November 1995

Chaos: Mapping - The Selection of Standard Data Elements

Sanda K. Paul

SKP Associates

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>

Recommended Citation

Paul, Sanda K. (1995) "Chaos: Mapping - The Selection of Standard Data Elements," *Against the Grain*: Vol. 7: Iss. 5, Article 17.
DOI: <https://doi.org/10.7771/2380-176X.1621>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Chaos

Mapping: The Selection of Standard Data Elements

by **Sandra K. Paul** (President, SKP Associates)

For more information about any of the concepts in this column, contact Sandy at the SISAC office, 160 Fifth Ave., NY, NY 10010. Internet 4164812@MCIMAIL.COM, phone 212/929-1393, fax 212/989-7542.

The following article was prepared by Bob Boissy of Faxon, Chair of the SISAC Technical Advisory Group. It describes the activity called "mapping," which — in fact — is the selection of standard data elements which "map" or reflect the business practices of a given community — in this case the serials community. It is intended to create thought and discussion. We look forward to seeing your letters to Katina — or to Bob or me. — SKP

Mapping X12 Industry Conventions: Rules of Thumb

by **Bob Boissy**, Chair,
SISAC Technical Advisory Group (TAG)

Introduction

The Serials Industry Systems Advisory Committee (SISAC) has been "mapping" X12 transaction sets for a number of years now. This work has been done in the hope of creating a format-neutral environment for automation development, thus lowering the unnecessary costs associated with accommodating multiple non-standard formats for multiple transaction types. Many mapping revisions have been made and many lessons have been learned in the course of generating consensus on conventions for implementing electronic data interchange (EDI) by the diverse and sometimes competing systems.

Some of the decisions made along the way have resulted in options that, with hindsight, we would not have allowed. But the resulting mappings for invoice, claims, claims response, and dispatch data, published June, 1995, are all serviceable, and will certainly be implemented by many library systems, agents, and publishers. It is my opinion that where some may say we have "erred," we have erred on the side of generosity; and that may be a fact of life when working in large groups seeking consensus. The very good thing about SISAC is that it is an organization that will keep revising and enhancing the standards infrastruc-

ture we all rely on, thus allowing further automation in the serials industry in the pursuit of cost control.

This then is a small set of general purpose rules for the design and communication of ANSI X12 industry conventions that I have gleaned from my SISAC mapping experience. No attempt is made here to outline all aspects of a great mapping or a great written specification. This is just a personal collection of my favorite rules of thumb. Some background in analyzing business data exchange needs and communicating methods for satisfying such needs is assumed. Familiarity with the X12 format is assumed.

The Functional Specification

No X12 mapping should take place without a clear functional specification and list of business data elements. The functional specification should not have to make any detailed reference to the X12 format. It is a mistake to try and let the ANSI formats dictate the qualities of a particular business transaction. A typical X12 transaction may allow slots for literally thousands of types of data.

Consider this. The 850 Purchase Order transaction at the Version 3 Release 2 level of the ANSI standards allows for the transmission of 923, 455, 362, 698 data elements in one X12 file. This number includes maximum segment repetition and maximum loop repetition. Remember, this is the number of slots allowed, not the maximum file length. To get the maximum file length you would have to add up all the maximum lengths of all the allowed elements. It would be a true Carl Sagan number.

Under these conditions, it is not really possible to let the X12 transaction tell you what data you need to exchange with your partners. The creation of a workable functional specification for an industry business document is normally accomplished by a committee. This committee should start with documentation on the current (non-X12) business transaction and eliminate data that is rarely or never used. The committee should provide a common name for every piece of

business data and a description of the function of the data. The committee should decide if the data should always be sent, if it should be sent under certain conditions, or if it is optional. Other important aspects of the specification include:

Background - context.

Intended use of the transaction - scope.

Any usage intentionally excluded from the scope.

Intended trading partner constituencies - audience.

Date and contact information for author(s).

Mapping

Macro-rules

The initial X12 mapping from a functional specification is best accomplished by either an individual, or at most two people. The review and polishing of a mapping is best accomplished by as many qualified technicians and business people as possible. Choose your transaction set first by its normal defined use, and second by its features. Seek to avoid using the same transaction set for more than one industry document type.

Avoid redundancy — map as much to the header (Table 1) and trailer (Table 3) as possible. Table 2 is by definition a loop designed to hold line item data. Allowed loop iterations for Table 2 vary by transaction, but it is safe to say that they are usually generous. Table 2 data should always be scrutinized the most closely before inclusion.

Never get into a mapping dilemma where it is necessary to send a distinct document for every detail line to be transmitted. Most X12 translation software programs provide audit control at the document level. An overabundance of documents of the same type, from the same partner, with little separation

continued on page 80

in time, is an audit control problem and an indication of poor engineering.

When mapping, put yourself mentally in the place of the recipient, not the sender of the data. Is the mapping sufficiently constrained to minimize the task of programming the upload? The more limited the options, the faster the implementation.

Setting limits is what industry convention specifications are supposed to do. In general, use as few segments and elements as possible to get the job done. More data means more communications costs. It is negligible for one transaction, but adds up for many transactions over a long period of time. It is harder to map a smaller set of segments and elements than to allow for the kitchen sink, but no one will be pleased with an industry convention with most ANSI options still in place.

X12 documents are intentionally designed for the exchange of transient, compact, business transaction data. X12 is not well suited for exchange of archival, free descriptive, or non-business documents. Provide as little opportunity for free-text description fields - such as REFO3, PIDO5, and CTTO7 - as possible. Such descriptive material should be limited to the types of data that are typically uploaded to free-text fields. Use numbers, not names. Use numbers, not addresses. Use numbers, not product descriptions. Etc. Limit the number of X12 codes used from the ANSI X12 code lists.

Concentrate on mapping the "match points", also known as "hooks." For every detail line that is going to post to a target system, there should be one or more data elements in the line that match to the recipient's target record. These should be clearly mapped, and mapped to only one place. If there is a return transaction planned that requires one or more hooks back to the originator's system, that should be clearly and cleanly mapped in the same way. There can be match point data in the header portion of a record, but this is not as typical. It is more typical for recipient systems to check header data to see if the document has already been processed, and if so, to reject it. Hook data does not necessarily post to a recipient system; it is data that identifies the record on the recipient system to which other transaction data will post.

SISAC and the International Committee on EDI for Serials (ICEDIS) have previously identified the following EDI match points:

LIBRARY

LSID Library's subscription identifier,

AKA the library's order line number. A unique number, for a given library, attached to a single order line.

LCN Library's claim number. A unique number, for a given library, which identifies a single claim it has made for one or more copies of a serial issue.

LCI Library's cancellation id. A unique identifier, for a given library, for each and every cancellation sent to agents.

SUBSCRIPTION AGENT

ACAT Agent's catalog identifier. Defines a single or grouped set of items for sale. Sometimes referred to as a title number.

ASID Agent's subscription identifier. A unique number, for a given agent, attached to a single order line. Some subscription agents create this number by combining an order reference which is unique for each library with a customer identifier or account number.

ACN Agent's claim number. A unique number, for a given agent, which identifies a single claim in its entire system. This may be derived from a combination of other descriptors, e.g. library claim number and account number.

ACI Agent's cancellation id. A unique identifier, for a given agent, for each and

actions will fail to load, and extra data will be needed for error reports. ("Door-to-door EDI.") For example, full addresses and contact information may be used. More free text description may be warranted. If mostly system to system exchange is expected, use either a number or code to carry such information, or do not include it at all. And frankly, it is probably better to stay with paper or other transaction techniques until post-receipt printing and load errors can be minimized.

Transactions that require real time or near real time responses are candidates for direct connectivity, rather than batch-oriented EDI. Such connectivity may still utilize the X12 format, but the mapping approach may be so streamlined as to differ significantly from richer, consensus-oriented, industry group mapping. It makes no sense to do real-time automated transaction processing. And to do such processing, it is necessary that transactions be very parsimonious and very heavily weighted towards numbers, identifiers, and codes. Generally, this type of inter-enterprise data processing is only done by trading partners with long-standing EDI relationships. In other words, you have to become very familiar with a partner's data before you can feel comfortable loading it in real-time.

The A in ANSI stands for American. The ANSI X12 standards were originally designed for American domestic trade. Trade with international partners is not

always perfectly served by the X12 format. For example, address elements (e.g. postal codes, state/province codes) tend to assume American practices, at least in the earlier versions of the standards. This situation is slated for change when the merger between X12 and EDIFACT occurs. In the meantime, we all live with this knowledge and learn to accept some work-arounds. The scientific/technical/medical serials and book industries will eventually migrate to EDIFACT standards. EDIFACT is so much like X12 that the transition will not be difficult, and no organization should delay entry into the EDI world on this account. Almost all commercially available translation software packages handle both X12 and EDIFACT.

Micro-rules

Start by filling all X12 mandatory elements.

Do not use Note (NTE) segments. If you have a transaction that is so specialized that it requires free-text explanation to process, it should not be sent via EDI. The recipient's only alternative is to print out all transmissions containing NTE segments.

Use of the Administrative Communica-

continued on page 81

"SISAC is an organization that will keep revising and enhancing the standards infrastructure ..."

every cancellation sent to publishers.

PUBLISHER

PCAT Publisher's catalog identifier (product id). Any identifier which uniquely identifies an item or grouped set of items available for sale in a single order for a given publisher.

PCRT Publisher's price criteria (price variation). A unique descriptor, for a given publisher and PCAT combination, that identifies the price rate applicable.

PSID Publisher's subscription identifier. A unique number for a given publisher attached to a single order line. As with most agents, this number may be derived from a combination of other descriptors.

ALL CONSTITUENCIES

ITID Item identifier. Preferred identifier is the machine and eye-readable serial item and contribution identifier known as the SICI. This string is defined in ANSI Z39.56-1991, soon to be re-released. This descriptor is capable of uniquely identifying a single serial issue within a run of issues for a title within a subscription. The string identifies the title as well as the issue.

Mappings are typically loaded up with data options when there is a fear that the majority of trading partners will be printing out received transactions, or that many trans-

Chaos from page 80

tions Contact (PER) segment is questionable but not unpardonable. Such contact data will usually be a constant across all documents of that document type. If it is a constant, that is a clue you are wasting your time and money sending it with every document. Give your trading partner all such constants ahead of time, and save the EDI overhead.

Put in dates in Date (DT) elements.

Put monetary amounts in Real or Name (N2) elements.

Provide unique REFO1 codes when using multiple References (REFs).

Provide unique qualifiers within any segment, e.g. IT1 or PO1, that contains many iterations of a qualifier and identifier pair.

The CTT07 element is bad for two reasons. It allows another slot for free text, and use of this element effectively washes out the entire document in which it appears. There is no way for a machine to know what kind of comment might be in this element, which by definition pertains to the entire document to which it is attached.

Use data elements that come later in segments to map the least used optional data. This shortens the typically transmitted segment.

It is not shameful to use X12 codes that are defined as one thing to mean something else, as long as it is clearly documented. The alternatives - using ZZ, waiting for ANSI-Accredited Standards Committee (ASC) X12 to accept your new code requests, and only using codes that carry accurate X12 definitions for your transaction - are more problematic and limiting than useful. The negative aspect of using the X12 code, element, segment, and transaction set definitions loosely is that it lessens the ability to use an industry specification with other industries. Therefore, the better the initial fit of transaction set to industry use, the better the chance of cross-industry trading. This is not sufficient reason to slow down the process of implementing X12 in an industry until the perfectly tailored transaction can be designed and passed through ANSI ASC X12 committees.

It is both possible and probable that industry groups will not use each data element mapped to carry exactly the piece of data intended by ASC X12. As long as the industry function is comfortably served by the element, the use is justified.

Updating/Upgrading

If no prior X12 industry mapping exists for the document type, choose the highest

level of the X12 standards currently published. There is no need to upgrade a good specification to a higher level of X12 once the specification is well into development or in use, unless there are demonstrable advantages in doing so. Some reasons for upgrading to a more current version of the X12 standards are as follows:

Movement to accommodate international trading data, e.g. ISO standards.

Increased X12 element length for a crucial piece of business data.

Increased loop counts for crucial repeating data.

Increased segment repetition in crucial area.

Switching entire mapping to a new, more suitable transaction set.

Rarely is a new special purpose segment added which so influences an industry as to make re-mapping a specification desirable.

Writing the Specification's Explanatory Text

Be definitive. Be prescriptive. It does not help anyone to suggest, recommend, guide, abjure, or cajole. Specifications are for telling how something is done. Define the allowed number of repetitions of a segment, and how each repetition may be used.

Always use requirement designations. Indicate both the X12 and industry requirements. (Optional, Conditional, Required.) Use as many mandatory statuses as possible. State exactly the conditions for the conditionals.

In comments and explanatory text:

- Use acronyms and initialisms as little as possible. X12 is code enough already.
- Use examples liberally.
- Explain more rather than less.
- Do not refer the reader to other places in the specification for an explanation of something that should be repeated for clarity in each instance of its occurrence.
- Avoid semantic confusion among trading partners: explain the role of key pieces of business data as well as giving them names.
- Explain industry practice for exchange of whole classes of data, e.g. money.

In scope statements (not in any particular order):

- State the version and release of X12 used for the mapping.
- Include all the key points from the functional specification.
- Include what cases the specification is and is not supposed to cover.
- Include the intended user group for the specification.

Profiles Encouraged:

Sandra Koodin Paul

Born: Sandra Koodin, 6/6/38, New York City, NY. I'm a native New Yorker with a "thing" for double numbers. Grew up in Mt. Vernon, a suburb of New York City just over the Bronx border.

Current Residence: New

York City. Chelsea Section, 2 blocks from the office. After years of commuting using the NYC transit system, this may be THE prime benefit of being a self-employed consultant.

First Job: Sales person at Macy's on Thursday nights and Saturdays; then waitress in a college dorm and a Catskill mountain resort.

Resume: Attended Cornell for two years and consider that my Alma Mater, although I received my BA from Hunter College (nights) and my MA and all-but-dissertation Ph.D. from the City University of New York. Was a Psychology major, but was trained as a management consultant by J.K. Lasser & Company, the accounting firm. Worked at Random House for over 11 years before establishing SKP Associates, which specializes in business systems analysis and development and the application of computer-based techniques, as well as association management. Bill Raggio, my Senior Associate, celebrated his 10th anniversary with the company in June, 1995! Our clients have included: the Library of Congress; Book Industry Study Group; Workman Publishing; the Association of American Publishers; Warner Books; the National Association of College Stores; and the Uniform Code Council.

Family: Immediate family includes: step-daughter Sarah Paul Bentley and husband Alden, parents of Lia Nicole (3 years old) and Paul Cody (1 year old), living 25 minutes away; step-son Adam and his wife Ann Marie Paul, who live in Buffalo and made me a grandma for the third time in July, 1995; Todd and Alice (the cats); Dad; and a variety of other important folks. Extended family includes Bill along with lots and lots of friends. Carol Nemeyer, who introduced me to the library community, continues to be a corresponding friend from her home aboard the sailboat, Rainbow, somewhere in the Caribbean.

Favorite pastimes: Dancing and spending summer weekends on Fire Island.

Spare time activities: Working on standards for books, serials and beyond. Writing for *Against the Grain*. Traveling.

What would I like to be doing five years from now: The same things, having as much fun, but making much more money doing so!