The multiple pidm repair process

The next few pages describe the multiple pidm repair process. The last page contains a summary of the steps involved.

Definition - multiple pidm situation

PIDM: personal identification master
A pidm is the underlying unique identifier (number) for a person in Banner. An ID (always linked to a pidm) is the visible identifier.

We speak of a multiple pidm situation if two or more records for the same person or non-person exist in Banner. Non-person situations will not be discussed in this document.
The consequence of a multiple pidm situation is that:

- information about the person is distributed amongst multiple IDs
- errors can arise from it (for example: the person gets paid under both IDs).

Multiple pidm situations are either one of two types:

- old cases that come to light by accident or because missing biographic information is added to one of the records that causes a computer match with another existing record. Many of these old cases came about when the student and HR legacy systems were brought into Banner. Although there was a matching routine then that could identify duplicate records, essential information (like birthdate, gender and SIN) was often not present.
- cases that are newly created. Forms through which a person can be added to Banner contain a matching routine that warns the user when it appears that the person may already exist in Banner. There are still situations however in which a newly entered record is not matched with an existing record: the user who enters the information makes a typing mistake, the information on the existing record is incorrect or incomplete, or the user chooses to ignore the warning that the person may already exist in Banner.

How can multiple pidm situations be fixed?

Forms have been developed to assist the user in 'repairing' multiple pidm cases.
The repair process consists of one of the two following options:

- a merge: in a merge situation the information of two records is combined. Information from what is designated as the 'source' record is added to the existing information of the 'target' record. Regardless of which record is designated the source or target record, one of the IDs for the records is designated as the primary ID: this is the ID that will be visible in Banner for this person. The other ID becomes an alternate ID after the merge. This means that the ID is still present in Banner, but when the user enters the alternate ID in the keyblock of a form, it is replaced with the person's primary ID. The choice for a merge is made when both records contain information that should be retained.
- a delete: in a delete situation the information for a specified ID is removed from Banner, as is the ID itself. The choice for a delete is made when one of the two records contains only a minimum of information (like personal data) that is present on the other ID as well. This happens for example if a record is created by accident by a user (the user realizes too late that a record already existed for this person).
  In situations where information other than just personal data is listed for an ID, it is sometimes possible to manually transfer information from this ID to the other ID, followed by a delete of the ID.

Note: the delete option of the multiple pidm repair process is not limited to just the multiple pidm scenario. It may be used to delete a record for any reason.

Detailed description of the multiple pidm repair process

The process of repairing a multiple pidm situation consists of the following steps, that will be explained in more detail hereafter:

- the reporting and registering of a multiple pidm situation
- the analysis or triage of the situation and Checklist items to be tended to
- the actual repair process (the merge or delete) or interim measures (like X'ing out an ID)

Reporting and registering of a potential multiple pidm situation

A potential multiple pidm situation comes to light when:

- it is reported on an ad hoc basis by Banner users
- it appears on the overnight *Report of Potential Duplicate PIDM Situations*. This report is based on a matching routine on the Banner database that runs on a daily basis and is sent automatically to people in the Registrar's Office (ARR), Human Resources (HR) and Information Systems Resources (ISR). The matching routine checks out records that were created or have undergone a change during the last 24 hours, to see if the newly provided or updated information causes a match with another record.

A registry is maintained by ISR in Outlook to keep track of multiple pidm situations and to be able to retrieve statistics about them. In this registry every reported potential duplicate pidm situation is listed with the date reported, a reason code and the decision made about the case. The registry is divided into sections *Active cases*, *Completed cases* and *False hits* (cases that turned out to not be real multiple pidm cases). All correspondence (by e-mail) about the individual cases is kept in the appropriate mailfolders in Outlook. This registry is available to the central office staff involved in the PIDM repair process.

Analysis or triage

When a potential multiple pidm case is reported, an analysis or 'triage' is done to investigate the situation further. By means of this triage we try to establish:

1. if it is a real multiple pidm situation (i.e. is it really the same person behind the IDs?)
2. the reason behind the creation of this multiple pidm situation
3. how or why it surfaced
4. if it can be repaired automatically and if additional action is required before or after the repair process
5. which office is responsible for doing the repair

With the development of forms to assist in these activities as well as this documentation, ISR is trying to position the users to do the 'triage' themselves.

1. Is this a real multiple pidm case?

There are different ways to find out if the person behind the IDs is really one and the same:

- Banner form GZIPRPI can be used to compare the data for ID 1 and ID 2.
  Sometimes the personal data displayed in this form gives enough evidence and confidence about this case being a real duplicate: a combination of matching or similar names, SIN, birthdate and gender are a strong indication that this is indeed one and the same person.

The form contains links to other forms that can be helpful in procuring more evidence about the credibility of the duplicate:
- the Address form (GYAADDR) can show for example that the same addresses and/or phone numbers exist for the two IDs.
- the Prior College form (SOAPCOL) shows information about a person's education. It can be used to compare the educational background for the two IDs under investigation.
IDs that were brought into Banner from the legacy systems often do not have a lot of information available in Banner. Therefore in multiple pidm cases in which a legacy record is involved it can be hard to determine if the two ID's really refer to the same person. In some of these cases the existence of an archived student record (shown on tabpage *Student* in form GZIPRPI) for a student legacy ID can be helpful. The archived record indicates the degree the person obtained (including the year and faculty) or the year the person was last active as a student at McGill.

- If the information provided in Banner does not suffice, it can be an option to contact the person in question and ask about their background at McGill. It often happens that former McGill students become employees at McGill and are given a new ID/record in addition to the already existing student record.

<u>2. What is the reason behind the creation of this multiple pidm case?</u>

Determining the reason why a duplicate has been created might help prevent it in the future. It can reveal errors or weak spots in programs or data entry forms or it can indicate that users need extra training.
Possible reasons for multiple pidm creation are:

- the multiple pidm situation is caused by a user error, for example when the user ignores the warning message that the person may already exist in Banner. In cases like these it should be investigated why the user chose to ignore the existing record.
- a wrong birthdate was entered for one of the records
- the name entered for one of the records has a spelling mistake
- the first and last name for one of the records are reversed
- a birthdate or gender is missing for one of the records
- the person has undergone a name change (for example maiden/married name)
- the multiple pidm situation is the result of an unmerged legacy duplicate

By tracing the operations a record has undergone, it often becomes clear why a duplicate was created. Forms SZHIDEN and SZHPERS show the audit trail for tables SPRIDEN (name information) and SPBPERS (personal data): they display when a record was created, changed or removed, the user who performed the operation and the batch program or form that was used for it. It happens for example that a record for a person is created with the first name and last name reversed. At a later date a user corrects the order of the names and this triggers a match with another record in Banner and the appearance of this case on the daily report. By navigating through the different entries in form SZHIDEN, one can see when the record was created (Action type *Insert*) and in what order the names were entered (information displayed in the *After* fields). One can also see when and by who the names were reversed by looking at Action type *Update* and by comparing the information listed under *Before* and *After* (information in red is information that has changed).
Form GZIPRPI contains links to forms SZHIDEN and SZHPERS.

<u>3. How or why did it surface?</u>

As mentioned earlier, a previously unidentified multiple pidm situation can come to light 'by accident' (a user discovers that multiple records exist for the same person) or by an action on one of the records that causes a match with another record and triggers the case to appear on the daily report. Some examples of actions that can bring a previously unidentified multiple pidm situation to light are: correcting a misspelled name, adding or correcting a birthdate or SIN or specifiying the gender.

4. Can it be fixed and/or is additional action necessary?

The possibility to repair a multiple pidm situation depends on:

- *the profiles held by a person*
  A 'profile' refers to what a person is in Banner, e.g. employee, student or admissions applicant. Details about these profiles are shown on the various tabpages in form GZIPRPI. In the case of a merge, conflicting profiles can exist for the records that will prevent the user from repairing the multiple pidm situation. An employee conflict exists when both records have an employee profile attached to them. A student conflict exists when both records have a student profile attached to them and the terms of these student profiles overlap. An admissions conflict exists when both records have an admissions profile and the target record has more than just a minimum of admissions information attached to it. These conflicts indicate that the analysis and programming to even consider doing a merge have not been done.
- *the data recorded on the record(s) to be merged or deleted*
  In some cases data exists for the record to be merged or deleted that should stay the way it is in Banner. The nature of this data is such that deleting or merging/changing it will cause serious problems in Banner or in related systems. For example, if an ID has a Web appointment form attached to it and the status of this appointment form is other than *CANcelled*, *COMpleted* or *BANner*, an attempt to delete this ID from Banner will fail. The reason is that a person with a pending appointment form should not be removed from Banner. The Checklist (see below) lists some examples where the data recorded for a record will prevent a merge or a delete.
- *the programming in place for a merge or delete*
  For each table being merged or deleted for a person, rules have to be defined and programmed that determine what should happen to the data in this table in the case of a merge or a delete.

To determine if it is possible to merge two records, the user should first consult tabpage *Profiles* in form GZIPRPI. On this tabpage the profiles for the two IDs are listed. The *Final Determination for a Merge* part shows the user if a merge of the two specified records is possible or not. It indicates if there is a profile conflict that will prevent the records from being merged. If a merge is possible, it will specify the direction of the merge (ID1 data merges with ID2 data, ID2 data merges with ID1 data or the merge can be done in either direction) and which ID will become the primary ID by default (i.e. remain visible in Banner).

Before doing a merge or a delete, the user should always obtain a Checklist (link *Checklist* in form GZRPRMD) to see if certain issues apply to this specific repair process.
A Checklist can contain warnings and action items. *Action items* require user intervention before the merge or delete can be performed. Failing to act upon these action items will result in the immediate failure of the merge or delete.
More serious however are the *warnings*. They must be heeded by the user and a course of action must be taken for each one. A good example is the warning that DARS has to be notified of a merge. The user should take care of saving the pre-composed e-mail message that is provided in the Checklist warning before performing the actual merge, because the warning will no longer appear on the Checklist after a successful merge. Once the merge is complete, the user should paste the text of the saved e-mail in the appropriate areas of a real e-mail message and send it off to DARS.

Checklist messages may vary depending on the radio button settings for ID priority on form GZRPRMD. Therefore the user must specify the same radio button setting when obtaining a Checklist as will be specified for a subsequent merge.
(A typical scenario is that the user starts off with a Checklist based on the initial defaulted radio button setting. Certain Checklist messages however, like the one informing the user about an ID-card issue, might persuade the user to override the ID priority. A new Checklist should then be obtained to see the messages related to this ID priority setting).

An optional step in the process to determine if a multiple pidm case can be repaired is the use of form GZIFCOL (link *Matching Columns* in form GZIPRPI) to check if all the tables in which the record to be merged or deleted appears are programmed for. This form lists the merge or delete readiness for the individual tables, as well as the overall readiness for a merge or a delete. If the specified record appears in tables that have not yet been programmed for the intended repair operation (merge or delete), the repair of the multiple pidm situation will have to wait until programming for these tables is in place.

After the triage has been done, the conclusion can be that:

- the case under investigation is not a multiple pidm case (it is a so called false hit). The records relate to two different people. Use form GZARVPP (Reviewed PIDM Pair) to enter this pidm pair as a False Hit and give an explanation in the Comments field. ISR will also register the case in Outlook for future reference.
- there is not enough evidence to confirm that this is a multiple pidm case. Use form GZARVPP to record this pidm pair with a review code of Undetermined.
- the case under investigation is a multiple pidm case, but repairing it is not possible (yet). The case is registered as an active case and as soon as the reason for not being able to perform the repair process is resolved, the repair operation is attempted again.
  In some cases it is required to take interim measures that will facilitate the future repair process. More details about these interim measures can be found in the next section.
- the case under investigation is a multiple pidm case and it looks like it can be repaired. The case is registered as an active case and as soon as a successful merge or delete is done it is moved to the completed cases section.

<u>5. Which office will do the repair?</u>

The profiles and the status of the profiles held by a person determine which office will undertake a merge or a delete or will X-out an ID:

- if a person has the same profile for both IDs (for example student vs student or employee vs employee) it is clear which office will do the repair (ARR in the case of student and admissions profiles, HR in the case of employee profiles)

- if a person has different profiles for the different IDs (for example student vs employee or admissions applicant vs employee), the status of the profile is decisive:
  1. if both profiles have a status of ‘Active’:
  - student vs casual employee: ARR will take care of the case
  - student vs regular employee: HR will take care of the case
  2. if both profiles have a status of ‘Inactive’:
  - the record that was created first determines which office will take action (if the student record was created first ARR will take care of the case, if the employee record was created first HR will do the repair)
  3. if one profile has a status of ‘Active’ and the other a status of ‘Inactive’:
  - the office of the ‘active’ record presses the button

  In these inter-office situations, the office doing the repair must notify the other office before doing the actual merge when the ID for the other office's "client" will no longer be the primary ID after the merge. This will allow the other office to change files, etc if necessary.

- in the case of X'ing out an ID (see details in the next section), the office of the ID to be X'ed out will X out the ID

Repair process (Actual delete or merge)

Up to this point, the following steps should have been taken:

- form GZIPRPI has been used to examine the personal and profile information for the record to be merged or deleted. In the case of a merge, a comparison of the information for the two records has been made by the user to determine if the two records apply to the same person. The form determines in which direction the merge should take place as long as no conflicts exist.
- the Checklist (link *Checklist* in form GZRPRMD) has been consulted to check if there are issues that will prevent a succesful merge or delete and to determine if any additional actions need to be taken before or after the merge or delete.
- form GZIFCOL (link *Matching Columns* in form GZIPRPI) might have been used to check if all the tables in which the record to be merged or deleted appears are ready for a merge or delete.

As said before, in some cases the repair process can not be performed right away, but interim measures or other actions can or should be performed to facilitate the pending repair process.

One such interim measure is X'ing out an ID with form GZAXOID. This will prevent the accumulation of information on a record that is marked to be deleted at a later date (because not all tables in which the ID appears are programmed for). An X'ed out ID (the first digit of an ID is replaced by an 'X' and the last name changed by the user to contain a reference to the ID that should be used instead) tells the user that this ID will eventually be deleted from Banner and that a different ID should be used to enter information for this person.

Another measure might be to flag the record to online users so that it is obvious that they should not be using the particular record. For example we have had cases where a person has 2 records which are not in a position to be merged automatically and it is desired to ensure that no further information accumulates onto one the records. In such a case one can change the name as follows:
First name = USE ID
Last name = 999999999 where the string of 9's is replaced by the ID number that should be used
By making such a name change, the original ID still sits on the record, and whether the record is retrieved by the user via ID number or via a name search using the person's proper name, the record which is subsequently placed on the desktop will contain the instruction in the components of the name to use the other ID.

If the reason for not being able to repair the situation is a profile conflict, it might be possible to manipulate or transfer data in Banner in such a way that the conflict is resolved. The department responsible for the data should be consulted to discuss the possibilities.
If the reason for not being able to repair the situation is that data recorded for a record prevents a merge or a delete, it might also be possible to manually manipulate data in Banner so that a merge or a delete will become possible. If for example the status of a Web appointment form is the only issue that prevents deleting an ID from Banner, changing its status from *PENding* to *CANcelled* will make the repair process possible.

To perform the actual repair process and to initially request a Checklist, it is recommended to start in form GZIPRPI. By navigating from there to form GZRPRMD using the link *PIDM Repair – Merge/Delete* in form GZIPRPI, the values for the merge/delete parameters as determined in form GZIPRPI will be used in form GZRPRMD (the user can override them if necessary). If the repair

process is started from form GZRPRMD, the user has to specify the values for these parameters and this can be risky!
The parameters that need to be specified are:

- *Source ID*: the ID of the record whose data will be merged or deleted.
- *Target ID* (applies to a merge only): the ID of the record to which the source ID's data is to be merged.
- *Action*: the user should specify if the repair process consist of a merge or a delete. If only a source ID is specified in the keyblock and the user performs a next block, the form determines that a delete is intended and automatically checks the delete option. If both a source and a target ID are specified and a next block is done, the merge option will be automatically checked.
- *Non Duplicate Override* (applies to a merge only): the user should check this checkbox when they know that the two records that will be merged do not have equal values for SIN or birthdate or if the SIN and birthdate are missing for both records.
  Background: before doing the actual merge, the merge process validates if the records to be merged appear to apply to the same person by comparing the values for the SIN and birthdate. If both the SIN and birthdate are missing or not equal, the program wants to be reassured that the user is aware of this situation and that these records nonetheless apply to the same person. By checking the *Non Duplicate Override* checkbox the program will omit this validation.
- *Test only*: this option is only possible in QA, not in Production. By checking the *Test only* checkbox the user can perform a test merge or delete in QA to see if a repair operation in Production will likely be successful. The program will go through the motions of a real merge or delete process, but the results will not be updated in Banner. The data in QA for cases like these should of course reflect the data present in Production.
  Background: this option was initially meant for developers so they could test their programming code. By using this option the developer can also use the same test cases over and over again, because the results of the repair operation are not updated in the database.
- *ID Priority* (applies to a merge only): the user has to indicate which ID will become the primary ID (i.e. remain visible in Banner) for this person. By default, the ID that was created first in Banner will have priority. In some cases it is requested to override the default ID priority: for example if a person has an ID-card for the ID that will become an alternate ID in Banner. Accepting the default ID priority would result in having to call the person in to obtain a new ID-card, because the existing one will no longer be valid after the merge. To avoid this situation the user can override the default ID priority.
  Another situation where the user might want to override the ID priority is when scanned documents exist for the ID that will become an alternate ID after the merge. Overriding the ID priority might avoid having to reindex these documents.
  The Checklist will notify the user about these situations and the consequences of accepting the default ID priority.
- *Comments*: because information about the repair process is stored in Banner, the user should put in a meaningful comment. One could for example make a reference to the e-mail in which it is agreed upon to merge or delete the record.
  Background: for each merge or delete process the activity date, the ID(s) involved, the user, comments and the list of tables that were involved are stored in Banner. This information might one day be used for auditing purposes or made available online so that users can easily obtain information about multiple pidm repair processes.

Clicking the link *Submit for Processin*g starts the actual repair process. As indicated on form GZRPRMD a merge or delete may take up to two minutes to complete. While the repair process is running, the hour glass symbol is be displayed.
The repair process can have two possible outcomes:

- the merge or delete was successful: a message indicating the time it took to perform the successful merge or delete is displayed to the user. The user should notify the others involved about the succesful repair operation. The e-mail correspondence about the case can be moved to the completed cases section.

- the merge or delete failed: a message advises the user to review the output for details about the failure. By clicking the link *Review Output* the user navigates to another part of form GZRPRM that will list the reason(s) for the failure of the merge or delete. Once the problem is fixed a new repair process can be attempted.

Summary

This summary gives an overview of the different steps which you, as a user, should take in the multiple pidm repair process:

1. When a potential multiple pidm situation is reported (ad hoc or on the overnight report), act on it as soon as possible as it is often easier to repair a multiple pidm situation in an early stage.
   Check the registry to see if the case was reported before.

2. Do the triage. Try to establish:
   - if this is a real multiple pidm situation
   - how or why it was created (the reason behind this multiple pidm situation) and why it came to light
   - if it can be repaired

   Use the forms available to you to get a complete picture of the case:
   - use GZIPRPI to compare the personal information for the two IDs to determine if the records are for the same person. Use it also to reveal the profiles attached to the IDs (eg. employee, student) and hence to determine which offices are affected. Form GZIPRPI also indicates if a merge will be possible.
   - use GYAADDR (Address form) and SOAPCOL (Prior College form) if necessary to obtain more evidence about the credibility of the duplicate.
   - use SZHIDEN (audit trail form for table SPRIDEN) and SZHPERS (audit trail form for table SPBPERS) to determine why, how and by whom the duplicate was created. It can also explain why the duplicate surfaced.
   - optionally use GZIFCOL to check if all tables in which the source ID appears are ready for a merge or a delete.
   - use the Checklist in GZRPRMD to see if there are any issues that can be acted on or otherwise prevent a merge or delete from proceeding.

   Send an e-mail with the outcome of the triage, including all the above mentioned aspects, to others involved in multiple pidm cases. ISR will review the triage and archive the e-mail in their registry.
   Enter the pidm pair in form GZARVPP in the case of a false hit or in the case of insufficient evidence.

3. Perform the merge or delete (if the case can be repaired). Start in form GZIPRPI and navigate to form GZRPRMD: the default settings for the repair process (merge or delete), the source/target ID and the ID priority (in the case of a merge) as determined in tabpage *Profiles* in GZIPRPI will be transferred to GZRPRMD. Override these defaults when necessary.
   Get a final Checklist before doing the actual merge or delete to see if any additional actions have to be taken as a result of specifying a radio button override or possibly because data has changed since the original review of the Checklist.
   Send an e-mail to ISR and other offices involved after the merge or delete (to let them know if the repair process was successful or not).

   If repairing the case is not possible (yet), see if any interim measures (like X'ing out an ID) can be taken to prevent more harm being done to the case.
   Send an e-mail to ISR about interim measures you have taken.