

---

*Mike Bixby*

Bixby Consulting Inc.

*Greg Alvord*

Gallagher Financial Systems

## DATA STANDARDS AND MORTGAGE CREDIT REPORTING

*Abstract: Data standards have been used in the mortgage credit reporting for the last decade. Their use has allowed mortgage lenders to more easily choose credit reporting services based on the best price, service and quality. At the same time data standards can have a dual effect on their users. Because data is provided in the same format, standards can lull the users into accepting that the quality of the data from different sources is also the same. Having data in a standard format can make it easier to identify differences in data provided from two sources. There can be many benefits from the adoption of data standards, but the need for clearly understanding the meaning of the data and its proper application never diminishes.*

### INTRODUCTION

“Data standards help the smaller companies play in the same ball field as the larger companies, and they help the larger companies be as nimble as the smaller ones.”

This statement was made by Michael Petree, Senior Vice President of First Franklin Direct, at a recent Mortgage Bankers Association conference panel session. For service consumers (lenders) like First Franklin Direct, data standards can also make it simpler to switch to, or add, service providers who support a common data standard. For service providers (credit report providers), supporting a data standard can expand their potential customer base to all service consumers that are also able to support the standard. The service providers who can offer the best price, quality, or service will still be able to distinguish themselves from their competitors.

In the mortgage credit reporting industry, the credit report data is the product being sold. When lenders are able to receive industry standard credit report data from multiple sources, they may naturally tend to assume that since the same format is used across the industry, the resulting credit report data will also be the same. This may not always be the case, especially for credit reports containing late payments, collections, or other adverse data. These differences can affect whether or not a particular lender offers

---

\* Mike Bixby; Bixby Consulting, Inc. & Greg Alvord; Gallagher Financial Systems  
© Copyright is held by Mike Bixby & Greg Alvord.

credit to an applicant and on what terms. Fortunately, the standardized data format can also make it easier for lenders to compare the quality and accuracy of the underlying data.

This article will explore these effects and others that mortgage credit reporting data standards have had on both the credit reporting service providers and the lenders who are the consumers of these services.

## BACKGROUND

This section will go into a fair amount of background information about credit reporting, because it helps bring out some of the significant, nontechnical issues regarding data standards.

### *Identifying the Players*

“Credit Bureau,” “Credit Reporting Agency,” “Credit Reseller,” and “Credit Repository” are all terms that are used inconsistently in legislation, data standards, and marketing materials to describe the players in mortgage credit reporting. It’s probably good to start by identifying these players with consistent labels and defining their roles.

The “**credit data repositories**” are the companies that collect, store, and report financial credit data on individuals in the United States. **Equifax, Experian, and Trans Union** are the three primary credit data repositories.

The “**credit bureau vendor**” is the agency that creates the credit report for the lender. They are also sometimes referred to as the “credit reporting agency” (CRA) or the “credit vendor.” For consistency, in this article we will refer to these companies as the credit bureau vendors. **Advantage Credit, Chase Credit Systems, Equifax Mortgage Services, Factual Data, First American Credco, FNIS Credit Services, LandAmerica Info1, LandSafe, and The Credit Network (TCN)** are examples of credit bureau vendors.

“**Borrower**” and “**applicant**” are often used interchangeably in mortgage industry discussions, but both are labels for the same person. When individuals first apply for a loan they are applicants. When they receive the loan they become borrowers. In the Mortgage Industry Standards Maintenance Organization (MISMO) standards, the borrower label won out over applicant because having the status of being an “applicant” only lasts a matter of weeks, whereas being a “borrower” can last for decades.

### *How a Credit Report is Prepared*

**The process of producing a credit report begins with the collection and storage of the credit data.** This is the function of the credit data repositories. Equifax, Experian, and Trans Union collect monthly financial and public record data reported by financial lenders, collection agencies, and the federal, state, and local courts. When the data is stored by the credit data repositories, it is organized into “credit files” — one for

each person. Over a period of time, as payments are made and account balances are updated, a “credit history” is built up within each credit file. This snapshot of the individual’s current credit status and total debt, combined with the payment history, provide the raw data that mortgage lenders use to determine whether or not an individual is a good, bad, or median credit risk.

**The next step in the process is to request the applicant’s credit report from a credit bureau vendor.** Not too long ago, the requests were sent by fax or mail or were submitted over the phone. Today, credit requests are generally sent as electronic transactions that use either proprietary or public data formats. When a credit bureau vendor receives the request, their software will generate and transmit one credit request to each credit data repository that was selected. Each of these requests must be in the proprietary data format specified by the credit data repository.

**The credit data repository locates the credit file data that matches the name, address, SSN, and date of birth that was submitted in the credit request.** If a “credit score” was requested by the lender then it is also generated at this point. It is calculated by looking at the borrower’s credit file and comparing its payment history and other factors with historical data from hundreds of thousands of past credit files. The “credit score” is a mathematical assessment of the borrower’s likelihood of making timely payments on future credit accounts. The higher the score, the better the credit risk the borrower represents to a future lender, a little bit like handicapping a horse race. When the credit data and credit score have been obtained, this data is converted into the credit data repository’s proprietary credit data format and returned to the credit bureau.

After the credit bureau vendor receives the requested credit file data from the credit data repositories, **additional processing is done by the credit bureau vendor’s software to produce a final credit report** that matches the lender’s specifications. Since a credit file is requested from each credit data repository for most mortgage credit reports, the data from the three sources is combined into a single report. Duplicate liability records (e.g., a “Sears” account reported by Equifax, Experian, and Trans Union) are identified and then removed and/or combined into a single, more complete liability record. The credit data is often converted into some type of printable or human-readable format, as well as reformatted into the public or proprietary credit data format used by the lender’s loan processing software.

**The resulting credit report transaction is returned to the lender** who made the original request. The entire request and response process normally takes less than 20 seconds, or more typically, less time than it took to read this paragraph.

### *Qualifying the Applicant*

How the lender comes to a decision to offer credit varies depending on their target market of potential borrowers. Lenders who offer credit only to the “prime” market may only look at the credit score and some summary data returned on the credit report as the basis for their decision. Lenders who are in the “sub-prime” market may look at the score, but focus much more on the details about the number of late payments and any collections, foreclosures, bankruptcies, liens, etc. The quality of the credit report data can be critical for the sub-prime market as we will see later in the article.

At some point, a decision is made to offer credit to the applicant, not offer credit to the applicant, or offer credit but with stipulations (e.g., applicant must pay off debt first or deliver proof of income, show tax returns, etc.). If credit is offered, the interest rate, points, and terms options are also usually set at this point and then the mortgage origination process begins. If credit is not offered or offered with stipulations, the lender and/or applicant will normally review the credit report to check if there are any errors that could affect the decision to issue credit. There are procedures for reporting errors to the credit bureau vendors so that they can be verified and then corrected by the credit data repositories in the applicant’s credit file. The updated credit file is then “rescored” and reissued to the lender so that the loan applicant can requalify using the updated credit report.

## DEVELOPMENT OF THE CREDIT REPORTING DATA STANDARDS

### *ASC X12 EDI Credit Reporting Standards*

The Accredited Standards Committee (ASC) is the U.S. standards body that coordinates the development, maintenance, and publication of Electronic Data Interchange (EDI) standards. The first public credit reporting data standards used in the mortgage industry were developed in the early 1990s. Volunteers from credit bureau vendors and lenders formed a Credit Reporting Work Group to develop data transactions that would allow the ordering of a mortgage credit report and the delivery of the credit report data.

The credit request transaction was released by ASC as the X12.833 Mortgage Credit Report Order. This transaction allowed the lender or mortgage broker to send the borrower data and other ancillary information needed to request a credit report from a credit bureau vendor.

There were two X12 transactions used for the response transaction from the credit bureau vendor back to the lender. The X12.200 Mortgage Credit Report transaction contained the credit report data – liabilities,

public records, credit scores, etc. The X12.864 Text Message transaction was used to transmit human-readable “text” versions of the credit report that could be viewed or printed by the lender.

Since ASC encourages the reuse of existing data structures, the Credit Reporting Work Group incorporated many of the common data structures that had already been developed. While reusing existing structures allowed the new credit transactions to be developed more quickly, it also meant that some of the existing data structures contained data elements and coded values that did not have any relevance to the credit reporting transactions. X12 Credit Reporting Implementation Guides were created to specify which segments should be built when creating one of the credit reporting transactions.

The X12 EDI transactions were originally designed for high volume “batch processing” commonly used in large mainframe computer environments. With the acceptance of the use of personal computers in small, medium, and large businesses in the 1980s and 1990s, data translation software that allowed the X12 transactions to be used in smaller volume environments became available. X12 data translation software was relatively expensive, and writing custom software to use or create X12 transactions was a labor-intensive effort. But because of the need and the benefit of using business to business data transactions that used a standard format, X12 was implemented enthusiastically by medium to large size lenders and a fair number of the larger credit bureau vendors.

Once the credit bureau vendor has developed software to accept the X12 833 Credit Request and deliver the X12 200 Credit Report and/or X12 864 Text Message transactions with a lender, it should be simple to add other lenders as customers whose systems also used the same X12 data formats. Like any standard, the X12 specification defines how the data should look in a data file. In actual practice, when lenders or credit bureau vendors implement a standard, they create a data file based on their interpretation of what the X12 specification says. This can lead to minor, but annoying, differences in expectations that need to be resolved in one way or another.

Of course, there have always been differences, sometimes big and sometimes small, between what is defined as a standard and the real world adherence to the standard. That’s as true with X12 standards as it is with any other standard such as the Ten Commandments, Roberts Rules of Order, or the English Language. But even with the time it takes to identify and resolve deviations from the standard, the X12 transactions greatly reduced the amount of time from the signing of a sales agreement to the first “live” credit transaction. Sometimes there is a religious fervor about the proper interpretation of the standard. Since each transaction in the realm of electronic commerce represents money being exchanged, it is usually wiser for implementers to focus on doing what it takes to begin exchanging and processing transactions as soon as possible.

Newer standards have evolved for credit reporting transactions, but the X12 Credit Reporting standards are still in use today and their use will probably continue for years to come. The compact X12 transaction files allow for fast transmission and efficient data storage. Companies that are successfully processing large volumes of X12 transactions may find it hard to justify converting to another data standard just because it's "new."

### *XML – The Extensible Markup Language*

The "buzz" regarding the Extensible Markup Language, or XML, began in earnest around 1999 when it was approved as a standard by the World Wide Web Consortium (W3C). Here is an example of XML data, which is obviously an address organized into Street, City, State, and Postal Code components.

#### XML Data Format:

```
<ADDRESS Street="401 9th ST NW" City="Washington" StateCode="DC"
PostalCode="20004"/>
```

"Washington" is clearly labeled as being a City. "DC" is clearly a State Code, and "20004" is a Postal Code. As you can see, XML is not rocket science. It provides a flexible, easy to build and maintain framework for describing and organizing data. The representation of this same data in an X12 format might appear as shown below. It's easy for anyone familiar with city names and common address formats to determine that this is address data containing a street name, city, state, and postal code.

#### X12 Data Format:

```
N3~401 9th ST NW
N4~Washington~DC~20004
```

Here are sample X12 and XML representations of the same court record from a credit report. This less common data is not easily deciphered in the X12 format as it is in the XML format

#### X12 Data Format:

```
RO~03~01~3~14~7541~LS~31~1099999~23
DTP~270~D8~20020104
DTP~621~D8~20020313
DTP~666~D8~20020313
DTP~234~CM~200203
```

#### XML Data Format:

```
<CREDIT_PUBLIC_RECORD                    _Type="Judgment"                    _DispositionType="Satisfied"
_DispositionDate="2002-03" _FiledDate="2002-01-04"
_ReportedDate="2002-03-13"   _PaidDate="2002-03-13"   _DocketIdentifier="1099999"
_LegalObligationAmount="7541" _PaymentFrequencyType="LumpSum"/>
```

For the X12 data, an implementation guides is needed to decode and understand its content. Even without looking at an implementation guide, it is as easy to understand the meaning of the XML data today as it will be ten or twenty years from now. This is an important trait for mortgage data that may be archived for up to thirty-seven years. Easy readability is also an asset during software development and testing, and whenever there are production issues that need resolution by a customer service representative.

XML has been widely embraced by the software industry. Software that supports XML is readily available and relatively inexpensive. New standards are being developed using XML as their definition framework. Some of these are vertical market standards like the MISMO XML Automated Underwriting System transaction, and many have horizontal market application like the Web Services standards, designed to simplify the access to software services across computer and operating system platforms.

So is XML the all knowing, all powerful, magic solution to all of the world's data processing problems? Not quite. For one thing, the XML representation of a set of data is considerably larger than the X12 representation of the same data, as the previous example showed. The XML format was clearly represented with 255 characters of data, while the X12 representation only required a compact 115 characters of data. In addition, because XML data files are so easy to create and implement, companies have been quick to implement their own custom XML data formats for use inside and outside their companies. XML standards can either become a *lingua franca* or a Tower of Babel, or both.

### *MISMO XML Standards*

For the mortgage industry, the promise and potential hazards of XML led to the creation of MISMO – the Mortgage Industry Standards Maintenance Organization ([www.mismo.org](http://www.mismo.org)). In Spanish, “mismo” means “the same.” The goal of MISMO is to create data structures that are “the same” and easily recognized between business partners. Like the X12 standards development effort, MISMO is made up of companies and individuals volunteering their time, effort, and business knowledge to create a data standard that hopefully is beneficial to all of the users of the standards.

After its creation in late 1999, one of the first tasks that MISMO tackled was to create a **Logical Data Dictionary** (LDD) for all of the business data that is exchanged between mortgage industry partners. This mortgage industry dictionary is probably the most valuable product that MISMO has produced for the simple reason that business data that has been clearly defined and agreed on by the industry participants is more likely to be used in a correct and consistent manner.

For example, the generic term “**Loan Amount**” could have a slightly different meaning depending on where it is used in the mortgage process. On the mortgage loan application form, the applicant is

requesting a “loan amount.” At the closing, the borrower signs a note agreeing to pay back a specified “loan amount” to the mortgage lender. The loan amount that the borrower requested on the loan application may not end up being the same as the amount that the lender is agreeing to loan to the borrower. Discussions among the participants in this early stage of MISMO led to the creation of two separate terms for loan amount in the MISMO LDD. The definition of each data element also includes its location on the Universal Residential Loan Application (URLA):

- **Borrower Requested Loan Amount** — the total dollar amount of the mortgage note that is being requested by the borrower. This amount may include financed PMI (private mortgage insurance), MIP (mortgage insurance premium), and Funding Fees. It is collected on the URLA in Section I (Amount) and in Section VII, line ‘o’.
- **Base Loan Amount** — the amount to be loaned to the borrower, not including PMI, MIP, or Funding Fee. It is collected on the URLA in Section VII, line ‘m’.

The MISMO LDD is an invaluable resource for companies who are well established in the mortgage industry and, even more so, for companies that are just beginning. Although the information present in the MISMO LDD is commonly known among the folks familiar with the business side of the mortgage industry, it is extremely educational for newcomers and can help them avoid expensive mistakes. Like any language dictionary, the MISMO LDD is always growing and evolving. It is the foundation for data structures that have already been developed and those that will be developed in the future, whether they use the XML framework or some other standard.

The first data structures produced, **MISMO Version 1**, revolved around the concept of an electronic loan package. The idea was that loan data from the various mortgage process areas could eventually be merged into an individual loan “package” like pieces of a pie. One segment of the “pie” might contain loan application data. Other segments might hold credit report data, mortgage insurance data, title data, appraisal data, loan data, etc. In the center of the “pie” was “core data” that was common to all mortgage process areas, such as borrower and property information.

(The industry’s response: “If we’re going to use XML to transact business, shouldn’t we build formats that are optimized for that purpose?”)

In mid-2001, the first **MISMO Version 2** standards were published. In MISMO Version 2, there was a deliberate move towards more “transactional” structures. Rather than focusing on the building of a single loan package, MISMO Version 2 was more oriented towards electronic commerce—exchanging business data as transactions. For example, there is a transaction for requesting an automated underwriting decision. There are others for requesting credit reports, appraisals, flood certifications, mortgage insurance,

and title insurance. Other transactions are built to provide the response data to the request transactions, such as the actual credit report data, appraisal data, etc.

(The industry's response: "Great! We've got transactions, but how are we going to consolidate all of this information into a single loan file, so that it can be recorded and archived?")

The architecture for **MISMO Version 3** is currently under development. In the computer industry, "Version 3" is a marketing label that translates as "this is the one that finally works as it's supposed to." It's hard to say at this point, but this version will probably end up as some type of blend of the "loan-centric" approach of MISMO Version 1 and the "transactional" view of MISMO Version 2. MISMO Version 3 will also correct some inconsistencies and errors, but mostly it will allow for a more precise definition of what is considered to be valid data.

#### THE EFFECTS OF IMPLEMENTING MISMO XML DATA STANDARDS

##### *Data Standards Have Leveled the Playing Field*

For companies like credit bureaus vendors whose product is the credit data, the use of data standards removes the uniqueness that distinguishes one company's product from a competitor's product. When the processed credit data is formatted according to the MISMO XML standards and implementation guides, there should normally be little difference between the content of a credit report produced by a two billion dollar company and that produced by a two million dollar company. **Using common data standards reduces the number of factors that distinguish one credit bureau vendor from another. Therefore, price, customer service, and the accuracy of the data become the most important selling points when a credit bureau vendor approaches a lender.**

The first wave of companies to implement the MISMO XML credit standards were mostly smaller companies who were looking to gain recognition and market share by demonstrating their ability to implement a leading edge, state of the art technology. Some of the established leaders in the industry also became early adopters to take advantage of the quick product development times that the steadily growing library of XML software tools offered. With the introduction of MISMO Version 2, the pace of adoption picked up when the MISMO Automated Underwriting System (AUS) standards were endorsed by both Fannie Mae and Freddie Mac. These two government sponsored enterprises purchase the bulk of home mortgages and mortgage related securities, and are vital to the success of the mortgage industry.

### *Data Standards Can Mask Data Quality Problems*

**Data standards can give the false impression that underlying data is more accurate just because it is presented in a standard way.** The standards may just be adding another layer of translation, which gives an additional opportunity for the incorrect or misleading conversion of information that often occurs in the translation process. There are a number of stages in which potential confusion and errors can be introduced as the credit data is converted from one data format to another. Two of the stages are especially critical.

**The first critical stage occurs when the original loan account data and public record data are reported to the credit data repositories,** Equifax, Experian, and Trans Union. This data is normally reported to the credit data repositories using a data standard, called “Metro 2.” In this stage, the credit grantors must convert their own internal account data and payment history data into the Metro 2 format. Whenever data is converted from one format to another, it frequently happens that for some old data there is not an “exact fit” in the new format.

For example, if a credit account balance is 120 days past due and a portion of the balance is charged off by the credit grantor, should “120 Days Past Due” or “Charge-Off” be recorded as the account’s current status? This is just one example of the numerous “decision points” that credit grantors deal with when writing software to convert their internal data into the Metro 2 format. The standards body that maintains the Metro 2 standard, the Computer Industry Data Association (CDIA), does an admirable job of addressing these types of questions in their documentation and their “Metro 2 Accuracy Workshops.” But with the sheer number of large and small credit grantors, collection agencies, and courts reporting data, along with the size and complexity of the Metro 2 standard, there are bound to be inconsistencies and errors in the data being reported to the credit data repositories.

Each of the credit data repositories then converts the Metro 2 format data into their own unique, proprietary data storage formats. When a credit report is requested by a credit bureau’s customer, another data translation takes place when the credit data repository retrieves the credit data from their databases and prepares the credit report format that is transmitted to the credit bureau vendor.

**The second critical stage of data transformation takes place during processing of the credit data by the credit bureau vendor.** Since the current trend in the mortgage industry is to request credit files from all three credit data repositories, there will be a lot of data from the three credit data repositories that will be duplicated. The credit bureau software first identifies which liability accounts and public records are duplicated. If all of the duplicates are not identified, then the borrower could appear to have more debt than they really have.

Once duplicates are identified, most credit bureau vendors begin a process called “blending” or “merging” of the data from the duplicates. This process is almost more of an art than a science and has a number of “decision points” that can have a significant effect on the content of the “merged” data record. Because of this, the final mortgage credit report data can have significant differences from one credit bureau vendor to another.

These effects are usually minimal for a credit report on a borrower that has a fairly “clean” payment history. But credit reports for borrowers with collections, large numbers of late payments, etc. are scrutinized much more closely by the lenders who offer credit to those with less than perfect payment histories. The differences from one credit bureau vendor’s credit report to another’s can affect whether or not the applicant is offered a loan, along with the interest rate and “points” options that are presented to the applicant.

Here’s an area where the data standards can help the lender ask the right questions about the credit report data and make the right choices about how that data is used in making their underwriting decisions.

#### *Data Standards Can Help Expose Data Quality problems*

For a lender, specifying the use of a standard data format among its credit bureau vendors can be helpful in uncovering data inconsistencies. A few months ago, a major sub-prime lender converted to exclusive use of the MISMO XML Version 2 Credit Reporting data format. Previously, they had used a proprietary credit report format from one credit bureau vendor and the MISMO XML credit report from another. When the lender began getting MISMO XML formats from both credit bureau vendors, **they were able to easily compare the credit data for the first time.** When requesting a credit report on the same individual from both credit bureau vendors, one credit bureau’s report indicated that there were six liabilities with a current status of “collection,” while the other credit bureau vendor's data showed 17 liabilities with a current status of “collection.” When the lender compared other credit reports, they consistently found these types of differences from the two different credit bureau vendors, even though the reports were generated from the same credit repository data.

The differences were significant enough that the lender’s Automated Underwriting System (AUS) might approve loans using one vendor’s credit reports, but reject them or trigger higher interest rates or “points” when using credit reports issued by the other vendor for the same credit repository data. This situation did not come to light until they started getting credit data from both credit bureau vendors using the MISMO Version 2 Credit Reporting standards.

Further investigation showed that when one credit bureau vendor's software merged the duplicate liabilities from the three credit data repositories into a single liability, it used the payment status from the most recent duplicate liability for the "merged" liability. The other credit bureau vendor's software used the most adverse payment status of all of the duplicates for the "merged" liability, regardless of the date that it was reported, which resulted in more liabilities being reported with a current status of "collection."

Having credit data available in a common data format allows lenders to identify these type of differences, determine their origins, and exercise the proper due diligence before using the credit data for their underwriting decisions.

## CONCLUSION

This isn't really the conclusion, because there are enough other issues that affect the content and quality of a credit report to fill a book. But for this particular article, this is the time to take inventory.

DO standards benefit us?

- Standards define how things should be.
- Standards enable us to see things the same way.
- Standards benefit those who need them.
- They impart understandable information.
- It takes discipline to adhere to them.
- Standards are a facade. They can hide what's underneath.
- Standards can make deviations, differences, and flaws more obvious.
- Focusing too much on standards can keep you from exploring better options.

Regardless of your industry, any of these answers may apply. Data standards are essential for any industry that combines data from multiple sources. Companies need to remember that just because data is presented in a standard format, it does not mean they can relax their due diligence in verifying the quality and meaning of the data before they use it in their own applications. At the beginning of this chapter, we included a quote stating that data standards level the playing field, bringing advantages to small and large companies alike. Companies and industries who recognize that data standards—or any standards—not only change the playing field, but significantly enlarge it, are destined to lead in the evolving marketplace.