

Metadata Use in the Commercial Banking Industry

by Joan Starr

Joan Starr is a graduate student in the School of Library and Information Science at San Jose State University. She is also a senior project consultant in the Information & Technology Services Department at the Federal Reserve Bank of San Francisco. She can be reached at joanstarr@earthlink.net. The views expressed in this article are those of the author and not necessarily those of the Federal Reserve Bank of San Francisco or any federal entity.

In the library and archive information professions, metadata is typically presented as an enabling tool for information resource retrieval and/or management. In the industrial and commercial worlds, however, metadata can serve an altogether different goal: information exchange, either in a transaction or report format. The exchange paradigm of information use is scarcely apparent in information science literature, perhaps, in part, because transactional information is, by nature, pre-archival. Yet, the World Wide Web has blurred the lines between that which is fixed and stored and that which is underway, so, perhaps, this is an appropriate time to examine metadata use in the commercial sector.

This inquiry will begin with a brief introduction into the information-related behaviors of the American banking industry and include a discussion of how the use patterns have changed in recent years. It will further describe the development of one particular metadata standard currently gaining momentum and endorsements in the financial sector. Finally, it will identify a number of critical considerations that may impact the future direction and acceptance of that metadata standard.

Classic Information Behavior of Banks

To gain an overall understanding of the role of metadata in banking operations, it is first useful to delineate the specific types of information banks typically handle, as well as what they do with that information.

Many Types of Information. Commercial banking is an enterprise with several distinct, but inter-related functions, each of which generates large amounts of information. Banks maintain information on deposit accounts, loan accounts, customer profiles, bank investments, market research and corporate accounting. This last category includes general ledger and balance sheet data and also reserve account balance information that the bank must report to regulatory agencies such as the

Federal Depository Insurance Corporation (FDIC) or the Federal Reserve Bank. Banks, like any modern organization, also have personnel-related information systems, and some very large banks support their internal and external communications with document-management systems as well. These systems have allowed banks to digitize signature cards, automate loan application processes and replace many other paper-based workflows.

To accommodate this range of content types, bank systems deal with several distinct information format types. Account information, either regarding individual deposits or loans or in connection with the corporation's accounts, tends to be almost exclusively numeric, although borrower's signatures are increasingly likely to be stored as graphic files, typically in TIFF format. Market research, on the other hand, may include numeric data as well as textual analysis, and perhaps even graphic representations. If the bank is using a document-management system, the system will be configured to handle a wide range of file types, including word-processed documents, Web pages, spreadsheets, slide presentations and perhaps even audio and video files. Where there is no document-management system in place, this full range of file types is typically stored on a shared (internal) network drive.

Three Principal Uses of Information. Banks, then, possess a wealth of information. And, the range of information types has a correspondingly broad range of information uses, falling into three categories: transaction, retrieval or reporting. For the purposes of this discussion, transaction is defined as information access in which content is changed. Retrieval is defined as information access in which content is *not* changed. Reporting, perhaps the most interesting category, presents an intersection of transaction and retrieval. That is, with reporting, the individual content elements are not changed, but they may be recombined and manipulated in such a way as to create a completely new and additional information packet.

Table 1. Use of Information in Banks

INFORMATION TYPE	DATA TYPE(S)	TYPICAL USES
Deposit and loan accounts	numeric, some graphic	Depositor or borrower access for view-only from corporate website, and for update from automated teller machines (<i>Retrieval or Transaction</i>) Employee access for view or update from internal systems (<i>Retrieval or Transaction</i>)
Customer profiles and loan applications	text, numeric	Internal automation systems access in a system-to-system relationship (may or may not include updates) (<i>Retrieval or Transaction</i>) Employees can access for view or update from internal systems (<i>Retrieval or Transaction</i>)
Investment files	numeric	Employee access for view or update from internal systems (<i>Retrieval or Transaction</i>) Shareholder reports (<i>Reporting</i>)
Corporate accounts	numeric	Shareholder and Regulator reports (<i>Reporting</i>)
Internal communications	text, numeric, graphic	Employee access for view or update from internal systems (<i>Retrieval or Transaction</i>)
External communications	text, numeric, graphic, multimedia	Public access for view-only from corporate website (<i>Retrieval</i>) Employee access for view or update from internal systems (<i>Retrieval or Transaction</i>)

Table 1 summarizes the discussion above and illustrates the distinctions made between the three information use types. Briefly, most bank systems for information access must allow both retrieval and transaction activity, depending upon the identity or status of the individual seeking access. The third use type, reporting, is critical to corporate accounting and investment information which must provide new compilations and combinations of data to shareholders and regulators.

Evolution of the Banking Industry

All the information described above, with the likely exception of document-like objects stored in a document-management system, has resided without descriptive metadata schemes in banks' computer systems. Most of the information (accounts, profiles and related graphics) has been retained in relational databases and accessed for viewing (retrieval), update (transaction) or extraction (reporting) using proprietary tools. This includes Web portal access tools provided by database vendors for intranet, extranet or Internet access. And, as indicated above, word-processed documents, Web pages, presentations and spreadsheets have been stored "loose" on network drives. Any access from outside the organization's network is increasingly provided using some variety of hypertext markup language (HTML) served up to an Internet or extranet site. Internal access for update (transaction) or extraction (reporting) activity is gained through the proprietary software tools used to create the files. Internal view-only access (retrieval) may use the company's intranet site or the proprietary software.

Needless to say, metadata projects are both complex and costly to implement and maintain. For the private sector, this effort is a program of such magnitude that it must proceed on the basis of a truly compelling cost-recovery (or cost-avoidance) case or a regulatory requirement. Interestingly, the banking industry happens to be poised at a convergence of both of these arguments.

The Internet Changes Commerce. It is widely acknowledged that the Internet radically alters the commercial environment. In 2000, Gary Gensler, undersecretary for domestic finance in the Clinton Administration, told a meeting of the Bank &

Financial Analysts Association, "There may be no part of our economy more suited to delivery in electronic form than financial services. . . . The Internet creates a 24-hour marketplace for financial services" (Gensler, 2000, p.50). Suddenly, not only do the bank's depositors and borrowers expect to see – and even manipulate – their account information on the Internet, but also the bank's suppliers would like to establish secure extranet relationships for data exchange. Indeed, the entire financial information community now seeks to access, exchange and research data in ways previously thought to be impossible. Bank executives are facing significant costs of *failing* to provide the new access with new tools.

Congress Changes the Rules. In 1999, Congress passed the Financial Services Modernization Act of 1999, commonly referred to as Gramm-Leach-Bliley. This act repealed the 1933 Glass-Steagall Act, which created statutory separation between commercial and investment banking. Gramm-Leach-Bliley authorizes the creation of financial holding companies to act as containers for commercial banks, securities firms, insurance companies and other financial services. Gramm-Leach-Bliley effectively introduced a wave of mergers and acquisitions. This has two effects: banks are now handling new kinds of information, and banks are challenged to combine, or at least reconcile, internal information from multiple pre-merger organizations.

The Next Generation of Banking

With all this activity, it may come as no surprise that there are many proposed financial information standards. A comprehensive list would include the following standards for information exchange and transactions: ACORD (insurance), FinXML (capital markets), FIXML (securities), FpML (financial derivatives), MDDL (financial markets), MISMO (mortgage industry), OFX (consumer finances), RETML (real estate), STPML (securities). Two more standards have been proposed specifically for financial industry messaging, IFX and SWIFTML. Finally, there are currently three proposed reporting standards: IRML and RIXML both for economic and investment research reporting, and XBRL for business reporting. For an explanation of all these acronyms, please see Table 2.

Table 2. Metadata Standards in the Financial Services Industry (Note: All Websites visited on February 8, 2003.)

ACRONYM	FULL NAME	WEBSITE
ACORD	Association for Cooperative Operations Research and Development	www.acord.org/home.aspx
FinXML	(na)	www.finxml.org/
FIXML	Financial Information Exchange Protocol	www.fixprotocol.org/cgi-bin/Welcome.cgi
FpML	Financial Products Markup Language	www.fpml.org
IFX	Interactive Financial Exchange	www.ifxforum.org/ifxforum.org/index.cfm
IRML	Investment Research Markup Language	www.irml.org (may be defunct)
MDDL	Market Data Definition Language	www.mddl.org/default.asp
MISMO	Mortgage Industry Standards Organization	www.mismo.org/mismo/index.cfm
OFX	Open Financial Exchange	www.ofx.net/ofx/default.asp
RETM	Real Estate Transaction Standard	www.rets-wg.org
RIXML	Research Information Exchange Markup Language	www.rixml.org
STPML	Straight Through Processing Markup Language	www.stpml.org
SWIFTML	SWIFT Markup Language	www.swift.com

Industry analysts warn that this glut of standards efforts, especially for transactions, will inevitably result in some fallout. The suggestion is that, ultimately, all transaction standards will coalesce around the ISO 15022 specification. ISO 15022 is the standard for message exchange in the securities industry. As noted, one area where there appears to be a *bona fide* standard is in business reporting, with XBRL.

A Brief History of XBRL. The genesis of XBRL is in the field of accounting, rather than banking. In 1999, the American Institute of CPAs gathered a group of accounting firms and technology vendors to work together on an XML-based standard focused on the requirements of accounting reports and aiming to serve all stakeholders for that information. The first XBRL specification was made available in July 2000. In early 2001, the XBRL consortium held its first international meeting, and, significantly, the FDIC joined the organization. Later that year, the federal government's Joint Financial Management Improvement Program recommended the use of XBRL to U.S. government agencies. At the close of 2001, the Bank of America conducted a pilot using actual client information, allowing customers to upload financial information from QuickBooks and then convert the data to XBRL format.

By December, 2002, the membership of XBRL International, the standard's governing body, had reached 170, and its steering committee included representatives from major accounting firms, such as Ernst & Young LLP, Morgan Stanley Dean Witter, and PriceWaterhouseCoopers, technology vendors such as Microsoft, regulators such as the FDIC, and academics. The XBRL US Domain Group chair and Microsoft executive, Rob Blake, has been quoted as saying, "The XBRL consortium is becoming more like a Who's Who list of vendors, banks and accountants" (Solomon, para. 13). This may be the best way to explain the standard's dominance in the financial reporting market.

Under the XBRL Covers. The XBRL specification consists of a document type definition (DTD) for a core bundle of ele-

ments and attributes, as well as a taxonomy syntax for creating new elements and attributes. To date, a number of taxonomies have been created, and XBRL International uses a ranking system to keep them organized. Only one, the Journal Taxonomy, is in a "Recommended, Final" state. Others, in "Recommended, Public Working Draft" or "Approved," include the Global Common Document Taxonomy, the U.S. Financial Reporting Taxonomy and several International Accounting Standards taxonomies. The

XBRL International website also hosts royalty-free, conforming taxonomies its members have created. At the present time, these include a Management Report Taxonomy and a US GAAP ("Generally Accepted Accounting Practices") Taxonomy.

XBRL does not include any explicit provision for a controlled vocabulary. One can argue that a cohesive community of practice, such as the accounting industry, with its national and international standards frameworks, supplies a *de facto* controlled vocabulary. As long as XBRL implementations and extensions are limited to this subject area, the lack of a formal vocabulary structure may not be discernible. The Version 2.0 specifications suggest that vocabulary problems may lie ahead by acknowledging that XBRL could be used outside the financial services arena. The specifications explicitly avoid an authority control role for XBRL International, but they include support for business entity authority control provided by some other organization, such as NASDAQ or CUSIP (a service bureau run by Standard & Poor's for the American Bankers Association). This takes the form of an identifier element with an attribute of scheme, as in the following example from the specifications:

```
<identifier scheme="www.nasdaq.com">SAMP</identifier>
One may translate this as "The company with the NASDAQ
ticker symbol SAMP."
```

For demonstration purposes, Reuters Group created a number of example financial reports, exposing the XBRL coding (Reuters Group, 2001). Here is the code for cells in a table entitled, "Reuters Group Revenue Analysis – Third Quarter 2001."

```
<td align="center" valign="middle"><p>
<xbrl:item type="ri:inc_exp_ord_ops_rev_rtrFinancial-rts" period
="P3M/2001-09-30" unit="ISO4217:USD">313</xbrl:item>
</p></td>
<td align="center" valign="middle"><p>
<xbrl:item type="ri:inc_exp_ord_ops_rev_rtrFinancial-rts" period
="P3M/2000-09-30" unit="ISO4217:USD">288</xbrl:item>
</p></td>
```

This example shows several notable characteristics of XBRL. First, it can be used within HTML documents. Second, its most basic unit is the “item,” and each item has sufficient metadata attributes to provide unique identification. In this example, both items are of the same type, an income category of operations revenue. In XBRL, every item has an associated time period. The first, 313, is being reported for the third quarter, ending September 30, 2001, and the second is being reported for the third quarter, ending September 30, 2000. Notice also the reference to an authority control, with the “unit” attribute pointing to ISO 4217. This is the International Organization of Standards list of currency type abbreviations.

Future Directions. XBRL looks very promising as an emerging standard. It has been designed for flexibility, extensibility and international portability, although, as noted above, it has made little provision for vocabulary control, and it presently allows for only two types of data, text and numeric. Also, its governing board, XBRL International, is a collaborative body of virtually every important stakeholder in financial reporting, with the exception of individual investors and borrowers.

Indeed, there are a number of key indicators that point to XBRL’s dominance. Numerous financial services software vendors now provide either add-on components to translate proprietary outputs into the XBRL format or completely new products, building XBRL compatibility into all the automated processes associated with financial reporting. This development is key, as borrowers, investors and back office clerks are not liable to perform XBRL encoding by hand. One assumes that there is no coincidence involved in the placement of the Microsoft representative to XBRL International, Rob Blake, at the head of the U.S. delegation.

The other interesting partner on the XBRL International Steering Committee is Phil Walenga, the XBRL project manager for the FDIC. This regulatory agency collects data on a quarterly basis from the majority of U.S. depository institutions in the “Report of Condition and Income,” more commonly known as the “Call Report.” In the summer of 2002, the FDIC, under the umbrella of the Federal Financial Institutions Examination Council (FFIEC), issued a request for proposals for the development of a system to automate the submission of call reports, using XBRL. The FFIEC is a consortium of the FDIC, the Board of Governors of the Federal Reserve System (FRB), the National Credit Union Administration (NCUA), the Office of the Comptroller of the Currency (OCC) and the Office of Thrift Supervision (OTS). This amounts to a wholesale endorsement of the standard, as it will force the financial institutions to adopt XBRL in order to comply with reporting requirements.

In the wake of the Enron scandals, some XBRL promoters are looking at the standard as a way to make corporate financial information more transparent and more easily and meaningfully analyzed. A first step toward the fulfillment of this promise is EDGAR Online. EDGAR Online is a proprietary online tool that is widely used to gain easy access to the information reported to the Securities Exchange Commission’s Electronic Data Gathering, Analysis and Retrieval (EDGAR)

For Further Reading

- Felber, R. (2000). Turn and face the strange changes: Business and finance librarians and financial services reform. *Business & Finance Bulletin*, 113, 45-48. Retrieved January 30, 2003, from www.slabf.org/FELBR113.pdf.
- Gensler, G. (2000) Internet model adds new dimension to financial services business strategy. *Bank Systems & Technology*, 50. Retrieved January 30, 2003, from EBSCOHost database.
- Hampton, L., & Vun Kannon, D. (Eds.). (2001). *Extensible Business Reporting Language (XBRL) 2.0 Specification*. Retrieved February 2, 2003, from www.xbrl.org/resourcecenter/specifications.asp?sid=22.
- Penler, P., & Schnitzer, M. (2001). *Web-Enabled Business Reporting for the Banking Industry* (Ernst & Young LLP and Morgan Stanley) White paper retrieved February 8, 2003, from www.xbrl.org/resourcecenter/whitepapers.asp?sid=21,
- Reuters Group (2001). *Reuters Group Third Quarter and Strategy Statement*. Retrieved February 8, 2003, from <http://about.reuters.com/results/2001-q3/xbrl/default.asp>.
- Solomon, M. (2002). XML technology: A standard for all to understand. *Bank Technology News*, 15(12). Retrieved January 29, 2003, from Factiva database.

system. In early 2001, EDGAR Online announced an XBRL repository for financial statements. The goal was to make available a single source of company financials for use by analysts, investors (individual and institutional), journalists, accountants and others. This will have the tendency of making financial and marketing analysis more consistent and probably speedier as well.

Overseas, several countries have actually endorsed XBRL as national standards or, at least, recommendations. These include Australia, Singapore and the United Kingdom. The incentive is apparently cost-avoidance, with one Italian bank making loans more affordable for companies that use XBRL-compliant accounting software. The discount reflects the bank’s belief that these borrowers are likely to have more consistently audited books, presumably making them a better credit risk. An organization of chief financial officers in Europe, CFO Europe, identified XBRL as number eight in its list of “Top 20 technology trends” for 2003, suggesting that the standard is likely to encourage the creation of many new and innovative financial reporting tools. These developments increase the pressure upon domestic banks to adopt XBRL as they compete for global investments.

Conclusion

With its many information constituents including federal regulators, the commercial banking industry has an increasingly powerful incentive to coalesce on a metadata standard for the exchange of information. While the field of contenders is still quite large for transactional exchanges, there is a clear leader for financial reporting. Economists and other social scientists should welcome this development, inasmuch as XBRL-formatted call reports will soon become the stuff of trend analyses and other statistical research. The metadata scheme designed for information exchange will turn out to enable efficient information retrieval.