Is It All About the Data?

As the mortgage industry looks to restore confidence and liquidity, lenders will have to **take a closer look** at the data.

BEFORE GRANTING A MORTGAGE THE LENDER WANTS to know if the value of the property is sufficient to cover the loan amount and is the borrower able to repay the loan? Information (data) must be collected, analyzed and validated to properly assess the risk. The capability to manage that risk and eliminate fraud in the loan process is the reason why data integrity is so important.

The first opportunity to control data quality starts with the initial collection of application data. Missing vital information is critical. So we build applications to identify required data and control the input with error trapping. If the information is not readily available, the user tends to enter invalid data as a workaround to move the process forward.

Transposition or misspelling adds another potential opportunity to impact the data quality. The next opportunity occurs as the processor interacts with other parties providing their respective information. This processor may unintentionally alter previously correct information and the intentional misrepresentation of information to perpetuate fraud is always a constant threat.

Once all of the information and supporting documentation is gathered the difficult task of evaluating and underwriting the loan begins. The underwriter examines that information and determines the risk using what is commonly referred to as the 4 C's:

1. Collateral: What is the value of the real estate property? The underwriter reviews the real estate purchase agreement, if applicable, the property appraisal (including the comps used for comparison) and takes into consideration the market area. Is it a new energy-efficient home or an older home facing possible repairs? What is the borrower's investment in the collateral?

2. Creditworthiness: What is the borrower's credit history? Can the borrower show the intent to pay? Is the borrower willing to pay?

3. Capacity: What is the borrower's ability to pay the monthly mortgage and other obligations? What is the borrower's employment history? What are the borrower's incidental expenses? What is the borrower's residual monthly income?

4. Capital: What is the source of the downpayment? What is the source of the funds to close? What is the borrower's cash on hand?

All of these questions are designed to evaluate the loan application. We collect information about the borrower and subject property. We check and double check the authenticity of that information. We use outside sources to check Social Security databases and analyze for possible fraud. We use data to control the workflow of the

loan process.

Clearly, knowing the information used to make that underwriting decision is accurate and truthful is crucial.

Let's look at how technology has helped the mortgage process. We use technology to automate the loan approval process and reduce the bottleneck of human intervention. But, in the haste to automate have we lost sight of the desired end result? Should we make this loan? Did we rely and depend on technology automation in the process?

Automating the ordering of appraisals and standardized responses has made a difference. Likewise, the process has benefited from automatic credit reports and the establishment of the standard FICO score. The ability to electronically import the liabilities from the credit bureaus into the loan origination systems has improved the data quality. Verifications of employment, income and deposits are still somewhat of a manual process but with the potential to become a more automated electronic exchange. The definition of industry standard data formats plays an important role in the exchange of information.

When we examine, what some feel is one of the industry's most significant technology solutions, Automated Underwriting Systems, we find the objective of AUS was to automate the analysis of the loan data, compare it to the parameters of the loan product and provide feedback to the underwriter as to how the loan complies with the investor requirements. There is no doubt that AUS had a great impact on the loan process.

However, we have to question whether we have relied on technology at the expense of reasonability. In the deployment of AUS, this "black box", the industry has bypassed some of the basics of underwriting in the evaluation of a loan application.

We were approving loans with back-end ratios of close to 60% compared to the industry benchmark of 36%. The elimination of the once-standard 20% downpayment, NINJA (no income, no job, no assets) loans and the acceptance of lower FICO scores combined with creative financing like the payment option ARMs, all in the pursuit of homeownership, has led to the current economic crisis. We missed all of the warning signs. Now we need to go back to the basics of mortgage loan underwriting.

So, why is this all about the data? We can't make an informed decision on fraudulent, erroneous or incomplete information.

I often quote a 2002 study from The Data Warehousing Institute that stated, "Current data quality problems cost U.S. businesses more than $600 billion per year." The study pertained to all businesses not just financial services and it is a little dated.

> "We can't make an informed decision on fraudulent ... information."

However, I would argue that the amount probably would be higher now and financial services would still be a significant contributor. Mortgage lenders have had difficulty in determining the true costs to originate a loan and close loans so it is understandable that the true costs for poor data quality are also difficult to ascertain. Suffice to say there is room for improvement.

There is a data processing term, GIGO, that has been around for probably 50 years. Garbage in, garbage out implies that computers will unquestioningly process the most nonsensical of input data and produce nonsensical output.

Let's go back to the original premise. Is the value of the property sufficient to cover the loan amount and is the borrower able to repay the loan?

Data, if properly managed, will be a strategic asset. The following excerpts from an article on business intelligence illustrate the four phases of how data is transformed into useful information.

Data: This stage is collecting raw business data from multiple sources, standardizing it and storing it. The objective of this level is to provide an integrated, cleansed, high-quality data repository.

Information: At this stage we start to leverage our data assets, put them in the right context and translate them into something meaningful.

Knowledge: This level is where we start to absorb and understand the information.

Wisdom: At this level, businesses will be able to make sound, timely and effective business decisions, improve their processes and gain a competitive advantage.

So it is not enough to just collect the data. We need to make a collective effort to quantify and qualify the data. Further, we have to have confidence in the data used to make that loan decision. The way to do that is to control and lock down the data used to grant that loan and ensure that data is delivered to the secondary market to restore the confidence.

So, is it all about the data? Yes, it most certainly is. MT

Roger Gudobba is the chief strategy officer at Compliance Systems. He currently is a member of the MBA's Residential Technology Providers Forum. Mr. Gudobba was a member of the MBA ResTech committee in 2002 and a member of the MISMO Governance Committee from 2000-2008. He has devoted his 20-plus years in the mortgage industry to improving the loan process. Mr. Gudobba led the effort of creating electronic laser form images and actively participated in the development of XML data standards.