



ISO PC 295 Audit Data Collection

Analysis of information requirements in CD1 1022
Proposal for modelling the requirements

Frans van Basten
EDI modelling expert
Member of NEN normcommissie Audit data collection (NC 380295)

Normalisation of data (1)

- *A lot of elements in CD1 are copied in multiple tables. The tables define a flat file exchange format. Maintenance and consistency of the information is very difficult if a lot of elements are copied around.*
- *When using ISO 17113 Method for development of messages, we maintain all data elements in a relational model without redundancy. From this model we can still generate the flat file exchange format in which a complete hierarchy of information can be exchanged in one table.
But we can also support object-oriented exchange formats like XML and JSON.*

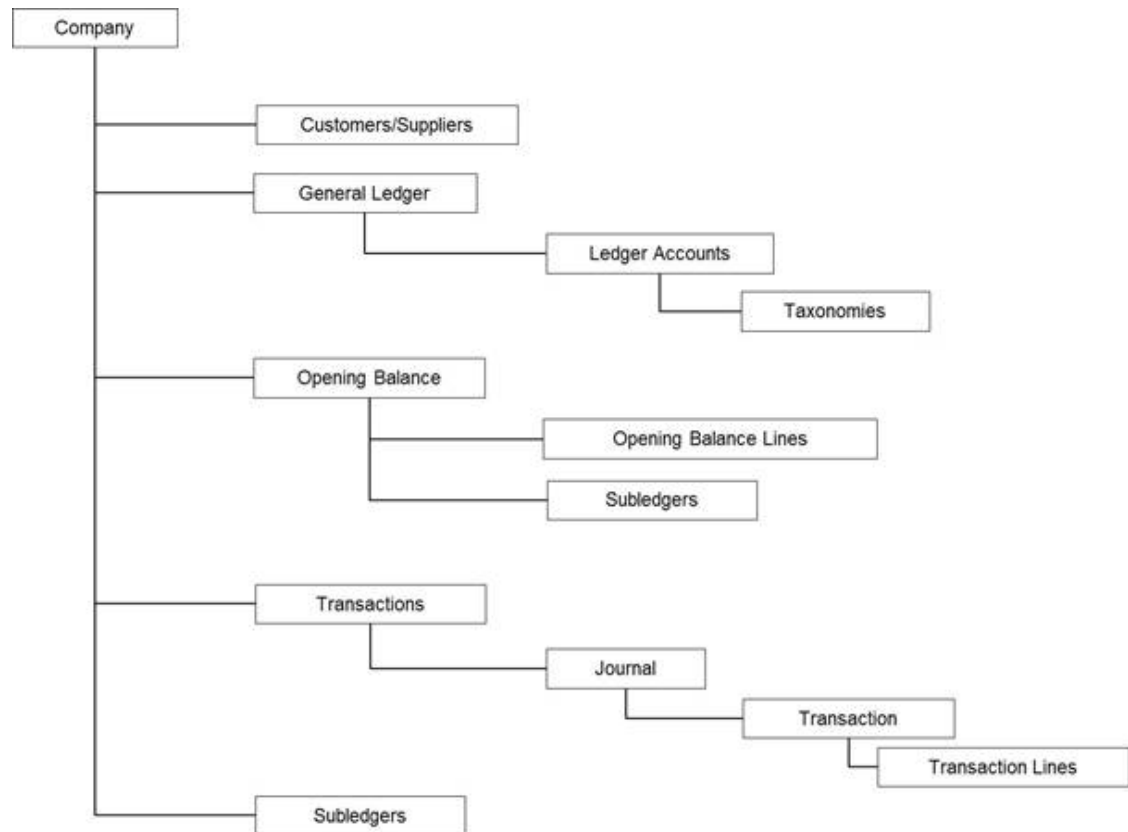
Normalisation of data (2) - Example 1

- **Trial Balance**

- * *GL Account Number*
- * *Amount currency*
- * *Amount reporting currency*
- * *Amount local currency*
- * *Amount beginning (3x)*
- * *Amount ending (3x)*

- **General Ledger Detail**

- * ...



CD1: Flat tables

redundancy, only flat structure support,
difficult to maintain

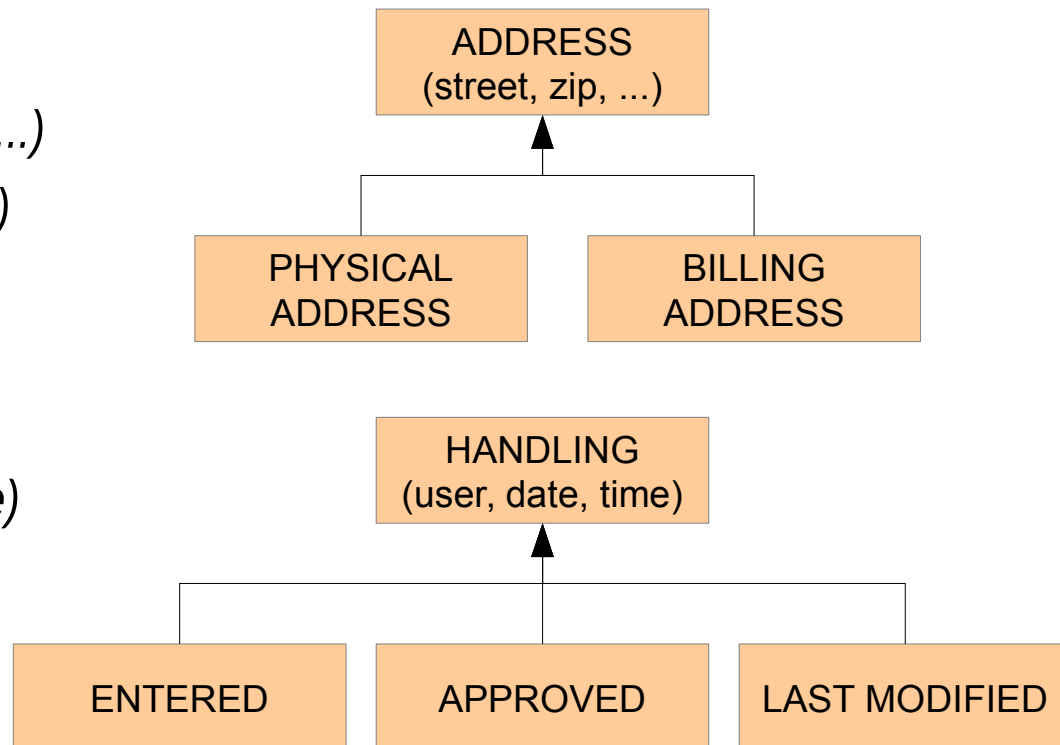
Dutch Autit File: Modelled, object oriented

no redundancy, support for object oriented exchange formats (XML,
JSON), easier to maintain, easier to create extensions

Normalisation of data (3) - Example 2

Customer/Supplier Listing

- * Physical Address (street, zip, ...)
- * Billing Address (street, zip, ...)
- * Entered (user, date, time)
- * Approved (user, date, time)
- * Last modified (user, date, time)



Reusable structures, less elements to maintain, support for other exchange formats (XML, JSON) using ISO 17113.

Normalisation of data (4) - First normalisation step

- First normalisation step on CD1, by creating reusable information blocks (entities):
 - * ADDRES
 - * AMOUNT AND CURRENCY
 - * CONTACT
 - * HANDLING
 - * TAX
 - * SUBSIDIARY ITEM

Number of fields defined in CD1 Tables: **1376**,

Number of fields (attributes) defined, after normalisation in datamodel: **893**,

Reduction of **483** field specifications = **35%** to maintain.

Normalisation of data (5) - Further possible steps

- Further possible normalisation steps on CD1, combining tables into reusable information blocks (entities):
 - * CUSTOMER / SUPPLIER
 - * SALES CONTRACT / PURCHASE CONTRACT
 - * SALES ORDERS / PURCHASE ORDERS
 - * AR ADJUSTMENT / AP ADJUSTMENT
 - * AR CASH APPLICATION / AP CASH APPLICATION
 - * INVOICE GENERATED / INVOICE RECEIVED
 - * CASH RECEIVED / PAYMENT MADE etc.
 - * PPE ADDITION / PPE REMOVAL / PPE CHANGE

This will remove another more than **200** field specifications to maintain!

Normalisation of data (6) - Reusability

Flat structure, currently used in ISO tables:

INVOICE GENERATED	1..*, R
INVOICE GENERATED LINE	1..*, R

Nested structure, normally used in XML messages:

INVOICE GENERATED	1..*, R	
TAX	1..4, R	(is 4 repeats enough?)
INVOICE GENERATED LINE	1..*, R	
TAX	1..4, R	

Datamodelling work done so far

- *All tables in CD1 have been modelled into **Audit Data Collection Datamodel**. **The complete model is there!***
- *First normalisation step has been performed (see slide 6)*
- *Naming conventions have been applied, for instance:*
 - * *Currency -> Currency code*
 - * *Employee Title -> Job Title*
- *Errors in datatypes have been corrected*
- *Tables from CD1 have been generated from the datamodel:*
 - * *AuditDataCollection_FunctionalTables (definitions, domains)*
 - * *AuditDataCollection_TechnicalTables (flat file spec, field sequence)*

AuditDataCollection_FunctionalTables (1) - Tables

BUSINESS SEGMENT		1..*, R
<i>BUSINESS SEGMENT</i>		
<u>Datamodel entity description</u>		
The portions of an organizational chart that is reflected in the business transactions as structural units (e.g. business unit, department, cost center, project, etc.)		
<u>Datamodel entity comments</u>		
02		
Business Segment ID	R	an..25
<i>domain: IDENTIFIER 25</i>		
<u>Datamodel description</u>		
Identifier that is unique for each business segment.		
Segment Name	R	an..25
<i>domain: NAME 25</i>		
<u>Datamodel description</u>		
Name of the business segment.		
Segment Reference Level	R	n..2
<i>domain: BUSINESS SEGMENT REFERENCE LEVEL</i>		
<u>Datamodel description</u>		
Relative level of the segment with 1 being the consolidated level and numbers increasing through lower levels.		
Organizational Name	R	an..60
<i>domain: NAME 60</i>		
<u>Datamodel description</u>		
Indicate the name of the organization type, can fill in Department, Business Unit, etc.		

AuditDataCollection_FunctionalTables (2) - Domains

Domain specifications

A

Name : **ACCOUNT HIERARCHY**
 Datatype : Signed Integer, total digits 2
 Format : n..2

Name : **ACCOUNT TYPE**
 Datatype : Alpha Numeric, maximum length 25 (variable length)
 Format : an..25

Name : **ACTIVE FLAG CODE**
 Datatype : Alpha Numeric, maximum length 1 (variable length)
 Format : an..1

Code list : **Active Flag Code**

Code(s)	Value	Name	Valid From	Valid To
	N	Inactive		
	Y	Active		

Name : **ADJUSTMENT INDICATOR**
 Datatype : Alpha Numeric, maximum length 1 (variable length)
 Format : an..1

Code list : **Adjustment Indicator**

Code(s)	Value	Name	Valid From	Valid To
	0	Transaction is the original shipment transaction.		
	1	Transaction is a shipment adjustment.		

Name : **AMOUNT**
 Datatype : Signed Monetary Amount, total digits 22, decimals 6
 Format : n..22,6

AuditDataCollection_TechnicalTables (1)

Inhouse file specifications:

Flat file type:	Variable Field Length Records
Decimal separator:	[.] (dot character, ascii=46)
Field delimiter:	[;] (semi-colon character, ascii=59)
Escape method:	Quote alphanumeric fields if it contains the field separator or quote character (quotes in the data will be doubled)
Text delimiter:	["] (double quote character, ascii=34)
Quote all text fields:	No

Record types in the inhouse file:

SEG , BUSINESS SEGMENT	1..*, R
REL , BUSINESS SEGMENT RELATIONSHIP	0..*, O
EMP , EMPLOYEE	0..*, O
USR , USER	0..*, O
CST , CUSTOMER TYPE	0..*, O
CUS , CUSTOMER	0..*, O

AuditDataCollection_TechnicalTables (2)

Record **CUS**: CUSTOMER

0..*, O

CUSTOMER

Sequence

1	'CUS'	R	an3
2	Customer ID	R	an..60
3	Customer Number	R	an..60
	...		
15	Terms Due Days	O	n..6
	PHYSICAL ADDRESS	1..1, R	
	<i>customer - PHYSICAL ADDRESS</i>		
16	Street Name Line 1	R	an..100
17	Street Name Line 2	O	an..100
18	City Name	R	an..100
19	State Province Code	O	an..6
20	Country Code	R	an..3
21	ZIP Postal Code	R	an..20
	BILLING ADDRESS	1..1, R	
	<i>customer - BILLING ADDRESS</i>		
22	Street Name Line 1	R	an..100
23	Street Name Line 2	O	an..100
24	City Name	R	an..100
25	State Province Code	O	an..6
26	Country Code	R	an..3
27	ZIP Postal Code	R	an..20

Next steps towards CD2

- *Verify the documents generated from the datamodel in the international working group*
 - * *AuditDataCollection_FunctionalTables*
 - * *AuditDataCollection_TechnicalTables*
- *Decide on using the datamodel as basis for generating the tables*
- *Decide on further normalisation steps for the datamodel*
- *Discuss and decide on the open issues:*
 - * *Issue 1 - Codes versus names*
 - * *Issue 2 - Abbreviation*
 - * *Issue 3 - Codelists are missing*
 - * *Issue 4 - Address structure*
 - * *Issue 5 - Other issues found while modelling*
- ***The Dutch team can support maintaining the data model!***

Issue 1 - Codes versus names

- *Normally a code is a small field, that identifies a certain name, but is not understandable for a human (no semantics).*
- *The name is a longer field which makes it understandable for a human (semantics).*
- *In CD1 the coded field is often longer than the name field!*
 - * *Customer Type **Code**, **an..100** - Customer Type **Name**, **an..80***
 - * *Supplier Type **Code**, **an..80** - Customer Type **Name**, **an..80***
 - * *Journal Entry Type **Code**, **an..60** - Journal Entry Type **Name**, **an..60***
 - * *Transaction Type **Code**, **an..60** - Transaction Type **Name**, **an..60***
 - * *(many more)*

Issue 2 - Abbreviation

- *An abbreviation of a name is smaller than the name itself*
- *In CD1 the abbreviation fields often have the same length!*
 - * *Customer Abbreviation, **an..100** - Customer (Account) Name, **an..100***
 - * *Supplier Abbreviation, **an..100** - Supplier (Account) Name, **an..100***
 - * *JE Type Abbreviation, **an..20** - JE Type Name, **an..60***
 - * *Measurement Unit Abbreviation, **an..60** - Measurement Unit Name, **an..80***

Issue 3 - Codelists are missing (1)

- *A lot of codelists in CD1 are not specified*
- *By specifying codelists, more semantics are added to the exchange formats, which makes it easier to generate and to process! Examples:*
 - * *Supplier Type Code*
 - * *Account Type, Account Subtype in 'Chart of accounts'*
 - * *Journal Entry Type Code*
 - * *Document Type Code*
 - * *Transaction Type Code*
 - * *Measurement Unit Code*
 - * *...*

Issue 3 - Codelists are missing (2)

- **TAX.Tax Type Code** no codelist specified!
 - * Possible codes:
 - » VAT (btw)
 - » Excise (accijns)
 - » Insurance tax (assurantie belasting)
 - * Or (in this case a repeat of 4 might nog be enough):
 - » VAT - high (btw hoog tarief)
 - » VAT - low (btw laag tarief)
 - » VAT - other (btw overige tarieven)
 - » Excise (accijns)
 - » ...

Issue 4 - Address structure

- *House number and House number extension are missing in the address structures used in CD1. This makes mapping from many national and international address standards very difficult.*
- *The address structure should be checked against other address standards.*

Other issues found while modelling

- Remarks previously reported (e.g. names, descriptions conform ISO 11179)
- Table 2: Length of Unique BS ID is 25, but the fields in other tables, that link to the BS ID has different length. F.i. Department Code has length 60 instead of 25.
- Table 16: Measurement Unit Code has length 80 in this table. But in other tables, measurement Unit Code has different length
- Table 23: File Content Hierarchy has representation string %2s, that should be integer %2d
- Table 10: Balance Debit Or Credit indicator (debit,credit) is different from the debit credit indicator in table 26 (D,C).
- Table 26: Cancellation Sign, Boolean field is not defined. I suggest to use string %1s, with values Y and N
- Table 28: Corresponding File, description not clear.
- Table 29: Field 15 and 16, debit and credit indication differs from debit credit indication f.i. in table 26
- Table 38: Sales Measurement Unit Code has length 60, in table 16 it has length 80. Description of "Unit Price Tax incl." should be "The unit price (including tax)". Description of "Amount Tax incl." should be "The amount (including tax)"
- Table 40: Dispatch Unit Code has representation %22.6f, this should be %80s. Sales Order Line Measurement Unit Code has representation %25s, in table 16, it has %80s.
- Table 42: Quantity Basic Measurement Unit has %22.6f, this should be %20.6f for Quantity. Field 33, 34, 35, has wrong descriptions (debit, credit).
- ... and many more ...