

Eliminate address duplicates for sure!

Duplicate-free address files guarantee immediate profit...

...in direct marketing...

Despite the growing significance of e-mail marketing, active telephone marketing and SMS advertising, addressed mailing continues to assume the clearly leading position in the array of all direct marketing instruments that are used. In doing so, high data quality of the addresses stands for more response, cost minimization and thus for more success.

...and in address data cleansing!

Everyone who is concerned with data cleansing within a CRM or a data warehouse project is inevitably confronted with the problem of merging address data without uniform reference numbers and hence with the topic of address duplicates.

A rational examination of the problem

A lot of time and effort is required to establish and maintain relationships. Companies today use diverse instruments (direct mailing campaigns, call center activities, Internet newsletters or sales) to build up and keep a continuous contact to the customer. The measures for building up a dialog are based mostly on the company's own address data, which are supplemented by external lists and which therefore show a great deal of overlapping in themselves and with each other (duplicates) which must be eliminated.

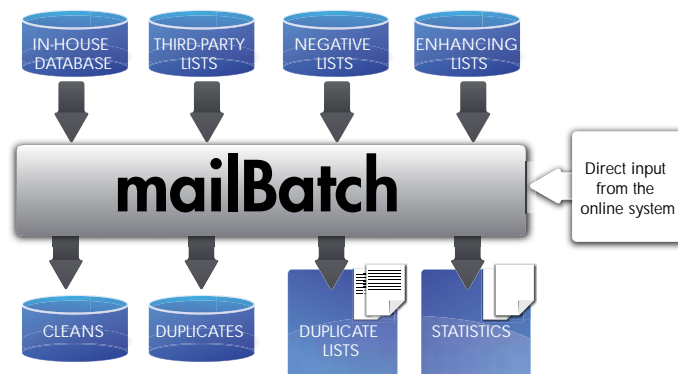
Often, however, several data sources exist alongside each other within a company. The items of information contained in them are not seldom redundant and differ in addition with respect to how the record is built up, the structure and the address quality. Here is where a lot of the duplicates occur unnoticed.

Address duplicates have an extremely negative effect. Addresses that are stored several times lead to completely unnecessary costs and to the fact that the same customer or prospect is correspondingly often confronted with identical matters of concern. Thus, direct dialog loses its credibility, and as a consequence, the relationship to the customer/prospect loses its stability.

Don't give duplicates a chance!

With our sequential software solution *mailBatch*, you achieve maximum success in detecting duplicates and ensure for yourself – in direct marketing as in address data cleansing – files that are duplicate-free to the greatest possible extent.

You will see the success directly in more profits through saved expenses for postage, printing and advertising costs per mailing campaign. And for CRM and data warehouse projects you create with *mailBatch* as a function module of the *Data Quality Batch Suite* the technical requirements for an integral view of your customer, the so-called *single view of the customer*, in that you can smoothly merge heterogeneous address data and customer information so that there are no duplicates even without a uniform reference number.



In both fields of application, *mailBatch* contributes decisively to strengthening your relationship to the customer and influences the ROI in a positive way.



Fields of application

For using *mailBatch*, the following practical applications appear to be typical; they have in common the processing of mass data.

■ Merging of address files

In the course of a central CRM project, during the creation or the regular data maintenance of a data warehouse or a customer information system, different data files need to be merged and consolidated (see also Listmix).

Important!

Of course in every duplicate check scenario the important thing is not the number but the quality and completeness of the detected duplicates.

■ File cleansing

In this application, duplicates in one's own files are tracked down and eliminated. In some applications it is recommended to consolidate the data linked to the duplicates to the address remaining in the file.

■ External cleansing

This is used for the duplicate check or comparison of one or more "foreign" files (potential addresses) against a reference file, often against one's own database, for instance to rent only those addresses for a mailing that are not already counted as one's own customers or to load only those records in the prospect-database that do not already exist.

■ Listmix

Several lists managed independently of each other are purged of duplicates both within each list and as a combined list. They are available for use without any overlapping.

■ View or cluster comparison

The objective of this duplicate check is not the elimination of the detected duplicates. Instead, automatically different "cluster views" are formed on the database. In many cases, the thus formed "cluster identifications" are entered as input into the database for further processing.

■ Negative comparison

The negative comparison plays an important role especially for mailings. With it, all addresses are eliminated that are in the negative file (examples for this are risk files or also Robinson lists, i.e. persons who do not wish to receive any advertising.)

■ Positive comparison (data enhancement)

The positive comparison, by contrast, pursues the goal of an "agreement". In a file the addresses found there in one or more files are to be "confirmed". As a rule, here the goal of data enhancement is pursued in order to transfer additional information from other files into one own customer/prospect file, for example telephone numbers, interests, etc.

Comparison in the international environment

In the course of multinational strategies address merging and duplicate elimination within international address databases and lists is becoming more and more important. That requires software solutions that can be used for different countries and language areas.

As intelligent comparison software, *mailBatch* fulfills this requirement. Using error-tolerant search methods, the system delivers first-class results even for complex tasks. For reliable duplicate detection in different countries *mailBatch* operates with a multitude of tested rule and term tables that are adjusted to the respective national singularities, which e.g. also master the conditions and requirements in multilingual countries (example: Belgium, Switzerland, Spain).

With *mailBatch*, we support all companies that must simultaneously process addresses from different countries in the context of data cleansing, CRM, data warehouse projects or projects in direct marketing.

For the respective country, *mailBatch* is equipped with extensive, integrated knowledge about the unique characteristics of the country and the language area. This knowledge is embedded in comprehensive rule and term tables and reflects the country-specific address know-how that has been tested in practice.

Performance features and functions

mailBatch is the optimal solution to really achieve maximal results – especially in the field of mass duplicate checking. The system is convincing due to a broad pallet of product advantages:

- ***mailBatch* is an expert system for error-tolerant duplicate identification** in existing data files. With it, addresses can be identified as duplicates even when the same address is differently represented because of reading, hearing and data entry errors and different abbreviations, the use of synonyms or different recording rules in the individual records (see also Technology).

Example of a duplicate group:

Sample & Son	LONDON	NW3 2QW	Sample Street 1
Sample & Sons	London	NW3 2QW	SAMPLE 1,
SAMPLE AND Son,	LONDON	nw3 2qw	sampl st
Sample and Sons	London	nw 2qw	smple Street 1
John Sample & Co	London	NW3 2QW	PO Box
Sample	LONDON		Sample Street1

- ***mailBatch* is designed for mass comparison (duplicate check) n:n**, in which the objective is on principle to check every incoming address against every other with respect to the suspicion of being a duplicate. That means, for instance, with a million incoming addresses that potentially about 500 billion (thousand million) error-tolerant comparisons must be conducted in the shortest of times. Even with today's powerful computers, this task can only be completed in an acceptable length of time with high and reliable quality only with especially designed software technology.
- ***mailBatch* provides excellent performance**, allowing the processing of the largest volumes of data in a justifiable length of time – and this with the highest quality of detection. Of course, the respective performance depends upon the hardware that is used and the respective task that is to be done. The performance spectrum ranges from more than a million addresses per hour in a “standard” duplicate check on a smaller INTEL server to far into the double-digit million range per hour on large servers and for special tasks.

- ***mailBatch* adjusts to your requirements!** With *mailBatch*, not only can all name-specific and address-specific elements be compared using error-tolerant methods, but fields related to addresses such as the telephone number, e-mail address, Web address, birth date etc. can be compared using them as well. In addition, there are also a number of blank fields with the aid of which you can also include other information in the comparison such as contract, invoice, or credit card numbers. All comparison fields can be assessed with special, error-tolerant checks specific to the respective information type.
- **The system is highly flexible.** You can adjust the system exactly to your requirements due to the extensive possibilities in the parametrization. You yourself determine under which conditions two addresses should be recognized as duplicates and in which cases these are certain or uncertain duplicates. You can determine individually which addresses should remain in the files (example: with regard to the list or the oldest, youngest or most complete record) and which addresses should possibly be removed.

Simultaneously, umpteen thousand lists (files, logical address tables etc.) can be processed in one run. *mailBatch* can differentiate between own, foreign, negative and enhanced lists and processes these correspondingly in different ways.

- ***mailBatch* makes you more productive!** The system is “ready to run”. It is capable of running by itself and, via the corresponding parametrization, it is adjusted to your requirements. The parametrization is implemented in the product line *BATCH.line* via corresponding script parameters and in the product family *INTEGRATED.systems* via the dialog interface.

Extensive presentations of the duplicate check results, e.g. via a special report writer for the duplicate list and extensive statistics on the comparison results are available.



Product variants and technology

mailBatch gives you investment security! You can use *mailBatch* as

- stand-alone tool for sequential processing sequences or as
- integrated solution with a Windows-supported user interface.

The software *mailBatch* supports the most important derivatives of the Unix world, Microsoft with Windows and mainframe systems with the operating systems z/OS, VSE, BS2000 and iSeries (AS/400).

The software is characterized by an extraordinarily exact selectivity and accuracy with remarkable performance. This is made possible due to a polished analysis of the address information with the aid of

- country-specific rule and term tables,
- language-area specific phonetics and
- address specific fuzzy logic.

Because of the availability of a multitude of country-specific analysis tables, *mailBatch* can also be utilized in an international context without any problems. Consideration of country-specific particularities is self-evident.

Product features

Processing of consumer or business addresses

Special country-specific and platform-dependent versions are available for consumer addresses (B2C) and business addresses (B2B). Compared with the B2C variant, the B2B version of *mailBatch* recognizes other company-specific components in addition to the conventional address elements and separates these from elements which relate to the contact person in a company. In addition to this, other entry fields such as the company name, legal form, home page, descriptive secondary company designators, geographical data and acronyms are recognized company-specifically and considered during the matching. Fields for information concerning the contact person in a company, such as the department or the title, are also recognized and matched. In this regard, *mailBatch-Business* contains a more extensive synonym table in comparison with the consumer variant. The B2B version is recommended for the matching of mixed B2C and B2B databases with a proportion of company addresses of approximately 30% and more.

Cooperation with the Uniserv solution *mailRetrieval*

For the integrated cooperation with the product *mailRetrieval*, special program features are available depending on the platform, making it possible for the address information managed by *mailRetrieval* to flow directly into *mailBatch*. For that reason, easier operation, higher sequence integration and an increase in performance not known until now are achieved.

Country-specific *mailBatch* versions for international use

mailBatch enables address comparisons even for international address lists and address files. For that, special country versions are available as expert systems with integrated knowledge about the unique characteristics of the country and the language area. The country availability is permanently enlarged.

Processing of relocation information

The duplicate identification of relocation information in German lists and address files and databases is supported by a service of Post Adress GmbH, which makes available the mail forwarding applications to Deutsche Post AG as current information about relocation addresses.

For the EDP-supported processing of relocation addresses, Uniserv offers the *mailBatch* supplementary feature 'relocation'.



Additional Information