P/C Actuaries: Happy with Your Data

Richard Marr

Biographical Sketch

Richard Marr is currently employed in the Data Management Division, Actuarial Department, ITT Hartford Insurance Group. He holds a doctorate in Industrial Psychology from Hofstra University, and CPCU and CIDM (Certified Insurance Data Manager) professional designations. He chairs the Education Committee and serves on the Executive Board of the Insurance Data Management Association (IDMA). He may be reached at

Data Management, T-1-86
ITT Hartford
Hartford Plaza
Hartford, Connecticut 06115

860-547-2300 (voice)

Rmarr@ITTHartford.Com

Abstract: Actuaries need good data to perform their jobs, yet find that it is often inaccessible and inaccurate within their companies. Advances in computer technology can help eliminate these data access obstacles, but initiation of the change requires activism on the part of the profession.

P&C Actuaries: Happy With Your Data

Things are not as good as they could be, at least not when it comes to insurance data. I hope to make you feel as dissatisfied as I do, not because misery loves company, but because you can shape the future in this area, where I cannot.

As an Actuary, you enjoy an impressive level of prestige and credibility in the industry. You have a working knowledge of mathematics that few people can claim. You are well compensated to sift kernels of information from the chaff of data.

If any one single profession in the industry could shape the future of how insurance data is collected, organized, stored, and retrieved, it would be yours. You know insurance data better than anyone.

Lets step out of the box for a moment. The CEO of a large Fortune 100 insurance company has made you an offer: You are to take no longer than one year to produce a workable information system master plan that will take the company into the 21st century. The plan is to encompass the standard premium, loss, and loss reserve data. Recommendations outside of that scope in areas such as interoffice communications, internet connectivity, and the like are nice, but not what you are being paid for. Neither are cost and time estimates of implementing the change. You are simply to present an overview of the ideal system that could be built using the technology and tools currently available (or which will be available within the one year).

If I were to envision the ideal insurance data collection, processing, and retrieval systems, it would not be the ones we currently have at The Hartford. Further, based on conversations I have had with other people in other companies, it would not be the systems at Aetna, or Travelers, or CIGNA, or Allstate. In fact, I would go so far as to say there is no insurance company today that has anything approaching the ideal system. There is no company that any other company should emulate in this regard, but it does not mean that things will stay this way.

Insurance is getting to be a pretty competitive business. Companies are constantly looking for opportunities to gain market share in profitable segments. Although this requires a fair share of creativity, it also demands good data on which to base marketing and underwriting decisions. The company that can get timely, accurate, and user-friendly business information to its employees will be in a position to gain the competitive edge.

Unfortunately, easy access to poor data is not much benefit, and your profession has been struggling with this issue for years. Why is it that once you obtain company loss data, you need to do some cleaning and scrubbing by matching it with premium data before you can apply statistical methods? Does the surgeon even put on the rubber gloves before the

patent is prepped, shaved, and disinfected? Isn't that somebody else's mess that you are cleaning up? Isn't that somebody else's job? Did your mother raise a custodian?

You are a consumer of information. You do not need to tell an auto mechanic how to tune an engine, but it is your responsibility to tell him when your acceleration is worse than expected. You are the customer of a service that someone else says serves you. You do not need to understand electron theory to tell the repairman that a couple of the switches in your new house don't do anything. You do not need to understand network packet protocol to tell someone that there is a coding problem in a front-end system giving you a spurious data field.

When you state that an additional data element is needed on a record layout, or that the quality of some of your company data is too poor to be relied upon, people in your company will listen. They may not have the knowledge or motivation to fix the problem. They may not be intelligent enough to remove the obstacle. No one, however, will doubt that you are correct in your assessment, that you are correct to ask the questions.

That credibility, along with the recent arrival of several interesting advances in technology, puts you in a position to legitimately ask of people in your company, "Why do we continue to do things this way?"

We can go out on the Internet and get information with a few clicks of the mouse that used to require physically going somewhere (library?) or making several phone calls. Often one did not know where to go for information, so a large portion of time was spent researching where the information was stored. One still needs to research where to look, but conducting a keyword search on the Internet is very simple. Of course, an Internet search can and often does fail to get us exactly what we need, because no person or organization found a need to put the information out there. It is the medium, however, that has captured our attention.

The fundamentals of information technology have reached the point where we should no longer think in terms of what is possible, but instead concentrate on what is desired. This is a pretty radical idea, and runs contrary to our natural tendency as adults in building mental boxes when thinking of problems and solutions. If we can just set aside what we think is possible, however, and clearly articulate what types of information we want, then it will happen.

The current transaction processing systems in a typical insurance company are like an assembly line in a factory. Applicant information comes in the front door, some of the applications are transformed into policy holders, and the premiums, losses, and loss reserves are tracked. All along the assembly line, records are collected together and batch processed for subsequent steps. Copies of records are siphoned off the assembly line and filed into databases, several of which serve the actuarial functions.

Obviously, the whole transaction processing assembly line for an insurance company did not come into being on one day but evolved over many years.

Systems were often developed independently within a company. For example, the policywriting system and the system for claims did not communicate with one another. The policy information for a claim, already in the policywriting system, had to be re-input into the claims system, often with resulting errors. Components of the process were converted from manual workflows and integrated into the remainder of the automated systems already in existence at that point.

Had everything stood still during the evolution of the transaction processing systems, there would be some kind of uniformity and logic to how they fit together. There would be standardized data element names, the modules would be written in a uniform style, the databases would be constructed with similar structures, and there would be one type of computer platform. If this had occurred, the data might still be difficult to access, but at least one would have the impression that there was a coherent grand scheme, and that easier access was just around the corner. Things did not, however, stand still during development of that assembly line.

Obviously, the technology changed over time. As new hardware and software become available, each subsequent development project tried to capitalize on those changes. The company had to find ways to integrate new demands into old code when complete systems rewrites could not be cost justified. More often than not, the old code was left behind in the modules. Programmers fear removing old code because it may still be needed for an odd record here and there. Unfortunately, the modules grow in size and become increasingly difficult for the next maintenance programmer to comprehend.

There may have been massive changes affecting large components of the systems. For example, the company may have reorganized its lines of business to better serve the ultimate customers. Personal lines systems may have either merged with life insurance systems or split apart had they already been as one. Systems that served small commercial accounts may have been merged with large commercial accounts, or split apart. Field offices may have been closed and new ones opened. New insurance products may have been developed while systems were maintained to support runoff from the old products.

As a result, your company's current systems, when viewed in their entirety, are a complex and convoluted hodgepodge of file structures and programming code, collectively referred to as the "legacy systems". OK, the systems designers that installed a new relational database in that other department would go apoplectic if you implied it is now a "legacy" system, but it is. Lets re-think the term, at least for technological systems, as pertaining to everything developed in the past, and I mean everything. After all, the proper definition of the word is "something received from an ancestor or predecessor or from the past". That

is a nice summary of what you have right now, including the recent relational data bases. Just because it is a legacy system does not mean that it is not good at what it was designed to do, but it also does not mean that it is the last word on the subject.

Data values found in legacy systems are not consistent in their use of flags, codes, numbers, dates, or text field sizes. There is little consistency in field names, which can vary from one system to the next even though they are carrying the same values. Two field names can be the same yet carry different contents. Values may be captured up front, but they are turned into codes in downstream systems, losing information, to pay homage to 80 column card restraints.

Even the very idea of an assembly line approach to data processing, although conceptually appealing, often creates its own problems. Downstream systems that obtain their feeds from different locations in the assembly line often do not balance due to timing differences. When an upstream job crashes and needs to be re-run, the downstream feeds need to be re-run as well.

So where does a company start in making the transition from what is to what could be?

The front-end.

The front-end is where human beings put the data into machine readable code. The front-end is a prime source of random errors. The front-end is the bottle-neck, because if the data does not make it at least that far, it certainly can't get any further. The front-end is where the biggest changes will come in the immediate future. Build an excellent front-end, and everything else will follow.

Insurance companies have designed their front-ends so that those who use it require thick manuals and coding charts. Because these positions were often pretty low in the corporate hierarchy, not much attention was paid to building graphical user interfaces, context sensitive help, and on-line manuals and field relationship edits. Things are changing, however. In the effort to streamline and redesign processes, the job of input technician is being replaced by that of a customer service representative or underwriter. Input technicians are going the way of the typing pool. In some applications, such as selling personal lines policies over the Internet, the input is done by the customer themselves.

And when the information is entered into the front end, lets pass it downstream without dropping half of it on the floor. It does the Actuaries and others no good if the front end input system captures information that stays in the front end files and never is passed to the MIS systems. This has been a widespread industry practice, but it is a good example of universal idiocy. For example, a company that will remain nameless used to do this with information such as which endorsements were attached to the policy. Since this often

did not involve discrete money amounts, the data was never passed. Even when data is passed along it is often materially changed. That is why we have policies at 25, 50, 100, 250, 500 and 1,000 limits that are passed along as limit codes 1,2,3,4,9, 9. Try using that data to do an analysis of your needed ILF to move from a 500 to 1000 limit.

Why do companies still do this? Because well-paid people in responsible positions continue to think in terms of the 80 byte IBM card. That card's layout was the cutting edge of technology when first applied in the 1890 United States Census. That card's image is still burned into the minds of many MIS managers who started their careers by sorting them 25 years ago. That card is why dollar limits are still condensed down to a single byte in many insurance systems today, even though the physical cellulosic medium that initiated it is at the bottom of the landfill. It is just plain embarrassing, and deserves the best ridicule you can muster.

The data entered in the front-end goes straight to a central database. All systems which used to get feeds off of the flow of records going down the assembly line now can pull the data off the database. As information ages from an active status to historical, it is moved off into data warehouses. Communications across workstations, databases, and platforms are via inter and intra nets.

The purpose is not to simply adopt something new, but to do something better. The constant reorganization of office field structure or the reshuffling of lines of business across marketing units isn't going to go away, but will probably increase as companies search for ways to serve the customer. To the extent that the information systems are designed for ease of update and change, companies will be able to pursue optimal customer service.

No question, there are hundreds of problems to solve along the way, such as security and response time, but there is nothing that cannot be overcome. The Actuarial profession can, and should have, an important role to play in this transition.

If you, as an Actuary, are not actively spending at least a little time each month to help bring your company's data technology into the twenty first century, then you deserve the data you get.