Data Driven Mail Production:
A Platform for ADF Integration

By William Hart, Chief Engineer, Lone Oak Technologies, LLC

Data Driven Mail Production is a powerful tool that has proven itself in hundreds of mailing applications. The Mailing Data File can be the backbone of a highly evolved Automated Document Factory. The current design of the files entangles statement application programming and mail production technology and creates security risks for the enterprise. A lack of standards has locked mailers to a single vendor, inhibiting the development of third party software, and unnecessarily complicating installation and support. Standardization has helped many industries work together to solve common problems and speed the implementation of new technology. Whether it is linking application programs and printers (PDF, AFP, and PCL), linking websites and browsers (HTML), or digital cameras and photo printers (JPEG, MPEG) data format standards have been the key to transitioning from a single vendor product to an industry. A standard for Data Driven Mail Production can do the same for the mailing industry.

The Production Mail Challenge

The statement mailing operation is the life blood of many enterprises, being the primary channel for their incoming cash flow. High volume statement production has all the challenges of conventional manufacturing: managing materials, personnel, and equipment, and the need to constantly improve productivity and quality. It also some very unique challenges: an extreme quality requirement and an unforgiving schedule.

The quality challenge stems from the fact that every mail piece is unique. Whereas a conventional high volume manufacturing operation will produce millions of *identical* widgets, a statement mail production shop must produce millions of *different* widgets. In conventional manufacturing environments defective products can be caught by a quality inspection resulting in, at worst, a loss of productivity. In a statement mailing operation every incorrectly assembled or unsent statement is an unsatisfactory customer interaction.

The scheduling challenge comes from the time critical nature of most statement mailings. A late statement can result in a dissatisfied customer and delayed incoming cash flow.

To meet these challenges, high volume statement mailers seek the best of new technology, in particular Data Driven Mail Production (DDMP). Although, it has proven to deliver significant improvements in quality and productivity, implementing DDMP has brought its own set of challenges, leaving many mailers looking for more:

- **Ability to buy from multiple machine vendors** – The existing technology is often vendor specific, creating artificial barriers to mixing or changing vendors.
- **Decoupling of business issues from paper handling** – Often material or machine issues require changes to the statement generation software. Not only are programming resources often dear, but changing a critical business application can be risky, mandating a complex re-qualification process.
- **Benefit from 3rd party produced products** – Mailers are often forced to do in-house implementations of system components such as quality control, reporting, and work flow management because the closed nature of the systems has suppressed 3rd party developments in the industry.

A key step toward achieving these goals would be the establishment of an open architecture for Data Driven Mail Production.

What is Data Driven Mail Production?

Data Driven Mail Production is simply the mailing industry's implementation of Computer Integrated Manufacturing (CIM), a technology that revolutionized manufacturing in the late 20th century. Manufacturing equipment, such as machine tools, were outfitted to receive electronic data which communicated the specific design and engineering requirements. By joining the design and engineering processes to the manufacturing process through this electronic link the result was a reduction in errors, improved quality and significantly improved productivity.

The mailing equivalent utilizes electronic data accompanying each mailing job to control the finishing process. Transmitting the data electronically to the finishing equipment has proven superior to the older technology of placing the finishing data directly on the statement pages with some type of printed code. The electronic format easily overcomes the problems of data capacity limitations and scanning errors inherent in large printed codes.

As shown in figure 1, a simple code is printed on each statement page linking that statement to a record in the accompanying Mailing Data File (MDF). The code can be a conventional barcode such as Code 39, a 2D code

or an OCR text string. It acts both as a unique identifier for the page and a pointer into the data file. The data file record contains the actual inserting instructions for finishing the mail piece.

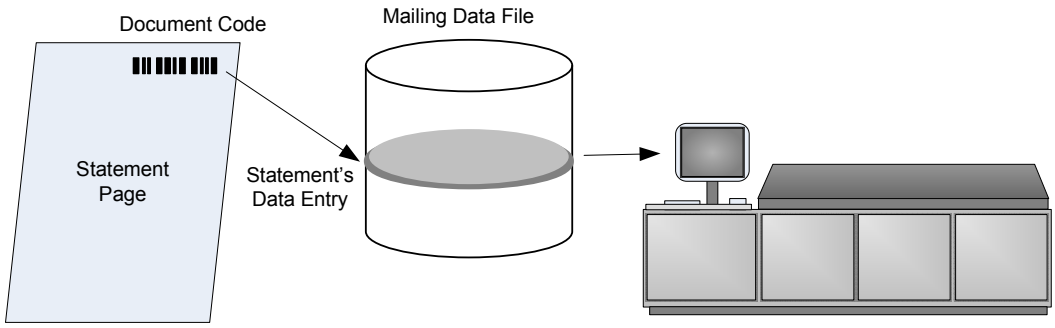

Figure 1 – Statement to Mailing Data File linkage

An electronic data flow, shown by the blue lines in figure 2, parallels the paper flow (in buff) to transmit the Mailing Data Files to the inserting operation. The electronic finishing instructions are normally created by the enterprise's data processing operation at the time that the print stream data is created either as part of the statement creation application program or by post processing the print stream data. The "Finishing Gateway" shown in the figure is a system, normally supplied by the inserter vendor that temporarily stores the MDFs and manages the distribution of them to the inserting equipment. Implementations of this gateway vary, depending on the vendor, and may include a file server, database server, workstation or a combination. In some installations, there is a reverse data flow back to the data processing center to automate reprinting of damaged mail or to feed mailing information to other enterprise systems (such as CRM).

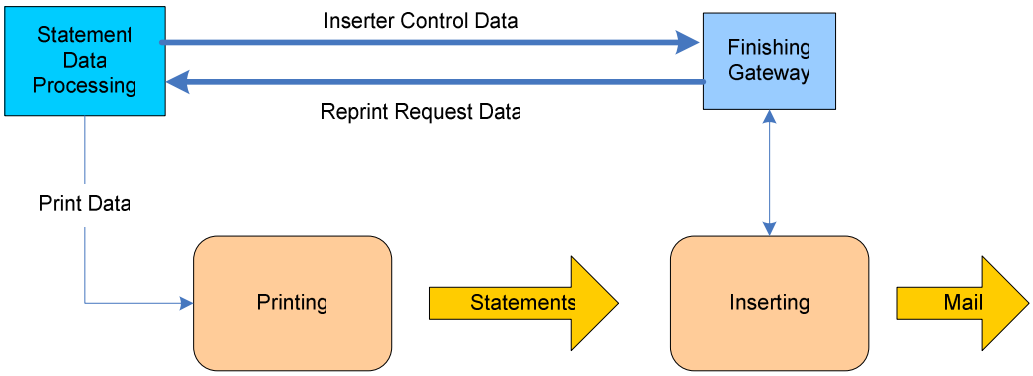

Figure 2 – Mailing Data Flow

The various inserter vendors have each developed their own flavor of the same basic technology, and of course each gave it their own terminology and product name. Different names include: "Database Driven Inserting", "File Based Processing", and "File Based Inserting." Adding to the confusion is a number of product names that lump together a variety of other functions, such as productivity reporting, or basic machine control. The result of this "name game" is that it can be difficult to compare competing solutions and understand where the biggest benefits can be gained.

(For a more complete description and history of Data Driven Mail Production see www.LoneOakTech.com/Content/WhitePapers/DDMPPrimer.pdf)

How is DDMP used today?

The basic application of the Mailing Data File is to control finishing operations:

- Collation of variable page count statements
- Selection of additional documents (inserts)
- Matching of personalized inserts, such as checks
- Applying correct postage to the mail piece
- Sorting out specific mail pieces for special handling (diverting)

Because of its larger data capacity, the MDF can also implement functions that are not possible using just a code on the statement:

- Printing the delivery and/or return addresses on the mail piece
- Printing personalized graphics on the envelope

- Providing account information to downstream operations

Moving the data electronically for these functions, rather than just in an on-document code, can yield improved productivity due to reduced code scanning problems and allow the instructions to be changed after the statements are printed (such as "late diverts).

More than Just Inserter Control

Data Driven Mail Production has shown its greatest strength addressing statement mail production's quality challenge. First, by moving the finishing instructions into the data file, it frees up statement code real estate to allow each statement page to be uniquely identified. Second the Mailing Data File is a complete check list for the mailing: it lists every piece that must be mailed and the contents of each piece. Combining a unique identifier on each page with the detail piece list in the Data file creates the foundation of a high integrity environment where every piece is either correctly processed or accurately identified for regeneration.

The Integration Opportunity

The use of the Mailing Data File marks a significant transition: from piece level processing instructions to manufacturing specifications for an entire mailing. The MDF, in a single portable electronic package, encapsulates an entire description of the mailing. It contains data for both the entire mailing and each piece of the mailing:

- The required output for each and every mail piece.
- Raw materials, such as statement forms stock, inserts, and postage
- The construction of each mail piece is specified
- SLA timings for each job

As a data file, the information is available in a compact form that can be easily processed for many applications in addition to controlling the inserter. Even if all of this information were printed on the statements, scanning them to capture the data would be impractical.

The MDF data is also timely. Since it is created in parallel with the print stream, it is available even before the statements are printed, allowing its data to be used for planning and work flow management.

Since the MDF can be a comprehensive and timely specification of the mailing, it can become the basis for a number of integration applications:

- **Work Flow Management** – The make-up and due dates for work are available in advance, facilitating planning and tracking.
- **Job Auditing** – The arriving MDF files create a "To-Do" list for the shop to help prevent jobs from getting misplaced. Especially helpful for shops with large number of small jobs.
- **Mail Piece QC Auditing** – The MDF provides the proofing data a QC inspector needs when making sample audits of a mailing.
- **Materials Management** – The MDF can be used to create a bill of materials for each mailing which can be used to direct material handlers and perform job cost accounting.
- **Enterprise Integration** – The data from MDF can be married to the inserter's results data and fed back to an enterprise CRM system adding statement mailing information to the customer database.
- **Postal Validation & Submission** –The MDF can easily provide the data needed for a manifest mailing system, improving the accuracy of this postage payment method.

What is wrong with current situation?

Few mailers are today taking advantage of these integration opportunities. The existing DDMP systems have primarily been focused on basic inserter control and have been built around vendor proprietary solutions. The benefits that can be derived from taking advantage of the integration opportunities inherent in DDMP require integrating multiple separate systems. The current Mailing Data File designs have hampered this integration

Closely Coupled to Paper Handling

In current implementations, the Mailing Data File's content is usually closely coupled to the inserter design. The file specifies when a particular device on the inserter should be activated, not what contents belong in the mail piece. So for example, the MDF will instruct the inserter to include an item from feeder 3 rather than specifying the inclusion a courtesy reply envelope (CRE).

The specification of machine operation rather than mail piece construction in the MDF complicates moving a job between different brands of inserters or even between different configurations of machines from the same vendor. For example, when feeding the CRE, while feeder 3 might be the best feeder on one model of inserter it may not even be present on another model.

This close coupling to the inserter's mechanics also complicates the relationship between application programming and mail processing. The application programmer has to know how the job will be setup on the inserter so that the correct device activations can be specified. Once programmed, the mail production operation cannot easily reconfigure the inserter or move the job to a different machine configuration.

Ideally there would be a clean separation between business rule issues and mail processing issues; just as today an application programmer is shielded from the details of how a printer operates by standardized printer languages such as PDF. A programmer can create a single print stream which can be sent to different brands and models of printers, a decision that can be made after the print stream is created. The business application programmer ought to be focusing on specifying what the contents of the mail piece should be, based on the business rules, not how it will be processed on the inserter.

Non-Standard Formats

The physical formats of Mailing Data Files are not standardized. Formats vary not only by vendor but often vary from site to site even for the same vendor. These uniquely formatted files make it difficult for mailers to mix inserters from different vendors or utilize 3rd party software tools.

Its not that the inserter vendors have created different formats capriciously. Rather they were driven by widely varying customer needs and the technology of the time. Simple "fixed field" technology was selected for most MDF designs because of its ease of specification and creation by legacy COBOL applications. The fields specified for the file varied by installation to meet the customers' needs without bloating the files with lots of blanks. The technology limitations that drove these decisions are for the most part no longer a factor.

Files are not Self-Descriptive

Modern data format designs contain "metadata": data that describes the other data in the file. This makes format designs much less brittle; programs processing the data can use the metadata to determine how to decode the rest of the file. Handing content variations becomes easier, as does migration to new standards.

The existing "flat-file" format MDF designs do not contain metadata. For a system to read an MDF it must have previously been configured to understand the particular file format. This configuration process is usually complex and sometimes requires customized programming.

Without metadata it is also difficult for systems to operate on mixtures of different format files, complicating support for mixed vendor shops and legacy statement application programs.

Open to Errors and Tampering

Existing MDF designs predate the current focus on data security.

- **Verification -** They have little or no built-in error checking (i.e. check sums) to ensure that a file isn't damaged or tampered with during transmission or storage.
- **Authentication -** There is no way to validate that the file is from the original source and has not been counterfeited.
- **Confidentiality -** There is no protection against unauthorized access, yet data in file is sometimes very sensitive (customer's names, account numbers, statement values, etc.).

In the past, MDF security has not been viewed as a significant problem because the data files were moved on the company's internal network, where public exposure and tampering were not expected. However, the industry is now learning that security protection must be implemented in-depth; at multiple layers in the organization's infrastructure. In-house networks are simply not as safe as they used to be. Implementing security features would also facilitate moving data across the public Internet, either to remote processing sites or to out-source vendors.

The increased legal demands for corporate data integrity and accountability also add to the security requirements for MDF data. This use of standardized, industry recognized protocols for verification, authentication, and confidentiality can simplify the certification process.

Time for an Industry Standard

Standardized data formats have helped many industries speed the implementation of new technology. Whether it is linking application programs and printers (PDF, AFP, and PCL), linking websites and browsers (HTML), or digital cameras and photo printers (JPEG, MPEG) data format standards have been the key to transitioning from a single vendor product to an industry

Modern information technologies make it much easier to deal with the challenges created by the large sizes and varied requirements of statement mailings. The XML (Extensible Markup Language) family of technologies is the obvious starting point for a standard. XML based formats are inherently flexible and self-describing (contain metadata), are supported by a large number of tools in virtually every programming environment, and are becoming increasingly familiar to programmers.

XML by it self, however is not enough of a standard. It only describes the basic structure of the data. A complete standard must specify the content and the context of the data.

A standard, similar to that needed for production mail, now exists for the graphical arts industry known as JDF (Job Definition Format). It is sponsored by the CIP4 (the International Cooperation for the Integration of Processes in Prepress, Press and Postpress, headquartered in Switzerland) and is a comprehensive XML-based file format used to exchange data between the various steps in the graphical arts printing processes, and is supported by a number of manufacturers. The JDF already has some support for finishing operations. A standard for Mailing Data Files would fit well as a component of or adjunct to the JDF.

Benefits of a Standard

The establishment of a mailing industry standard for Mailing Data Files would bring about a number of benefits:

- Make it easier for mailers to mix and match different models and vendors of inserting equipment.
- Facilitate the development of off-the-shelf systems to provide integrated manufacturing functionality, such as work flow, materials, and QC management.
- Provider a cleaner interface between business application programming and mail processing, reducing the number of change requests going to IT due to material issues and equipment issues.
- Facilitate the emergence of 3rd party tools for MDF creation and management.
- Improve security and confidentiality
- Improve reliability.
- Simplify the installation of new Data Driven Mail Production systems, by reducing the amount of site specific design.

These benefits would help production mailers to achieve greater productivity and higher levels of quality.

Characteristics of a Good Standard

Developing a good standard, as with any data processing system, should start with a set of clear requirements. What should the industry expect from of a standard design?

- Facilitate the decoupling of business applications from paper handling implementation.
- Based on modern standards, such as XML.

- Flexible enough to cover a range of applications and technology growth in the future. To this end it must be based on metadata.
- Create data files that can be communicated and stored efficiently.
- Allow for basic security capabilities of verification, authentication, and confidentiality.
- Be well documented.

Conclusion

Data Driven Mail Production is a technology that is still in its infancy. The formats used to move the mailing data have over complicated the interface between application data processing and mail production, and present a security risk to the enterprise. The lack of any standards has resulted in vendor specific implementations limiting flexibility for the mailer and stunted the growth of value-add applications. The establishment of a modern standard can spark the development of more secure and effective solutions for production mailers.

Bill Hart, Chief Engineer, Lone Oak Technologies, LLC is a nationally recognized expert in production mail technology. He has over 30 years experience in the management, design and installation of engineered solutions in manufacturing environments. He was the author of Pitney Bowes' *Direct Connect Inserter Control System*[1] and has designed comprehensive mail facility solutions for some of the largest mailers in the telecommunications, insurance, retailing, and banking industries. He founded Lone Oak Technologies in 2003.

Lone Oak Technologies, LLC is a production mail focused software technology firm offering a suite of state-of-the-art products aimed at productivity management, mailing data automation, and production modeling and simulation.

[1] *Direct Connect Inserter Control System* is a trademark of Pitney Bowes, Inc.