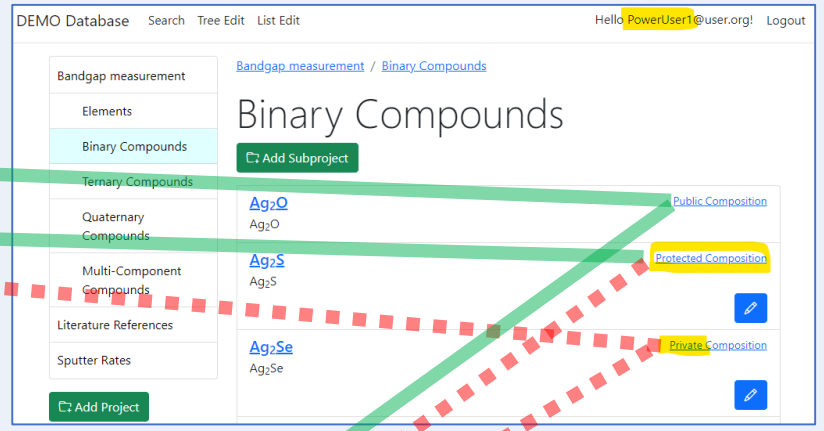
RDMS: Adjust Functionality According to User Permissions

User credentials (role) affects user interface: https://demo.mdi.ruhr-uni-bochum.de/rubric/binary-compounds

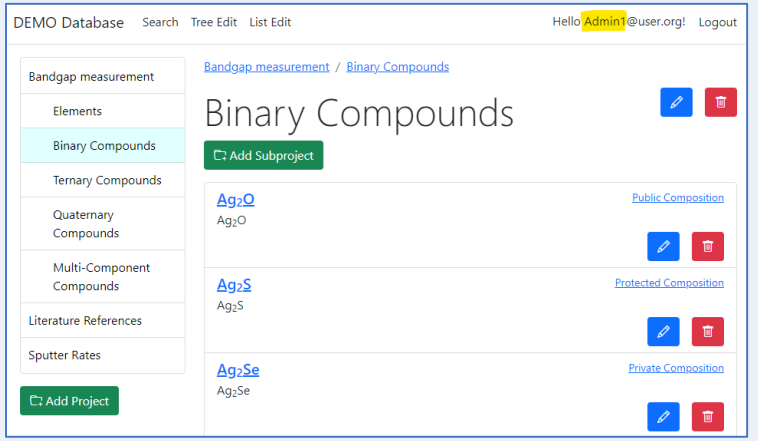
User1 (User role)



PowerUser1 (PowerUser role)

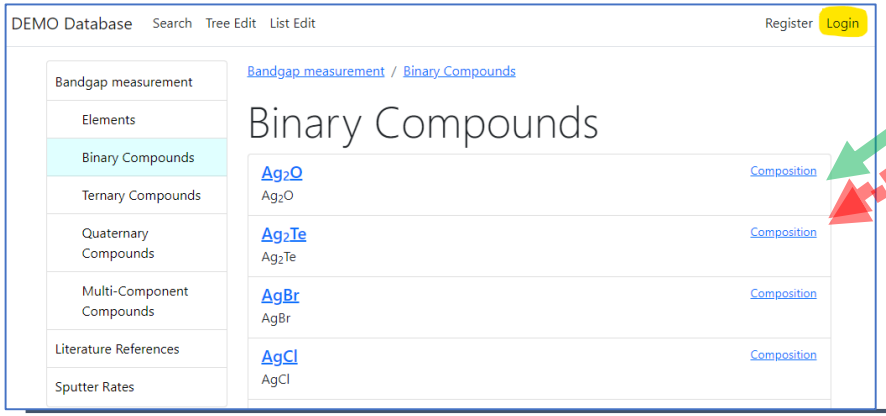


Admin1 (Administrator role)

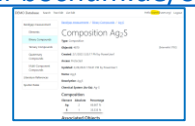


Authorized CRC Community

User (no role) = Anonymous (search engine)



Visit
https://demo.mdi.ruhr-uni-bochum.de/object/ag2s-4870



- User => access granted
- Anonymous => access denied

```
access denied
```

Preconditions:

Ag_2O created (public) by Admin1
Ag_2S created (protected) by PowerUser1
Ag_2Se created (private) by PowerUser1

Accessible (visible)

Denied (hidden)

Login and Password (User): User1@user.org
Login and Password (PowerUser): PowerUser1@user.org
Login and Password (Administrator): Admin1@user.org

RDMS: Upload and Store Documents (with minimalistic mandatory metadata)

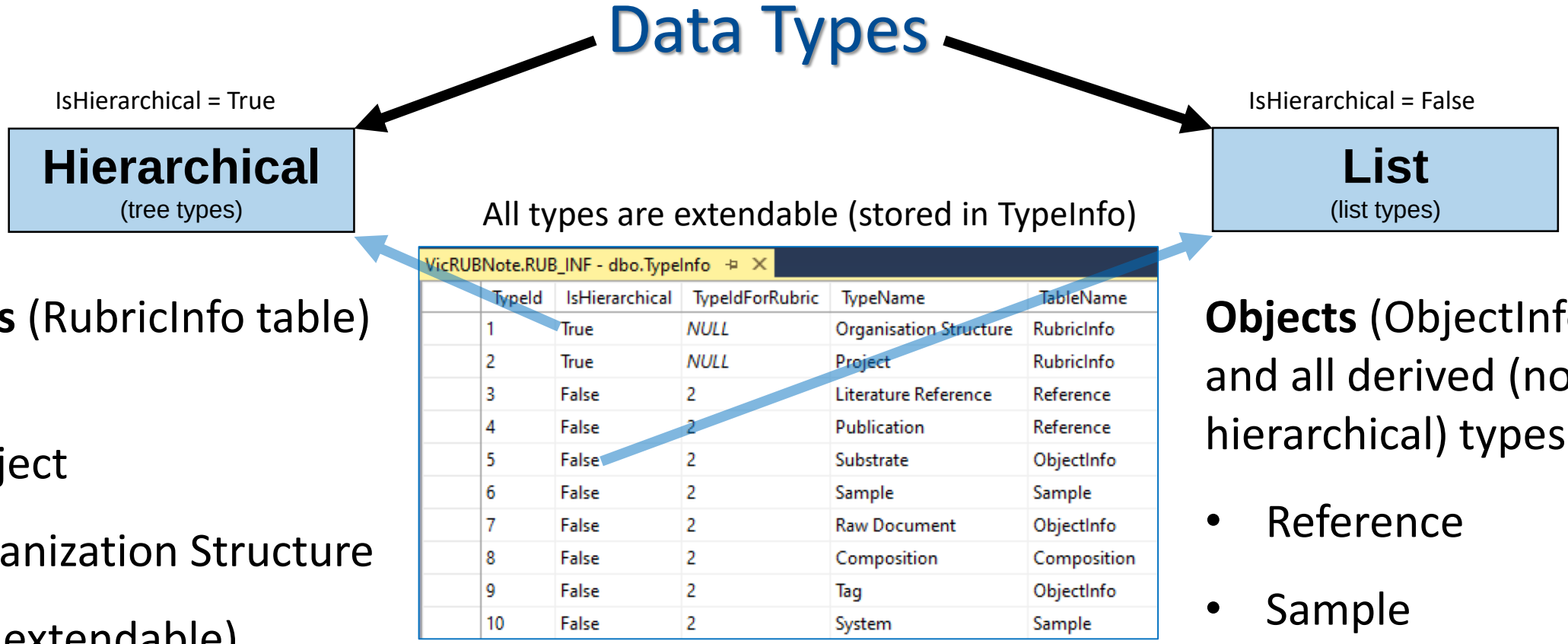
INF(N)

ObjectInfo

	Column Name	Data Type	Allow Nulls
🔑	ObjectId	int	☐
	TenantId	int	☐
	_created	datetime	☐
	_createdBy	int	☐
	_updated	datetime	☐
	_updatedBy	int	☐
	TypeId	int	☐
	RubricId	int	☑
	SortCode	int	☐
	AccessControl	int	☐
	IsPublished	bit	☐
	ExternalId	int	☑
	ObjectName	varchar(128)	☐
	ObjectNameUrl	varchar(256)	☐
	ObjectFilePath	varchar(256)	☑
	ObjectFileHash	varchar(128)	☑
	ObjectDescription	varchar(256)	☑
			☐

- Unique Object Identifier (assigned automatically), primary key
- Tenant Identifier (assigned automatically based on URL address)
- User who created the object and date with time (automatically)
- User who modified the object and modification date and time (auto)
- Object type reference (determined by user on object creation)
- Project (tree-structure) reference (can specified by user on object creation)
- Code to define a certain order in list of objects (auto, can be changed)
- One of predefined Object Access Levels: **public**$_{def}$ / protected / private
- **Name of the object** (**must be specified**)
- Relative URL for page to access object (auto formed; can be changed)
- Reference to Data File uploaded to system (uploading by choice)
- SHA-256 file hash (must be unique within every tenant)
- Object Description (can be empty, as all nullable attributes above)

RDMS: Data Types (hierarchies and lists)



VicRUBNote.RUB_INF - dbo.TypeInfo

TypeId	IsHierarchical	TypeIdForRubric	TypeName	TableName
1	True	*NULL*	Organisation Structure	RubricInfo
2	True	*NULL*	Project	RubricInfo
3	False	2	Literature Reference	Reference
4	False	2	Publication	Reference
5	False	2	Substrate	ObjectInfo
6	False	2	Sample	Sample
7	False	2	Raw Document	ObjectInfo
8	False	2	Composition	Composition
9	False	2	Tag	ObjectInfo
10	False	2	System	Sample

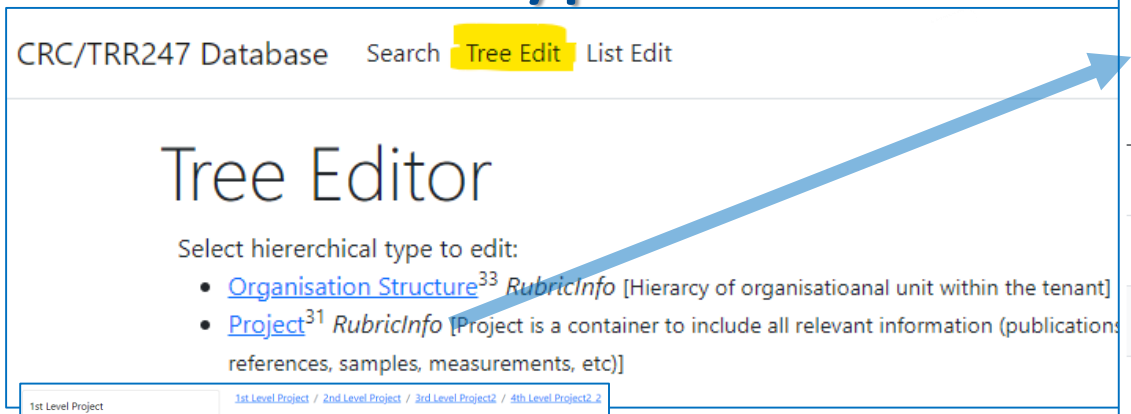
Rubrics (RubricInfo table) types:

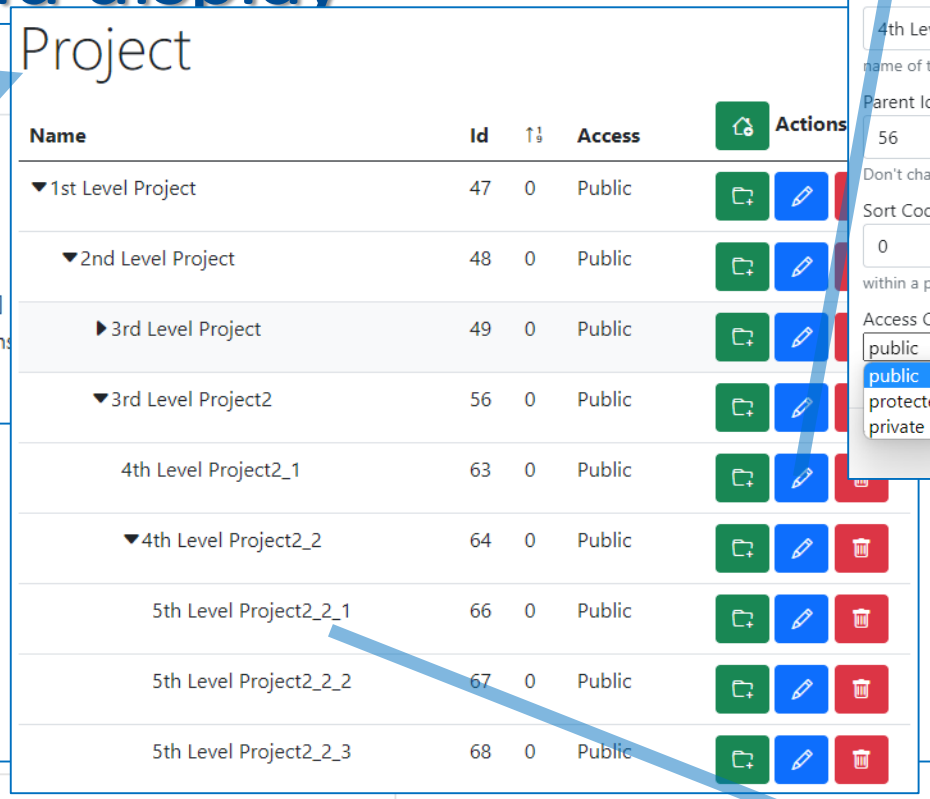
- Project
- Organization Structure
- ++ (extendable)

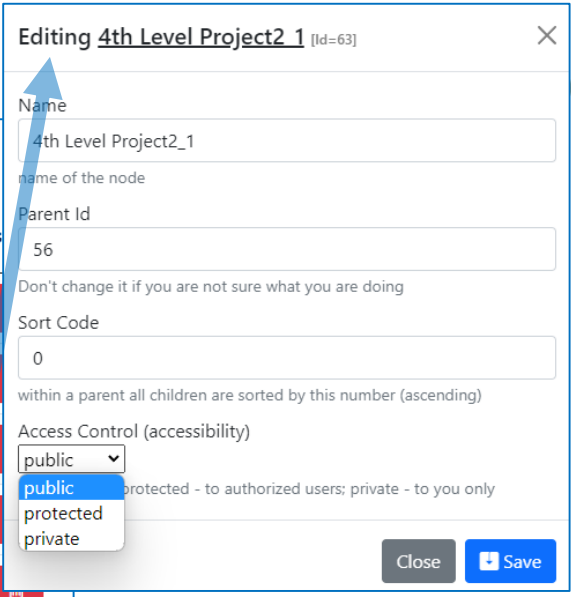
Objects (ObjectInfo table) and all derived (non-hierarchical) types:

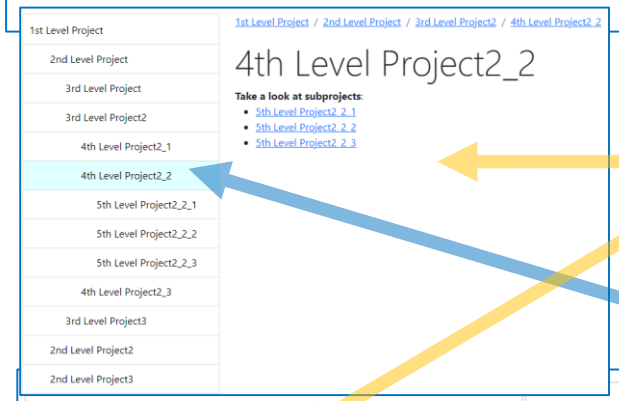
- Reference
- Sample
- Composition
- ++ (extendable)

RDMS Tree Types: control and display





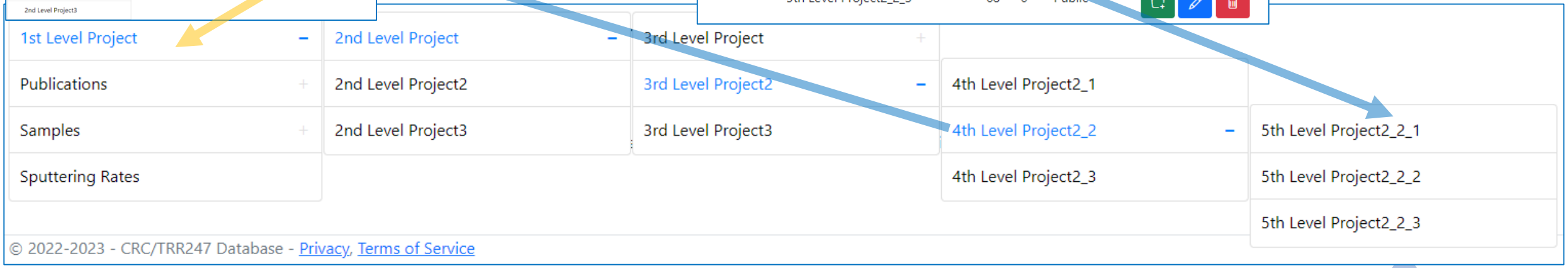




Display for users in Web Application

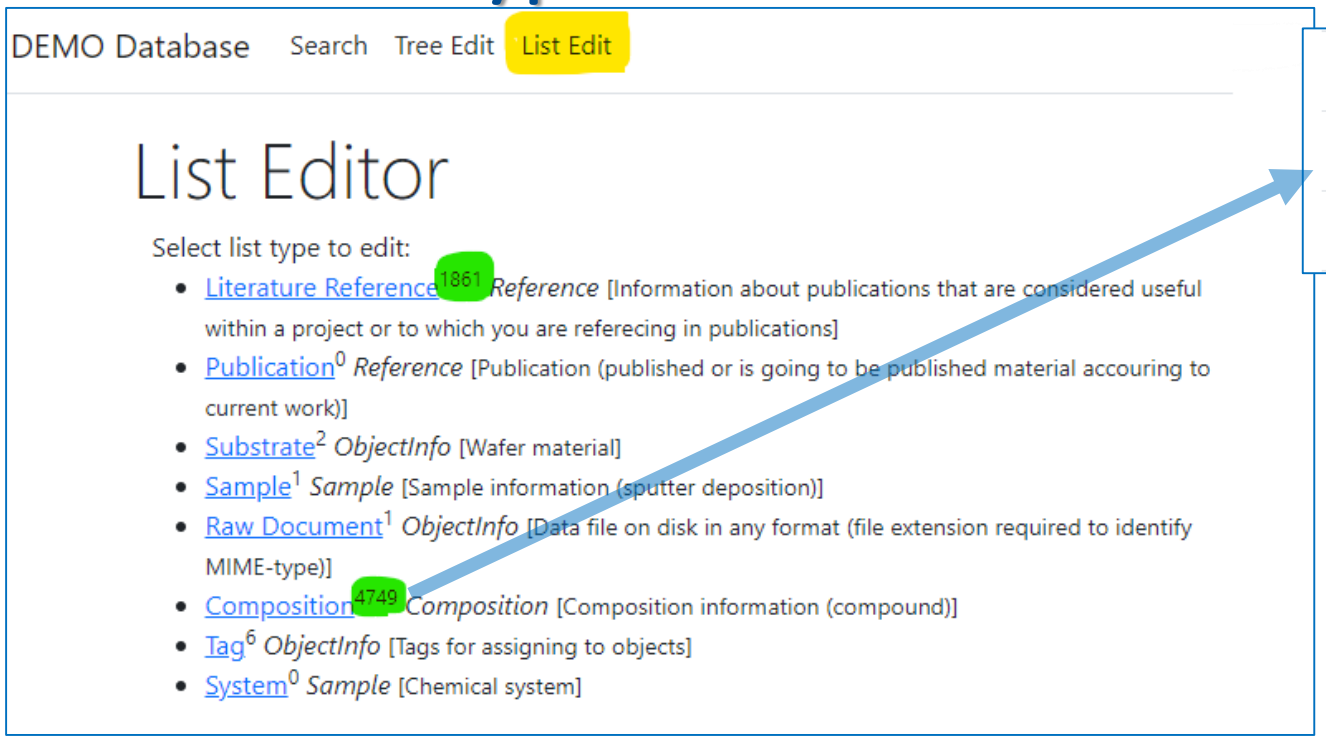
Control (admin):

- Name
- Parent Id
- Sort Code
- Access

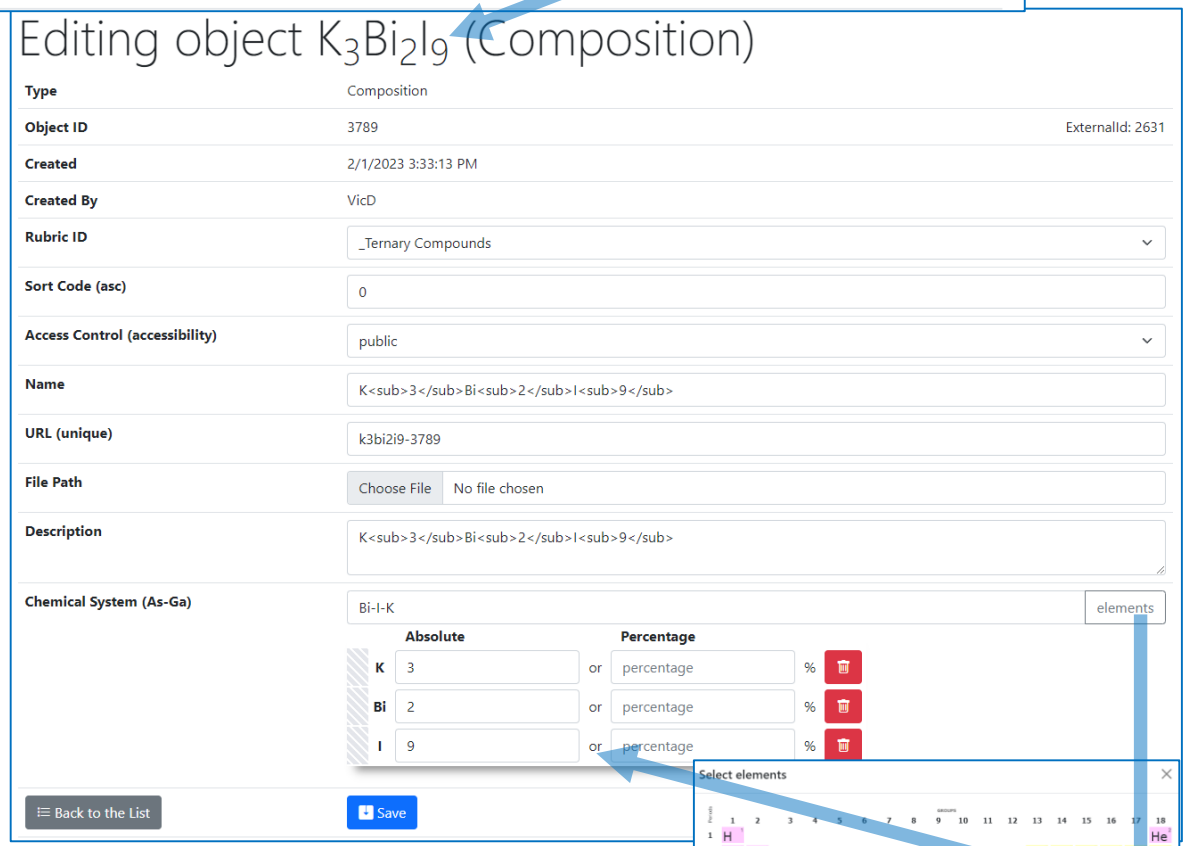


https://demo.mdi.ruhr-uni-bochum.de/

RDMS List Types: control



List of all objects (of type Composition):



Control (common to all objects):

- Rubric Id;
- Sort Code;
- Access;
- Name;
- URL (adjustable*);
- Attach File (document);
- Description.

1

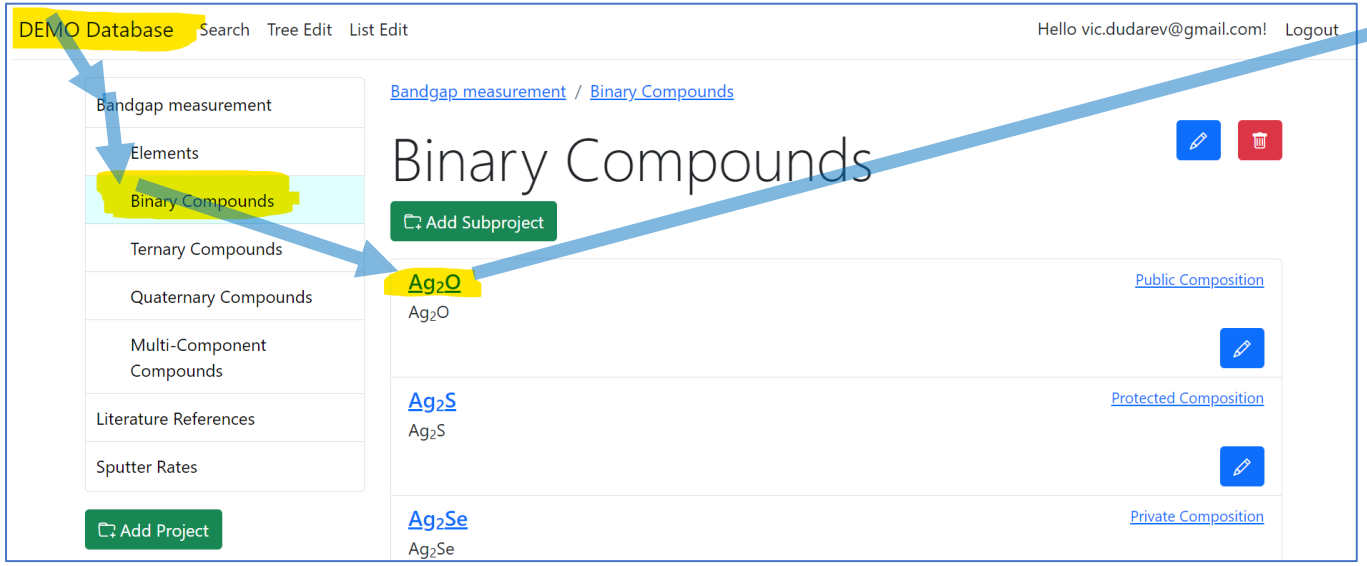
2

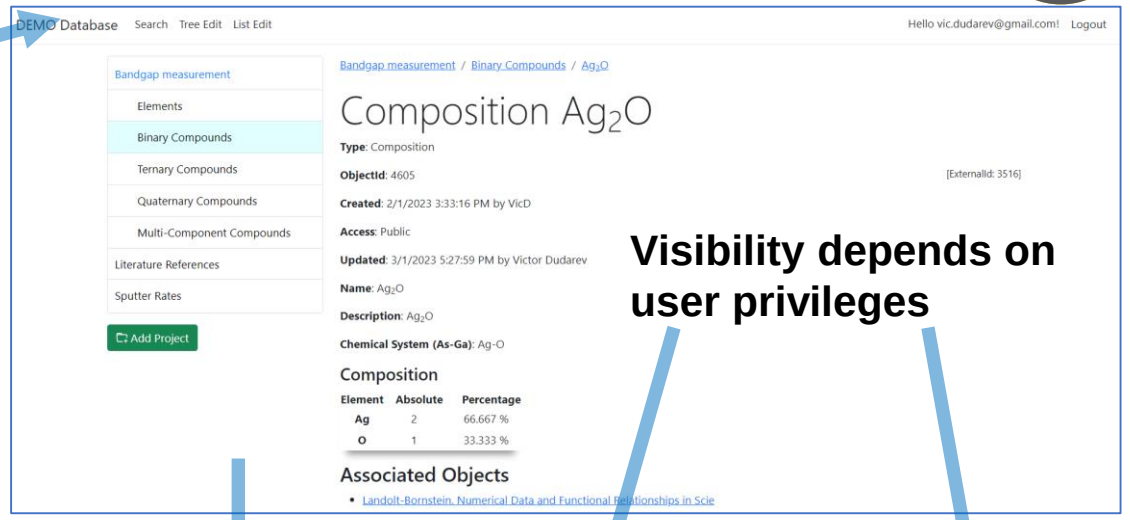
Type specific control (e.g. Composition):

- Chemical System;
- Composition (elements containment).

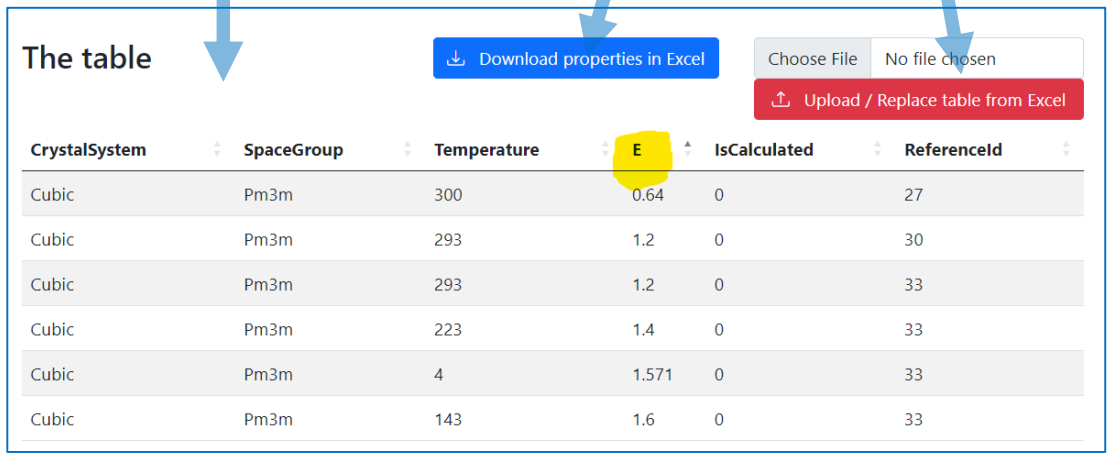
https://demo.mdi.ruhr-uni-bochum.de/

RDMS List Types: display





Visibility depends on user privileges



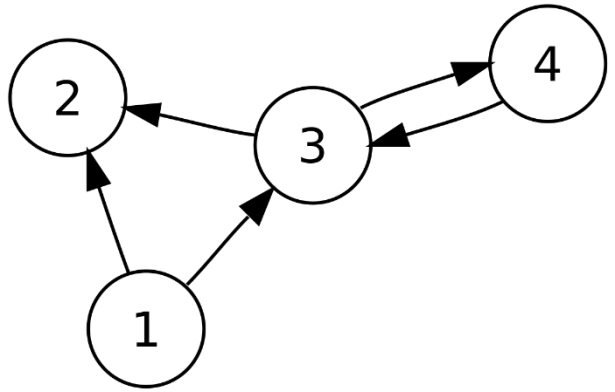
CrystalSystem	SpaceGroup	Temperature	E	IsCalculated	ReferenceId
Cubic	Pm3m	300	0.64	0	27
Cubic	Pm3m	293	1.2	0	30
Cubic	Pm3m	293	1.2	0	33
Cubic	Pm3m	223	1.4	0	33
Cubic	Pm3m	4	1.571	0	33
Cubic	Pm3m	143	1.6	0	33

UI Experience:

- Select project (rubric with objects)
- Select object of interest
- Follow the link and explore object data and properties

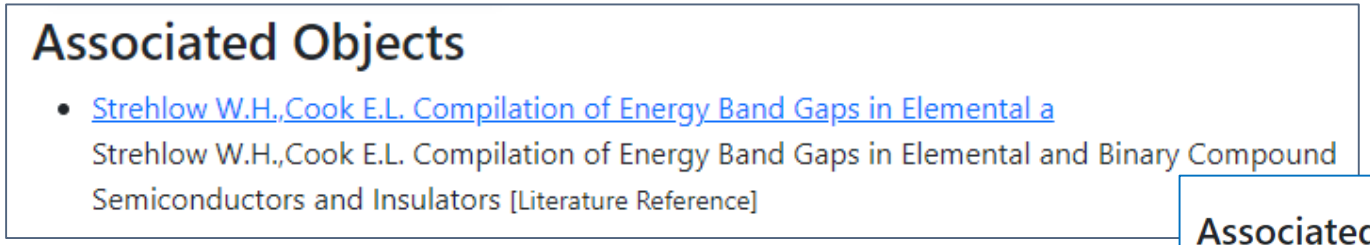
https://demo.mdi.ruhr-uni-bochum.de/object/ag2o-4605

RDMS: Interlink Objects



INF(N)

Example: literature reference for published data on composition:



To add associated objects:

0) Go to “List edit” or select object to edit

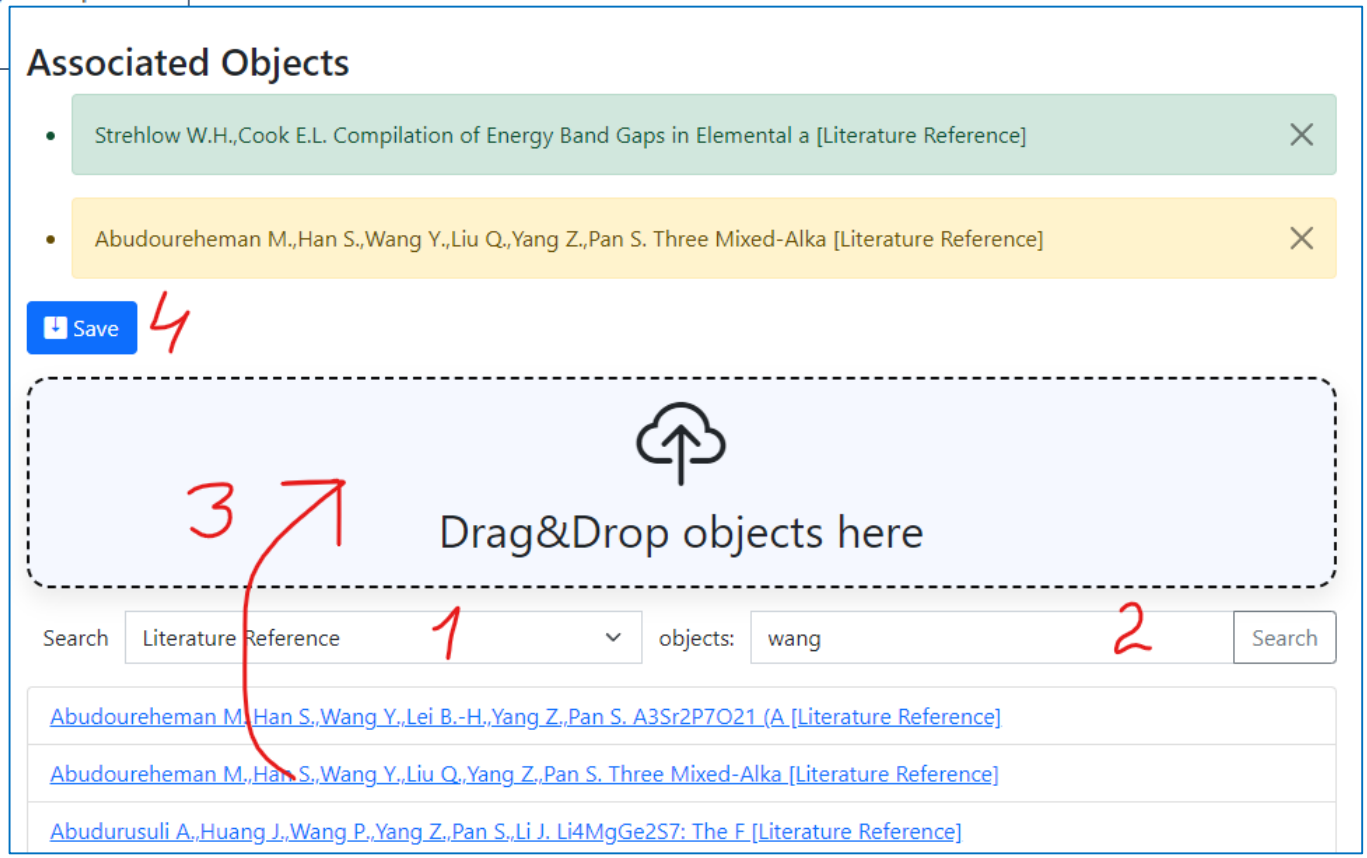
1) Select associated object type (optional filter)

2) Input search phrase, contained in the Name of desired object to be associated

3) Drag & Drop desired object(s) from search result list to the area / adjust list

4) Save changes

Important: **Reverse associations** allow to browse reverse-directional associations, such as finding all compounds mentioned in a particular literature reference.



User view: https://demo.mdi.ruhr-uni-bochum.de/object/ag2s-4870
Admin view: https://demo.mdi.ruhr-uni-bochum.de/adminobject/edititem/4870

RDMS Object's Flexibility: Extended Properties

INF(N)

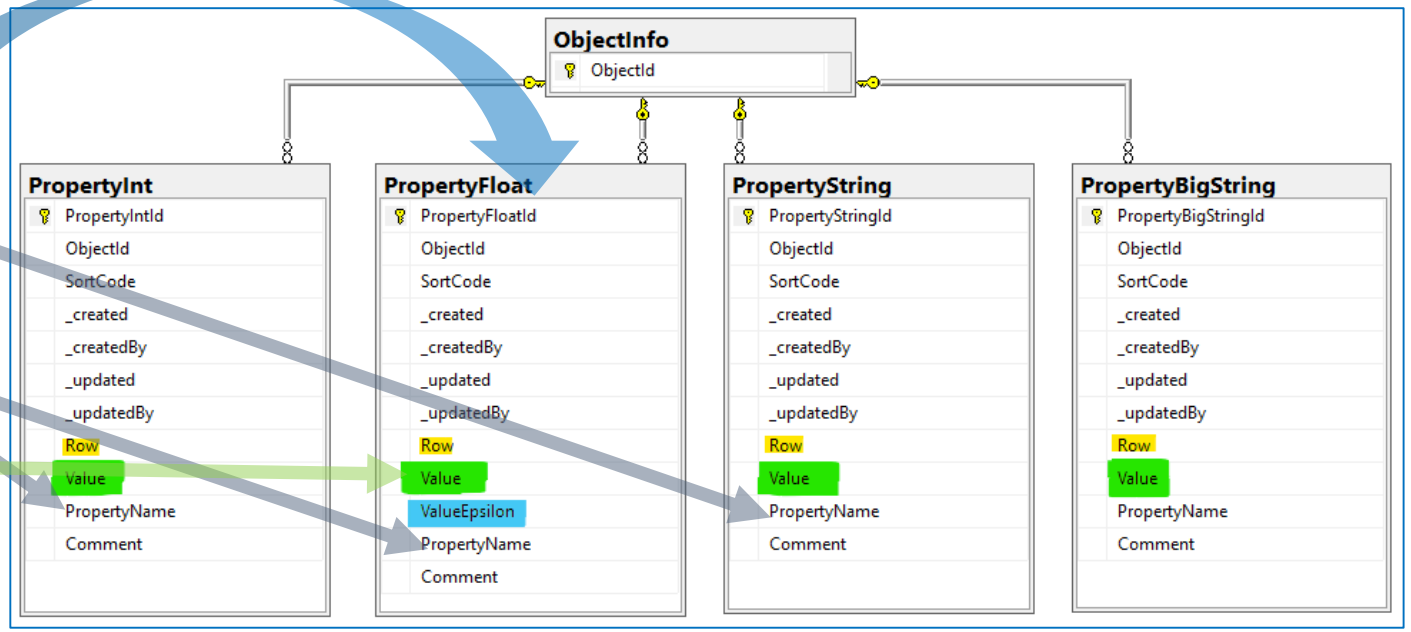
Use case: add additional properties values (or even a table) to object and afterward make search on them.

OR

When the existing fields set is not enough…

Task: associate table with object

Modification	CrystalSystem	SpaceGroup	Temperature	E	IsCalculated	ReferenceId	Comment
α	Monoclinic	$P2_1/c$	300	0.87	0	581	Optical absorption spectroscopy. Direct transition. Thin film, single crystal
Row=1							
β	Cubic	Im3(-)m	77	1.23	0	581	Photoconduction. Thin film, polycristalline
Row=2							
β	Cubic	Im3(-)m	296	1.03	0	581	Photoconduction. Thin film, polycristalline
Row=3							
String x3			Float x2		Int x2		String



Solution: Decompose the table so, that each cell value is stored in Property* table according to cell value type.

Important attributes:

- **PropertyName** – contains column name
- **Value** – contains cell value
- **Row** – contains row number (1, 2, 3, …). If no row number specified – value is threated as object's property

User view: https://demo.mdi.ruhr-uni-bochum.de/object/ag2s-4870
Admin view: https://demo.mdi.ruhr-uni-bochum.de/adminobject/edititem/4870

RDMS Extended Properties Control

Extended Properties Control for object is available according to the object access managements in administration panel for objects.

Property values

Floating-point properties

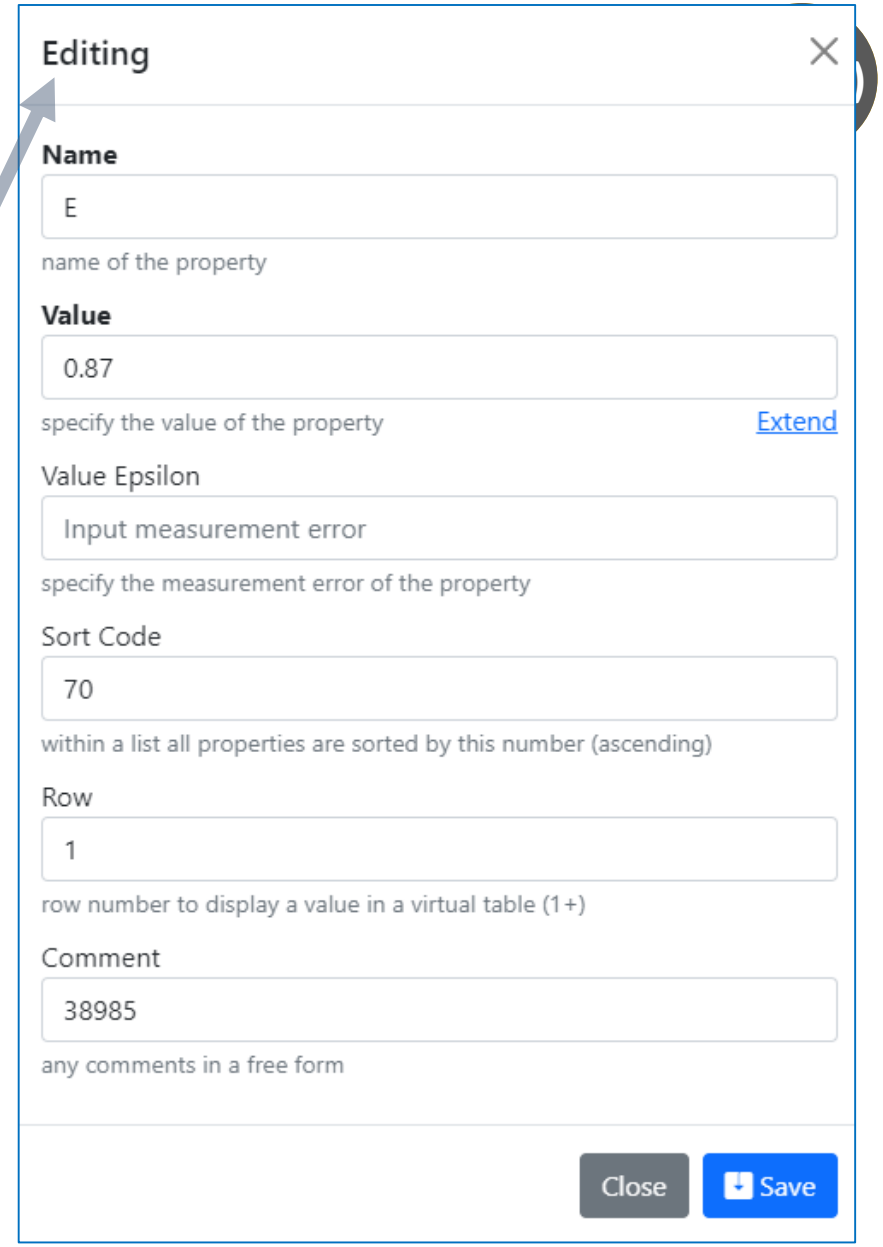
Row	Name	Value	Epsilon	Comment
1	Temperature	300		38985
1	E	0.87		38985
2	Temperature	77		38986

Integer properties

Row	Name	Value	Comment
1	IsCalculated	0	38985
1	ReferenceId	581	38985

String properties

Row	Name	Value	Comment
1	Modification	α	38985
1	CrystalSystem	Monoclinic	38985
1	SpaceGroup	P2₁/c	38985
1	Comment	Optical absorption spectroscopy. Direct transition. Thin film, single crystal	38985



User view: https://demo.mdi.ruhr-uni-bochum.de/object/ag2s-4870
Admin view: https://demo.mdi.ruhr-uni-bochum.de/adminobject/edititem/4870

RDMS Extended Properties: Export and Import

Task: make properties modification easy

Steps:

1) **Download** data
2) Open in Excel and **edit**
3) **Upload** changed data back to the RDMS

The table

Download properties in Excel | Choose File | No file chosen | Upload / Replace table from Excel

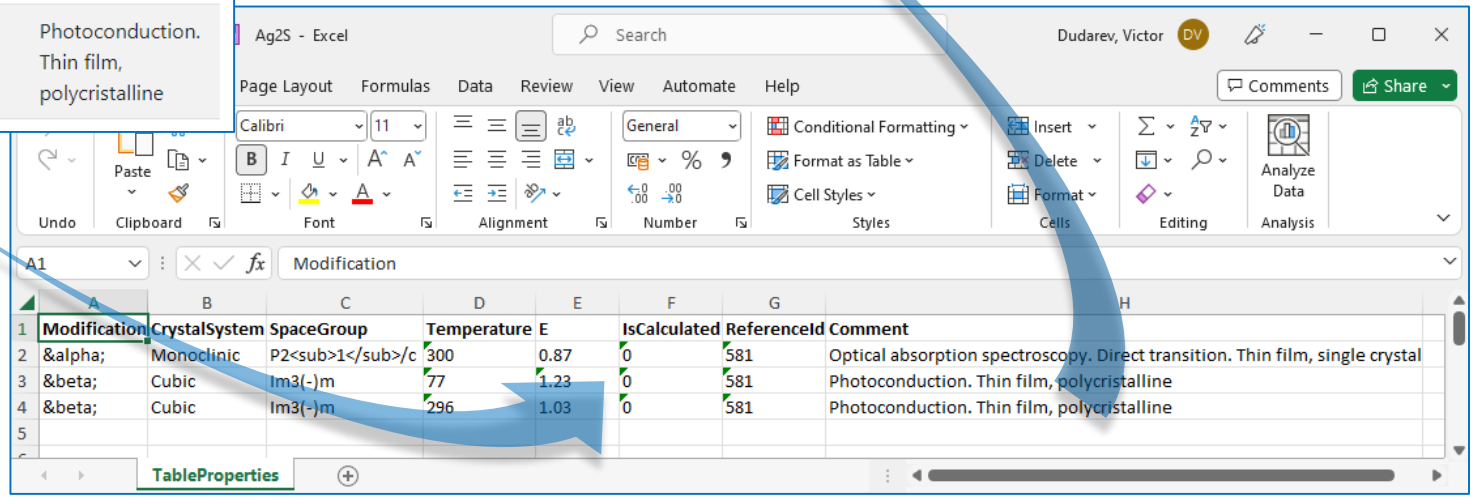
Modification	CrystalSystem	SpaceGroup	Temperature	E	IsCalculated	ReferenceId	Comment
α	Monoclinic	$P2_1/c$	300	0.87	0	581	Optical absorption spectroscopy. Direct transition. Thin film, single crystal
β	Cubic	Im3(-)m	77	1.23	0	581	Photoconduction. Thin film, polycristalline
β	Cubic	Im3(-)m	296	1.03	0	581	Photoconduction. Thin film, polycristalline

https://demo.mdi.ruhr-uni-bochum.de/object/ag2s-4870

Export

Import

(requires permissions)

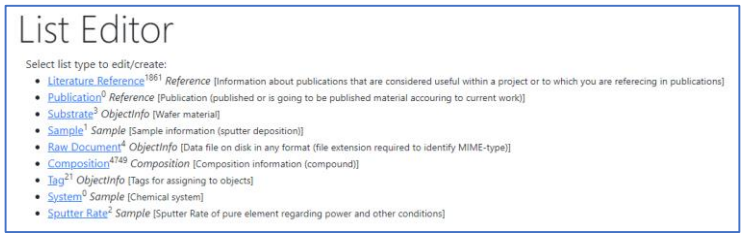


Modification	CrystalSystem	SpaceGroup	Temperature	E	IsCalculated	ReferenceId	Comment
α	Monoclinic	P2₁/c	300	0.87	0	581	Optical absorption spectroscopy. Direct transition. Thin film, single crystal
β	Cubic	Im3(-)m	77	1.23	0	581	Photoconduction. Thin film, polycristalline
β	Cubic	Im3(-)m	296	1.03	0	581	Photoconduction. Thin film, polycristalline

RDMS Extended Properties: Templates for data types

INF(N)

Task: define Template for properties with respect to particular type



Steps:

1) After creating a data type (e.g. based on existing table) **create a "_Template" object** of given type: https://demo.mdi.ruhr-uni-bochum.de/adminobject/list/11

2) **Add Columns** of desired table to corresponding Properties (w.r.t. data type: Float / Int / String / BigString) **with Row=-1** – predefined template: https://demo.mdi.ruhr-uni-bochum.de/adminobject/edititem/6650

3) Create a new object and in UI **press "Download properties in Excel" button**: https://demo.mdi.ruhr-uni-bochum.de/object/all-sputter-rates-for-test-6651 to get template in Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Material	Trust	Deposition Date	Chamber	LogbookId	Rate, nm/sec	Time, sec	Thickness, nm	Normalized Rate, nm/W*sec	Power, W	Power Supply	Cathode	Pressure, mTorr	Argon Flow, sccm	Working Distance Cathode Tilt, mm	Working Distance Table Height, mm	Comment
2	String	Int	String	Int	String	Float	Float	Float	Float	Float	String	Int	Float	Float	Float	Float	String

4) **Modify** template in Excel file **and upload** the data: https://demo.mdi.ruhr-uni-bochum.de/object/all-sputter-rates-for-test-6651 (test passed: 16 columns and 468 rows)

RDMS: Search

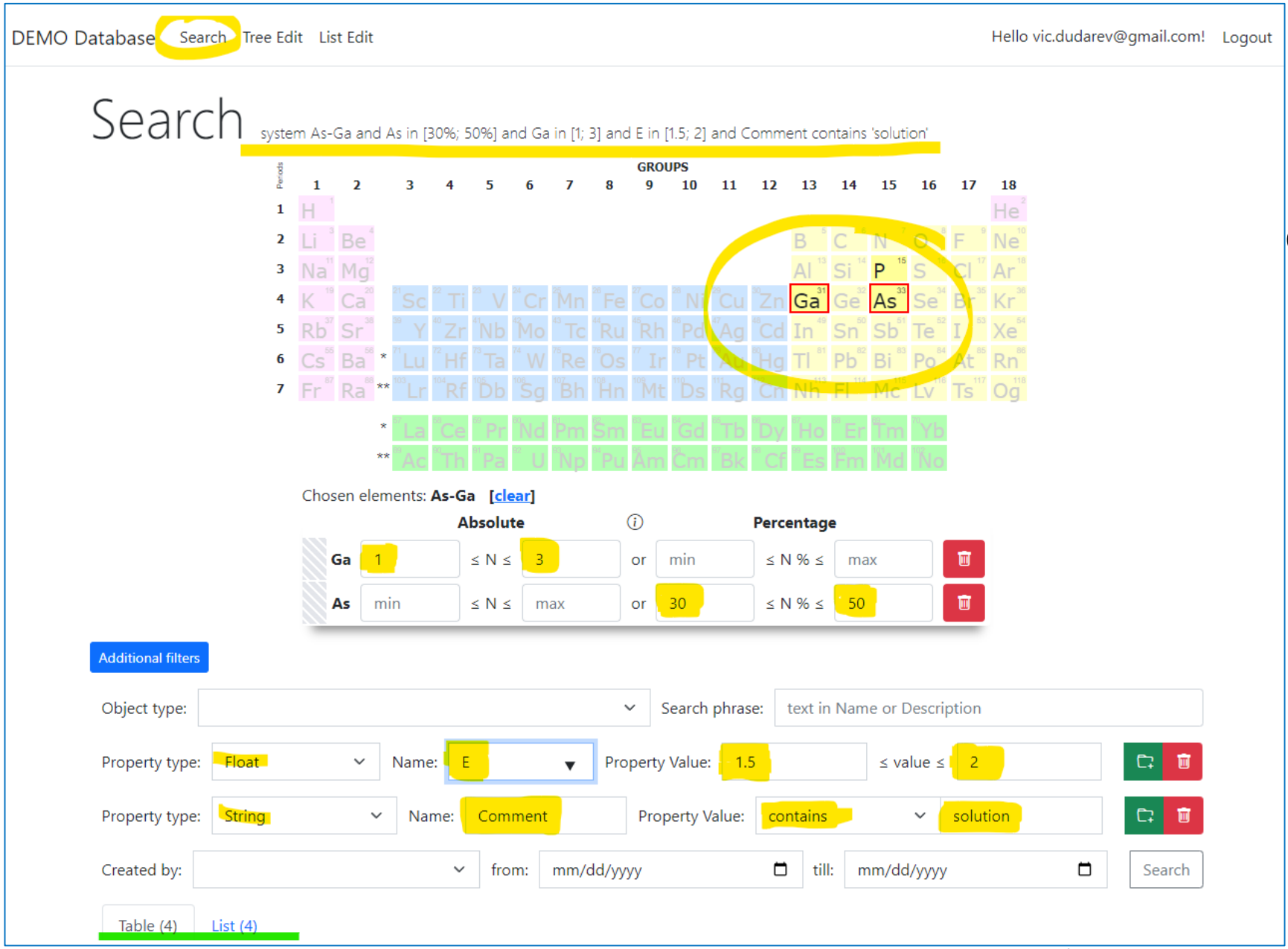
Search on:

1) Chemical system
2) Composition
3) Object type
4) Phrase in object's Name or Description
5) Properties values (all available within tenant)
6) Person created
7) Creation date

Important feature:

persistent URL on search (share search results easily by URL*).

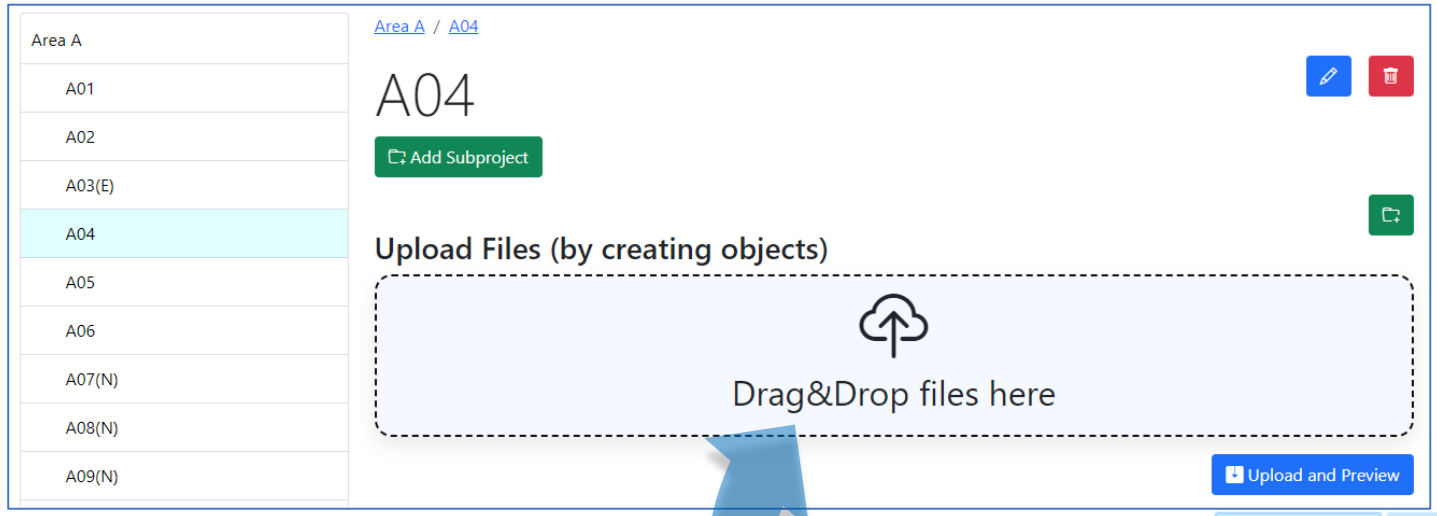
* - respecting current security context



(1) https://demo.mdi.ruhr-uni-bochum.de/search/
(2) https://demo.mdi.ruhr-uni-bochum.de/search/?system=As-Ga&Aspctmin=30&Aspctmax=50&Gaabsmin=1&Gaabsmax=3&pr0name=E&pr0type=Float&pr0min=1.5&pr0max=2&pr1name=Comment&pr1type=String&pr1max=3&pr1str=solution&prcnt=2
(3) https://demo.mdi.ruhr-uni-bochum.de/search/?system=As-Ga&Aspctmin=30&Aspctmax=50&Gaabsmin=1&Gaabsmax=3&pr0name=E&pr0type=Float&pr0min=1.5&pr0max=2&prcnt=1

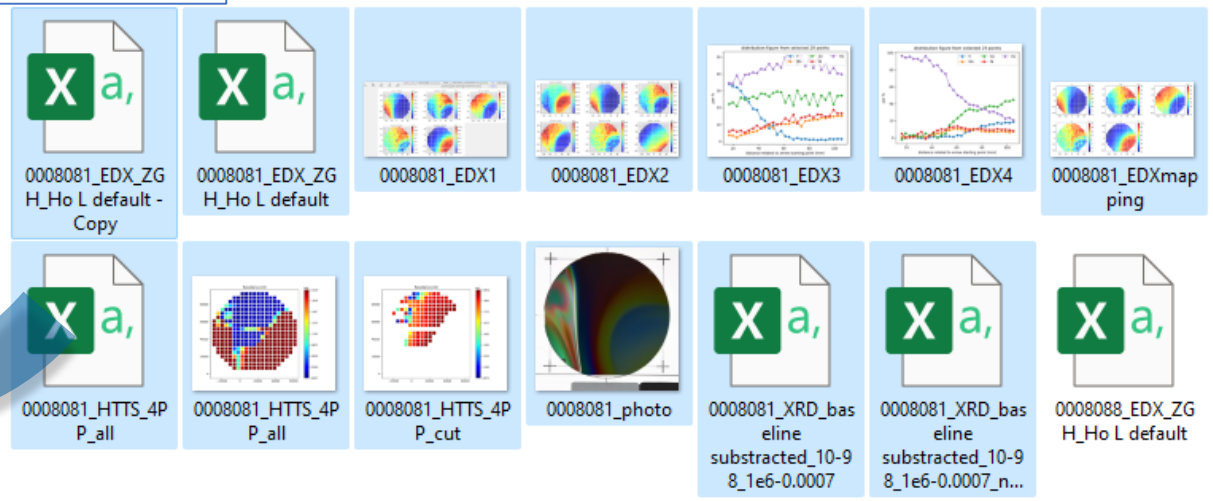
RDMS DropZone Tasks: Data Validation & Import

Task: validate documents & import data



Implemented Features:

- Automatic type detection;
- Uniqueness check;
- Data **validation** (build-in + external);
- Data **import**.



RDMS DropZone: Staged Files (before import)

INF(N)

Objects to create

Files to process

0008081_EDX1.png

Type: EDX Image | Name: 0008081_EDX1 | Sort Code: 10

System.Exception column 6: dfg (unknown column name) 0008081_EDX_ZGH_Ho L default - Copy.csv

Type: EDX CSV | Name: 0008081_EDX_ZGH_Ho L default - Copy | Sort Code: 20

0008081_EDX_ZGH_Ho L default.csv

Type: EDX CSV | Name: 0008081_EDX_ZGH_Ho L default | Sort Code: 30

0008081_HTTS_4PP_all.csv

Type: HTTS CSV | Name: 0008081_HTTS_4PP_all | Sort Code: 40

0008081_HTTS_4PP_all.png

Type: HTTS Image | Name: 0008081_HTTS_4PP_all | Sort Code: 50

0008088_EDX_ZGH_Ho L default.csv

Type: EDX CSV | Name: 0008088_EDX_ZGH_Ho L default | Sort Code: 60

Common properties:

Access Control (accessibility): public | Rubric: _A04

Link as Associated Objects with:

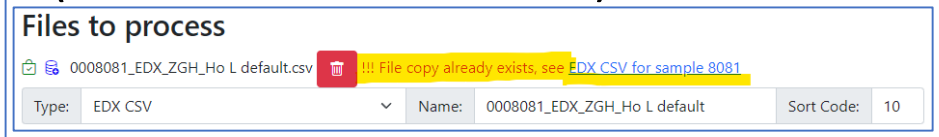
Sample 8081 (V-Mn-Co-Ni-Ho) [Sample]

Search: Sample | objects: 8081 | Search

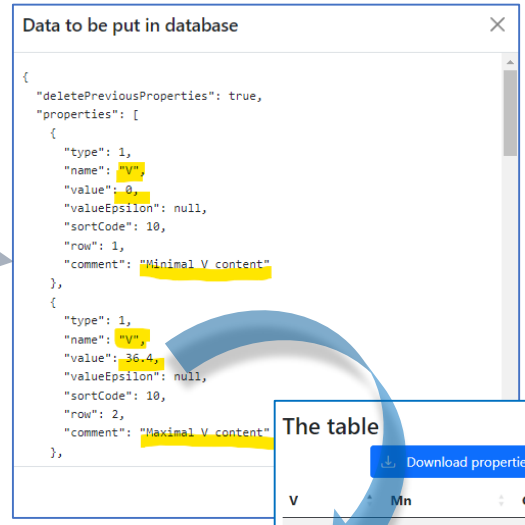
Successful validation of all files required

Create Objects from Files

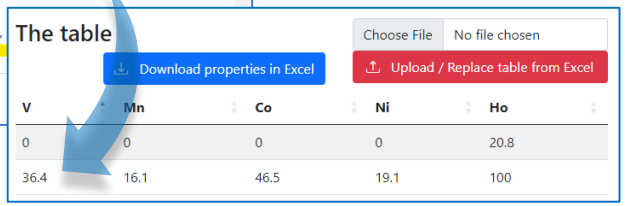
Check for unique documents (based on SHA256 hash):



Check data to import:



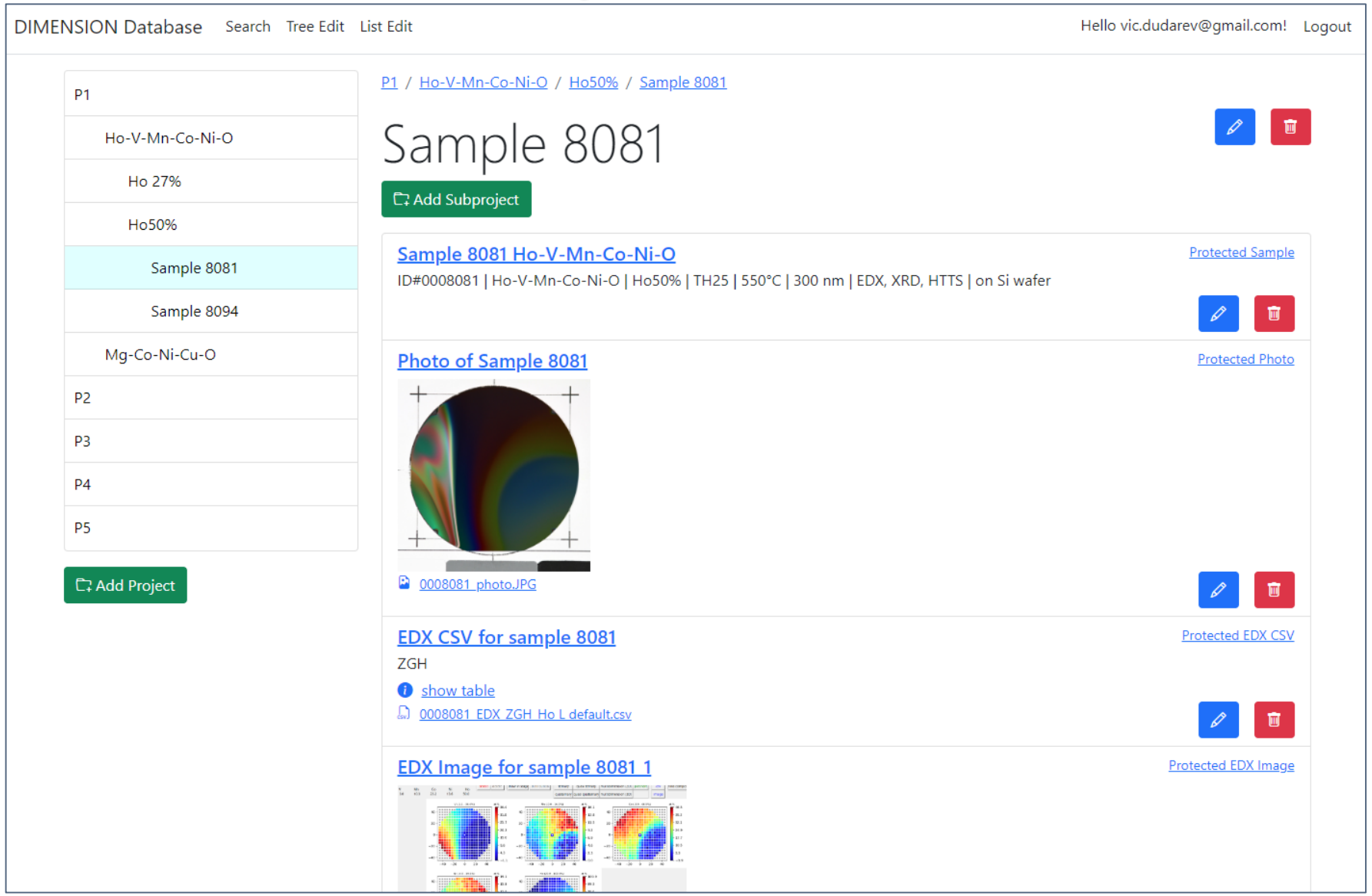
Properties (after import):



Workflow:

- validate every file;
- check/adjust all properties;
- **Import files**.

RDMS: Created Objects

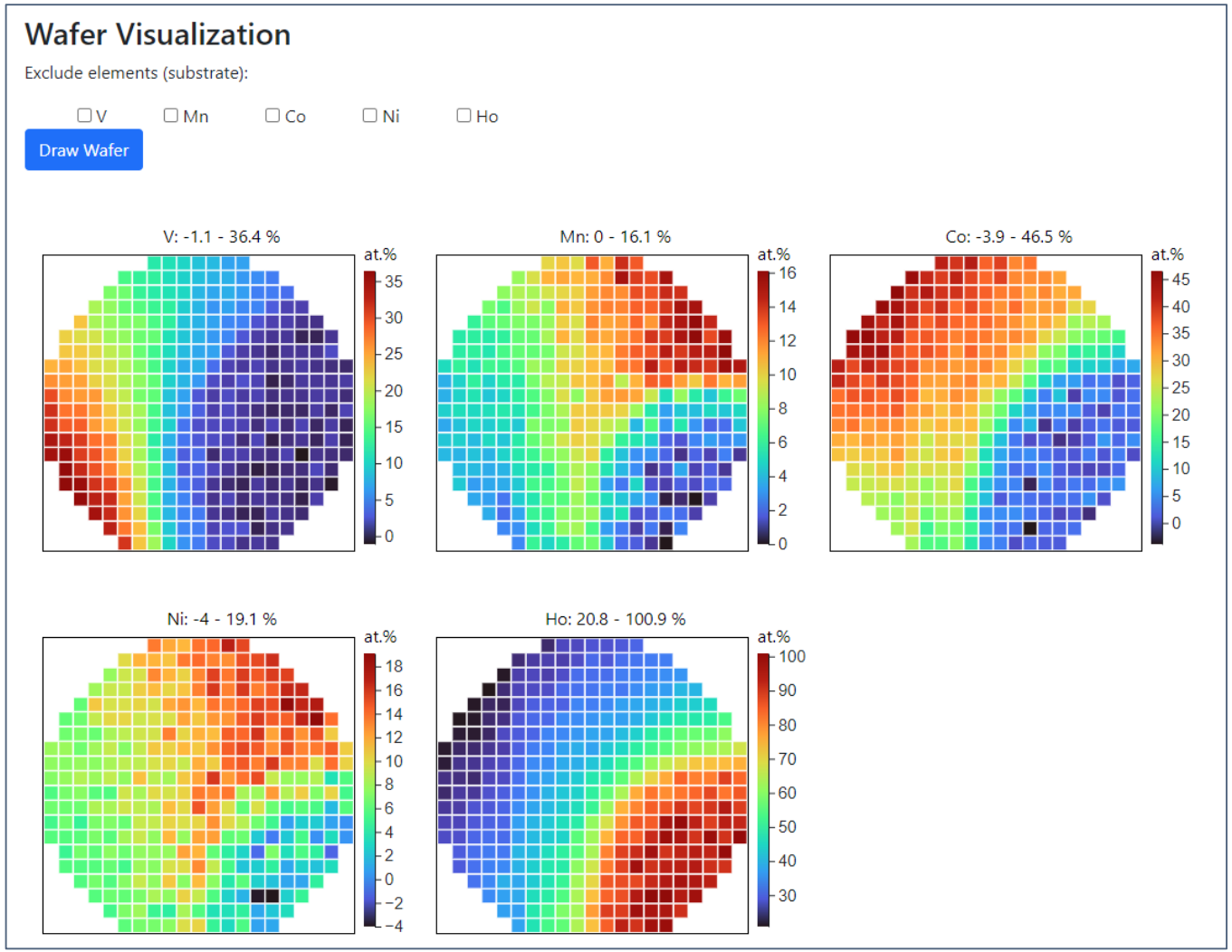


After import:

- check all created objects (names, access level, etc…);
- share a link.

RDMS: CSV Visualization

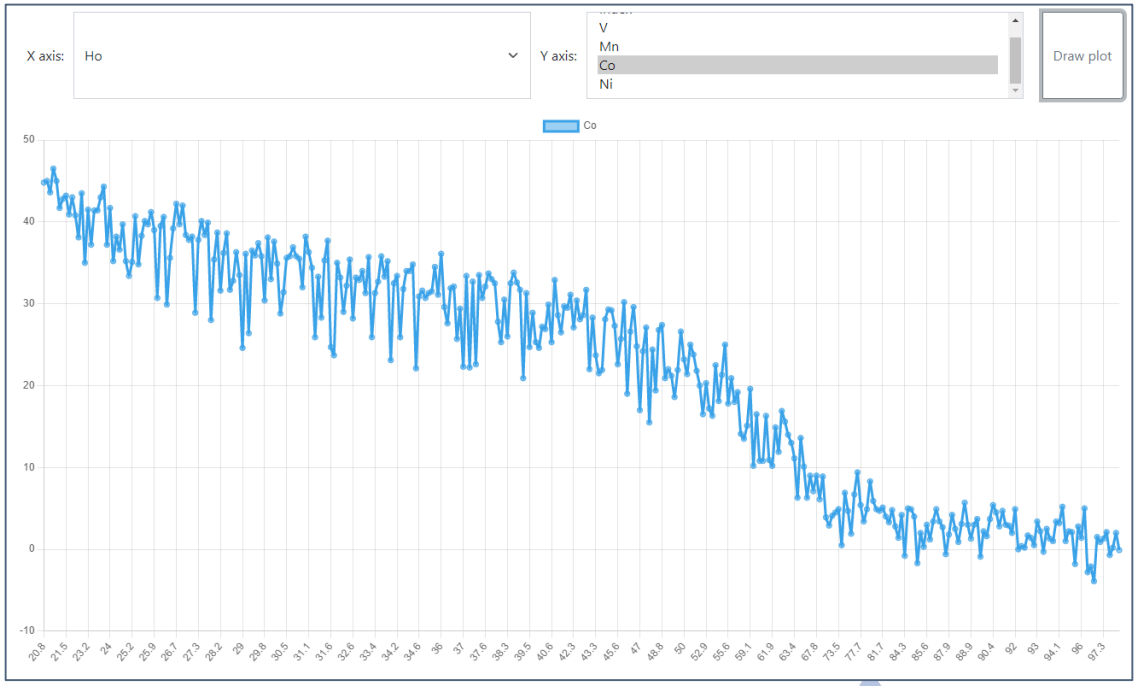
Wafer Visualization



CSV Viewer: EDX CSV for sample 8081

Index	V	Mn	Co	Ni	Ho
297	18.8	7.5	44.8	8	20.8
261	22.4	5.6	45	6.2	20.9
280	23.2	6	43.6	6.2	21.1
312	16.7	8	46.5	7.7	21.1
262	21.4	5.8	45	6.6	21.2
221	24.3	5	41.7	7.8	21.3
243	22.9	5.9	42.8	7	21.4
242	23	5.7	43.2	6.6	21.5

Configurable XY-charts



https://dim.mdi.ruhr-uni-bochum.de/file/csv/6691

RDMS Extensibility: Additional Object Types

INF(N)

Task: introduce new object types

New Data Type

Based on existing tables
(hierarchical and non-hierarchical)

To be done manually by admin

Approximate time: 3 minutes

Just add a new table row with object specification required (assign validator & data extractor)

Add new type to TypeInfo table

VicRUBNote.RUB_INF - dbo.TypeInfo

	TypeId	IsHierarchical	TypeIdForRubric	TypeName	TableName
	1	True	*NULL*	Organisation Structure	RubricInfo
	2	True	*NULL*	Project	RubricInfo
	3	False	2	Literature Reference	Reference
	4	False	2	Publication	Reference
	5	False	2	Substrate	ObjectInfo
	6	False	2	Sample	Sample
	7	False	2	Raw Document	ObjectInfo
	8	False	2	Composition	Composition
	9	False	2	Tag	ObjectInfo
	10	False	2	System	Sample

New Data Structures
(new table & code to fit particular requirements)

To be done manually

Approximate time: 1 week

Tasks: develop and add new table; define read / write logic with respect to mandatory (or optional) fields and types; reflect to OOP model and provide consistency check (validators); develop import / export and search facilities, etc.

Question: It seems that all can be done with extended properties.
Why to use new data structures?

Short Answer: **Performance** (especially if number of new type objects is expected to exceed 10K)

RDMS Extensibility: Type Configuration and Flexibility

INF(N)

Task: introduce new types and configure them (validation and data import)

All types

For Hierarchical data types Table must be RubricInfo

Id	Name	Table	Description	Validation Schema / Data Schema	Actions
8	**Composition**	Composition	Composition information (compound)	type:TypeValidationLibrary.TypeValidator_Ok /	
13	**EDX CSV**	ObjectInfo	CSV of EDX with Header row (e.g. "Index,V,Mn,Co,Ni,Ho")	type:TypeValidationLibrary.TypeValidator_EDX_CSV / type:TypeValidationLibrary.TypeValidator_EDX_CSV	
15	**EDX Image**	ObjectInfo	Image (bitmap) that refers to EDX	type:TypeValidationLibrary.TypeValidato	
14	**HTTS CSV**	ObjectInfo	CSV of HTTS with Header row (e.g. "Index,Resistance")	type:TypeValidationLibrary.TypeValidato	

Responses

Code	Description	Links
200	Success	*No links*

Media type: application/json

Controls Accept header.

Example Value | Schema

```
{
  "code": 0,
  "message": "string",
  "warning": "string"
}
```

Data to be put in database

```
{
  "deletePreviousProperties": true,
  "properties": [
    {
      "type": 1,
      "name": "V",
      "value": 0.0,
      "valueEpsilon": null,
      "sortCode": 10,
      "row": 1,
      "comment": "Maximal V content"
    },
    {
      "type": 1,
      "name": "Mn",
      "value": 30.0,
      "valueEpsilon": null,
      "sortCode": 10,
      "row": 2,
      "comment": "Maximal V content"
    },
```

Close

Editing type EDX CSV

TypeId	13
Type Name	EDX CSV
Table Name (Data Structure)	ObjectInfo
Hierarchical classifier	Project
Validation Schema	type:TypeValidationLibrary.TypeValidator_EDX_CSV
Data Schema	type:TypeValidationLibrary.TypeValidator_EDX_CSV
Description	CSV of EDX with Header row (e.g. "Index,V,Mn,Co,Ni,Ho")

Back to the Type List | Save

Validation & Data Schema Protocols:

- **type**: built-in type (object instantiated by class name);
- **https:** URL to a REST Web Service (OpenAPI specification).

```
66 references
public class TypeValidatorResult
{
    12 references
    public int Code { get; set; }
    11 references
    public string? Message { get; set; }
    7 references
    public string? Warning { get; set; }
}
```

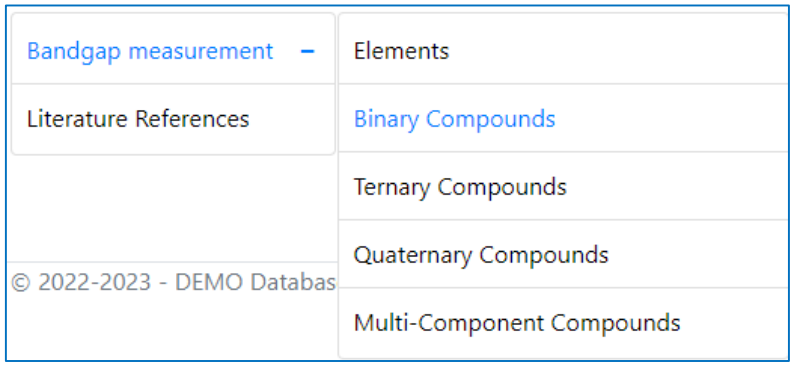
RDMS Demonstration: Bandgap Data

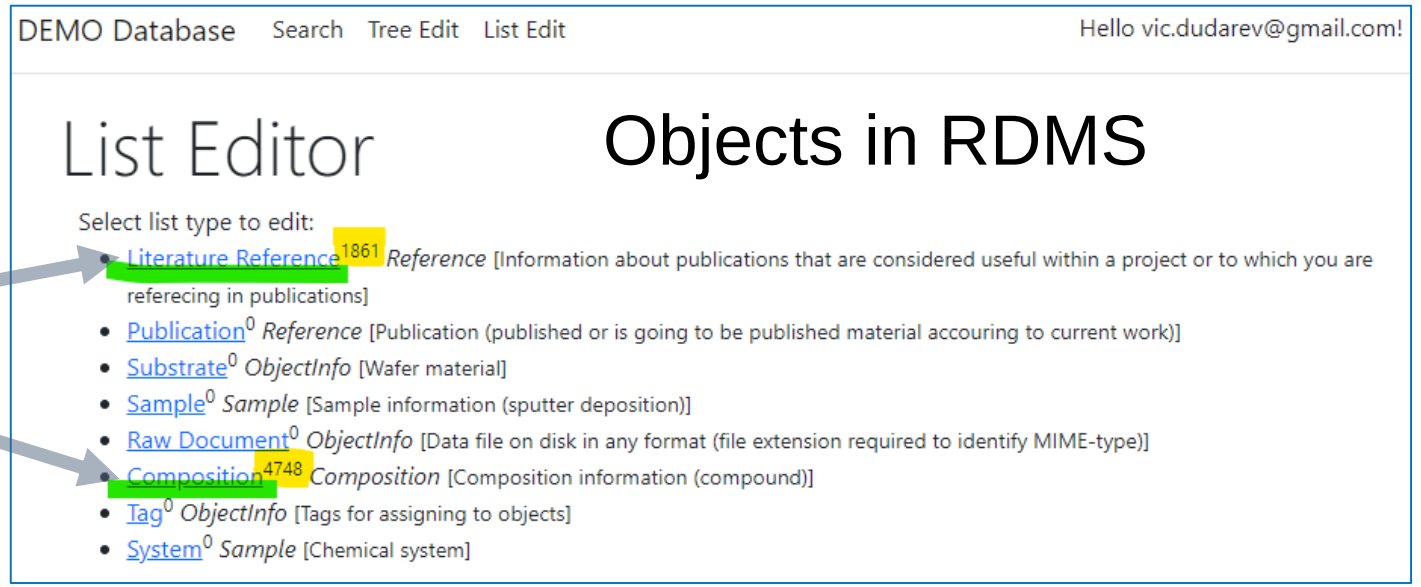
Task: check RDMS on Bandgap data from bg.imet-db.ru to test functionality and features.

Source data in numbers:

- Substances: 4748
- Bandgap records: 10264
- Literature references: 1861

Projects structure:





Bandgap values (~10K) are stored in extended properties tables:

- BigString: 0 records
- String: **28992** records
- Int: **20527** records
- Float: **15149** records

Total: **71 277** records in demo tenant for bandgap data.

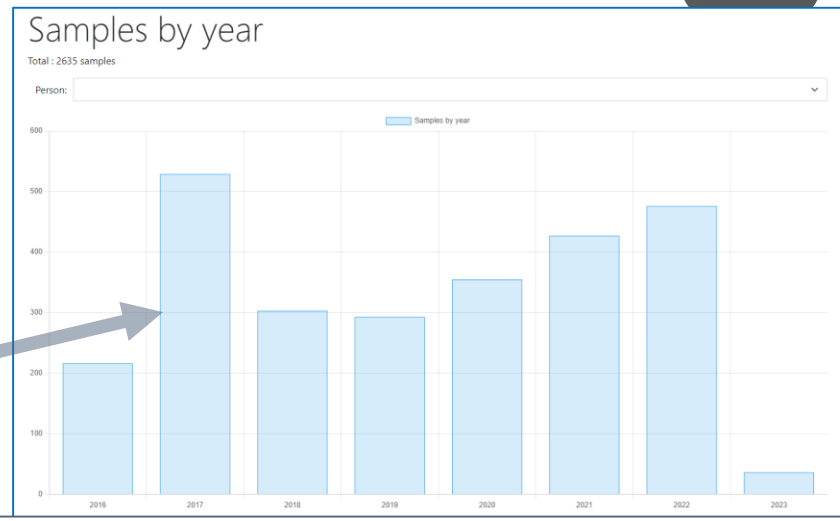
(1) https://demo.mdi.ruhr-uni-bochum.de/
(2) Source data: https://bg.imet-db.ru/

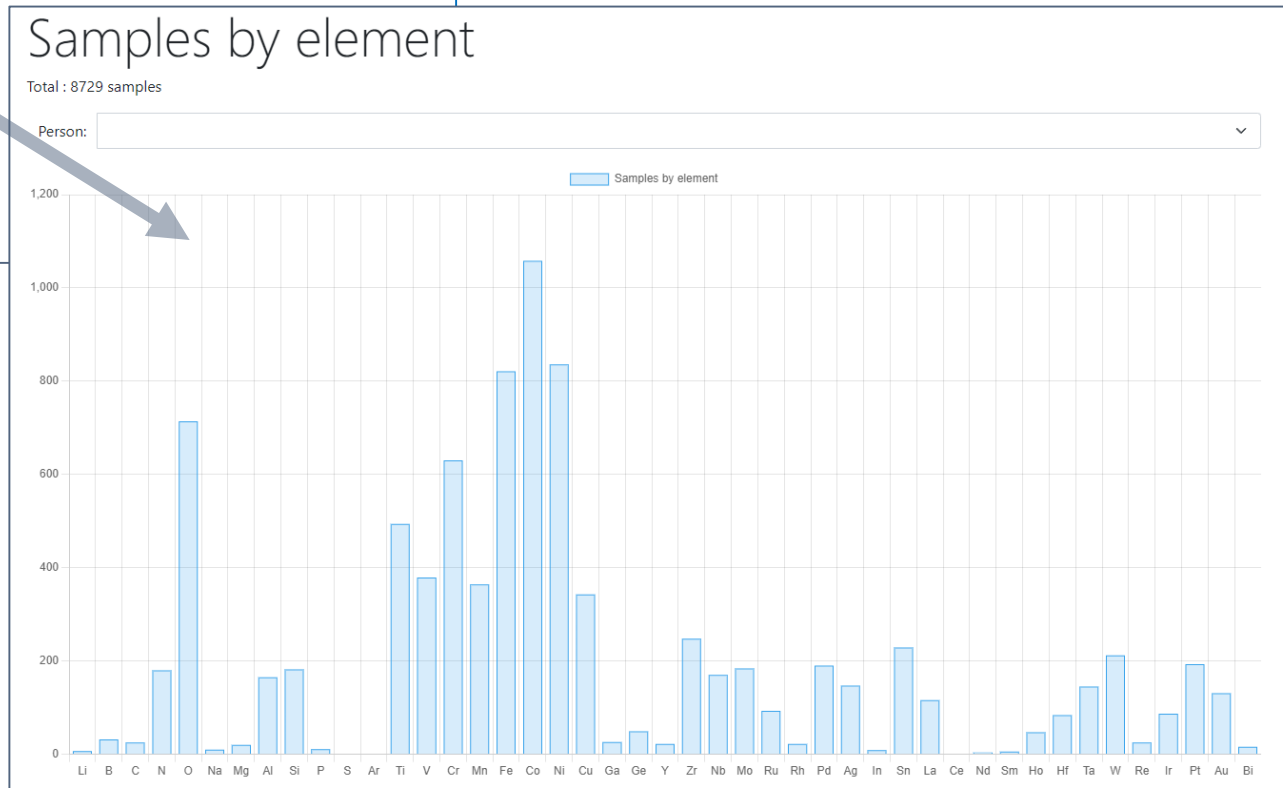
RDMS: the Next Steps

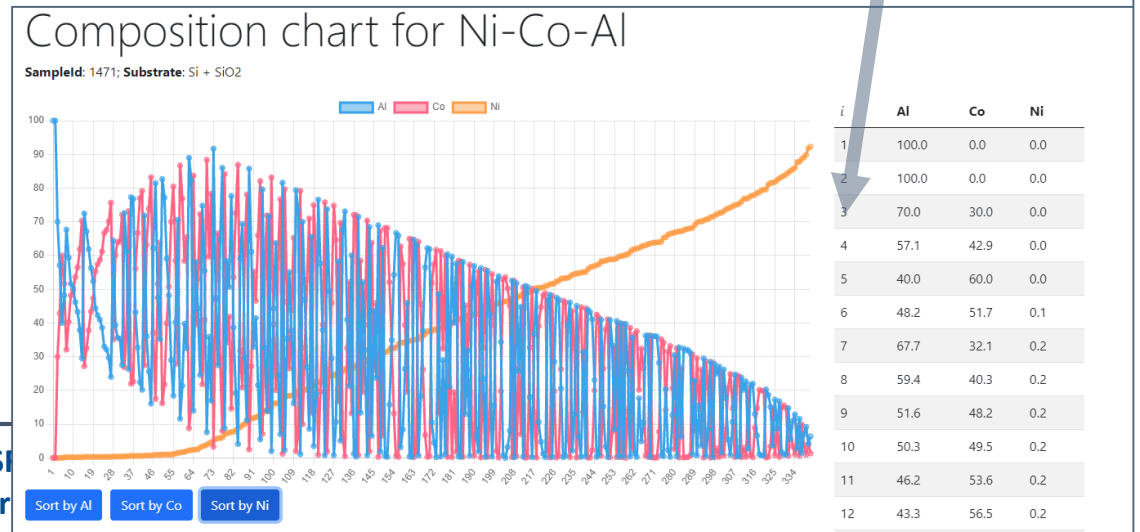
INF(N)

Tasks (to be done):

- refine UI for PowerUsers, incorporated in regular (end-user) interface + UI w.r.t. data types
- **API to upload data from measurement devices (Bandgap, Resistance)**
- Implement Reports / Charts / Diagrams on stored data in various aspects (like in WebCompact)
- Existing object types Bulk Import from CSV files (with CSV-schema validation) and Export







i	Al	Co	Ni
1	100.0	0.0	0.0
2	100.0	0.0	0.0
3	70.0	30.0	0.0
4	57.1	42.9	0.0
5	40.0	60.0	0.0
6	48.2	51.7	0.1
7	67.7	32.1	0.2
8	59.4	40.3	0.2
9	51.6	48.2	0.2
10	50.3	49.5	0.2
11	46.2	53.6	0.2
12	43.3	56.5	0.2

Source Control (GitLab)

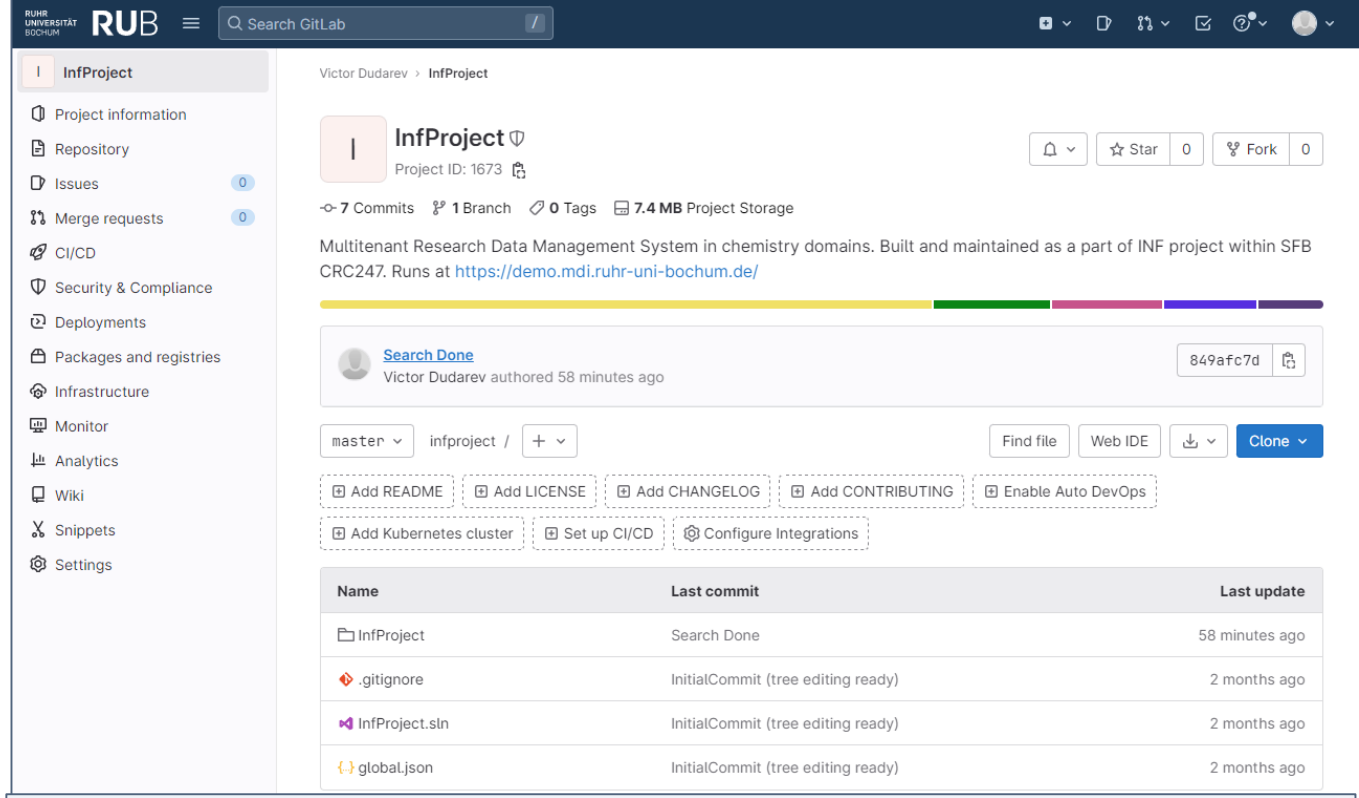
Everybody's welcome!

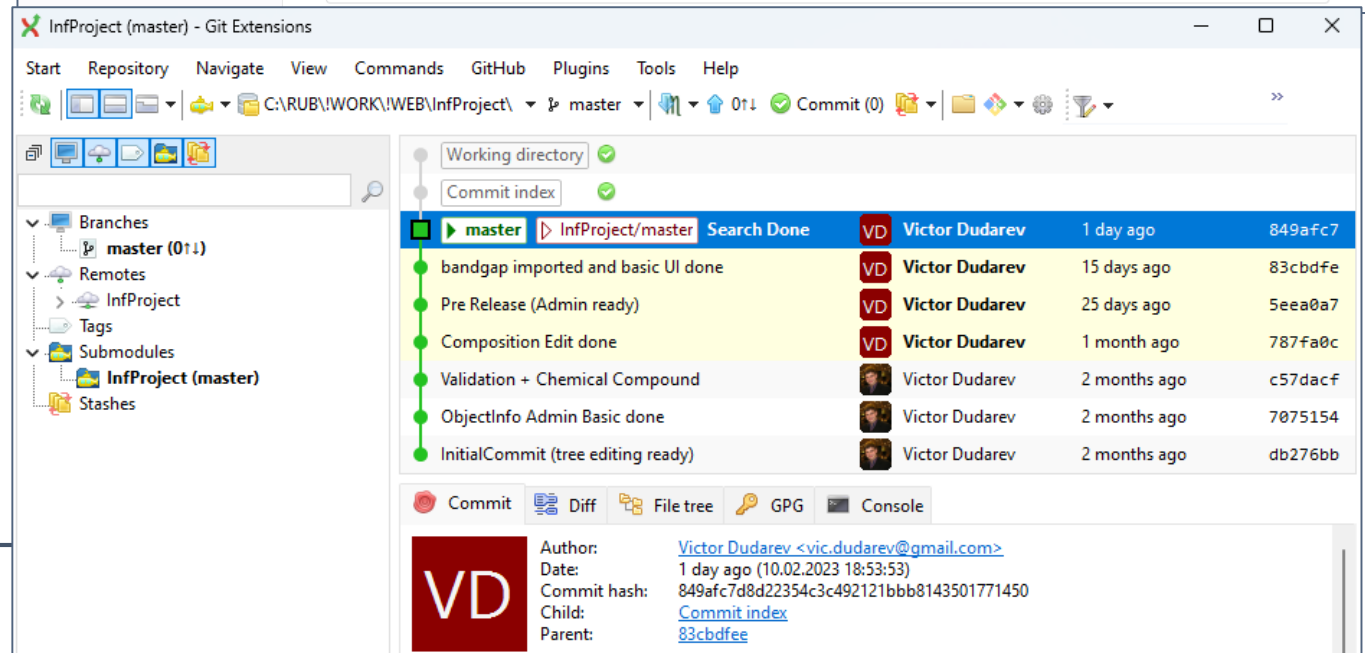
Core INF project:
https://gitlab.ruhr-uni-bochum.de/vic/infproject

Web Compact:
https://gitlab.ruhr-uni-bochum.de/vic/WebCompact

Shared projects (INF & WebCompact):
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





RDMS Playground

INF(N)

https://demo.mdi.ruhr-uni-bochum.de/

https://inf.mdi.ruhr-uni-bochum.de/

Playground to test RDMS (multiple tenants with shared users list):
https://demo.mdi.ruhr-uni-bochum.de/
https://inf.mdi.ruhr-uni-bochum.de/

Login and Password (**User** role): User1@user.org
Login and Password (**PowerUser** role): PowerUser1@user.org
Login and Password (**Administrator** role): Admin1@user.org

Thanks for your kind attention

INF(N)

Demonstration

Q&A

You are welcome to write your questions (and I'll answer) and provide feedback:
https://shorturl.at/ikNST

Victor Dudarev
Victor.Dudarev@rub.de

FAIR: **Findable**, Accessible, Interoperable, Reusable

INF(N)

FAIR Principles for scientific data management

1) Findable

First step in (re)using data: Find them (humans and computers)

- F1. (Meta) data are assigned a **globally unique and persistent identifier** → **URI**
- F2. Data are described with **rich metadata** (as defined by R1 below) → **JSON Schema**
- F3. Metadata clearly and explicitly include the identifier of the data they describe → **… of course**
- F4. (Meta)data are registered or indexed in a searchable resource → **schemastore.org/json**

(1) https://www.go-fair.org/ - **FAIR Principles for scientific data management**
(2) URI - Uniform Resource Identifier
(3) JSON - JavaScript Object Notation

FAIR: Findable, Accessible, Interoperable, Reusable

INF(N)

FAIR Principles for scientific data management

2) Accessible

Once the user finds the required data, they need to know how they can be accessed, possibly including authentication and authorization.

- A1. (Meta)data are retrievable by their identifier using a standardized communications protocol
- A1.1 The protocol is open, free, and universally implementable → **HTTPS, Open API**
- A1.2 The protocol allows for an authentication and authorization procedure, where necessary → **OAuth + OpenID Connect**
- A2. Metadata are accessible, even when the data are no longer available → **keep it in mind… Metadata are never deleted!**

(1) HTTPS - HyperText Transfer Protocol Secure
(2) OAuth 2.0 - industry-standard open protocol for authorization
(3) OpenID Connect - industry-standard open authentication protocol (works on top of OAuth 2.0)

FAIR: Findable, Accessible, Interoperable, Reusable

INF(N)

FAIR Principles for scientific data management

3) Interoperable

The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.

- I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. → **OWL**
- I2. (Meta)data use vocabularies that follow FAIR principles → **reference**
- I3. (Meta)data include qualified references to other (meta)data → **URI**

(1) OWL - W3C Web Ontology Language
(2) URI - Uniform Resource Identifier

FAIR: Findable, Accessible, Interoperable, Reusable

INF(N)

FAIR Principles for scientific data management

4) Reusable

The ultimate goal of FAIR is to optimize the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings.

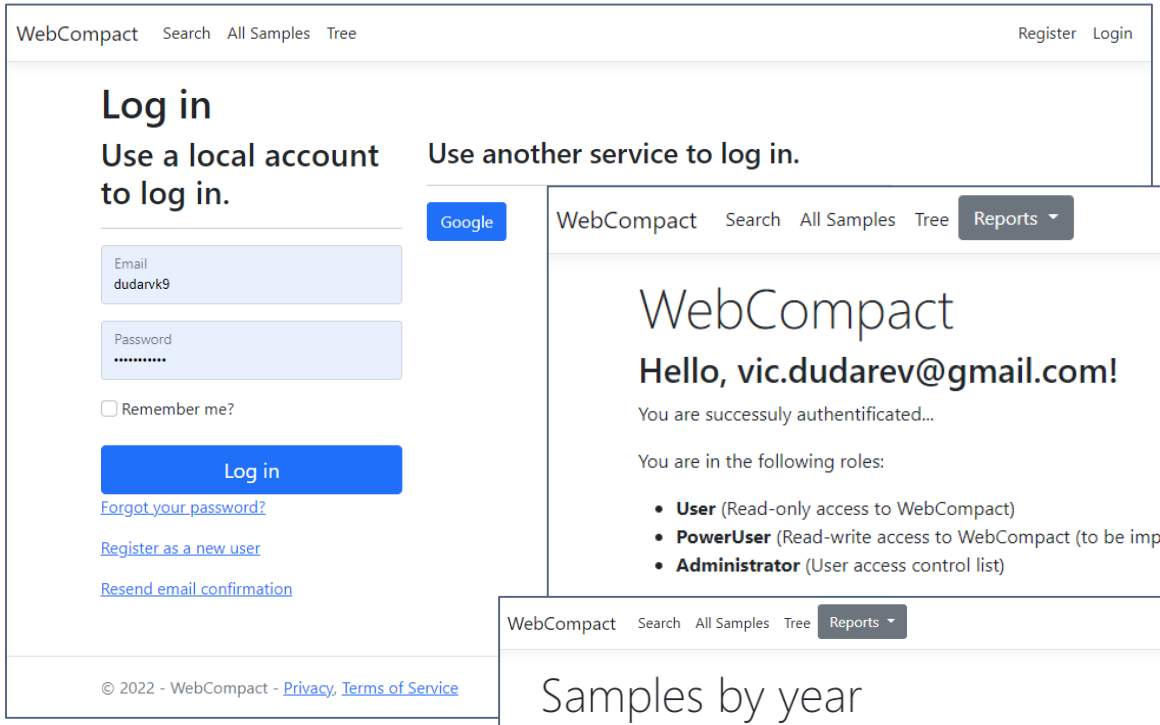
- R1. (Meta)data are richly described with a plurality of accurate and relevant attributes → **JSON Schema**
- R1.1. (Meta)data are released with a clear and accessible data usage license → **GPL**
- R1.2. (Meta)data are associated with detailed provenance → **reference source by URI**
- R1.3. (Meta)data meet domain-relevant community standards → **JSON Schema Validator + Git Flow**

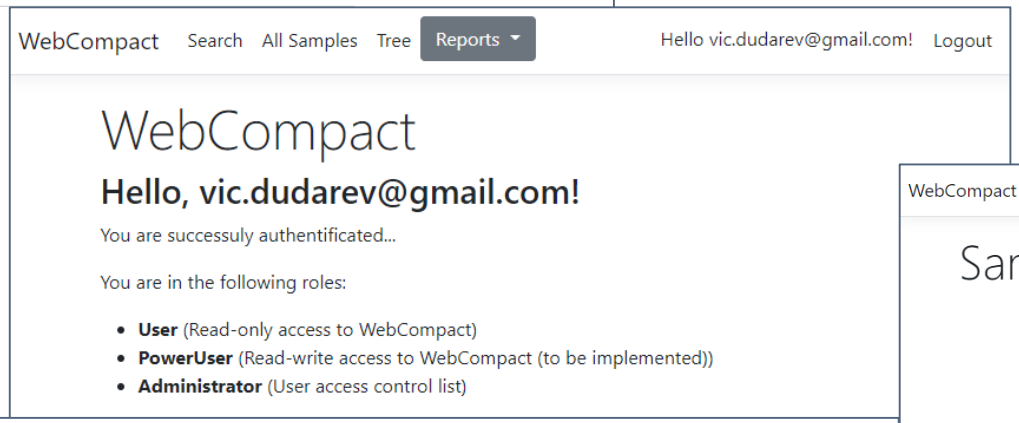
(1) GPL - General Public License
(2) JSON Schema - specification for JSON-based format for defining the structure of JSON data
(3) Git - free and open source distributed version control system

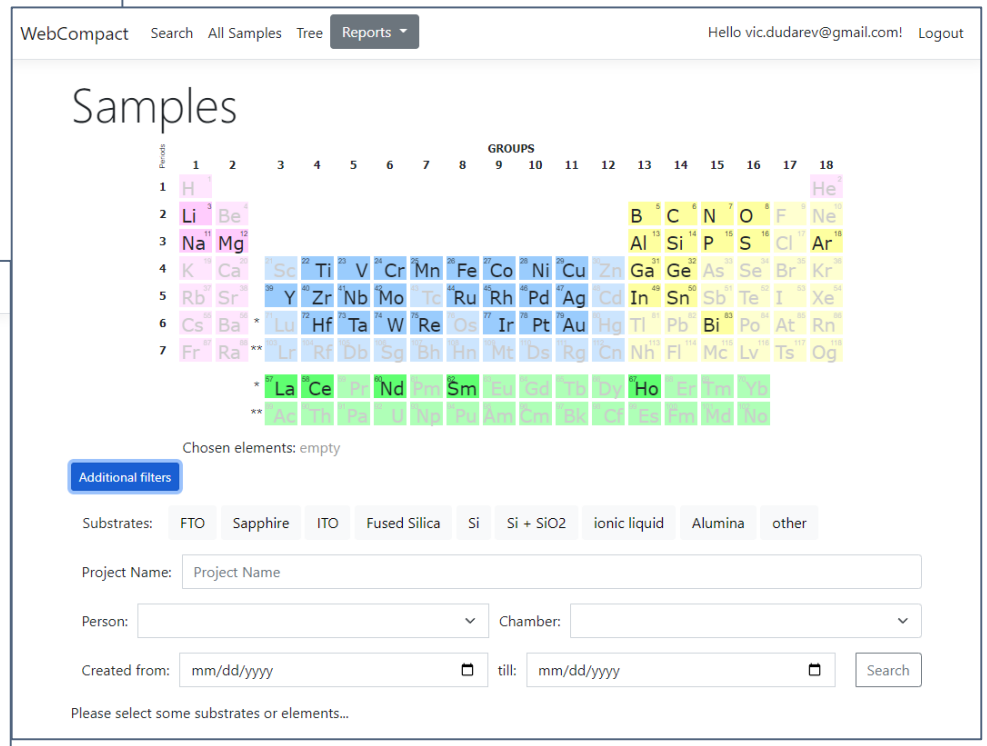
WebCompact https://webcompact.wdm.ruhr-uni-bochum.de/

INF(N)

Runs under .NET 7







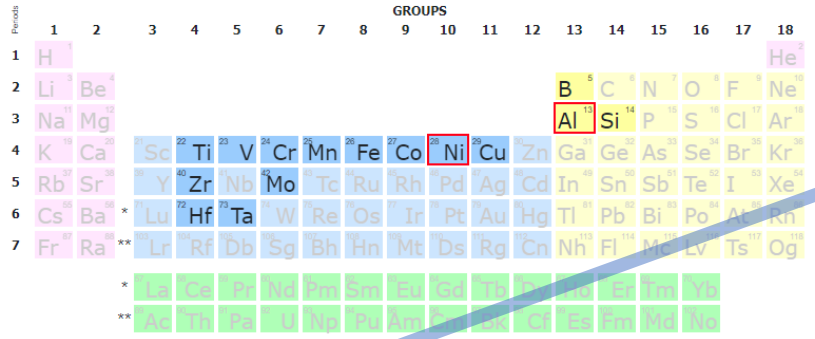
Al-Ni-Co (TIFF + JPG + no tree): /sample/samplebyid/1471
Cr-N (JPG, BMP, TIFF + tree): /sample/samplebyid/3699

WebCompact Search

INF(N)

Samples

system Al-Ni



Chosen elements: **Al-Ni** [clear]

Additional filters

Table (14) List (14)

Id	Created	Project Name	Person	N	Sample Material
1471	2016-11-02	Ni-Co-Al based materials	Decker Peer	3	Al-Co-Ni (Ni-Co-Al)
1124	2016-04-20	Fachlabor (Dennis N.)		2	Al-Ni (Ni-Al)
1261	2016-07-14	Fe-based SMAs	Decker Peer	4	Al-Fe-Mn-Ni (Fe-Ni-Mn-Al)

Sample Project Ni-Co-Al based materials

Sample Information

Name	Value
Project Data	
Project Name	Ni-Co-Al based materials
Sample Name	Ni-Co-Al full ternary - Ni-Co-Al
Sample Description	Ni-Co-Al full ternary
Sample Material	Ni-Co-Al
Composition Description	ternary
Substrate Material	Si + SiO2
Project Comment	
Measurement Grid	
Project dates	
Project Id	
Creation Date	

Sample Documents

Document Id	Processing State	Characterization	CadRefFileName	Data
9522	Raw Data	chem. composition	141128K1-2_annealed.txt	000008958_00.00_9522.txt process & show compositions
9523	Raw Data	chem. composition	141128-K1-2_annealed_standard.txt	000008959_00.00_9523.txt process & show compositions
	Processed Data	chem. composition	141128-K1-2_Co-Al-Ni_EDX_standard_correct.opj	000008960_00.00_9524.opj
	Processed Data	chem. composition	141128-K1-2_EDX_Al-Differenz.tif	000008961_00.00_9525.tif
	Processed Data	chem. composition	141128-K1-2_EDX_Vergleich.tif	000008962_00.00_9526.tif
	Processed Data	chem. composition	141128-K1-2_EDX_Vergleich_mit Standard.tif	000008963_00.00_9527.tif

9519	Processed Data	photo	141128-K1-2_bearbeitet.tif	000008955_00.00_9519.tif
9520	Processed Data	photo	141128-K1-2_mit Nummern.doc	000008956_00.00_9520.doc
9521	Raw Data	photo	IMG_8731.JPG	000008957_00.00_9521.JPG
9536	Raw Data	synchrotron	141128-K1-2_Synchrotron_raw_data.zip	000008972_00.00_9536.zip
9537	Processed Data	synchrotron	141128-K1-2_Peak fits_Synchrotron.zip	000008973_00.00_9537.zip
9547	Processed Data	synchrotron	141128-K1-2_Synch.txt	000008979_00.00_9547.txt
9548	Processed Data	XRD	141128-K1-2_XRD.txt	000008980_00.00_9548.txt
9549	Processed Data	XRD	141128-K1-2_XRD_reduced.txt	000008981_00.00_9549.txt
9550	Raw Data	XRD	141128-K1-2_XRD.opj	000008982_00.00_9550.opj
9552	Raw Data	XRD	141128-K1-2_XRD_raw.zip	000008983_00.00_9552.zip

Al-Ni-Co (TIFF + JPG + no tree): /sample/samplebyid/1471

WebCompact: Sputter Rates

- **List of digitalized sputter rates** (User group access)
- **Download (export) data** (Administrator group or user with **`_SputterRate.read`** claim)
- **Upload (import) data** (Administrator group or user with **`_SputterRate.write`** claim)
- **Person in charge**: Alan Savan

ompact Search All Samples Tree Reports Search Rates All Rates Hello vic.dudarev@gmail.com!

All Sputter Rates

Download rates in Excel

Choose File No file chosen

Upload / update rates from Excel

Table (468) List (468)

Material	Trust	LogbookId	Rate, nm/sec	Normalized Rate, nm/W*sec	Time, sec	Thickness, nm	Cathode	Power, W	Power Supply	Comment
Ag	1	061228-K1-3	0.09506	0.000951	1800	171.1	6	100	RF	Book 3: April 2006-Jan 2007
Ag	0	061229-K1-1	0.34848	0.001742	1800	627.27	6	200	RF	Book 3: April 2006-Jan 2007
Al	0	080319-K1-1	0.145	0.000725	1800	261	6	200	RF	
Al	0	080424-K1-2	0.056	0.000373	2700	151.2	5	150	RF	
Al	0	080516-K1-3	0.02107	0.000351	2700	56.9	5	60	RF	Book 6: March 2008-April 2009
Al	0	080519-K1-2	0.01421	0.000237	12000	170.5	5	60	RF	Book 6: March 2008-April 2009
Al	0	080529-K1-1	0.04711	0.000337	5400	254.4	5	140	RF	Book 6: March 2008-April 2009
Al	0	100228-K1-1	0.058	0.000387	2700	156.6	5	150	RF	

Hints: LogbookId – primary key; Trust = -1000 – delete record on upload.

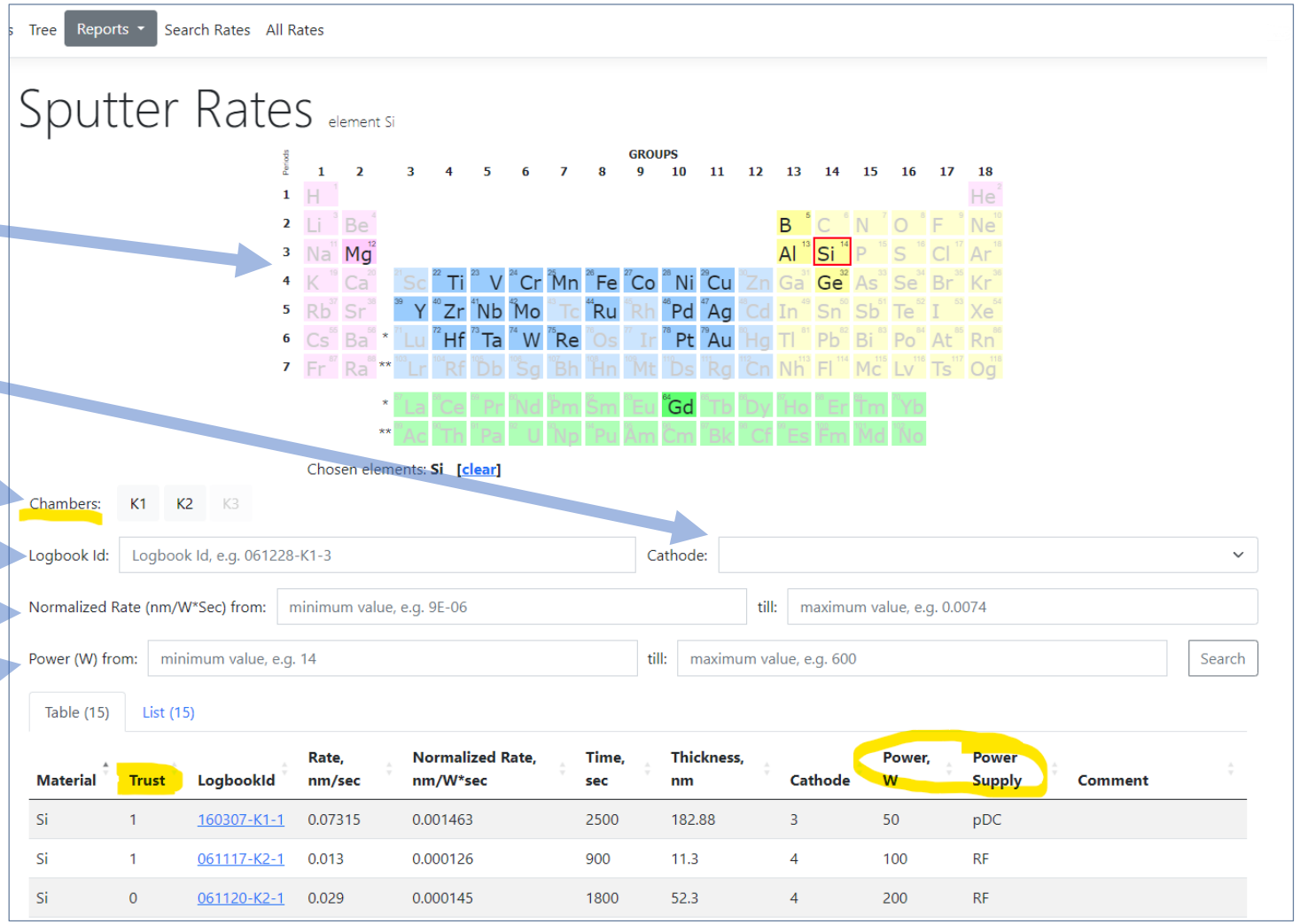
https://webcompact.wdm.ruhr-uni-bochum.de/sputterrate/allrates

WebCompact: Search for Sputter Rates

Search on (+ sort results):

- Elements
- Cathode
- Chambers
- LogbookId (Date-Chamber-N)
- Rate range
- Power range

It provides a good “starting point” to suggest sputtering mode settings and time to achieve desired thickness.



Material	Trust	LogbookId	Rate, nm/sec	Normalized Rate, nm/W*sec	Time, sec	Thickness, nm	Cathode	Power, W	Power Supply	Comment
Si	1	160307-K1-1	0.07315	0.001463	2500	182.88	3	50	pDC	
Si	1	061117-K2-1	0.013	0.000126	900	11.3	4	100	RF	
Si	0	061120-K2-1	0.029	0.000145	1800	52.3	4	200	RF	

https://webcompact.wdm.ruhr-uni-bochum.de/sputterrate/

Detailed Information

Id
Material
Trust
Deposition Date
Chamber, №
Logbook Id
Rate, nm/sec
Normalized Rate, nm/W*sec
Time, sec
Thickness, nm
Power, W
Power Supply
Cathode
Pressure, mTorr
Ar-Flow, sccm
Cathode Tilt, mm
Table Height, mm
Comment

Sputter Rate for W (111118-K3-4)

Name	Value
Metadata	
Id	506
Material	W
Trust	0
DepositionDate	2011-11-18
Chamber, №	3
Logbook Id	111118-K3-4
Rate	
Rate, nm/sec	0.245 nm/sec
Normalized Rate, nm/W*sec	0.00098 nm/W*sec
Sputtering Parameters	
Time, sec	
Thickness, nm	441.5 nm
Power, W	250 W
Power Supply	DC
Cathode	2
Pressure, mTorr	10 mTorr
Ar-Flow, sccm	40 sccm
Working Distance	
Cathode Tilt, mm	7.5 mm

https://webcompact.wdm.ruhr-uni-bochum.de/sputterrate/ratebyid/506

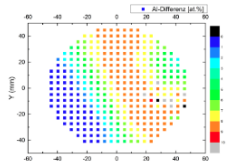
WebCompact and Data Visualization

INF(N)

Data processing with external software (written in Python) + visualization

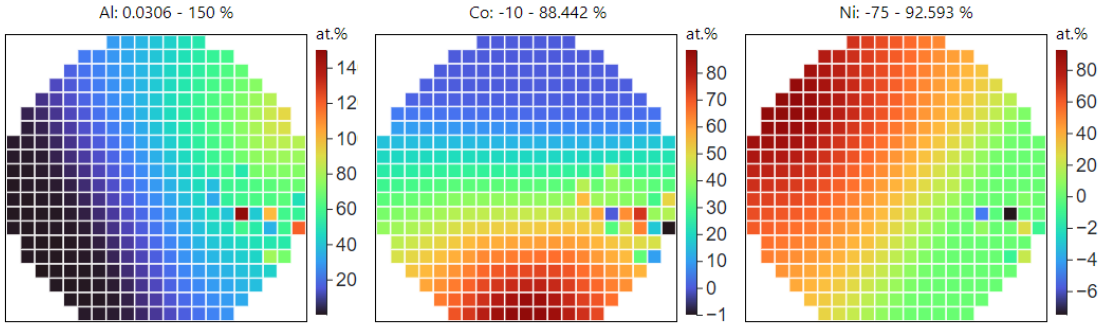
Sample Documents

Document Id	Processing State	Characterization	CadRefFileName	Data
9522	Raw Data	chem. composition	141128K1-2_annealed.txt	000008958_00.00_9522.txt process & show compositions
9523	Raw Data	chem. composition	141128-K1-2_annealed_standard.txt	000008959_00.00_9523.txt process & show compositions
9524	Processed Data	chem. composition	141128-K1-2_Co-Al-Ni_EDX_standard_correct.opj	000008960_00.00_9524.opj
9525	Processed Data	chem. composition	141128-K1-2_EDX_Al-Differenz.tif	000008961_00.00_9525.tif



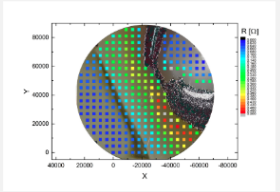
Composition for Ni-Co-Al

SampleId: 1471; **Substrate**: Si + SiO2



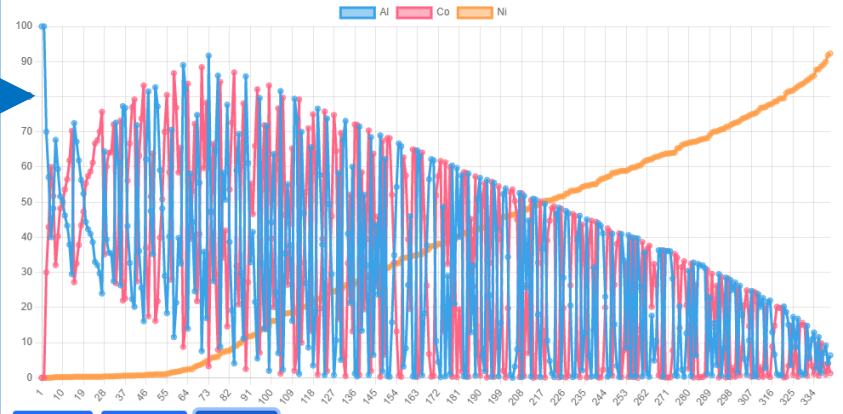
Simple composition visualization (Chart.js)

9545	Processed Data	chem. composition	141128-K1-2_EDX.txt	000008977_00.00_9545.txt show chart
9546	Processed Data	chem. composition	141128-K1-2_EDX_Clustering_Standard.txt	000008978_00.00_9546.txt show chart
9529	Processed Data	el. resistance	141128-K1-1_R_neu.opj	000008965_00.00_9529.opj
9530	Processed Data	el. resistance	141128-K1-1_R_neu_ausreißer korrigiert.opj	000008966_00.00_9530.opj
9531	Processed Data	el. resistance	141128-K1-2_R_neu.jpg	000008967_00.00_9531.jpg



Composition chart for Ni-Co-Al

SampleId: 1471; **Substrate**: Si + SiO2



i	Al	Co	Ni
1	100.0	0.0	0.0
2	100.0	0.0	0.0
3	70.0	30.0	0.0
4	57.1	42.9	0.0
5	40.0	60.0	0.0
6	48.2	51.7	0.1
7	67.7	32.1	0.2
8	59.4	40.3	0.2
9	51.6	48.2	0.2
10	50.3	49.5	0.2
11	46.2	53.6	0.2
12	43.3	56.5	0.2

https://webcompact.wdm.ruhr-uni-bochum.de/sample/samplebyid/1471