# Feature Summary: Automated Duplicate Management

## Module: Cross Platform

Release: Summer 2010
Document Owner: Thomas Bressie

---

IMPORTANT NOTE: THIS DOCUMENT INTENDS TO PROVIDE A COMPLETE AND ACCURATE REPRESENTATION OF WHAT CONVIO FULLY EXPECTS TO DELIVER IN SUMMER 2010 (AS OF SEVERAL WEEKS PRIOR TO LAUNCH) BASED ON COMPLETED SOFTWARE DEVELOPMENT. THIS DOCUMENT AIMS TO SERVE AS THE BASIS FOR PLANNING AND PREPARATION WELL IN ADVANCE OF ACCESS TO THE NEW FEATURES. GIVEN THIS DOCUMENT WAS DRAFTED PRIOR TO COMPLETED QUALITY ASSURANCE, ALL FEATURES, CAPABILITIES, SCREENSHOTS, ANSWERS, CLIENT IMPACTS, SITE DATA, PARAMETERS, AND LIMITATIONS LISTED HEREIN, WHILE REASONABLY CERTAIN FOR INCLUSION IN THE RELEASE, ARE SUBJECT TO CHANGE AT ANY POINT UP UNTIL ROLLOUT ON YOUR CONVIO SITE.

---

**New Concepts or Terms**

- Automated Duplicate Management –introduced in the Aspen release of COM, new constituent records are automatically merged with high-scoring duplicate records.
- High-Scoring Duplicate – Existing constituent records that the duplicate matching algorithm identifies as exceeding the automatic merge threshold.
- Automatic Merge Threshold – duplicate matching score indicating a high probability that the scored record is a duplicate of the comparison record. The default automatic merge threshold score is 100 and can be configured in the range of 90-100 by Convio Support.
- Not a Match/Low-Scoring Threshold – duplicate matching score indicating a low probability that the scored record is a duplicate of the comparison record. The default 'Not a Match'/low-scoring threshold value is 60 and can be configured by Convio Support.
- Offline Database of Record – the client's master database of record. This database is commonly referred to as the 'offline' database because it usually cannot be accessed by the client-facing website (such as, for online fundraising).

---

**Overview**

In the current version of Convio Online Marketing (COM), duplicate management is performed by the client system administrator on a record-by-record basis.

Duplicate Management Process – Current COM Version

1. As new constituent records are received by COM (through online event registrations or otherwise), the duplicate matching algorithm is applied to them to determine if any potential matches exist in either COM or the client's offline database of record (such as Convio Common Ground)
2. All potential duplicate records found are assigned a numeric score (from 0 to 100) to indicate their level of similarity
3. All matches exceeding the low-score threshold ('Not a Match') are made available in the 'Resolve New Registrations' user interface
4. The client administrator reviews all potential matches, and resolves each match as either to be merged or to be ignored.

With the Aspen release, COM adds a new mechanism – Automated Duplicate Management – to enhance duplicate management by automatically merging high-scoring duplicate records.

With the Aspen release, COM will place duplicates in the following categories based on potential duplicate constituent record scores:

- If any existing constituent records score at or above the Automatic Merge Threshold
  - The new constituent record is merged with the existing record (applying the criteria for determining the master record and the individual field merge criteria).
  - If multiple existing constituent records exceed the Automatic Merge Threshold relative to the new constituent record, the new constituent record is placed in 'Resolve New Registrations' for the client administrator to reconcile.

- If any existing constituent records score below the Automatic Merge Threshold but above the 'Not a Match' threshold
  - The new constituent record is queued in 'Resolve New Registrations' with details on the existing constituent records that match along with their matching score
  - The client administrator will periodically review the matches in 'Resolve New Registrations' and merge or ignore them.
- If no constituent records score above the 'Not a Match' threshold
  - The new constituent record is inserted into the Constituent 360 database with no administrator intervention required.

New constituent records are automatically merged with duplicates as follows:
1. The master record is identified with two criteria: Constituent Record Source and Specific Field Comparison
   a. Constituent Record Source
      i. If one of the two records is from the offline database of record, the offline database record is master.
      ii. If both records are from COM and not the offline database of record, use the specific field comparison to determine the master record (detailed below).
   b. Specific field comparison (Note: the following comparisons are listed in priority order – each subsequent check is performed only if a master record cannot be selected based upon the preceding criteria)
      i. Member ID: If either record has a Member ID, that record becomes the master. If neither record has a Member ID value, use next criterion.
      ii. Login Date: the record with the most recent login date is master. If the login date is null for both records, use next criterion.
      iii. Modified Date: the most recently modified record is master. If neither record has a modified date, use next criterion.
      iv. Constituent ID: the record with the higher Constituent ID value is the master record.

After the master record is identified, the field merge operation takes place. See Appendix A for complete field merge criteria.

**Manual Unmerge Operation and Criteria**

A new user interface provides a summary view of all duplicate management activity within a specified time period.

Data summarized within the Duplicate Management Activity Dashboard:
- Snapshot
  - Number of potential duplicates to review in 'Resolve New Registrations'; number of merges in last two weeks.
- Last Activity
  - When the last run of the duplicate processing activity took place (Automated Duplicate Management only)
  - Suspected Duplicate Records
  - Duplicate Records Merged
  - Records not duplicates
  - Suspected duplicate records not processed
  - When the next duplicate processing activity will take place (Automated Duplicate Management only)
- Activity over Time (Timeframe Configurable – Last week/month/6 months/year)
  - Duplicate Records Merged – Manually/Automatically/Unmerged
  - Duplicate Records Matched with External Database – Manually/Automatically
  - Constituent 360 profiles reconciled with records in external database – Manually/Automatically

**Screenshot - Duplicate Management Activity Dashboard**

**Manual Unmerge Operation and Criteria**

In addition to Automated Duplicate Management, the client administrator also has available a new user interface to unmerge any records merged through automated duplicate management (Duplicate Merge Undo Management, see screenshot below) . If, after review, records that were automatically merged are deemed to be unique constituent records (and should therefore remain separate constituent records), the client administrator will be able to uno the merge manually.

Records can be unmerged under the following conditions:
- Automated merge operation occurred within two weeks of the unmerge date.
- The merged constituent record has not been edited or subject to any new modifying activity (such as, new donation, contact information update).

**Screenshot – Duplicate Merge Undo Management Client Administrator Interface**

**Reporting**

A report available to the client administrator summaries all automated merges within the specified time period. The constituent record information available in this report is as follows:

- Standard Constituent Fields (please see Appendix B of this document for a complete list of fields)
- Master Constituent ID
- Duplicate Constituent ID
- Merge Operation Started
- Merge Operation Completed
- Undo Started ( has a value only if the client administrator requested a merge undo operation )
- Undo Completed (has a value only if the client administrator requested a merge undo operation )

**Key Features & Benefits**

Features
- Automation of the merge operation for high-scoring duplicate records

Benefits
- Significant reduction in time needed for duplicate constituent record management. High-scoring duplicates online registrations are automatically merged without manual intervention from the client administrator. The amount of manual duplicate management required in COM is significantly reduced.

**Considerations for Clients**

- Your client admin administrator must enable this feature, as it is disabled by default.
- Clients should consider their business processes for duplicate management prior to enabling. To increase the chance of successfully unmerging records that were merged erroneously, client administrators should review all automated merge operations at least every two weeks.
- Automated merge operations occur when only one potential duplicate record meets or exceeds the automatic merge threshold. If more than one potential duplicate constituent records score at or above the high-scoring threshold, they are not automatically merged and are instead placed in 'Resolve New Registrations' for the client administrator to reconcile.

**Feature Configuration**

- The client administrator must enable this feature. To do so, the administrator clicks the Edit Settings link on the Duplicate Management Activity main page. On the Edit Settings page, the administrator clicks the Enable Automatic Duplicate Management link.
- This feature requires duplicate checking to be enabled. Please ensure that duplicate checking is enabled prior to enabling Automated Duplicate Management.
- The default automatic merge threshold score is 100, indicating the new constituent record and a potential duplicate must be identical to trigger an automatic merge. The threshold score can be changed by contacting Convio Customer Support to any value between 90 and 100 (indicating the % match of the new constituent record and the potential duplicate record required to trigger a merge). It is strongly recommended that all clients review their merge history to ensure that their merge results will be satisfactory with an automated merge threshold value of less than 100.

**Appendix A – Field Merge Criteria**

**Table 1 - Email Address**

| Master Email | Duplicate Email | Master Status | Duplicate Status | Result Email | Result Status |
|---|---|---|---|---|---|
| *null* | *null* | *any* | *any* | *null* | *Master Status* |
| *Master Value* | *null* | *any* | *any* | *Master Value* | *Master Status* |
| *null* | *Duplicate Value* | *any* | *any* | *Duplicate Value* | *Duplicate Status* |
| Master Value | Duplicate Value | Good | Good | Master Value | Good |
| Master Value | Duplicate Value | Good | Unknown | Master Value | Good |
| Master Value | Duplicate Value | Good | Soft Bounce | Master Value | Good |
| Master Value | Duplicate Value | Good | Hard Bounce | Master Value | Good |
| Master Value | Duplicate Value | Unknown | Good | Duplicate Value | Unknown |
| Master Value | Duplicate Value | Unknown | Unknown | Master Value | Unknown |
| Master Value | Duplicate Value | Unknown | Soft Bounce | Master Value | Unknown |
| Master Value | Duplicate Value | Unknown | Hard Bounce | Master Value | Unknown |
| Master Value | Duplicate Value | Soft Bounce | Good | Duplicate Value | Soft Bounce |
| Master Value | Duplicate Value | Soft Bounce | Unknown | Duplicate Value | Soft Bounce |
| Master Value | Duplicate Value | Soft Bounce | Soft Bounce | Master Value | Soft Bounce |
| Master Value | Duplicate Value | Soft Bounce | Hard Bounce | Master Value | Soft Bounce |
| Master Value | Duplicate Value | Hard Bounce | Good | Duplicate Value | Hard Bounce |
| Master Value | Duplicate Value | Hard Bounce | Unknown | Duplicate Value | Hard Bounce |
| Master Value | Duplicate Value | Hard Bounce | Soft Bounce | Duplicate Value | Hard Bounce |
| Master Value | Duplicate Value | Hard Bounce | Hard Bounce | Master Value | Hard Bounce |

**Table 2 – Accepts Email Communication Status**

| Master Accepts Email | Duplicate Accepts Email | Result |
|---|---|---|
| null | null | null |
| TRUE | null | TRUE |
| FALSE | null | null |
| null | TRUE | TRUE |
| TRUE | TRUE | TRUE |
| FALSE | TRUE | TRUE |
| null | FALSE | null |
| TRUE | FALSE | TRUE |
| FALSE | FALSE | FALSE |

**Table 3 – Username/Password Status**

| Master Login Date | Duplicate Login Date | Master SSN XREF | Duplicate SSN XREF | Result |
|---|---|---|---|---|
| NEVER | NEVER | *any* | *any* | Master Record |
| NEVER | OLD | *any* | *any* | Duplicate Record |
| NEVER | RECENT | *any* | *any* | Duplicate Record |
| OLD | NEVER | *any* | *any* | Master Record |
| OLD | OLD | *any* | *any* | Master Record |
| OLD | RECENT | *any* | *any* | Duplicate Record |
| RECENT | NEVER | *any* | *any* | Master Record |
| RECENT | OLD | *any* | *any* | Master Record |
| RECENT | RECENT | *any* | *any* | Master Record |

**Table 4 – Accepts Postal Mail Status**

| Master Accepts Postal Mail | Duplicate Accepts Postal Mail | Result |
|---|---|---|
| null | null | null |
| TRUE | null | TRUE |
| FALSE | null | null |
| null | TRUE | TRUE |
| TRUE | TRUE | TRUE |
| FALSE | TRUE | TRUE |
| null | FALSE | null |
| TRUE | FALSE | TRUE |
| FALSE | FALSE | FALSE |

**Table 5 – Mailing Address**

| Master Value | Duplicate Value | Results |
|---|---|---|
| null | null | null |
| null | Partial | Duplicate Value |
| null | Complete | Duplicate Value |
| Partial | null | Master Value |
| Partial | Partial | Master Value |
| Partial | Complete | Duplicate Value |
| Complete | null | Master Value |
| Complete | Partial | Master Value |
| Complete | Complete | Master Value |

**Table 6 – Membership Info**

| Master fields | Duplicate fields | Result | Master Status | Duplicate Status | Result | Master Expiration | Duplicate Expiration | Result |
|---|---|---|---|---|---|---|---|---|
| all null | all null | Master Value | sac disabled | sac disabled | *check expiration =>* | sac disabled | sac disabled | Master Value |
| all null | not null | Duplicate Value | *Unknown(0)* | *Unknown(0)* | *check expiration =>* | null | null | Master Value |
| not null | all null | Master Value | Unknown(0) | Never a Member(1) | Duplicate Value | null | not expired | Master Value |
| *not null* | *not null* | *check status =>* | Unknown(0) | Current Member(2) | Duplicate Value | null | expired | Master Value |
| | | | Unknown(0) | Lapsed Member(3) | Duplicate Value | not expired | null | Duplicate Value |
| | | | Never a Member(1) | Unknown(0) | Master Value | not expired | not expired | Master Value |
| | | | *Never a Member(1)* | *Never a Member(1)* | *check expiration =>* | not expired | expired | Master Value |
| | | | Never a Member(1) | Current Member(2) | Duplicate Value | expired | null | Duplicate Value |
| | | | Never a Member(1) | Lapsed Member(3) | Duplicate Value | expired | not expired | Duplicate Value |
| | | | Current Member(2) | Unknown(0) | Master Value | expired | expired | Master Value |
| | | | Current Member(2) | Never a Member(1) | Master Value | | | |
| | | | Current Member(2) | Current Member(2) | *check expiration =>* | | | |
| | | | Current Member(2) | Lapsed Member(3) | Master Value | | | |
| | | | Lapsed Member(3) | Unknown(0) | Master Value | | | |
| | | | Lapsed Member(3) | Never a Member(1) | Master Value | | | |
| | | | Lapsed Member(3) | Current Member(2) | Duplicate Value | | | |
| | | | *Lapsed Member(3)* | *Lapsed Member(3)* | *check expiration =>* | | | |

**Appendix B – Standard Constituent Fields**

| Field Name | Data Category/Type |
|---|---|
| Constituent ID | Biographical |
| Origin Reference ID | Biographical |
| Created By | Biographical |
| Creation Date | Biographical |
| Modify Date | Biographical |
| Last Change By | Biographical |
| Accept Mail | Address |
| Address - City | Address |
| Address - Country | Address |
| Address - State/Province | Address |
| Address - Street 1 | Address |
| Address - Street 2 | Address |
| Address - Street 3 | Address |
| Address - ZIP/Postal Code | Address |
| Congressional District (Automatic) | Advocacy |
| Congressional District (Manual Override) | Advocacy |
| County (Automatic) | Advocacy |
| County (Manual Override) | Advocacy |
| State House District (Automatic) | Advocacy |
| State House District (Manual Override) | Advocacy |
| State Senate District (Automatic) | Advocacy |
| State Senate District (Manual Override) | Advocacy |
| Active Detail | Biographical |
| Active Status | Biographical |
| Contact ID | Biographical |
| Created By | Biographical |
| Creation Time | Biographical |
| Date of Birth | Biographical |
| Donor Status | Biographical |
| Exchange ID | Biographical |
| First Name | Biographical |
| Gender | Biographical |
| Last Modified By | Biographical |
| Last Name | Biographical |
| Matched Donor | Biographical |

**Appendix B – Standard Constituent Fields (continued)**

| Field Name | Data Category/Type |
|---|---|
| Member ID | Biographical |
| Middle Name | Biographical |
| Modification Time | Biographical |
| Origin Application Reference ID | Biographical |
| Origin Application Type | Biographical |
| Origin Source Code | Biographical |
| Origin Subsource Code | Biographical |
| Password Changed | Biographical |
| Password Type | Biographical |
| Professional Suffix | Biographical |
| Suffix | Biographical |
| Title | Biographical |
| User Name | Biographical |
| Accept Email | Email |
| Email | Email |
| Email Status | Email |
| Hard-Bounce Count | Email |
| Opt-Out Date | Email |
| Opt-Out Method | Email |
| Opt-Out Reference ID | Email |
| Preferred Format | Email |
| Employer | Employment |
| Occupation | Employment |
| Phone | Phone |
| Engagement Factor 1 | Transaction |
| Engagement Factor 10 | Transaction |
| Engagement Factor 2 | Transaction |
| Engagement Factor 3 | Transaction |
| Engagement Factor 4 | Transaction |
| Engagement Factor 5 | Transaction |
| Engagement Factor 6 | Transaction |
| Engagement Factor 7 | Transaction |
| Engagement Factor 8 | Transaction |
| Engagement Factor 9 | Transaction |
| Frequency PCTL | Transaction |
| Lifetime Alert Response Count | Transaction |
| Monetary Value PCTL | Transaction |
| Previous Year Alert Response Count | Transaction |
| Recency PCTL | Transaction |
| YTD Alert Response Count | Transaction |