

## How to get an Information Quality Program Started: The Ingenix<sup>SM</sup> Approach

(Practice-Oriented Paper)

### Beverly K. Kahn

Suffolk University  
Sawyer School of  
Management  
8 Ashburton Place  
Boston, MA 02108  
Tel: (617) 573-8642  
Fax: (617) 573-8345  
bkahn@acad.suffolk.edu

### Raïssa Katz-Haas

Ingenix  
450 Columbus Blvd  
Hartford, CT 06115  
Tel: (860) 702-8721  
rkatzh@uhc.com

### Diane M. Strong

Worcester Polytechnic Institute  
Management Department  
100 Institute Road  
Worcester, MA 01609  
Tel: (508) 831-5573  
Fax: (508) 831-5720  
dstrong@wpi.edu

### Abstract

Ingenix included an information quality function in its project plan for its new data warehouse. Recognition that an information quality function is important to the success of the data warehouse, however, is only a first step. This paper describes how Ingenix moved from a plan to an established Information Quality group, and established the foundation for its first and future information quality improvement projects. The lessons learned in this process will provide insights for other organizations attempting to start information quality improvement initiatives.

### INTRODUCTION

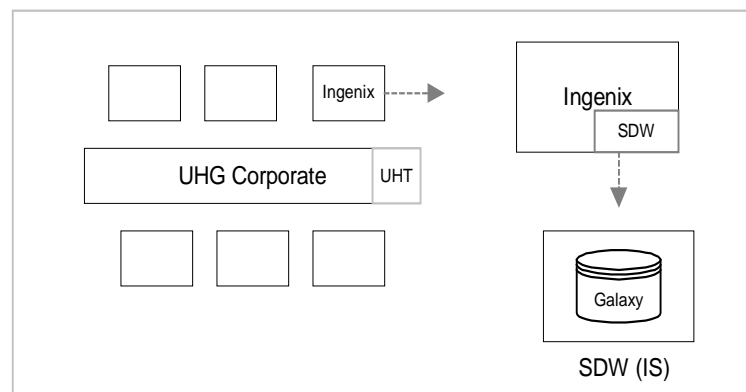
The literature about information quality provides a wealth of knowledge about defining quality, measuring quality, and quality problems. Often overlooked, however, is the essential first step of an information quality (IQ) improvement endeavor—getting the project started. This is much more difficult than one might expect. It takes far more than recognizing that data quality anomalies exist. This paper describes one company's approach to starting a data quality improvement program.

### OVERVIEW OF INGENIX

Ingenix was formed in January 1997. In reality, its legacy began many years beforehand. Ingenix was created when many of the industry's leading health data and information firms—highly reputable companies with long histories of their own—merged into one new company. This unprecedented union within the health care community brought a new opportunity to the industry. For the first time, all sides of the health care community could find information solutions from one source. Ingenix is a wholly owned subsidiary of UnitedHealth Group (UHG), a multi-billion dollar diversified health care company headquartered outside Minneapolis. Ingenix is one of the six business segments of UHG. (For more information on UnitedHealth Group, see <http://www.unitedhealthgroup.com>.)

Ingenix has become one of the largest health care information and research companies in the industry, providing a comprehensive line of products and services—many of which are rooted in proprietary databases. More than 100,000 providers, 1,500 payers and 100 FORTUNE 500 companies rely on Ingenix software, services and expertise to improve their health care delivery and operations. Operationally, Ingenix serves multiple health care market segments through internal business segments that provide specific tools and services for actuarial consulting, large employers, health care providers, and health care payers.

The **Ingenix Shared Data Warehouse (SDW)** is a department within Ingenix. It is comparable to departments commonly called Information Systems. SDW contains four interdependent areas: 1) Development and Legacy Support, 2) Customer Analysis and Support, 3) Database Design and Delivery, and 4) Technical Infrastructure and Data Management. Customer Analysis and Support works with users/customers and source systems to resolve issues and deliver information and metadata, training, and support. Customer Analysis and Support focuses on continuous improvement activities from the viewpoint and needs of its information customers. A high-level organizational chart is in Figure 1.



**Figure 1**  
**Organizational Structure of UnitedHealth Group**

In the summer of 1998, SDW was charged with developing a new data warehouse known as Galaxy. The purpose of this data warehouse was to integrate data from UnitedHealth Group's two major data stores (MARS and DSS) into one repository. This would make reporting across the enterprise easier for customers/users. As part of the Galaxy charter and to support the needs of its direct and indirect users, SDW committed to deliver data of equal or better quality than current standards. This project initiated SDW's current interest in information quality. From the beginning of the Galaxy project, Ingenix recognized the need to consider information quality issues as an important facet of the data warehouse project.

## UNDERSTANDING THE PROBLEM

While many companies recognize that they have problems with the quality of their information, few take a structured approach to understanding those problems well enough to

design effective solutions. Ingenix's first step was to understand the nature and scope of the areas of concern within MARS and DSS - anomalies manifest in part by the varied nature and quality of data submitted from disparate sources.

Recognizing how critical quality information is to Ingenix and its customers, a structured process was developed for Galaxy. Consistent with its customer orientation, the Customer Analysis and Support section of SDW surveyed MARS and DSS users, totaling approximately 700. Understanding imperfections is complex because the vast majority of users are employed by other UHG business segments, not Ingenix. (Current MARS and DSS users will be future Galaxy users.)

A short questionnaire focusing on metadata was e-mailed to the 700 users. More than 200 responses were received. Follow-up interviews to obtain more detailed information were conducted with 18 users. In addition, observations of users as they worked with MARS and DSS data and metadata were made. Finally, a short questionnaire focusing on data quality was emailed to the 700 direct users. Approximately 10% responded. These data collection activities took place during the spring of 1999.

The scenarios that follow focus on the results of the first three activities, i.e., the metadata survey, the interviews, and the observations of use. The data quality survey responses have not been analyzed in detail. Analysis of the metadata survey and interviews determined that improved data quality in the new data warehouse, Galaxy, was a most important goal. To focus attention and requirements on key user concerns, scenarios were developed that illustrate users' typical data quality experiences. Several of these are described below.

**Scenario 1:** A UHG business unit used MARS and DSS to retrieve membership information. Internal queries revealed more than 1000 members who were 1000 years old. This data was used to understand differing health care needs by age and to provide special services to senior citizens. Such errors skew the analyses. As a result, the internal business unit questioned the reliability of information obtained from MARS and DSS.

**Scenario 2:** During the development of a MARS/DSS sourced data mart, business analysts needed to know how MARS and DSS source data was transformed. The business analysts needed this metadata to correctly interpret data extracted from MARS and DSS. Research to determine the data transformation rules took several months because MARS transformation rules had not been documented. This resulted in a large loss of productivity and was costly to UHG.

**Scenario 3:** Ingenix offers several products and services that aid in detecting potentially fraudulent, abusive billing practices within the health care industry. One business segment within Ingenix also provides investigative services. Some of the investigators analyze data from MARS and DSS directly against the mainframe. To do this, these investigators needed to know SQL for MARS and SAS for DSS, skills not necessarily common to fraud investigators. This is a very cumbersome process. Additionally, data errors, such as invalid field values, were obtained in many cases. The nature and dates of data problems were not always documented and available to users, but as data inconsistencies were discovered, they were addressed by the SDW Steering Committee. Another investigator had problems due to inconsistency of data field names and/or

descriptions. Better quality data and easier data access would shorten and improve the investigation process. In summary, the investigators need accurate metadata, accurate data, more denotation of data that is not free of error, and a user-oriented (non-SQL, non-SAS) way to access data. Currently, some types of metadata are available on UHG's Intranet. Because this intranet site was developed with Usability Engineering methods, the metadata is easily accessed, but users need more types of metadata. This metadata will be expanded to include Galaxy transformation rules, legacy mapping, efficient joins, and much more. The scenario above illustrates a combination of problems related to ease of access, data accuracy, and metadata availability.

In conclusion, users need data and metadata that is accurate, easily accessible, and consistent across systems and databases. Specifically, "Users need a system that is easy to learn and use, enables quick fact finding, and provides high value utility, functionally, and information. The information/metadata must be accurate, timely, complete, easy to find and the system needs to increase user productivity and satisfaction." (Information Directory: Part I, Raïssa Katz-Haas, 6/11/99) Additionally, no single entity/group was in charge of data quality at that time. Such a group should be able to assist in the resolution of these problems.

## **GETTING STARTED ON SOLUTIONS**

The SDW Ingenix Vice President and SDW Directors included a data quality function in the Galaxy project plan. Developing the new data warehouse was so complicated and resource-intensive, however, that by the summer of 1999, the data quality function had not yet been solidified—illustrating a key problem in getting a data quality initiative started. Even though Ingenix had initiated studies regarding the nature and scope of those problems, and had included data quality in its data warehouse project plan, it still did not have an active data quality initiative. What the Galaxy project needed was an information quality champion.

From within SDW, an information quality champion stepped forward to find a solution. The champion understood the need for organizational commitment to data quality. This was especially critical for the Galaxy project since both users and sources are spread throughout all UHG business segments. Individuals throughout UHG must understand why information quality improvement is so critical. Data quality improvement requires cooperation among the companies and segments of UHG, not just departments within Ingenix. Obtaining organizational commitment requires the support of senior executives at UnitedHealth Group. Therefore, the cornerstone of a data quality initiative for Galaxy is top management support and commitment.

In November 1999, the information quality champion and the SDW Vice President discussed the possibility of sponsoring an organization-wide seminar focusing on data quality issues. This would be a first step toward organizational commitment and, hopefully, would involve some visible support from UHG's executive management.

## **MAKING THE BUSINESS CASE**

To obtain top management commitment, they needed to be informed of the extent and impact of inconsistent data on business costs and risks. The VP of SDW at Ingenix understood that high-quality data was critical to the success of Galaxy. This Ingenix VP asked UHG's

President to support an organizational seminar focusing on data quality. UHG's President asked for some specific examples supporting the criticality of data quality for UHG. Making the business case for data quality improvement efforts is typically an important part of obtaining top management support and thus an important part of any successful organization-wide data quality initiative.

In response to the request from UHG's President, the Ingenix information quality champion produced a document, part of which quoted published facts about the high cost of poor data quality in the health sector. For example:

"The U.S. Attorney General's office has stated, "Approximately \$23 billion or 14% of the health care dollar is wasted in fraud and inaccurate billing." [*Nashville Business Journal*, Sept.1997]

"In 1997, data quality problems cost Oxford Health Plans Inc. \$69.3 million. In addition, its stock dropped 62% and New York Insurance Department fined Oxford \$3 million." [Wall Street Journal, 12/4/97]

The second part of the document focused on how much poor quality data could cost UHG. The figures provided are indicative of comparable problems throughout the health care industry. Raïssa Katz-Haas of Ingenix cited examples of excess labor costs, missed opportunities, and misguided decisions based on the previously described scenarios. The document did not include a calculation of total cost, but illustrated sample costs in several areas. Noted in the third part of the document was that poor data quality is an enterprise-wide problem/opportunity that required top management commitment and resources, and that good information quality can give the whole enterprise a significant competitive advantage in the marketplace.

## **THE "KICK-OFF" MEETING**

As Ingenix was obtaining top management support, it was simultaneously making plans for an enterprise-wide seminar on Information Quality. In addition to securing organizational commitment, a goal of the seminar was the establishment of an Information Quality Initiative. The seminar plans were in place when the Ingenix VP discussed the meeting with UHG's President who offered to help. This led to his giving the keynote address at the seminar.

On June 1, 2000 Ingenix sponsored an enterprise-wide seminar on Information Quality. UHG employees outside of Connecticut attended the seminar via video conferencing. This seminar was the "kick-off" of the Information Quality Initiative. An expert, from MIT's Sloan School of Business, led the seminar on Information Quality and Continuous Improvement. UHG's President stressed the importance of information quality to the entire enterprise and offered to provide support and business sponsorship for a company-wide Information Quality Process Improvement (IQPI) initiative. He challenged the seminar participants to make visible progress in improving information quality within twelve months. This had a very positive impact on the participants and helped get users on board the IQPI Initiative. Additionally, Ingenix had accomplished the necessary factor of obtaining top management commitment.

During the seminar, participants in roundtable discussions created a list of the Top Ten information quality issues/needs at UHG. These are summarized and are listed in no order of

importance in Table 1.

- Form an information quality steering group and dedicate a high-level person, who should report to the UHG Executive Council, who has enough authority to ensure such things as cross-functional activities are possible.
- Develop accurate member counts, eligibility information, and consistent membership information across systems.
- Provide training and education about systems (e.g., claims systems), business processes and practices, etc.
- Centralize a list of topic experts.
- Create process maps/flows of core information chains (e.g., member, claims, etc.)
- Establish a centralized location for information quality issues and caveats.
- Develop and share across the enterprise code and code grouping management: a consistent, common way to manage codes and groupings.
- Maintain good enterprise-wide metadata.
- Integrate revenue information with claims and membership information.
- Centralize consistent information that the business creates about Market, OHS, Region, Site, Product, etc. across enterprise. This information should be 'owned' & sponsored by UHG.

**Table 1**  
**Top Ten Information Quality Issues/Needs (UnitedHealth Group)**

A list of action items was also developed and prioritized. Based on this plan, information quality issues/needs are currently being addressed and implemented.

## **CURRENT STATUS**

The **Information Quality Process Improvement (IQPI)** group was formed in early July 2000 (one month after the kickoff meeting and one year after initial data collection from users). The Ingenix SDW VP and the information quality champion have recruited members and sponsors for the IQPI group. There are fifteen members representing all six of the major business segments of UnitedHealth Group (UHG). All IQPI members are volunteers. In their first meeting on August 8, 2000, the group prioritized and addressed the Top Ten List (Table 1). The IQPI group agreed to meet monthly.

In addition to the formation of the IQPI group, several supportive processes and resources have been put in place. An infrastructure consisting of a Data Quality Strategy specifically for Galaxy, a Galaxy Data Quality team, supporting technical staff, and appropriate tools has been built. Processes are in place to continuously improve information quality for UHG. Standards for data names and descriptions have been established. User-oriented metrics for information have been developed.

Mapping data flow for a core business process is under development. The mapping includes relationships between business events, people (e.g., providers, members, employees, etc.), data, and systems. As part of continuous improvement of data quality, a questionnaire to determine user satisfaction with the data dictionary and other meta-data, training, data access, etc. will be sent out once or twice a year.

The model for Galaxy's permanent metadata repository is complete. The metadata repository is in construction. Using backward engineering, the MARS transformation rules were determined and used in the development of Galaxy. Galaxy transformation rules will be included in the metadata repository. Users will be able to easily use and access the majority of Galaxy's metadata via the UHG Intranet. For now, interim metadata is available through the Intranet. Galaxy's permanent metadata repository is scheduled for production on October 31. Four subject areas have been implemented in Galaxy. Data in these subject areas, such as the Provider Subject Area, are defined in the interim metadata repository. Completion of the Galaxy project is scheduled for mid-2001.

The Information Quality Process Improvement (IQPI) Group had their kick-off meeting in early August. The sponsor search has been more than successful with four sponsors onboard. These sponsors are CEOs of UHG business segments and the President of UHG. To provide more clout, visibility, and resources, the IQPI initiative has been moved to the UnitedHealth Technologies group (UHT) (see Figure 1). UHT 'invests' in initiatives such as IQPI, providing resources and technical expertise. Currently, the creation of a full-time, permanent Director of Information Quality position is being considered. If this happens, the IQPI group would work under this Director. Currently, all IQPI members volunteered for this initiative; they were not assigned.

At the conception of a project, a subgroup will be formed from the whole IQPI group. "Develop accurate member counts, eligibility information, and consistent membership information across systems" from the Top Ten List (Table 1) was selected as the basis for the first project. Due to its complexity, this objective was divided into four information quality issues. The first project will be one of these four issues. These 'sub-projects' have started with the task of scoping out the four issues. Scoping is concerned with determining project feasibility. All aspects of feasibility are examined including technical, political/organization and business issues.

The scoping activity is scheduled to be completed by the beginning of October so that the results can be presented to the whole IQPI group at its October meeting. The first IQPI project will be decided at this meeting. This first project should be accomplished within a reasonable time frame, and should not have many obstacles or negative organizational impacts.

## **CONCLUSION**

From the title of this paper, "How to get an Information Quality Program Started," readers might expect very specific instructions—a "how to" guide, like a recipe—that will work in almost all circumstances. Unfortunately, such a "magic bullet" for getting an organizational information quality program started does not exist. One of the difficulties of implementing IQ programs for companies is that, while there are some broad guidelines, there is not, and probably never will be "one way" or an "easy way" to do it.

The Shared Data Warehouse (SDW) department of Ingenix provided the impetus that resulted in UHG's IQPI group. This group is implementing an enterprise-wide IQPI Initiative. The interest in information quality began with the Galaxy Charter, written in mid-1998. It was mid-1999, however, before data quality activities began and mid-2000 when the kick-off meeting for the data quality initiative was held.

While progress may seem slow, several key accomplishments have laid the foundation for improving data quality. First, a document that identifies concerns about UHG's data quality from a business perspective was prepared for top management. Second, top management commitment was obtained. The IQ seminar rallied support from other executives who heard UHG's president and the IQ expert address the group. Third, to help set priorities and create common goals, a Top Ten List was developed. Fourth, the IQPI group was formed with volunteers including many from UHG senior management. Last, sufficient resources were allocated for the Information Quality function to build the necessary infrastructure.

One of the most valuable lessons learned by the information quality champion is that implementing IQ for a data warehouse—never mind across a whole enterprise—is an incredibly complicated and difficult endeavor. Adding to the complexity are relationships among various components including company culture, politics, technology, skills, planning, resources, measurement, diplomacy, timing, behaviors, reward systems, people's availability, senior management involvement, source system involvement, customer involvement, consciousness/education, and commitment and buy in. It is critical that any person or group undertaking a data quality improvement initiative not underestimate the complexity, time, and effort required.

The two most important factors to getting the data quality initiative started at UHG were the SDW VP's belief in the importance of IQ and in using a continuous process improvement approach, and the perseverance on the part of the information quality champion. The Ingenix VP was highly supportive, but his energies and time were constantly being pulled elsewhere. Constant communication (phone calls, e-mail, and plane trips) between the information quality champion and the VP and between the VP and others over a six-month period were necessary to obtain seminar approval and funding, to work out all the logistics, and to attract the right people to the seminar. The continual perseverance and persistence of the information quality champion and support from SDW's VP, UHG's senior management, and UHG's President are requisite to continue the data quality improvement process.