



Toolbox for Querying, Matching and Scoring Data against GND

Selected License
Attribution-NonCommercial-
ShareAlike 4.0 International



Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

ThuLB, Friedrich-Schiller-Universität Jena

Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

Agenda

- Background for the Toolbox Development
- Design & Development Concept for Toolbox
- Technologies Implemented
- User-flow for Toolbox
- Core and Current Features and Functions
- Live Demo

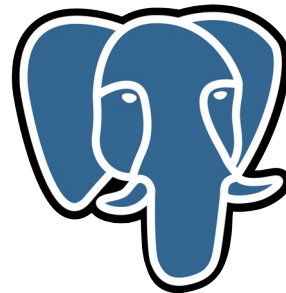
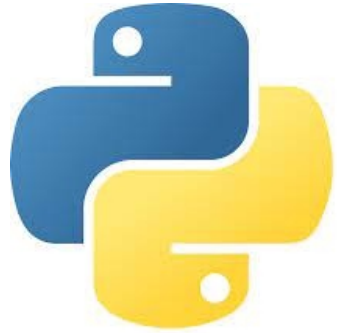
Background for the Toolbox Development

- Project Team for Development:
 - 2 Employees from DigiCult: Query Standards, Scoring Algorithm Development and Source Data Provider
 - 2 Employees from ThuLB: Software development & Implementation of Query Standards and the Scoring Algorithm via software
- Main Purposes: query, match, score the objects against GND and validate the data sets for GND import
- Types/Nature of Source Data: Person, Bauwerke, Geografika, Sachbegriffe via NDS or Private Data (New)
- Querying & Matching against: GND via Lobid, Wikidata, Geofabrik
- Expected Outcomes: easily customizable, expandable, sustainable Toolbox app.

Design & Development Concept for Toolbox

- Open-source
- Open Access Features (statistics, visuals, etc.)
- User-Friendly Interface (usability tests)
- Flexible (in function adding for querying, matching and scoring and improving front-end features)
- Compatible with multiple devices (PC, Tablet, Cell Phone)
- Both Static and Standalone via Docker technology (New)

Technologies Implemented



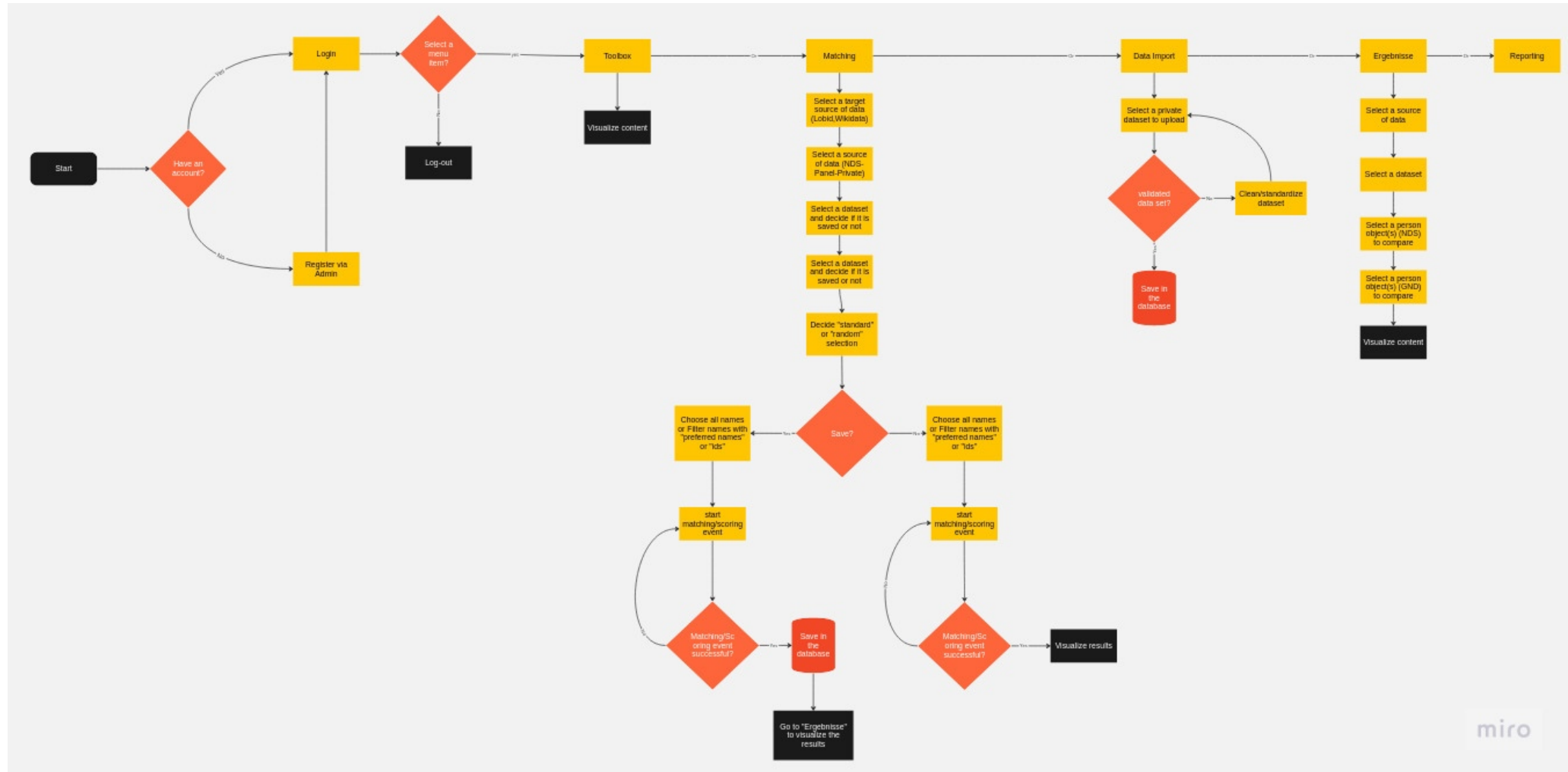
Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

ThuLB, Friedrich-Schiller-Universität Jena

Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

User-flow for Toolbox



Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

ThuLB, Friedrich-Schiller-Universität Jena

Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

Core and Current Features and Functions

Menu & Dashboard

The screenshot shows a web browser at the URL `https://gnd4c.thulb.uni-jena.de`. On the left is a sidebar menu with the following items: Home (highlighted in green), Über GND4C, Personen, Nachrichten, Blogposts, Tutorials, Veröffentlichungen, **Toolbox-Login** (highlighted with a pink box), Kontakt, and Impressum. A pink arrow points from the label "Validated User Login" below to the "Toolbox-Login" item. The main content area is titled "Home" and contains a paragraph of Lorem Ipsum text. A pink arrow points from the label "Dashboard" above to the main content area. Below the text is a grid of four data selection boxes, each with a dropdown menu. The boxes are labeled: "Entität: Person", "Entität: Bauwerk", "Entität: Geographika", and "Entität: Sachbegriff". Each box contains the text "Wählen Sie einen [Entität]datensatz". At the bottom of the main content area, it says "LAST UPDATE: 19/04/2023 19:31:34 By Erdal Ayan".

Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

ThuLB, Friedrich-Schiller-Universität Jena

Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

Core and Current Features and Functions

Menu for Toolbox Usage

The screenshot displays a web application interface. On the left is a vertical navigation menu with items: Home, Über GND4C, Personen, Nachrichten, Blogposts, Tutorials, Veröffentlichungen, **Toolbox-Login** (highlighted in green), Kontakt, and Impressum. At the top is a horizontal navigation bar with items: **Toolbox** (highlighted in green), Matching, Data Import, Ergebnisse, Reporting, Personal Liste, Data Visualisierung, Blogpost, and Benutzer Verwaltung. Below the menu is a login form titled "Melden Sie sich bitte in Ihrem Konto an." with fields for "Benutzername" (containing "eayan") and "Password" (masked with dots), a "Login" checkbox, and a "Logged in as eayan" status bar. To the right of the login form is a "User-flow in Toolbox" diagram, a complex flowchart with yellow and red nodes and arrows, illustrating the process flow within the toolbox.

User-flow in Toolbox

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nam consequat tortor quis lorem volutpat, venenatis finibus tortor luctus. Mauris nec velit et purus tristique vestibulum nec a turpis. Cras libero odio, veh non dictum tristique, dapibus ac mi. Quisque convallis dolor vitae ipsum egestas placerat. Aenean pellentesque sit amet mauris sed blandit. Phasellus tellus sapien, pellentesque sit amet mattis et, molestie sit a arcu. Ut rutrum finibus ligula, facilisis pharetra augue eleifend bibendum. Quisque eget neque volutpat, eleifend libero vitae, dictum neque. Cras blandit pulvinar nulla, at faucibus lectus hendrerit ac. Vivamus n dolor, tristique ullamcorper lacinia vitae, feugiat at ante. Sed venenatis ante et velit ultricies, at tincidunt dolor facilisis. Morbi condimentum sagittis placerat. Sed eu pharetra quam.

Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

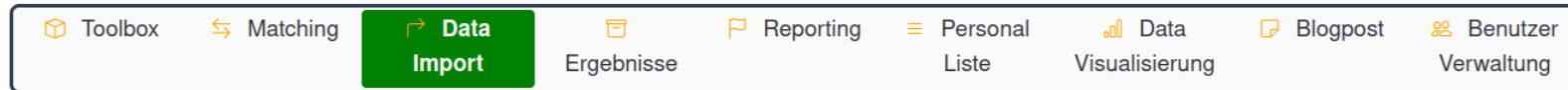
ThuLB, Friedrich-Schiller-Universität Jena

Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

Core and Current Features and Functions

Private Data Import



Data Import

This site is under construction

Sie können die JSON-Beispieldaten für Person Data herunterladen und Ihre eigenen privaten Daten entsprechend organisieren.

Download Person Data

Sie können die JSON-Beispieldaten für Bauwerke Data herunterladen und Ihre eigenen privaten Daten entsprechend organisieren.

Download Bauwerke Data

Data Type
 Person Bauwerke

Choose a file

Drag and drop file here
Limit 200MB per file

Browse files

A screenshot of the 'Data Import' page. The page is mostly white with blue and yellow highlights. A pink box highlights the 'Data Import' button in the navigation bar. Another pink box highlights the 'Data Import' section header. A third pink box highlights the 'Data Type' section, which has radio buttons for 'Person' (selected) and 'Bauwerke'. A fourth pink box highlights the 'Browse files' button. There are also pink arrows pointing to the 'Download Person Data' and 'Download Bauwerke Data' buttons.

Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

ThuLB, Friedrich-Schiller-Universität Jena

Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

Core and Current Features and Functions

Matching

The screenshot shows a web browser window with the URL <https://gnd4c.thulb.uni-jena.de>. The navigation bar includes a 'Matching' button (highlighted in green) and other options like 'Data Import', 'Ergebnisse', 'Reporting', 'Personal Liste', 'Data Visualisierung', 'Blogpost', and 'Benutzer Verwaltung'. The main content area is titled 'Matching' and has sub-links for 'Matching für Personen', 'Matching für Bauwerke', 'Matching für Geografika', and 'Matching für Sachbegriffe'. The 'Matching für Personen' section is active. It contains several form elements: a 'Target systems' section with radio buttons for 'Lobid' (selected) and 'Wikidata'; a 'Source data systems' section with radio buttons for 'NDS Data' (selected), 'Private Data', and 'Panel Data'; a 'Dataset list' section with a dropdown menu showing 'Abgebildete Personen ohne GND - Stiftung Schloss Friedenstein Gotha'; and a confirmation message 'Sie haben "Abgebildete Personen ohne GND - Stiftung Schloss Friedenstein Gotha" gewählt'. At the bottom, it shows 'Entität: Person'.

Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

ThuLB, Friedrich-Schiller-Universität Jena

Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

Core and Current Features and Functions

Matching

Bitte wählen Sie, ob Sie die Ergebnisse speichern möchten.

Ergebnisse speichern.
 Speichern als Panel speichern Nicht speichern

Matching-Ereignis definieren

Erstellt am: 19/04/2023 19:47:02
Zielentität: Person
Datenlieferung: ter02660
Default Title: ter02660Person19/04/2023 19:47:02
Geben Sie ein alternative Titel ein
Geben Sie hier Ihre Kommentar ein

Matching-Ereignis ist nur für Sie sichtbar

Limit
 Standard Random

Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

ThuLB, Friedrich-Schiller-Universität Jena

Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

Core and Current Features and Functions

Matching

Standard Random

Die Gesamtanzahl der Zeilen in diesem Datensatz beträgt: 1157

Range Start: 0

Range End: 1157

double ended slider: 0 to 1157

slider range: (0, 1157) 0 1157

Bitte wählen Sie Ihre Limit: 1

	nds_id	id	forename	surname	name_addition	personal_name	prefix	counting	display	preferred	type	entity	uri	gender	so_profession	so_year_birth	so_placeOfBirth	so_year_death
0	220940	300203	Hugo	Abensperg-Traun	None	None	None	None	None	True	None	person	http://terminology.lido-schema.org/actor_type/person	male	abgebildete Person	-9999-01-01	Not Described	-9999-01-01

Bitte Wählen: Preferred Names nach Personennamen filtern nach Objekt Id filtern

Load Data

Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

ThuLB, Friedrich-Schiller-Universität Jena

Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

Core and Current Features and Functions

Querying, Matching & Scoring

220940 Hugo Abensperg-Traun

Protokoll zum Matching

que...	query_id	totalItems	query_url
0	https://lobid.org/gnd/search?q=preferredName:(Hugo~2+Abensperg~2+)&filter=(type:RoyalOrMemberOfARoyalHouse+OR+LiteraryOrLegendaryCharacter+OR+DifferentiatedPerson)&size=20&format=json	1	
1	https://lobid.org/gnd/search?q=variantName:(Hugo~2+Traun~2+)&filter=(type:RoyalOrMemberOfARoyalHouse+OR+LiteraryOrLegendaryCharacter+OR+DifferentiatedPerson)&size=20&format=json	3	
2	https://lobid.org/gnd/search?q=variantName:(Hugo~2+Abensperg~2+)&filter=(type:RoyalOrMemberOfARoyalHouse+OR+LiteraryOrLegendaryCharacter+OR+DifferentiatedPerson)&size=20&format=json	1	
3	https://lobid.org/gnd/search?q=preferredName:(Hugo~2+Traun~2+)&filter=(type:RoyalOrMemberOfARoyalHouse+OR+LiteraryOrLegendaryCharacter+OR+DifferentiatedPerson)&size=20&format=json	7	
4	https://lobid.org/gnd/search?q=preferredName:(Abensperg~2+Traun~2+)&filter=(type:RoyalOrMemberOfARoyalHouse+OR+LiteraryOrLegendaryCharacter+OR+DifferentiatedPerson)&size=20&format=json	23	
5	https://lobid.org/gnd/search?q=preferredName:(Hugo~2+Abensperg~2+Traun~2+)&filter=(type:RoyalOrMemberOfARoyalHouse+OR+LiteraryOrLegendaryCharacter+OR+DifferentiatedPerson)&size=20&format=json	1	
6	https://lobid.org/gnd/search?q=variantName:(Abensperg~2+Traun~2+)&filter=(type:RoyalOrMemberOfARoyalHouse+OR+LiteraryOrLegendaryCharacter+OR+DifferentiatedPerson)&size=20&format=json	24	
7	https://lobid.org/gnd/search?q=variantName:(Hugo~2+Abensperg~2+Traun~2+)&filter=(type:RoyalOrMemberOfARoyalHouse+OR+LiteraryOrLegendaryCharacter+OR+DifferentiatedPerson)&size=20&format=json	1	

Core and Current Features and Functions

Querying, Matching & Scoring

Alle Wertungsergebnisse

gndIdentif...	id	preferredName	variantName	dateOfBirth	dateOfDeath	place...	pl...	professi...	total_score	max_name_score	pref_name_score
143410067		Abensberg und Traun, Hugo	Traun, Hugo Abensberg und,...	1828-09-20	1904-08-03			[object Ob...	0.26845	3	2
133145999		Abensberg-Traun, Otto	Traun, Abensberg Otto-, Abe...						-5.81655	2	2
1155486676		Abensberg und Traun, Carl	Traun, Carl Abensberg und, ...	1947-10-12				[object Ob...	-5.81655	2	2
125808321		Abensberg-Traun, Carl							-5.81655	2	2
173238300		Abensberg und Traun, Ernst ...	Traun, Ernst von Abensberg ...						-5.81655	2	2
120691515		Abensberg und Traun, Ernst ...	Traun, Ernesto di Abensberg ...	1608	1668			[object Ob...	-5.81655	2	2
130852678		Abensberg-Traun, Otto Ehre...	Traun, Otto Ehrenreich von ...	1644	1715	[object ...	[obje...	[object Ob...	-5.81655	2	1
130066419		Abensberg und Traun, Otto ...	Traun, Otto Ehrenreich Mari...	1848-09-23	1899-02-12		[obje...		-5.81655	2	2
115912342X		Abensberg und Traun, Franz ...	Traun, Franz Anton von Abe...	1674-06-04	1740				-5.81655	2	2
1132279852		Abensberg und Traun, Eva M...						[object Ob...	-5.81655	2	2
1132281342		Abensberg und Traun, Eleon...	Rueber von Pixendorf, Eleon...	1645	1719				-5.81655	2	2
120751704		Abensberg und Traun, Judit...	Tättenpach, Judith Sidonia v...	1628	1668-07-07		[obje...		-5.81655	2	2
143504797		Abensberg und Traun, Otto L...	Otto Lorenz, Abensberg und ...	1638-11-04	1695-04-12		[obje...		-5.81655	2	2
1059447614		Abensberg und Traun-Meiss...	Traun-Meissau, Rudolf von A...	1872-09-21	1954-10-27				-5.81655	2	2
1229988238		Abensberg und Traun, Ferdi...		1647	1685			[object Ob...	-5.81655	2	2

1 to 15 of 23 |< < Page 1 of 2 > >|

Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

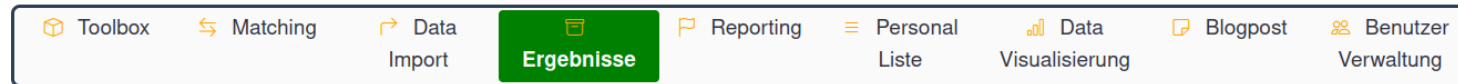
ThuLB, Friedrich-Schiller-Universität Jena

Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

Core and Current Features and Functions

Visualizing the saved results



Ergebnisse

Ergebnisse für die Personen Ergebnisse für die Bauwerke Ergebnisse für die Geografika Ergebnisse für die Sachbegriffe

Scoring Type
 NDS-Data Panel Data Private Data

Bitte Wählen Ihre Ergebnisse

Ergebnisse
Choose a Dataset
Abgebildete Personen ohne GND - Stiftung Schloss Friedenstein Gotha

Objekt Wählen
ter02660Person19/04/2023 20:09:26
Bitte Wählen
 Lobid Wikidata

Sie haben "Abgebildete Personen ohne GND - Stiftung Schloss Friedenstein Gotha" gewählt

Entität: Person

Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

ThuLB, Friedrich-Schiller-Universität Jena

Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

Core and Current Features and Functions

Visualizing the saved results

Bitte wählen Sie eine NDS-ID aus

NDS Id	Person Id	Data History	Source System	Entity Type	Forename	Surname	Personal Name	Name Addition
<input checked="" type="checkbox"/> 220940	300203	be added	ter02660	person	Hugo	Abensperg-Traun	None	None

1 to 1 of 1 < > Page 1 of 1 > >|

Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

ThuLB, Friedrich-Schiller-Universität Jena


Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

Core and Current Features and Functions

Visualizing the saved results

Bitte definieren Sie Ihre neuen Bewertungsparameter.

Bearbeitbare Werte anzeigen 

Neue param_max_name_score einfügen	Neue param_job_score einfügen	Neue param_O_birthplace_score einfügen
1.30 - +	2.78 - +	0.86 - +
Neue param_max_name_score: 1.298	Neue param_job_score: 2.778	Neue param_O_birthplace_score: 0.857
Neue param_istMx_x_Anz einfügen	Neue param_G_birthdate_score einfügen	Neue param_P_deathplace_score einfügen
-0.28 - +	1.14 - +	2.02 - +
Neue param_istMx_x_Anz: -0.283	Neue param_G_birthdate_score: 1.139	Neue param_P_deathplace_score: 2.024
Neue param_istMaxNameScore einfügen	Neue param_H_deathdate_score einfügen	Neue param_activity_period_ijklmn_score einfügen
5.07 - +	1.20 - +	0.18 - +
Neue param_istMaxNameScore: 5.07	Neue param_H_deathdate_score: 1.2015	Neue param_activity_period_ijklmn_score: 0.178

Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

ThuLB, Friedrich-Schiller-Universität Jena
















Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

Core and Current Features and Functions

Visualizing the saved results

Bitte wählen Sie eine GND Id aus

gndidentifizier	GND-URL	new_total_score	total_score	new_param_max_name_s...	new_param_lstMx_x_Anz	new_param_lstMaxNameS...	new_param_job_score	new_param_G_birthdat
<input checked="" type="checkbox"/> 143410067		0.27	0.27	3.894	-0.283	5.07	0	
<input type="checkbox"/> 120751704		-5.82	-5.82	2.596	0	0	0	
<input type="checkbox"/> 1199597198		-5.82	-5.82	2.596	0	0	0	
<input type="checkbox"/> 1064901344		-5.82	-5.82	2.596	0	0	0	
<input type="checkbox"/> 11740599X		-5.82	-5.82	2.596	0	0	0	
<input type="checkbox"/> 117405973		-5.82	-5.82	2.596	0	0	0	
<input type="checkbox"/> 1032932279		-5.82	-5.82	2.596	0	0	0	
<input type="checkbox"/> 1041592833		-5.82	-5.82	2.596	0	0	0	
<input type="checkbox"/> 1229988238		-5.82	-5.82	2.596	0	0	0	
<input type="checkbox"/> 1059447614		-5.82	-5.82	2.596	0	0	0	
<input type="checkbox"/> 143504797		-5.82	-5.82	2.596	0	0	0	
<input type="checkbox"/> 1132281342		-5.82	-5.82	2.596	0	0	0	
<input type="checkbox"/> 133145999		-5.82	-5.82	2.596	0	0	0	
<input type="checkbox"/> 1132279852		-5.82	-5.82	2.596	0	0	0	
<input type="checkbox"/> 115912342X		-5.82	-5.82	2.596	0	0	0	

1 to 15 of 23 |< < Page 1 of 2 > >|

Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

ThuLB, Friedrich-Schiller-Universität Jena

Event: Projektübergreifendes Arbeitstreffen nicht-bibliothekarischer Agenturen

Date: 20.04.2023

Core and Current Features and Functions

Visualizing the saved results

NDS Source Data		GND Target Data	
	0		0
NDS Id	220940	gndidentifier	143410067
Person Id	300203	id	https://d-nb.info/gnd/143410067
Data History	Url	preferredName	Abensberg und Traun, Hugo
Source System	ter02660	variantName	['Traun, Hugo Abensberg und', 'Traun, Hugo Abensberg und', 'Abensberg u
Entity Type	person	dateOfBirth	['1828-09-20']
Forename	Hugo	dateOfDeath	['1904-08-03']
Surname	Abensperg-Traun	placeOfBirth	nan
Personal Name	None	placeOfDeath	nan
Name Addition	None	professionOrOccupation	[{'id': 'https://d-nb.info/gnd/4012401-0', 'label': 'Diplomat'}]
Counting	None	new_total_score	0.27
Prefix	None	total_score	0.27
Non-Preferred Name	None	match_status	None
Preferred	True	comment	None
Gender	male	max_name_score	3.0
Birthdate	None	pref_name_score	2.0
Deathdate	None	var_name_score	3.0
Birthplace	Not Described	non_pref_name_score	0.0
Deathplace	Not Described	non_pref_var_name_score	0.0
Profession	Not Described	1stMaxNameScore	1.0

Erdal Ayan, M.A.-M.Sc. (E-mail: erdal.ayan@uni-jena.de)

ThuLB, Friedrich-Schiller-Universität Jena

Live Demo

- Private data import
- Querying, matching and scoring private data and NDS data
- Analyze scoring results

**Thank you very much for your
interest & patience**

Contact:

Erdal Ayan, M.A.-M.Sc. → erdal.ayan@uni-jena.de

Dr. Michael Markert → michael.markert@uni-jena.de

Toolbox → <https://gnd4c.thulb.uni-jena.de/>