# Exemplifying Business Opportunities for Improving Data Quality From Corporate Household Research

Stuart Madnick, Richard Wang, Krishna Chettayar, Frank Dravis, James Funk, Raïssa Katz-Haas, Cindy Lee,Yang Lee, Xiang Xian, Sumit Bhansali

Working Paper CISL# 2004-03

February 2004

Composite Information Systems Laboratory (CISL)
Sloan School of Management
Massachusetts Institute of Technology
Cambridge, MA 02142

This page is blank

# Exemplifying Business Opportunities for Improving Data Quality From Corporate Household Research

Stuart Madnick, Richard Wang, Krishna Chettayar, Frank Dravis, James Funk,Raïssa Katz-Haas, Cindy Lee,Yang Lee, Xiang Xian, Sumit Bhansali

**Abstract:** *Corporate household* (CHH) refers to the organizational information about the structure within the corporation and a variety of inter-organizational relationships. Knowledge derived from this data is becoming increasingly important for improving data quality in applications, such as Customer Relationship Management (CRM), Enterprise Resource Planning (ERP), Supply Chain Management (SCM), risk management, and sales and market promotion. Extending the concepts from our previous CHH research, we exemplify in this paper the importance of improved corporate household knowledge and processing in various business application areas. Additionally, we provide examples of CHH business rules that are often implicit and fragmented – understood and practiced by different domain experts across functional areas of the firm. This paper is intended to form a foundation for further research to systematically investigate, capture, and build a body of corporate householding knowledge across diverse business applications.

**Key Words:** Corporate Householding, Data Quality, Organizational Structures, Inter-dependence, Name Matching, Entity Aggregation, Information Quality, Account Consolidation, Conflict of Interest, Risk Management, Customer Relationship Management (CRM), Supply Chain Management (SCM), Regulation and Disclosure.

## 1. Introduction

The business environment has witnessed widespread and rapid changes in corporate structures and corporate relationships. Regulations, deregulations, acquisitions, consolidations, mergers, spin-offs, strategic alliances, partnerships, joint ventures, new branches, bankruptcies, and franchises all make defining organizations and understanding corporate relationships an intimidating task and a source of many serious data errors. Furthermore, the relationships can be quite complex involving multi-level subsidiaries, joint ventures, and such. Yet this is precisely the kind of knowledge corporations need today to further their competitive advantage and improve the quality of their information processing.

Context plays a large role in how entities should be known and understood. A corporate customer for example, can be viewed in multiple ways from within the same corporation, depending on context. Units within a corporation have different relationships, perspectives, and concerns related to a customer or supplier, partner, competitor, etc. A few examples of these perspectives and concerns include financial credit risk, marketing, and legal liability.

These perspectives represent different views of the customer. The number of touch points between two corporations can easily reach into the hundreds or thousands. To be of use to business a tremendous amount of Corporate Household (CHH) information needs to be understood and organized in a clear and meaningful way.

Customer Relationship Management (CRM), Supply Chain Management (SCM), Business Risk Management (BRM), Business Intelligence (BI) and other business practices have

allowed organizations to make significant progress in understanding and managing corporate data and relationships. Previous research on corporations has mainly focused on organizational knowledge management (Becerra-Fernandez and Sabherwal, 2001; Gold et al, 2001). However, because of the complexity and the rapid speed of change in today's business environment, better knowledge of CHH is needed.

## 2. Corporate Householding

Actionable knowledge about organizations and their internal and external relationships is known as *corporate household knowledge* (Madnick and Wang, 2001; Madnick et al, 2001). The process of capturing, analyzing, understanding, defining, managing, and effectively using corporate household knowledge is known as *corporate householding*.

Corporate householding provides a way to identify, understand, organize, and use certain types of information and knowledge in a manner that allows the corporation to strategically harness this knowledge and to answer critical business questions. This knowledge includes:

1. Knowledge of corporate relationships and structures, such as:
   - Structures within the corporation, such as departments, divisions, and branches
   - Legal ownership relationships with other corporations, such as subsidiaries
   - Relationships with business customers, partners, suppliers, competitors, etc.
   - Relationships with third-party intermediaries, such as dealers, distributors, brokers, agents, and resellers
   - Relationships with governing and regulatory bodies
2. Knowledge of how, where, when, and why these relationships operate
3. Knowledge of the ways in which a corporation conducts business with related organizations and their related entities, such as intermediaries
4. Knowing in which context, that is with which lens (e.g., marketing, finance, legal, procurement, etc.), entities and relationships are defined and define themselves

### *2.1 Importance of Corporate Householding*

Knowing the who, what, where, when, and why of every corporate relationship in a way that enables the corporation to strategically harness this knowledge can be a daunting task. Corporate householding addresses this need.

### *2.2 Example Question*

Consider answering this question. *How many employees does IBM have?* In other words, how should we define the "IBM" corporate household? In responding to this seemingly simple question, consider the following list (Madnick and Wang, 2001):

| | |
|---|---|
| International Business Machines Corporation | IBM |
| IBM Microelectronics Division | IBM Global Services |
| IBM Global Financing | IBM Global Network |
| IBM de Columbia, S.A. | Software Artistry, Inc. |
| Lotus Development Corporation | Dominion Semiconductor Company |
| MiCRUS | Computing-Tabulating-Recording Co. |

These names are all related in some way to each other and to International Business Machines Corporation. The names include abbreviations, divisions, wholly or partially owned subsidiaries, companies that were acquired by IBM, companies that were acquired and later sold by IBM, and companies in which IBM has a majority or minority joint venture interest. The list also includes IBM's original name, Computing-Tabulating-Recording Co.

To complicate things even further, we have to consider the *purpose* of the question, i.e. the *context* in which the question is asked. Consider this more complex form of our example question. *An insurance company needs to set premium rates for business owner protection insurance for IBM. Which entities listed above should be included in IBM's employee count?*

There is yet another layer of complexity: changes over time, which we refer to as *temporal context*. At one point Lotus Development Corporation was a separate corporation from IBM; it is now a wholly-owned subsidiary. When comparing historical growth or decline in "number of IBM employees from 1990 to 2000" (i.e. before and after IBM acquired Lotus), should the Lotus employees be counted in the total as of 1990 or 2000? How should a meaningful comparison be made?

We can see how even a seemingly simple question can have significant complexity. Incorrect processing can lead to serious errors – which can be regarded as poor data quality. We want to avoid these problems.

## 3. Three Categories of Challenge

Following are three of the most common types of challenge that Corporate Householding addresses (Madnick et. al., 2002; Madnick, 2003).

### *3.1. Entity Identification*

Part of the complexity comes from the often ambiguous naming of a particular entity. As in the IBM example, many names can refer to the exact same entity (i.e., International Business Machines Corporation, IBM, I.B.M., IBM Corp, IBM Corporation). In other words, one entity can appear to be multiple entities, making it difficult to identify an entity correctly and efficiently. This kind of challenge is known as *Entity Identification*.

### *3.2. Entity aggregation*

Even after we have determined that "IBM", "I.B.M.", and "International Business Machines" all refer to exactly the same entity, we need to determine what other entities should be included. That is, depending on the context, what other unique entities, such as Lotus Development Corporation, should be included or aggregated into the definition of "IBM."

Consider another example. The MIT Lincoln Lab, is defined as "the Federally Funded Research and Development Center of the Massachusetts Institute of Technology." It is physically separated from the main campus of MIT. Problems arise when trying to answer business questions such as *How much was MIT's budget last year?* and *How much did MIT buy from IBM last year?* In which cases should the Lincoln Lab employees, budget, or purchases be included in "MIT" calculations and in which cases should they not be?

Answers to the questions above will differ depending on the context in which they are asked – under some circumstances, Lincoln Lab should be included, whereas in other circumstances it should not be. This type of challenge is referred to as *Entity Aggregation*.

### *3.3 Entity Transparency*

Relationships between entities often involve complex multi-layer relationships. For example, MIT purchases computers from IBM both directly and through local computer stores (e.g., CompUSA). This is the classic case where a seller sells its products both directly and through a broker. So what is the answer to the question: *How much did MIT buy from IBM last year?* Are only direct purchases to be counted or should indirect ones be included also?

Whether an organization is interested in only its direct interfaces, such as with the CompUSA broker/dealer, only the original source, or both, depends upon the context — different answers will be appropriate for different circumstances. Knowing when these interfaces are important, and how to capture and organize knowledge about them is a challenge known as *Entity Transparency*.

## 4. Corporate Householding Application Areas

In this section, we explore how Corporate Householding applies to some business concerns, or applications, in a more detailed manner. Most of these concerns are not industry-specific — rather they often span multiple industries. We have organized some of these application areas into the eight categories below, which are elaborated upon in the following sections.

**Some key Corporate Householding Application Areas**:
1. Account Consolidation
2. Risk Management
3. Licensing
4. Customer Relationship Management
5. Sales & Marketing
6. Supply Chain Management
7. Conflict of Interest
8. Regulations & Disclosure

These examples provide deeper insights into the nature of corporate householding.

### *4.1. Account Consolidation*

The need for corporate householding comes into play in the consolidation of financial statements. The Securities and Exchange Commission (SEC) has laid out rules concerning

consolidation in Regulation S-X, Article 3A, Section 210.3A-02 (SEC, 2002). The commission presents several criteria for establishing the most meaningful presentation of a company's financial position in its year-end statements.

For example, consider a large company like IBM: *How should IBM prepare its financial statements? Should its financial statements be consolidated with those of Lotus, a company acquired by IBM?* To answer this question, one needs to evaluate Lotus's relationship with IBM. According to the criteria set by the SEC, IBM should consolidate its accounts with Lotus if it has majority ownership (i.e., if IBM owns more than a half of Lotus.) Since IBM owns 100% of Lotus' stock, it should indeed consolidate its financial statements.

Suppose that the situation were a little more complicated and that IBM only directly owns 40% of Lotus but then owns 20% of Lotus indirectly, such as through another subsidiary. Should consolidation of financial statements still occur? Under the SEC regulations, the existence of a parent-subsidiary relationship in a way other than majority ownership of voting stock still requires consolidation of accounts, given that the consolidation is necessary in presenting a fair view of IBM's financial position. This can be quite complex since there are often multiple levels of subsidiaries involved in large international corporations, such as IBM.

While majority ownership is the major criterion concerning consolidation, the SEC has three other conditions listed in Regulation S-X.

1. Suppose IBM and Lotus differed substantially in their financial periods, the SEC regulations stipulate that consolidation should not occur. Instead, earnings/losses from such entities should be recorded in IBM's financial statements using the equity method of accounting. However, differences in fiscal periods is not a sufficient reason for avoiding consolidation of accounts; instead, for consolidation purposes, entities should try to prepare financial statements that generally coincide with the parent company's fiscal periods. Thus, if IBM and Lotus did indeed have different financial periods, it would be necessary for Lotus to make changes to its financial system, so that consolidation of its accounts with IBM could eventually occur.

2. Another condition set by the SEC is that if a company is a bank-holding company, it should not consolidate its accounts with any subsidiaries that are subject to the Bank Holding Company Act of 1956. This is the case if either a decision requiring divesture has already been made or there is a strong likelihood that divesture will be required to comply with the Bank Holding Company Act. So, if a company like Citibank wholly acquires a company that is involved in activities that are not financially related, and thus, subject to the Bank Holding Company Act, it should not consolidate its financial statements with this company; eventually this acquired company, or sections of this company, will have to be divested under the provisions of this act.

3. Another criterion set by the SEC concerns the consolidation of accounts with foreign subsidiaries. In such cases, the SEC leaves the decision mainly up to the company. The SEC does however urge that a company give due consideration to consolidation with any foreign entities, given that foreign subsidiaries operate under different political, economic, and currency restrictions. If a company does decide to consolidate accounts, proper disclosure should be made about foreign exchange restrictions on the consolidated financial position. Thus, given that IBM owns companies outside of the US, the SEC gives it leeway in deciding whether or not it would like to consolidate its accounts with these foreign subsidiaries.

A summary of these rules is depicted in Figure 1. To save space, we will not be as detailed in presenting the corporate householding rules for the subsequent application areas.

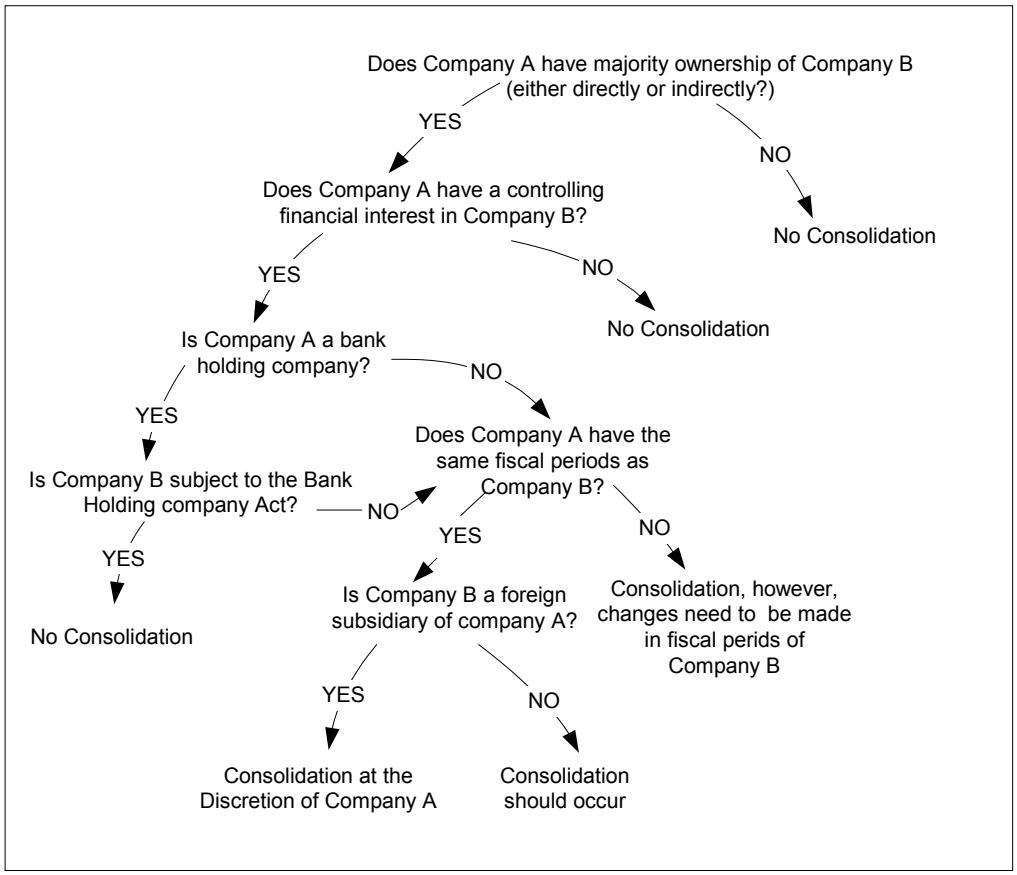


**Figure 1: Example, Should Company A consolidate its accounts with Company B?**

## *4.2 Risk Management*

### *4.2.1 Credit Risk*

Credit risk, a crucial consideration in many financial transactions, can be defined as "the possibility that a contractual counter-party does not meet its obligations stated in the contract," in other words defaults, "thereby causing the creditor a financial loss" (Ammann, 2001). In a more general sense, credit risk is "the risk associated with any kind of credit-linked events, such as: changes in credit quality (including downgrades or upgrades in credit ratings), variations of credit spreads, and the default event" (Bielecki and Rutkowski, 2002). Because of the complexity of corporate structures and relationships and the importance of credit risk, corporate householding in this field requires a significant amount of effort and attention. Several types of problem may occur in the process of credit risk evaluation as shown below.

***Multiple Instances of the Same Entity:*** A financial institution can extend credit to a number of different domestic as well as global organizations. Suppose CIBC is considering extending credit to IBM. To evaluate the overall risk involved, CIBC would need an aggregate

report showing all branches of CIBC and their business relationships with all branches and subsidiaries of IBM. This process becomes complicated quickly. Different CIBC branches may maintain information on IBM in different ways, such as using different names (IBM, I.B.M., or International Business Machines). Even if the names the CIBC branches use are the same, the IBM contact information that CIBC branches maintain is likely to differ by region – this is very similar to a common type of problem in customer information management.

***Relationships Between Entities:*** It can be important to know about all the relationships between two entities. For example, a firm planning to extend a large credit line to Hewlett Packard Puerto Rico may find it useful to know that Hewlett Packard only has a rating of AA, though Hewlett Packard Puerto Rico has a credit rating of AAA. Both entities must be considered in its decision to extend credit.

***Temporal Factors***: Looking at credit risk evaluation from another angle, we encounter problems caused by the dynamics of corporations over time. Suppose a bond held by a bank a year ago had an investment grade rating, but since then the company's bonds have been downgraded to junk bond status. The bank will have to know about the change in status and re-evaluate the bond even though it is still the same bond with the same company.

The relationship among corporate entities can normally be represented by a multi-dimensional tree structure. When banks evaluate a corporation's risks, they draw a tree to represent the corporation and its surrounding entities. The bank considers the credit-worthiness of every entity, and assigns credit limits to those entities. Banks keep their internal risk rating system to themselves because they can use some of the information obtained from their private channels to achieve competitive advantages. However, in some large deals, major credit risk is performed by a group of financial institutions pooling their information as part of the decision-making process. By merging their views of the entity hierarchy, the financial institutions have a more accurate view. Given the complexity of varying relationships and points of view, it is common for risk managers and analysts to have disagreements on questions such as ownership relationships and what the credit limit should be. Without complete and timely corporate household knowledge, the banks' decisions could result in significant negative outcomes.

### *4.2.2 Bankruptcy Risk*

Bankruptcy risk is closely related to credit risk. Bankruptcy normally results either in "liquidation of debtor's nonexempt property" or "debtor rehabilitation" or "reorganization of the debtor's assets" (Epstein, 2002). When deciding whether to issue loans to a particular company, banks need to know who will be responsible if the company bankrupts. For example, if a subsidiary goes bankrupt, how much liability, if any, does the parent company have? One concept that plays a significant role in the bankruptcy rules is *affiliate*. Specifically, an affiliate is defined as: (a) any entity that owns or controls 20% or more of the outstanding voting securities of the debtor, (b) any subsidiary for which the debtor parent corporation owns or controls more than 20% of the outstanding voting shares, or (c) any sister subsidiary of the debtor where a common parent corporation owns or controls 20% or more of the outstanding voting shares of both. Depending on whether or not a corporation is an affiliate, different Bankruptcy Code may apply. There are also rules to determine if the corporation is an "insider" under the Bankruptcy Code.

Typically, if the operation and management of the parent and the subsidiary are totally independent of each other (i.e., the parent is merely an investor in the subsidiary), the

corporations, attorneys, and bankruptcy court usually agree that the parent company does not have any liability if the subsidiary goes bankrupt. Additionally, bankruptcy laws and regulations vary from country to country, increasing the need for Corporate Householding knowledge.

*4.2.3 International or Country Risk*

As companies develop increasing global reach, risks caused by differences in business protocols need to be considered. This risk is called "international/country risk." Consider a company located in the US, whose parent company is in Japan, such as a Toyota manufacturing plant in the US. When should this plant be considered a "US company" and when should it be considered simply a branch/subsidiary of a Japanese firm?

Another type of international risk involves the constant changes in exchange rates. When a company conducts business with foreign or multinational corporations, it is highly likely that the currencies are those of the host country, e.g. US dollars for American companies and Japanese Yen for Japanese companies. Exchange rates fluctuate constantly which affect hundreds or thousands of companies.

### *4.3 Licensing*

An issue many software vendors face is keeping their licensing information current regarding the consolidation of customers through mergers and acquisitions. This issue is not unique to the software industry, but we will use an example from that industry to draw out the application of householding with regard to licensing. Consolidation of customers is a particular challenge for vendors of enterprise-wide solutions, and/or those who sell enterprise-wide licenses.

For example, an issue arises when two customers merge. A software vendor experienced this when two of its printer-publisher customers merged to form one of the world's largest printer-publishers. After the merger the two printers compared databases of software licenses to determine licensing overlap. A manual householding process was used to compare the hierarchies of vendor→ product→ license→ restrictions. Where products overlapped between the two printers, but licensing restrictions did not prevent usage at new facilities or "sites" the printers were able to consolidate licenses.

Another example is if MIT buys a company-wide license for Windows XP from Microsoft, is Lincoln Laboratory authorized to use the software too? Must the license specifically state the cases that would apply to Lincoln Lab? If not, what is the common guideline that could resolve ambiguity that is not addressed in the license?

Understanding Corporate Householding concepts is valuable not only to purchasers of product licenses, but to vendors as well. Consider another example scenario. A vendor is selling an Extraction, Transformation, Loading (ETL) solution to a multi-national corporation. The corporate family tree for that organization is extensive, encompassing many subsidiaries, divisions, branches, and field locations. The vendor, through householding, is able to map which entities in the specific corporate family tree they have sold licenses to, and under what restrictions. Not only does this allow them to roll-up all revenues for that one global customer for a single customer view, but it also allows the vendor to propose the optimal licensing configuration.

Corporate Householding provides a tool and an information source to use in software license structuring. The same kinds of issues involve other licensing arrangements, such as licensing of patents.

### *4.4 Customer Relationship Management (CRM)*

The use of householding practices within CRM allows a vendor to efficiently structure contact records of its customers (retail consumers or corporations).

Consider this example scenario of a retail consumer Tom Jones who has a vegetable garden. Kim Jones, Tom's wife, is a flower gardener. Linda Jones, Tom's daughter, while not a gardener, loves to plant pine trees. All three of the Jones' have ordered products from Sweet Flowers seed catalog. Sweet Flowers have learned the Jones' try to consolidate orders when possible, to save on shipping and handling fees. Sweet Flowers is also sensitive to the issue of "contact fatigue" where sending multiple copies of the catalog to the Jones residence – one copy per customer – annoys the customer. Not only do the Jones's have more trash to throw away, they also question if Sweet Flowers really "knows" them. Reducing duplicate catalog mailings has the additional appeal of cutting postage and printing fees. The CRM-oriented solution is to identify all customers who live at the same residence, regardless of last name, and to household them.

The seed catalog example is not unique. The practice of householding for consolidation of account communications is spelled out in the following letter illustrating an actual financial services example:

> "Dear Shareholder:
> In an effort to reduce printing and mailing expenses for the Smith Barney Mutual Fund family, we have adopted a policy known as "housekeeping." Starting in September, shareholders who are members of the same family, share the same address and have multiple accounts in the same Smith Barney Mutual Fund will receive only one copy of a fund's annual prospectus. …"

As another example, a product vendor communicating with a multi-national corporation might have hundreds, if not thousands, of unique contact records (individuals) in its CRM system. For example, if the product vendor wants to promote a specific electronic component to engineers at Ford Motors, the vendor would do well to target a specific type of Ford engineer in a specific subsidiary or division. The vendor needs to household their customer data in various ways, depending on the circumstances. The first circumstance or context is job category. A unique identifier corresponding to a job category must be applied to all contact records. Examples of categories are senior management, IT line management, mechanical engineer, electrical engineer, etc. The job category unique identifier represents one contextual view of the data. Then, another identifier must be applied to all contacts in the Ford, Lincoln Mercury, Jaguar, etc. divisions and subsidiaries. The vendor now has a second contextual view of the same data. By comparing the intersections of these two views, job category and division, the vendor has a subset of customer contact records which can be used to conduct a promotional campaign for new products.

The commonality between these examples is that householding is used to both segment and draw relationships between customer records in order to improve the effectiveness of customer communications. Those communications can take the form of one-to-one marketing,

aggregation of mailings, or the consolidation of customer feedback. For example, householding can be used to consolidate all similar customer inquiries or feedback on a specific product. This allows product managers to sort customer incidents based on product and call type.

### *4.5 Sales & Marketing*

The sales and marketing functions of the corporation serve as the stewards of demand management, the heart of all business activity. Demand management can be thought of as the intersection of product and service, customer and channel. Unfortunately, customer management is sometimes thwarted by the inability to define and accurately record interactions with customers. Consequently, customer identification systems supporting sales and marketing efforts is a growing need (Kunz and Shinnebarger, 2002). These systems are used in the development of an integrated view of business-to-business customers. An integrated view of the business-to-business customer is elusive to attain due to myriad contexts for how the business-to-business customer must be viewed, reinforcing the need for Corporate Householding. The establishment of Corporate Householding supports sales and marketing activities to identify existing or high-potential customers, assign resources to penetrate them, and report on the performance of these efforts.

#### *4.5.1 The Marketing Organization*

Consider the marketing organization, which in the business-to-business world supports the sales organization. Marketing teams work to integrate and analyze customer information to identify the customers that provide the most revenue and profit, to identify any predictive variables that may determine future purchase or signs of attrition, and to use the profile to guide customer penetration and acquisition efforts. The first step in this effort is the customer integration process. This process presents major challenges for marketers because customer information is usually collected and maintained in discrete information systems across the enterprise and often stored in inconsistent formats representing different views of the same customer.

For example, accounting may view the customer from a perspective of a bill-to; the service group sees the customer from the point of a ship-to, and the sales team is focused on the economic buying unit. In this situation the customer takes on three different faces but the true identity is in fact expressed by all of these views as part of a complex Corporate Household that has different members responsible for different activities. The marketer has the delicate job of piecing together all of these views to form a story or a picture of demand that can form the basis of actionable information (Chettayar, 2002).

The customer can be an economic buying unit in a large organization, often referred to as a strategic business unit (SBU). These businesses within a business can purchase directly. In this context, the customer can be an SBU or the corporate parent. Defining the customer from the view of corporate parent would conceal an important view of the SBUs in a larger organization, preventing insights into additional revenue and profit opportunities. The marketing department needs to be able to look at the customer at all levels of decision-making authority to identify the needs of the small operations within a large organization as well as the overall organization itself. By doing so, the marketing organization uncovers the best opportunities and learns how to effectively relate and penetrate the complex organizations.

As mentioned above, different sub-units of an organization often have different perspectives of the customers. Due to the different views, the sub-units often define the semantics of key business terms related to customers differently. An example that we encountered in a financial services firm involved ontological conflicts when both the Marketing and Finance departments used the same term “number of active accounts”, but attached different semantics to it. Marketing defined "number of active accounts" as the number of accounts with active status. Finance, on the other hand, defined "number of active accounts" as the number of accounts with active status plus the number of accounts with closed/deceased status but with outstanding debt (i.e., the accounts have not been written off). The definitions make intuitive sense in the contexts in which they are used; however use of the same term with conflicting semantics can create confusion and present householding issues. For example, what number should be released to the outside world? What might happen when an employee transfers from marketing to finance and is unaware of the different meanings of the same term in a new setting? Each of these situations, if not handled correctly, could lead to erroneous data processing.

*4.5.2 The Sales Organization*

The need to manage Corporate Households also appears in the sales organization. Companies often assign sales people to businesses in a variety of ways. Sales people can be assigned contacts at a specific location, a branch, headquarters, subsidiary, or even the corporate parent. In addition, such contexts as geography, industry affiliation, channel, and current and potential revenue contribution may be used to further segment the market for the purpose of assigning sales people.

Some sales organizations are organized to serve specific customers based on the total spent and the location of the customer. In such an organizational structure it is possible to have multiple sales reps assigned to one large business customer. What may be one organization viewed from top sales executives is really considered several accounts at the field sales level. Coordinating the information exchange between different levels of sales teams and managing enterprise sales activities to further penetrate customers is complex without some way to link all the accounts together.

Other examples of Corporate Householding needs and applications in the sales department occur when companies assign sales people to industry verticals or to specific distribution channels. Sales people assigned to specific industry verticals may be responsible for driving demand for the entire sector, in which case two direct competitors may be lumped together in the seller’s view of the demand chain. This view becomes important in understanding the total demand from the sector and hence the total revenue opportunity.

Managing the channel of distribution also poses challenges addressed by a better multidimensional customer identification framework. Sales people are often aligned to specific channels such as value-added retailers, wholesalers and retailers. Yet, the true customer is the user down the demand chain who purchases for use. The definition of the customer in this context may take different meanings, which we referred to as Entity Transparency, based on the relationship with the third-party distributor and the applicable accounting rules (Conneighton, 2002). Different parts of management may want to view demand from the channel as well as the end user, requiring the association of the end user information to the third-party distributor, which creates a unique Corporate Household along the lines of the demand chain. The goals and objectives of each level in the sales management hierarchy force the need to manage multiple

views of the business based on the relationship of contacts, locations, headquarters, and corporate parents in the contexts of total and potential revenue, geography, channel, and industry.

*4.5.3 Business Activity Monitoring & Reporting*

Reporting on the results of sales and marketing operations has taken on new prominence largely because sales and marketing organizations are under increased pressures to be accountable (Herschel et. al., 2002). Sales and marketing people must therefore report on activities and investments. Yet, these organizations struggle to understand the sales for their products and services and the impacts of marketing campaigns on customers. The largest companies struggle with understanding the total relationship they may have with any one customer, in particular the total revenue. The difficulty lies in the disparate information systems managing different views of the customer (Wang and Madnick, 1988) without any common attribute or identifier that integrates all of the views under a common Corporate Household framework. It is common in many large companies that there is the lack of a single coherent aggregate view as a result of the many organizational silos that operate largely autonomously.

A common customer directory or Corporate Household construct is needed to support sales and marketing needs and applications to manage customers. Furthermore, Corporate Householding must concurrently support other departments requiring still different views – there is usually no one "right view." These views must be specific to different internal constituencies and yet relate to each other. Of course, the customer may have defined their own representations, as evidenced by how they identify themselves, and may expect that the company remember this view as the basis for interacting with the customer in the future. This requires companies to develop multiple views for how to relate to customers in different contexts, such as sales and service, at different touch points. The sales and marketing needs and applications for Corporate Householding are quite complex.

***4.6 Supply Chain Management (SCM)***

Corporate householding issues also exist in the realm of supply chain management. For example, an information executive of a global manufacturing company was interested in global sourcing, which involves identifying manufacturing sites that could produce a particular product with the lowest costs, including the cost of manufacturing as well as transportation. A large part of the manufacturing cost comes from raw material cost. Therefore, identifying and maintaining relationships with material vendors are critical in order to achieve cost reduction. However, due to localized information systems, two manufacturing sites of her company might have two different, independent relationships/contracts with the same vendor for the same material. The situation becomes even more complicated when a vendor has different relationships with different functional areas within the same organization, such as manufacturing, financial, and accounting. Therefore, it becomes difficult to have a single, consistent view of a vendor globally.

Inconsistencies among information systems that are maintained locally make it difficult for a company to understand its relationships with its business partners in a uniform way. In the case of the firm mentioned above, it is impossible for the company to know how much raw material is used on a global basis. In addition, the company cannot take advantage of the lowest price across all of its manufacturing sites from a particular vendor.

Consider another example**.** A company that works with the USA Defense Logistics Agency (DLA) often needs to help the agency identify when sources (i.e. manufacturers) of a product or parts of a product are the same entity under different names. Usually DLA will have a record of accepted manufacturers for a given part. The company obtains information about that part from a web catalog, which also contains information about the manufacturer. The company needs to find out if the manufacturer that the web catalog is referring to is in fact the same company that is on the record of accepted manufacturers. In short, a problem occurs when trying to determine when two different names and addresses represent the same entity. Is Acme Manufacturing in Iowa part of the same company as Acme Parts in New York?

As illustrated in the examples above, if companies can maintain better quality data on their suppliers and/or buyers, and perform corporate householding activities more efficiently, they will be able to avoid an enormous amount of extra work, possible errors, and benefit from significant cost savings.

### *4.7 Conflict of Interest*

In certain industries, especially in accounting and consulting, there can be legal and/or professional obligations to avoid conflicts of interest. These can be of various forms, such as having ownership interests or non-audit business activities in a company that is being audited or simultaneously consulting for a competitor.

Many large accounting firms have found it desirable to sell non-audit services to their clients in order to compensate for declining audit revenues. However, such practices have resulted in clashes with the Securities and Exchange Commission (SEC), who has accused these firms of failing to maintain independence as auditors, resulting in a conflict of interest. In this area, an understanding of corporate householding issues will help firms avoid such violations of SEC rules.

#### *4.7.1 Auditor Independence – Relationships*

In Regulation S-X, Article 2, the SEC provides a general definition of auditor independence, which is followed by a complex set of rules to reflect the application of this definition in specific situations. The general standard states that an accountant is independent from its audit client only if it is "capable of exercising objective and impartial judgment on all issues encompassed within the accountant's engagement" (SEC, 2002). The SEC then presents specific rules on maintaining auditor independence in the areas of financial relationships, employment relationships, business relationships, and non-audit services.

The rules for maintaining independence in employment and business relationships are relatively straightforward. An accounting firm is clearly not independent from its audit client if the client currently employs a partner or shareholder. In terms of business relationships, the SEC declares a firm as not independent if it engages in any direct or material indirect business relationship with an audit client. An example of an inappropriate business relationship between an accounting firm and its audit client has recently been pointed out by the SEC between Ernst & Young and PeopleSoft Inc, where E&Y had a marketing agreement with PeopleSoft to sell and install their software (Cassell, 2003).

#### *4.7.2 Auditor Independence – Financial Interests*

In the area of financial activities, the rules defining auditor independence are much more complex, and thus, firms can find it difficult to determine if they are violating the rules.

For example, suppose a firm is auditing the financial reports of IBM. Is it independent if it owns shares of Lotus (assuming that Lotus was not 100% owned by IBM)? Is it independent if the firm is invested in IBM through an intermediary, such as through an investment portfolio? What about if IBM owns stock in the accounting firm? These are clearly questions that an accounting firm must be able to answer to determine its independence from a client.

According to the SEC rules, an accountant is not independent if s/he has a direct financial interest or a material indirect financial interest in the audit client. Thus, owning stocks, bonds, notes, or other securities in an audit client would render an accountant not independent. Likewise, owning securities in a subsidiary of the audit client would also be a conflict. Thus, having an investment in Lotus would make an accountant not independent of IBM. In fact, owning securities in any entity where IBM exercises significant financial influence would render the accountant not independent. Vice versa, if IBM were to own shares in the accounting firm, auditor independence would still be violated according to SEC regulations. However, auditor independence is maintained if an accountant holds 5% or less of the shares of a diversified management investment company that invests in the audit client. Thus, in most cases, indirect investment in IBM through an investment company would not compromise an accountant's independence. These are a few of the many rules and complexities that govern accountants in this area.

*4.7.3 Auditor Independence – Non-audit Services*

Also, accounting firms have restrictions in the types of non-audit services that they can provide to their clients, if they wish to maintain independence. A list of possibly prohibited activities is defined in the Sarbanes-Oxley Act of 2002 and includes bookkeeping of audit clients' financial statements, actuarial services, broker services, and financial information systems design. Once again, following these rules can be difficult if they involve separate corporations that are related.

Here's yet another example question. *Can the accounting firm provide non-audit services to Lotus while being the auditor of IBM? What if IBM only owned a small percentage of Lotus?* Corporate householding can increase understanding of, and compliance with, the complex rules in this area.

### *4.8 Regulations and Disclosure*

The Securities and Exchange Commission (SEC) has numerous rules concerning disclosure of information by publicly-traded companies. For example, it requires an entity, whether an individual or corporation, to disclose information if it acquires beneficial ownership of more than 5% of a class of a company's equity securities registered under Section 12 of the Securities Exchange Act of 1934. Beneficial ownership is defined under SEC rules as an entity who directly or indirectly has voting power or investment power (i.e., ability to sell the securities). In such a case, it must file a Schedule 13D reporting such ownership. This information is then reported to the issuing company and to the exchanges where the securities are traded. The general public can obtain this information through the SEC's online database called EDGAR (Electronic Data Gathering, Analysis, and Retrieval System). The SEC makes such information

available and allows the public to view information about the individuals who exercise control in a company.

Corporate householding issues play a role in disclosure when one tries to determine if an investor has beneficial ownership of 5% of a company's securities. The definition given by the SEC for a beneficial owner states that it is "any person who, directly or indirectly, through any contract, arrangement, understanding, relationship, or otherwise has or shares" voting power or investment power. Clearly, there must be a series of rules in this area to objectively define beneficial ownership. One of these rules is that any person who has power of attorney over a trust fund that owns 5% of a company's equities should be declared the beneficial owner. However, this is just one of many rules that govern the area of disclosure.

This issue is further complicated by the fact that different countries define, "beneficial ownership" differently and have different requirements for disclosure. To maintain some sort of uniformity, the International Organization of Securities Commissions issued a set of International Disclosure Standards with a definition of "beneficial ownership", which the SEC has incorporated into its own policy. However, it should be noted that the document acknowledges that the laws defining persons who are required to file a disclosure statement varies from country to country.

This is especially the case with the new amendment proposed by the Sarbanes-Oxley Act of 2002, which requires corporate insiders, i.e., those who have 10% beneficial ownership of a company's securities, to file a disclosure report within two business days instead of the much longer deadline granted before (the tenth day of the month following the month in which the transaction occurred). With the use of corporate householding, companies can improve their ability to quickly determine beneficial ownership, thus helping them comply with these new SEC laws.

## 5. Conclusion

In this paper, we have described a major challenge to data quality that is faced by corporations in today's rapidly evolving business environment, that is, how to correctly interpret and make full use of the information and knowledge related to the complex corporate structure and entity relationships in business activities. We defined the concept of Corporate Householding, described categories of corporate householding problems, and the application areas that align with many functions in a corporation, including financial, legal, sales and marketing and operations. Many examples have shown that corporations have substantial needs to better manage their corporate household data and to improve data quality. Corporate householding is aimed to solve the existing problems and bring efficiency and cost reduction to businesses. An important conclusion is that there is not one single view of a company's corporate household that correctly serves all the different application needs (Madnick, 1999). Thus, a flexible corporate householding process is needed that has sufficient and relevant "corporate household knowledge" and "application knowledge" to adapt to satisfy specific applications, and more importantly, to answer critical and changing business questions. Our related research efforts to produce a Corporate Household Knowledge Processor (Bressan et. al. 2000; Goh et al, 1999; Madnick et al, 2003) is an attempt to address this need.

**Acknowledgements**

Work reported herein has been supported, in part, by Cambridge Research Group, D & B, Firstlogic, MIT Total Data Quality Management (TDQM) Program, Naval Inventory Control

Point (NAVICP), and the Singapore-MIT Alliance (SMA). Thanks are due to Xing Ping Chen, Chris Haywood, Pat McCoy, Ahmad Shuja, Wei Zhang, and Harry Zhu for providing helpful information and suggestions.

**References**

Ammann, M. *Credit Risk Valuation: Methods, Models and Applications*. 2nd Ed. Springer-Verlag, 2001.

Becerra-Fernandez, I. and Sabherwal, R. Organizational knowledge management: a contingency perspective, *Journal of Management Information Systems*, 18, 1 (Summer 2001), 23-56.

Bielecki, T. R. and Rutkowski, M. *Credit Risk: Modeling, Valuation and Hedging*. Springer-Verlag, 2002.

Bressen, S.; Goh, C. H.; Levina, N.; Shah, A.; Madnick, S.; and Siegel, M. Context Knowledge Representation and Reasoning in the Context Interchange System, *Applied Intelligence: The International Journal of Artificial Intelligence, Neutral Networks, and Complex Problem-solving Technologies*, 12, 2, (2000), 165-179.

Cassell, B-L. More Ernst Nonaudit Services Have Come Under Fire at SEC. *Wall Street Journal*, (March 10, 2003).

Chettayar, K. *Creating Institutional Memory: The Key to CRM.*
http://www.integratedsolutionsmag.com/Articles/2002_12/021211.htm, (December 2002).

Conneighton, C. *The Venture Management Handbook*. Venturebooks, 2002.

Epstein, D. *Bankruptcy and Related Law in a Nutshell*. 6th edition. West Group, 2002.

Gold, A. H.; Malhotra, A.; and Segars, A. H. Knowledge management: an organizational capabilities perspective. *Journal of Management Information Systems*, 18, 1 (Summer 2001), 185-214.

Goh, C.H.; Bressan, S.; Madnick, S.; and Siegel, M. Context interchange: new features and formalisms for the intelligent integration of information, *ACM Transactions on Office Information Systems*, 17, 3, (1999), 270-293.

Herschel, G.; Janowski, W.; Marcus, C.; and Sarner, *A. Predicts 2003: Marketing Will Become Accountable*. Gartner Group, November 27, 2002.

Kunz, B. and Shinnebarger, D. *Understanding Unique ID Solutions: Strategic and Operational Approaches for Identifying Customers*. Pepers & Rogers, 2002.

Madnick, S. Metadata Jones and the Tower of Babel: The Challenge of Large-scale Semantic Heterogeneity, *Proceedings of the 1999 IEEE Meta-Data Conference*, (1999), 1-13.

Madnick, S. Oh, So That is What you Meant! The Interplay of Data Quality and Data Semantics. *Proceedings of the Entity Relationship Conference*, (October 2003), 3-13.

Madnick, S. and Wang, R. Y. Corporate Household Knowledge Processing: Challenges, Concepts, and Solution Approaches, (August 2001). *Sloan Working Paper* #4222-01, and *Center for Information Systems Laboratory Working Paper* #2001-09.

Madnick, S.; Wang, R.; Dravis, F.; and Chen, X. Improving the Quality of Corporate Household Data: Current Practices and Research Directions, *Proceedings of the Sixth International Conference on Information Quality*. Cambridge, MA, (November 2001), 92-104.

Madnick, S.; Wang, R. Y.; and Zhang, W. A Framework for Corporate Householding, *Proceedings of the Seventh International Conference on Information Quality*. Cambridge, MA, (November 2002), 36-40.

Madnick, S.; Wang, R.; and Xian, Xiang. The Design and Implementation of a Corporate Householding Knowledge Processor to Improve Data Quality. *Journal of Management*

*Information Systems*, forthcoming 20, 3, (Winter 2003-04).
United States Securities and Exchange Commission (SEC). Regulation S-X. (August, 2002). Available at http://www.sec.gov/divisions/corpfin/forms/regsx.htm.
Wang, R. Y. and Madnick, S. Evolution towards strategic applications of data bases through composite information systems, *Journal of Management Information Systems*, 5, 2, (Fall 1988), 5-22.