Pseudonymization and personal data export (GDPR)

Webdesk provides two data protection (GDPR) tools: the **pseudonymization** of all persons in the system (e.g. before a database copy is handed over for troubleshooting) and the **export of all personal data of a single person** (right of access, Art. 15 GDPR). This page is aimed at administrators and consultants. Pseudonymization is run from the administration menu or under *Extended Tasks*; the data export is a report in HR-Expert (prerequisite: HR module, see General Information on HR-Expert[1]).

Warning: Pseudonymization affects **all persons in the system** and irrevocably deletes the audit trail data (see section "Warnings and risks"). It is intended for test and copy systems, not for the live system.

Overview and delimitation

Tool	Purpose	Where
Pseudonymization	Replaces or removes identifying data of all persons. The original values are stored encrypted in the system and can be restored.	Mask "Pseudonymization (GDPR)"; actions under "Extended Tasks"
Data export of one person	Delivers all stored personal data of one person as a report (access request under Art. 15 GDPR). Does not change any data.	HR-Expert, menu "Reports", entry "GDPR Personal Data"

Pseudonymization is **not deletion**: persons are kept with their structure (group assignments, bookings, workflows); only the fields named in the scheme are changed. Deleting persons after a retention period is done by separate jobs (see Job PoDeletePersonsAfterYears[2]; terminations are handled by the job ProcessTerminatedEmployees[3]). In the code there is no coupling between these jobs, the trash bin and pseudonymization.

Pseudonymization

Purpose and use

According to the description of the mask, pseudonymization is used to pseudonymize person data in a GDPR-compliant way before a copy is handed over for the purpose of bug fixing. Typical procedure: copy the live database to a test system, pseudonymize it there, and only then hand over the copy.

What pseudonymization does:

- According to the scheme, names, user name, contact data, date of birth, bank accounts, images and further HR master data are replaced or removed.
- The action can be repeated: persons that are already pseudonymized are skipped, provided the scheme has not changed.
- The original values are stored encrypted in the system (pseudonymization log, per person and entity). This makes a restore possible.

What it does **not** do:

- Only the fields listed in the scheme are handled. All other data remain unchanged, e.g. passwords (not part of the default scheme), workflow data and form contents, logs, files and documents. Additional handling exists only for notes and the time system (see below).
- It is not anonymization: the original data are still stored in the database in encrypted form.

Default scheme: which fields are changed and how

The schemes are delivered as the Spring beans PoPersonDefaultPseudonymizationConfig (PO) and HrPersonDefaultPseudonymizationConfig (HR). A scheme saved in the mask takes precedence over the default bean.

Person (PoPerson)

Field	Result	Strategy
userName	Random string of letters (15 characters) that does not yet exist in the system. If the client has a user name prefix, it is prepended (bean PoPseudonymizeUsernameWithPrefix). Afterwards, login is only possible with the new user name.	CUSTOM_HANDLER
firstName, lastName	Random first or last name that occurs for another person in the system. First and last name are drawn independently. The result are names that actually exist in the company.	RANDOMIZE
cellPhoneNumber	Constant +43 676 123 4567	CONSTANT
dateOfBirth, email	Value is cleared	DELETE
bankAccounts, personImages	All bank accounts and the person image are deleted	DELETE

HR person (HrPerson, only with HR module)

The following fields are all cleared with DELETE, or their entries are deleted: birthName, ethnicCode, privatePhoneNumber, privateCellPhoneNumber, comment, privateEmail, socialSecurityNumber, contactAddresses, contactInstantMessagings, contactWebProfiles, relatives, internalRelatives, secondaryJobs, priorEmployments, personMaritalStatuses.

The PoPerson belonging to the HrPerson (same ID) is processed along with it using its own scheme (table above).

Additional tasks

Task (bean)	Effect	When active
"Pseudonymize person's quick notes" (PoQuickNotePseudonymizationTaskHandler)	The text of each note of the person is replaced by placeholder text (Lorem ipsum) of the same length.	Part of the default PoPerson scheme
"Pseudonymize persons in TA system" (TaPersonPseudonymizationTaskHandler)	If the person has a TA ID, the first name and last name of the person record in the time system (TA connector) are set to the pseudonymized values. Other time and attendance data are not changed. If the connector is not available, this step is skipped with a warning in the log. On restore,	Added automatically when pseudonymization is started and the TA module is present

	the name in the time system is synchronized again.	

Obfuscation strategies

In the mask, one of the following strategies can be chosen per field.

Property	Description	Possible value(s)
CONSTANT	The field receives the given constant value (column "Custom Bean/Constant Value"; for date fields a date in the format dd.MM.yyyy).	any value
RANDOMIZE	The field receives a value that occurs in this field for another person of the same type (up to 10 attempts to find a value different from the current one; otherwise random text for text fields, or an empty value).	-
RANDOMIZE_UNIQUE	The field receives a random string of letters (15 characters) that does not yet exist in the system. Text fields only. If no free value is found after 10 attempts or the field is not text, the value becomes empty.	-
DELETE	For lists (e.g. bank accounts) all entries are deleted; for simple fields the value is cleared.	-
CUSTOM_HANDLER	The value is calculated by a Spring bean whose name is entered in the column "Custom Bean/Constant Value".	bean name (e.g. PoPseudonymizeUsernameWithPrefix)

Running it

There are two entry points. Both process **all persons**; with the HR module the processing starts at the HrPerson entities, without the HR module at the PoPerson entities. Processing is done in batches of 100 records.

Option A: Mask "Pseudonymization (GDPR)"

Menu Administration > System > Security > "Pseudonymization (GDPR)" (in the menu template full-admin directly under "System"). At the top the mask shows the status "System is pseudonymized" or "System is not pseudonymized".

Property	Description	Possible value(s)
Entity Type	Selects the scheme that is displayed and edited.	PoPerson; HrPerson (only with HR module)

Additional Pseudonymization Tasks	List of additional tasks (column "Additional Tasks") for the selected entity.	person's quick notes; persons in TA system
Field Policies	Table with the columns "Field Name", "Obfuscation Strategy" and "Custom Bean/Constant Value". Rows can be added and deleted.	see Obfuscation strategies

Buttons: Save, Save and close, **"Save and Apply Pseudonymization"** (warning icon) and Back.

1. Make sure this is a test or copy system and that a database backup exists.
2. Open the mask, select the entity type and adjust the field policies if necessary.
3. Click "Save and Apply Pseudonymization". The scheme of the currently selected entity is saved; afterwards the pseudonymization of all persons runs in the background with a progress display.
4. When finished, reopen the mask and check that the status reads "System is pseudonymized". The status only applies if a log exists for every person. If it still reads "not pseudonymized", individual batches have failed (see warnings).
5. Re-enable the audit trail if it is needed (see warnings).

Note: The scheme saved in the mask replaces the delivered default bean for this entity. If nothing has been saved, the default bean is used.

Option B: Action under "Extended Tasks"

Menu Administration > System > "Extended Tasks", section "Module specific actions". In the selection field the actions appear in the group "Human Resources":

- **"Apply default pseudonymization to HrPerson entities"** (class HrApplyDefaultPseudonymizationOnHrPerson)
- **"Restore pseudonymized HrPerson entities"** (class HrApplyRestoreOnPseudonymizedHrPerson)

Select the action, click "Execute Action" and confirm in the dialog "Run module-specific action". The dialog shows the description of the action. For these two actions no separate warning for destructive actions is shown. Despite the name "default", the action uses the currently effective scheme (a scheme saved in the mask, otherwise the default bean). Here too, all persons are processed; there is no progress display.

Restoring

Restoring is only possible with the action "Restore pseudonymized HrPerson entities" (the mask has no restore button). It reads the encrypted original values from the log, writes them back into the fields of the scheme and then deletes the log of the person. Nothing happens for persons without a log.

Data	Behavior on restore
Simple fields (names, user name, phone, email, date of birth, ...)	The original value is written back. Changes made to these fields in the meantime are overwritten.
Deleted lists (bank accounts, contact data, relatives, secondary jobs, ...)	The entries are created from the saved values as **new records** (new internal IDs).
Notes, names in the time system	Text or name is reset.
Audit trail	Cannot be restored (deleted during pseudonymization).

Without the HR module or without the TA module

- Without the HR module only the PoPerson entity is processed (default scheme above). The HR-specific fields do not exist then. The mask offers only the entity PoPerson. Opening the mask on systems without an HR license was fixed in version 4.59.27 (see Release Notes 4.59.27[4]).
- Without the TA module the synchronization with the time system does not take place.

Export of all personal data of one person (Art. 15 GDPR)

Authorization and location in the menu

- Report **"GDPR Personal Data"** in HR-Expert, menu **"Reports"**.
- By default the report (configuration exportAllPersonalGDPRData) is assigned to the default role HR_ADMIN, i.e. to HR-Expert users (e.g. the role $HRX_User). On existing installations the default role was added by an update script.

Procedure

1. In HR-Expert open the menu "Reports" > "GDPR Personal Data".
2. Select the person concerned as parameter (PersonUID) (selection from the persons; the full name is displayed).
3. Choose the output format. HTML view, PDF, Excel and print are configured. In the Excel export every table is on its own sheet.
4. Run the report and check the result. The result should be passed on to the data subject via a secure channel.

Contents of the report

The report starts with the name of the person and, if available, the person image.

Section	Contents	Prerequisite
"Personal Data"	All filled fields of the HR person, in the order of the person mask. Key values and selection values appear translated, dates as yyyy-MM-dd.	HR module
Tables of the HR person	For each list of the HR person (e.g. contact data, relatives) one table with the columns configured for this list in the person mask. Empty lists are omitted.	HR module
"Logging Information"	"Logged Actions": Action, Mobile Browser, Start/End, IP, Browser User Agent, URL. "Login Attempts": Date, IP, Result, Browser User Agent.	always; entries only if present
"Time Acquisition"	"Attendance Information": Attendance, Full Day/Half Day, Interval, from the start of time recording until today.	TA module; person with TA ID

"Travels Information"	Travels: From, To, Travel ID, Travel Type, Destination, Total Without Taxes, Total, Status.	TM module
"Workflow Requests"	Workflow requests started by the user: Process Name, Process Definition Name, Process Instance ID, State, Created, Last Activity Date.	WF module
"Projects"	Time efforts from the valid-from date of the person until today: Date, From/To, Duration (Minutes), Description, Project, Activity, Customer, Billable.	PTM module
"Reschedules"	Plannings: Date, From/To, Duration (Minutes), All Day Reschedule, Planning Title, Plannings (structure).	RS module

Note: The scope of the sections "Personal Data" and the tables of the HR person depends on the form definition of the HR person for the client of the person. Fields not provided in this mask do not appear in the report.

Data for which the report has no section (e.g. document management, groupware/calendar, notes, audit trail entries) must be collected separately if required.

Warnings and risks

- **Warning:** Do not run it on the live system unless that is explicitly intended. The action always affects *all* persons; individual persons cannot be selected. Create a database backup beforehand.
- **Warning:** At the start of pseudonymization the audit trail is **deactivated** (system parameter HibernateAuditTrailInterceptor, property active is set to false) and its tables are dropped and recreated because they may contain personal data. The existing audit trail entries are irrevocably lost. The restore does not bring them back and does not re-enable the audit trail. If needed, set the property back to true afterwards.
- **Warning:** Errors in individual processing batches (100 records each) are only noted in the log at debug level; processing continues and a success message may still appear. Always check the status in the mask afterwards and spot-check persons.
- **Warning:** Do not change the scheme between pseudonymization and restore. The restore works with the currently effective scheme. Fields that were added to the scheme later have no saved original value and may therefore be cleared; fields that were removed from the scheme are not written back. If the scheme is changed while persons are already pseudonymized, pseudonymization first restores these persons and then applies the new scheme.
- **Warning:** Pseudonymized names are real names of other employees (RANDOMIZE). This is intended for tests but can cause confusion in the test system. User names are replaced randomly; passwords remain unchanged. For copy systems "Reset all passwords" is a useful addition (see Erweiterte Funktionen[5]).
- **Note:** A database copy that is handed over still contains the encrypted original values (pseudonymization log). It is therefore pseudonymized, not anonymized.
- **Note:** List entries re-created by a restore have new internal IDs. References to the old IDs (e.g. from imports) point to nothing.

1. /daisy/webdesk-manual-en/hr-expert/general.html
2. /daisy/webdesk-manual-en/admin/po-overview/3731-dsy/3732-dsy/8632-dsy/9259-dsy.html
3. /daisy/webdesk-manual-en/admin/2675-dsy/3767-dsy/7475-dsy.html
4. /daisy/webdesk-manual-en/2455-dsy/g6/11219-dsy.html
5. /daisy/webdesk-manual-en/admin/po-overview/3731-dsy/463-dsy.html