

Manufacturing the **PERFECT MORTGAGE:** An **INNOVATION** Challenge

THOUGHT PAPER ON PORTABILITY STANDARDS FOR DOCUMENTS & DATA

Data Validation Sub Committee Team Members:

Lead: Limin Hu – Ellie Mae, Inc.

Parvesh Sahi – Ellie Mae, Inc.

Jonas Moe – Ellie Mae, Inc.

Brett Benson – Corelogic, Inc.

Felice Kesselring – Corelogic, Inc.

Tom Booker – Corelogic, Inc.

Christian Leask – FNC, Inc.

Shawn Telford – FNC, Inc.

Ginger Ricchetti – Regions Bank

Darius Bozorgi – Veros Real Estate Solutions

Brent Chandler – FORMFREE

January 1, 2014, Version 1.5



Table of Contents

1	Executive Summary	3
2	Demonstrating Compliance, Increasing Efficiency & Mitigating Costs	4
3	Current State Inefficiencies and Issues	4
4	Proposed Solutions.....	6
5	What is the Core Set of Data and Documents?.....	8
5.1	Data Validation Begins at Data Origination	8
5.2	Portability Standards.....	8
5.3	Data Verification and Quality.....	9
5.4	Data Ownership and Accountability	10
5.5	Validation Workflow Components	10
5.6	Data Validation Tools.....	10
5.7	Testing.....	11
6	Summary.....	11



1 Executive Summary

The purpose of this Thought Paper is to recommend approaches and steps to achieving long term, sustainable data validation and data/document sharing improvements to the mortgage lifecycle processes. This paper:


- Elaborates on the initial proposed steps
- Focuses on the details of the residential loan data set and the validation/quality assurance information exchanges
- Shares our approach, methodology, and perspectives on some of the high-level issues regarding Portability Standards for Data and Documents, and
- Suggests best practices as possible alternatives for addressing the challenges of our mission.
- Provides evidence that demonstrates the need for a common framework to facilitate industry participation in the packaging and transport of disparate loan data elements as well as the need to improve data validation and verification throughout the entire loan lifecycle.

The goal of the proposed common framework is to promote efficiency, manage risks, and minimize friction. To that end, the validation and portability recommends that:

- Data be automatically captured directly from/within the documents themselves
- A unique Loan ID is created to travel with the package
- Validation is preformed pre-delivery against the core system of record, and
- Access is managed and information sharing protocols are used throughout the chain of custody.

In addition, industry standardization is a critical factor in creating a portable package for data and documents. Therefore, we recommend promoting standardization through e MISMO and industry efforts such as Fannie Mae ULDD.

Likewise, assurance of data quality and validation throughout the loan lifecycle is a mission critical requirement. This requirement can be met by either a common protocol or an industry utility to ensure that validated data and documents are unchanged, complete, and populated correctly. The proposed



solution includes “certification of data” that tracks the chain of custody, the digital signatures, and the associated audit trails, all of which must be complete and understood.

2 Demonstrating Compliance, Increasing Efficiency & Mitigating Costs

Increased demand for mortgage transaction data, whether to adhere to new industry regulations around quality standards or to drive greater efficiencies, has placed significant pressure on ensuring data accuracy. For too long the mortgage industry has only placed value on the ability to send and receive imaged documents. There has been little in the way of requiring lenders to validate the data or authenticate the documents being transmitted. Now, with new regulations and additional requirements placed on the secondary market, as well as the rising costs of manual data validation, lenders now need processes and infrastructure to move data validation from a “nice-to-have” to a core competency of origination.

In 2008 the Mortgage Bankers Association (MBA) estimated that the average cost to originate a loan was \$3,500. Three years later, that average figure jumped over 40 percent to \$5,000. To address the new regulatory requirements, lenders “threw bodies” at the problem in the hopes of producing higher-quality loans and satisfying the changing investor guidelines. And yet, one of the biggest quality issues that investors reported was incomplete files and incorrect and/or missing documentation for correspondent loans. Defective files are commonly reported as a top operational problem and a prime source of repurchase risk. Clearly, these labor cost increases have not gotten the industry to where it needs to be nor are they not sustainable, particularly as gross margins revert to normal levels.

3 Current State Inefficiencies and Issues

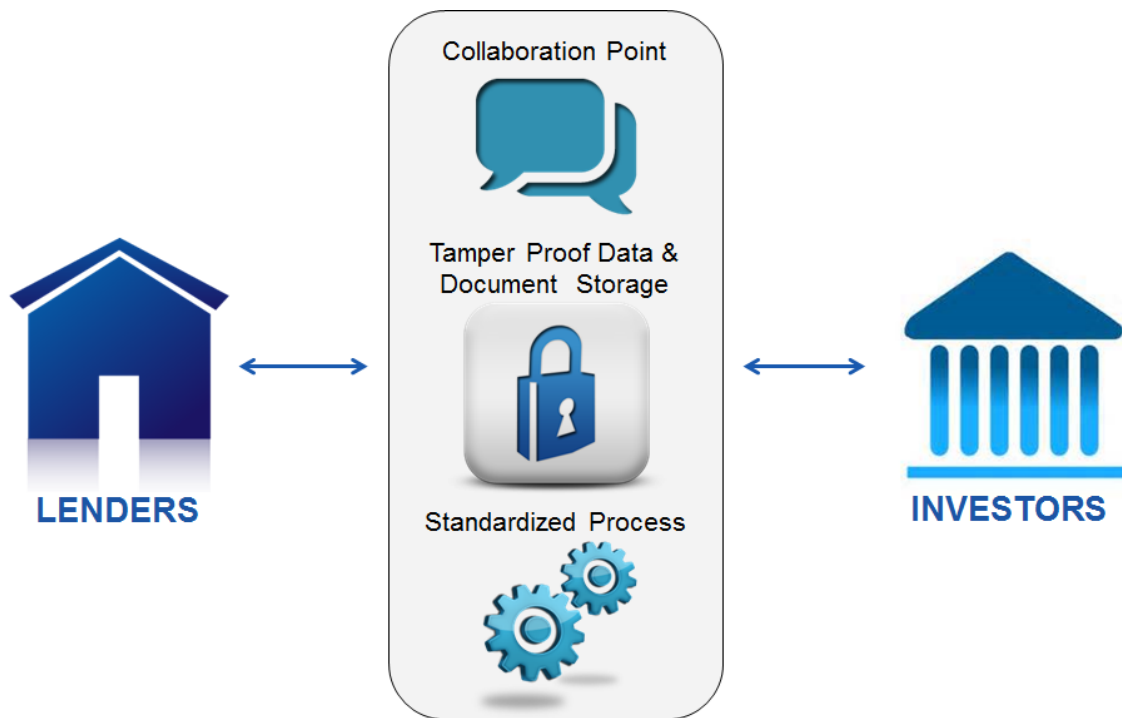
Currently there are process inefficiencies when a loan moves through the chain of custody due to several factors:

- Inaccurate or missing data
- Missing documents
- Invalid documents due to the change of source data
- Inconsistent data definition and mapping
- Compliance issues
- Manual re-work/Review of file
- Duplicative costs and work

With existing standards (UCDP, UMDP, ULDD), failures often occur at delivery. At this point, it is usually too late or very expensive to address. Therefore, early validation, indication of quality, and portability of this package is very important.

But even early validation by the lenders is not enough, as investors are almost required by regulation not to trust the information or documentation that is provided by the lenders. Compounding this challenge is the fact that investors do not have insight into how the data was used in the lending decision as there is:

- No communication between lenders and investors regarding the interpretation or application of the data in the origination process,
- No Trusted custodian of validated data and documents to ensure elements have not been altered or tampered with.



Therefore, we propose a validation process with greater insights into the lending decision and the ability to ensure the authenticity of the information used. Such a process will provide an unparalleled level of benefits to both lenders and investors alike and lead to greater trust and willingness to invest in the secondary mortgage market.

This type of validated package will:

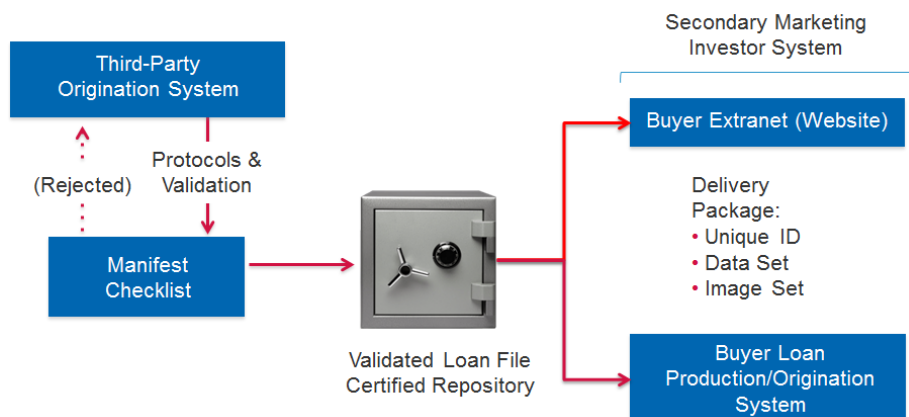
- Reduce costs
- Minimize repurchase risks
- Increase accuracy
- Increase compliance
- Reduce cycle times
- Improve pull through rates
- Deliver actionable data, and
- Improve transparency


Leveraging the standardization efforts of MISMO and GSEs, will promote the use of industry protocols and the portability of standardized data and documents. The standardization of the data, documents and workflow protocols will enable the vendor community to support lenders and Fannie Mae's requirements.

However, the lack of conformity in how this data standardization is applied today, as well as the lack of a utility to ensure the authenticity of the information, will continue to create significant friction in the supply chain. This must be changed.

4 Proposed Solutions

The proposed solution involves the creation of **data validation** and **portability protocols** as well as the creation of an industry utility to house, certify and then share the data and documents. The definitions of the *validations* and *protocols* are needed to support submission to the industry utility or eVault, in a manner that creates both delivery standardization and the ability to demonstrate that certain regulatory governances were met.





To address validation and portability issues:

- Data should be automatically captured directly from/within the documents themselves
- A unique Loan ID is created to travel with the package
- Validation is preformed pre-delivery against the core system of record, and
- Access is managed and information sharing protocols are used throughout the chain of custody

The goals are to create data-driven standards and mechanisms to ensure:

- Data and documents are standardized and data are machine-readable. MISMO standards should be explored and packaging/indexing content should be defined
- Validated data and documents are unchanged, complete, and populated correctly
- Certification of data, including tracking of chain of custody that is complete and understood
- Inclusion of digital signatures, tamper-proof seal, and associated audit trail
- Data validation with respect to the sources. We need a way to make sure that the source data in the system of record has not be changed to invalidate reports, documents, and data in the package
- Data in the chain of custody needs to be reconciled with either the data used in the closed loan docs or the images
- Compliance is monitored and the result is transparent from the beginning of the loan life cycle
- Pre-packaged documents are presented in a format consistent with how regulatory bodies will evaluate compliance

The team has identified two options necessary to accomplish these goals, but will defer to other Innovation groups to recommend the final solutions.

1. ***Industry Utility or Universal eVault.*** (As proposed by the eMortgage team.)

The industry needs a central repository that serves as the system of record for all data and loan documents. There should be a way to authenticate and certify that data and documents are consistent and accurate, and verified. Furthermore, there must be a linkage back to the source system or a third party source to verify that this is the correct, un-tampered version of data or reports. The proper level of access will be managed and granted appropriately to the authorized parties.

2. ***Industry-Adopted Transfer Specification.***

An alternative is to use technology to ensure that data are unchanged, and the ensured quality can be verified easily. The exchange of data and documents while maintain all quality and validation aspects are preserved.

5 What is the Core Set of Data and Documents?

The data validation and portability standards should include a flexible definition of the minimum data and loan documents needed to provide a trusted exchange between originator and investor. The proposed solution requires stakeholders to establish a manifest of data and documents, so a participating party can set up a workflow and rules to manage missing data or documents.

Below is the list of critical opening and closing documents, as determined by MISMO, to establish the minimum dataset:

- Standard industry documents such as URLA (1003), closing instructions, first payment letter, affidavit of occupancy, signature/name affidavit, borrower's certification and authorization, hazard insurance authorization, 4506T, W9, flood insurance, etc.
- Legal documents such as promissory notes, security instruments, and assignment of mortgage
- Federal Truth in Lending Documents such as TILs, Itemization, and ARM disclosure
- RESPA documents such as GFE, HUD-1, initial escrow account statement, and notice of assignment, sale, transfer of servicing rights
- HOEPA/PMI documents such as PMI disclosure and amortization schedule
- ECOA documents such as fair lending notice and notice of right to copy of appraisal

5.1 *Data Validation Begins at Data Origination*

Data is either gathered or generated. Every data point originates from someplace in the Mortgage Ecosystem. "*Data Origination*" is the point where the data enters the Mortgage Ecosystem for the very first time. The data could range from a borrower signature to a calculated rate for financing. All data must have a measurable level of accuracy and verification. In our proposal, the focus is to ensure that data is unchanged, complete, and populated correctly. Furthermore, any validation and quality measurement can be guaranteed and verifiable through the loan cycle.

5.2 *Portability Standards*

The **first time data** appears, it should have:

- A Unique identifier across the entire Mortgage Ecosystem
- Timestamp
- Originator: Entity where data first appears, ie lending institution.
- Data and images payload: The data stored securely. There should exist and audit trail for all data access and modification. Data is never deleted. LEGACY data may be pulled separately

Additionally, there are other components of the portable package::

- Language Transfer: MISMO/XML specification
- System of record certification: the originating system provides certification that the controlled data and documents have not been altered or tampered with and that the chain of custody has not been broken.
- Security: data and documents in motion
- Business rules: origination workflow on when data and documents are fed into the vault.
- Access credentials
- Digital signatures
- Version/dates
- Indexing of images
- Data Type:
 - Retail to Investor
 - Broker to Wholesale to Investor
 - Retail to Investor & Servicer (retained/released)
 - Servicer to Servicer
- Even if a data item changes, it will keep the same unique identifier across the entire Mortgage Ecosystem with an updated timestamp and new data quality level and new data payload.

5.3 Data Verification and Quality

Ultimately, it is imperative to understand and verify the quality of the data, especially when it is comes from so many areas and is generated by so many systems. Ideally, the investor needs to know:

- *“What is the quality of a loan?”*
- *“What is the borrower's ability to repay a loan?”*
- *“How accurate and truthful is the borrower's data?”*

In addition to substantiating the borrower's ability to repay and authenticating the quality of the loan, the process must also ensure that the loan complies with the various regulatory requirements. If the process is designed correctly, it will enable investors to quickly reconcile information and approve loans for purchase.

5.4 Data Ownership and Accountability

The Originator of any data takes on the responsibility of “Data Ownership and Accountability of how data is gathered or generated” on behalf of the entire Mortgage Ecosystem. This means the originator introduces new data and makes every reasonable effort to provide Best Possible Data. This aspect is critical because the data can be traced back to where it started and any changes can be propagated back up the change, with all parties notified in the process. The Originator of the data is the authority of data exactness and correctness, therefore accountable for data in regard to the entire.

It is important to note that data elements pulled from outside the Mortgage Lending Ecosystem may be incomplete or inaccurate. In these cases, there must be a notification and correction mechanism in place. Ideally, this function will reside in the platform that the Originator uses for its origination process. That platform must be able to handle the standardization requirements as well as protect the chain of custody and validate the information.

5.5 Validation Workflow Components

- Requirement 1. A Unique Loan ID is issued
- Requirement 2. Stakeholder deposits data and documents and provides proof (seal) that data and documents are consistent – unique loan ID is attached to the file
- Requirement 3. Validation and quality preservation is started
- Requirement 4. Data validation and quality metrics are guaranteed and re-verifiable either with a tamper-evident seal or by an eVault service provider
- Requirement 5. Alerts are issued if data and docs are out of sync with source data
- Requirement 6. Audit trails of all changes are created and access to the history is made available
- Requirement 7. “Digital signatures” are provided to ensure correctness
- Requirement 8. Data and documents remain authentic so no redraw or recheck is needed
- Requirement 9. Verification is done to ensure captured data is consistent with images
- Requirement 10. The validation process is automatic

5.6 Data Validation Tools

- Upload tool with built-in protocol support
- Validation check
- Access tool to view and download into a destination system
- Quality check such as EarlyCheck
- Compliance tool
- Verification tool



5.7 Testing

Out of scope for this paper.

6 Summary

Catching defects before a loan is offered for sale streamlines and enhances the interaction between lenders and investors. However, the greatest value is in the reduction of defects identified before a loan is funded by the lender. Today, investors spend time and incur expenses to resolve post funding conditions that should have been prevented or corrected upfront during the origination process. This leads to a lack of trust between lenders and investors. Instead, if lenders are able to provide actionable data, trusted and validated documents, and greater insight into the lending decision, it will reduce the number of errors and the associated costs while enhancing the relationship and level of trust between Lenders and Investors. Likewise, it will slash the funding constrictions that lenders face to resolve post-funding investor conditions.

When the loan is offered and available to the investor through the “e-vault”, quality control teams can access and or download copies of the original images and data instead of re-ordering it. Our proposed workflow ensures that the investor and sellers both work off identical images, data, reports, etc. This eliminates the need for re-keying and re-orders. Both originators and investors save time and money as a result. More importantly, actionable data can be leveraged by the lender and they know what checks to run to accelerate the process of selling a loan to a specific investor.