



Cleansing Data Using Ontologies

PANEL

Dr Gerald Radack
Dr SALOMON DE JAGER

Who saw the ECCMA white paper on data cleaning?

ECCMA

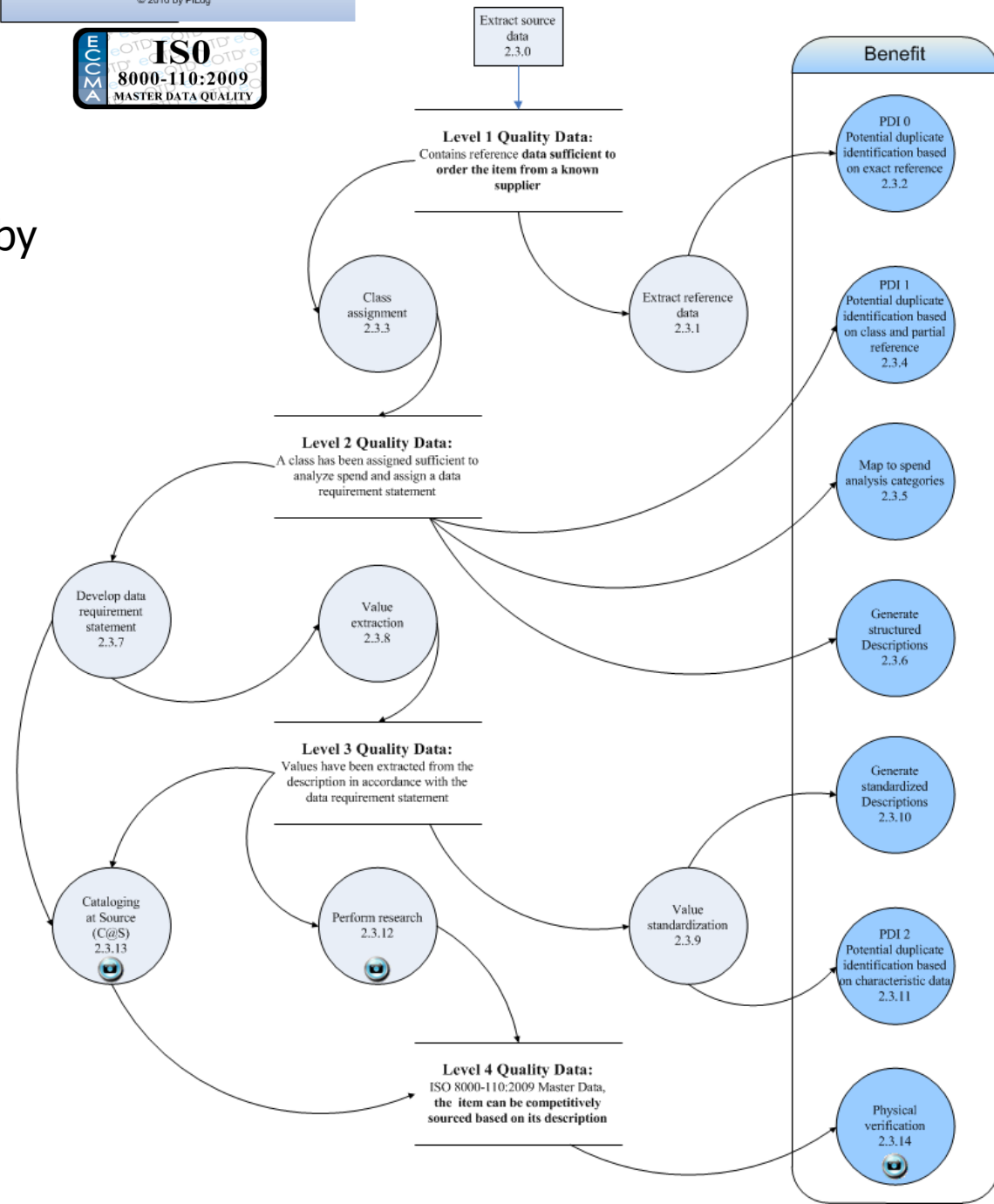
October 2010

© Copyright Protected



PiLog provides state of the art data cleansing services backed by rigorous quality control.

All cleansed data delivered by PiLog is ISO 8000-110:2009 compliant.



Levels of Cataloguing

- Level I: The item is identified in terms its source of supply and has a reference sufficient to successfully place an order for the item (sufficient for contracting with a specific supplier).
- Level II: The item is identified and has been assigned a class (Noun & Qualifier) sufficient to allow the classification of the item for the purposes of enhanced spend analysis
- Level III: The item is identified and partially described, a class has been assigned and some of the properties specified in the Identification Guide for the class have been provided utilizing the client original data.
- Level IV: The item is identified and completely described (enrichment) in accordance with the buyer's Identification Guide (sufficient for competitive bid or contracting on generic items).

Level 1 & 2 Cataloguing

Original Client Description

MOTOR 400KW 6 POLE 525VOLT MAKE WEG P/N 1234EF: SF1,0 CODE G:IP65:IL/IN 6.6:DUTY SI: NR.88695 11 00:AMB 40DEGREE C:DELTA T 80DEG: COS 0,86:COOLING IC 411:ALT1000M:MPD SEC DREDGE MAIN MOTOR

Class

MOTOR ELECTRIC

Part No

Mnfr

1234EF

WEG

Properties

- Power Rating
- Electrical Rating
- Frame
- Mounting
- Speed
- Insulation Class
- Poles
- Service Factor
- Enclosure

FFT

400KW, 6 POLE, 525VOLT, SF1,0 CODE G:IP65:IL/IN 6.6:DUTY SI: NR.88695 11 00:AMB 40DEGREE C:DELTA T 80DEG: COS 0,86:COOLING IC 411:ALT1000M:MPD SEC DREDGE MAIN MOTOR



Level 3 Cataloguing

Original Client Description

MOTOR 400KW 6 POLE 525VOLT MAKE WEG P/N 1234EF: SF1,0 CODE G:IP65:IL/IN 6.6:DUTY SI: NR.88695 11 00:AMB 40DEGREE C:DELTA T 80DEG: COS 0,86:COOLING IC 411:ALT1000M:MPD SEC DREDGE MAIN MOTOR

Class

Properties

Values

UOM

MOTOR ELECTRIC

Part No

Mnfr

1234EF

WEG

- Power Rating
- Electrical Rating
- Frame
- Mounting
- Speed
- Insulation Class
- Poles
- Service Factor
- Enclosure

400

KW

525

VAC

6

1.0 CODE G

IP65

FFT

IL/IN 6.6: DUTY SI: NR.88695 11 00:AMB 40 DEGREE C:DELTA T 80DEG: COS 0,86:COOLING IC 411:ALT 1000M: MPD SEC DREDGE MAIN MOTOR



Level 4 Cataloguing

Original Client Description

MOTOR 400KW 6 POLE 525VOLT MAKE WEG P/N 1234EF: SF1,0 CODE G:IP65:IL/IN 6.6:DUTY SI: NR.88695 11 00:AMB 40DEGREE C:DELTA T 80DEG: COS 0,86:COOLING IC 411:ALT1000M:MPD SEC DREDGE MAIN MOTOR

Item Class

MOTOR ELECTRIC

Part No

Mnfr

1234EF

WEG

FFT

Properties

- Power Rating
- Electrical Rating
- Frame
- Mounting
- Speed
- Insulation Class
- Poles
- Service Factor
- Enclosure

Values

400
525
HGF355E
FOOT
988
F
6
1.0 CODE G
IP65

UOM

KW
VAC
RPM

IL/IN 6.6: DUTY SI: NR.88695 11 00:AMB 40 DEGREE C:DELTA T 80DEG: COS 0,86:COOLING IC 411:ALT 1000M: MPD SEC DREDGE MAIN MOTOR



Description Generation

Original Client Description

MOTOR 400KW 6 POLE 525VOLT MAKE **WEG P/N 1234EF**: SF1,0 CODE
G:IP65:IL/IN 6.6:DUTY SI: NR.88695 11 00:AMB 40DEGREE C:DELTA T 80DEG:
COS 0,86:COOLING IC 411:ALT1000M:MPD SEC DREDGE MAIN MOTOR

Catalogue PO Description

MOTOR, ELECTRIC: POWER RATING 400 KW, ELECTRICAL RATING 525 V, FRAME HGF355E, FOOT MOUNTING, SPEED 988 RPM, INSULATION CLASS F, 6 POLES, SERVICE FACTOR 1.0 CODE G, ENCLOSURE IP65, Mnfr P/N: 1234EF Mnfr: WEG, FFT: IL/IN 6.6: DUTY SI: NR.88695 11 00:AMB 40 DEGREE C:DELTA T 80DEG: COS 0,86:COOLING IC 411:ALT 1000M: MPD SEC DREDGE MAIN MOTOR

Catalogue Short Description

MOTOR, ELECTRIC: 400 KW, 525 V, 988 RPM

What is an Ontology?

What is an Ontology?

a formal representation of knowledge as a set of concepts within a domain, and the relationships between those concepts. It is used to reason about the entities within that domain, and may be used to describe the domain.

¹Wikipedia

ECCMA

October 2010

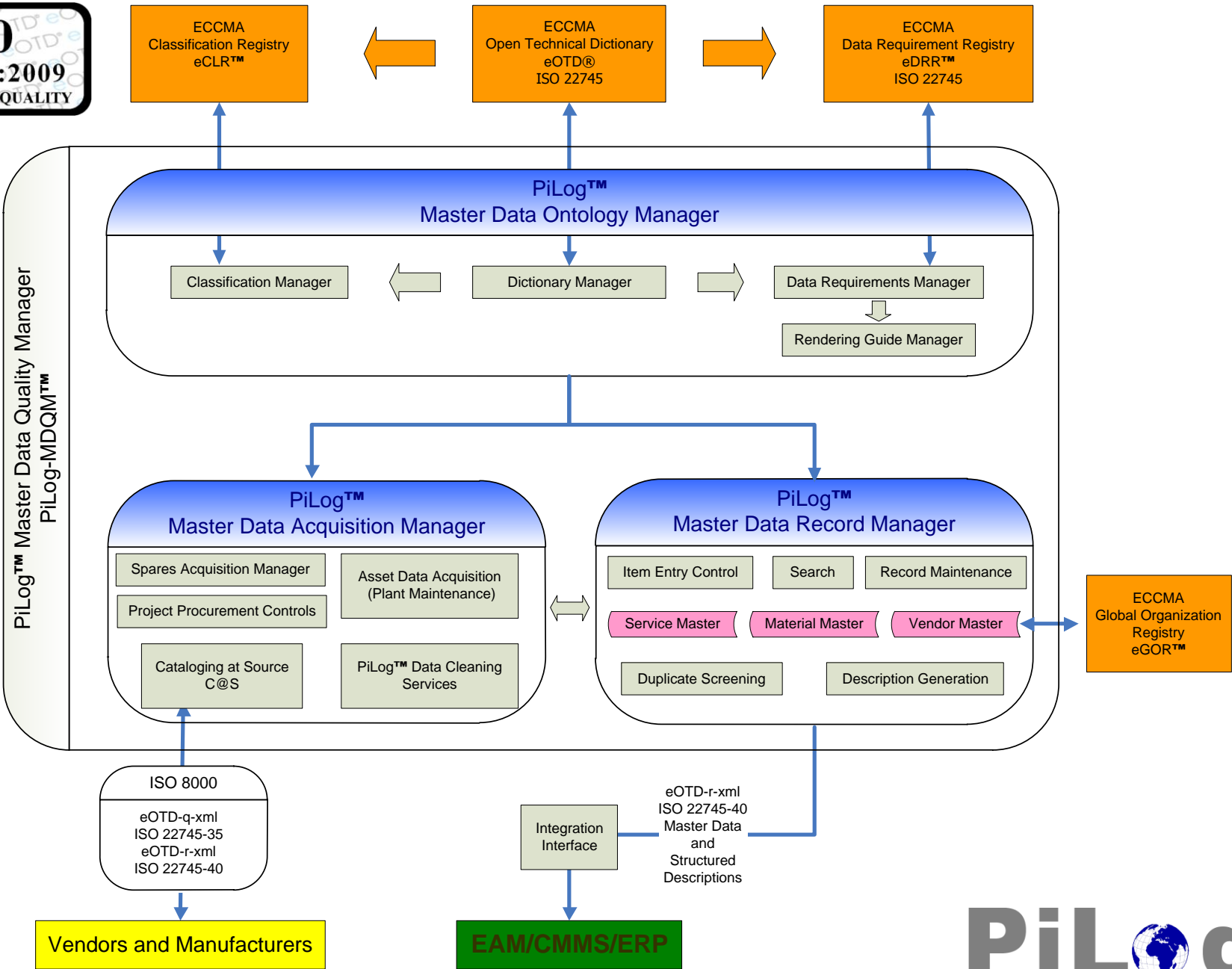
© Copyright Protected

What is an Ontology?

In practice, the name ontology covers a spectrum of useful artifacts, from formal upper-level ontologies expressed in first order logic...to the simple lists of user-defined keywords.... In between the two extremities of the ontology spectrum are taxonomies and controlled vocabularies...often used for information indexing and retrieval, and whose organization is mostly hierarchical. Finally, there are ontologies which represent not only subsumption , but also other kinds of relationships among entities (e.g., functional, physical), often based on formalisms such as frames or description logics.¹

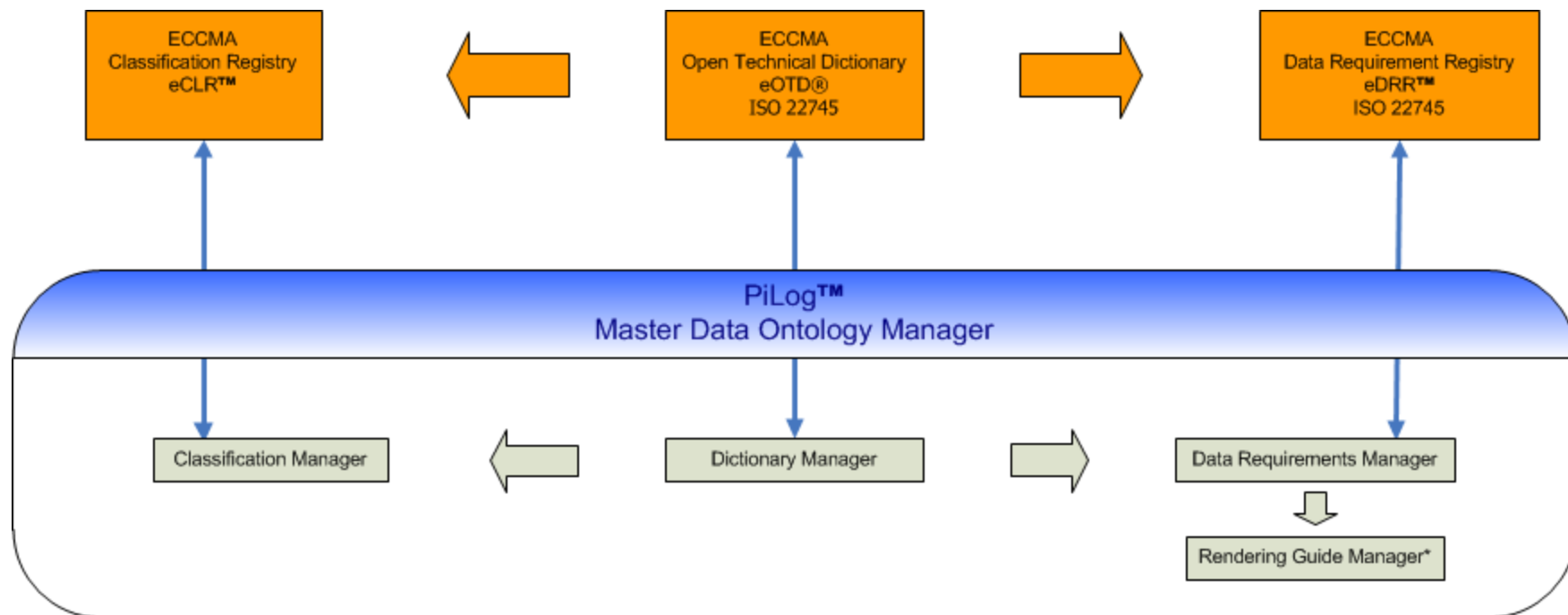
¹OntologySummit2007 Communiqué , version 1.0.0, 2007-04-24

http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2007_Communique



<http://www.pilogint.com/>

Commercial Off The Shelf (COTS) Master Data Ontology Manager integrated with the ECCMA registries



Available under a Software as a Service (SaaS)

- Create corporate subsets of the eOTD
- Create and manage data requirement statements (Identification Guides)
- Create and modify classifications

* Rendering Guide Manager available as an optional extra

- Ontology Manager
- OM1, Dictionary Manager
 - Concepts
 - Terms
 - Definitions
- Dictionary Manager
- OM2, Data Requirement Manager
 - Data Requirement Manager
- OM3, Classification Manager
 - Classification Manager

eOTD Terminology - Full eOTD Terminology - PiLog Preferred eOTD Terminology - Corporate Preferred

Concept Type	Concept ID	Term	Term ID	Originating Organization	Originating Reference
Class	0161-1#01-000009#1	ADAPTER,RACK,STORAGE	0161-1#TM-001135#1	DLIS	FIG=A23900;INC=51472
Class	0161-1#01-000010#1	ADAPTER,ARMAMENT,POD,AIRCRAFT	0161-1#TM-001136#1	DLIS	FIG=A23900;INC=61516
Class	0161-1#01-000011#1	ADAPTER,GUN,MOUNTING	0161-1#TM-001137#1	DLIS	FIG=A23900;INC=50745
Class	0161-1#01-000012#1	ADAPTER,MUNITION CHUTE	0161-1#TM-001138#1	DLIS	FIG=A23900;INC=62353
Class	0161-1#01-000016#1	ADAPTER,GUN,GRIIP ASSEMBLY,CONTROLLER,WEAPON	0161-1#TM-001142#1	DLIS	FIG=A23900;INC=41303
Class	0161-1#01-000019#1	ADAPTER,MUNITION STOWAGE RACK,AIRCRAFT	0161-1#TM-001145#1	DLIS	FIG=A23900;INC=52963
Class	0161-1#01-000020#1	ANTENNA FLANGE,CAMOUFLAGE SCREEN	0161-1#TM-001146#1	DLIS	FIG=A23900;INC=53286
Class	0161-1#01-000021#1	APERTURE,SIGHT	0161-1#TM-001147#1	DLIS	FIG=T102-E;INC=22224
Class	0161-1#01-000022#1	ARM,HANDLE,BREECHBLOCK	0161-1#TM-001148#1	DLIS	FIG=A23900;INC=62405
Class	0161-1#01-000027#1	CARTRIDGE,12.7 MILLIMETER BLANK	0161-1#TM-002402#1	DLIS	FIG=T007-C;INC=33370
Class	0161-1#01-000027#1	CARTRIDGE,27 MILLIMETER	0161-1#TM-002403#1	DLIS	FIG=T007-H;INC=33246
Class	0161-1#01-000027#1	CARTRIDGE,28 GAGE SHOTGUN	0161-1#TM-002404#1	DLIS	FIG=T007-D;INC=32935
Class	0161-1#01-000027#1	CARTRIDGE,9 MILLIMETER PRACTICE	0161-1#TM-002406#1	DLIS	FIG=T007-D;INC=35399

Queried Records: 1711381

Documentation

Definition	Definition ID	Originating Organization	Originating Reference
...ous designs with provisions to facilitate connection of a gun or guns to a mounting to obtain a desiredporate gun-firing, recoil-shock, absorbing, or vibration dampener mechanisms.	0161-1#DF-000011#1	DLIS	FIG=A23900;INC=50745
...ous designs with provisions to facilitate connection of a gun or guns to a mounting to obtain a desiredporate gun-firing, recoil-shock, absorbing, or vibration dampener mechanisms.	0161-1#DF-412026#1	DLIS	FIG=A23900;INC=50745

Select from eOTD Registry

Select from PiLog Preferred Registry

Add to Corporate Preferred Registry

**PiLog™ MDOM
Master Data Ontology Manager
Dictionary Manager**

Add Selected Item to Corporate Preferred List

- Ontology Manager
- OM1, Dictionary Manager
 - Concepts
 - Terms
 - Definitions
- Dictionary Manager
- OM2, Data Requirement Manager
- Data Requirement Manager
- OM3, Classification Manager

- eDRR Registry - Full
- eDRR Registry - PiLog Preferred
- eDRR Registry - Corporate Preferred

Data Requirement Registry

Data Requirement Name	Class Name	Data Requirement Ref	Submitted By	Status	Submission Date	Schema
SELF-LOCKING,HEXAGON	NUT,SELF-LOCKING,HEXAGON	0161-1#IG-009235#1	SSC2	APPROVED	2009-10-05:00:...	i-xml version 1.0
HOLDER,KEY	HOLDER,KEY	0161-1#IG-007486#1	SSC2	APPROVED	2009-05-05:00:...	i-xml version 1.0
BOLT,MASSIVE...	BOLT,MASSIVE...	0161-1#IG-006044#1	SSC2	APPROVED	2009-04-07:00:...	i-xml version 1.0

Select from eOTD Registry

Select from PiLog Preferred Registry

Add to Corporate Preferred

View Data Requirement and edit

Display i-xml

DR Viewer

WINCH,DRUM,VEHICLE MOUNTING (0161-1#IG-000203#1)

XML View Validate Setup

Class: WINCH,DRUM,VEHICLE MOUNTING

- Property: ITEM NAME
- Property: OPERATION METHOD
- Property: PRIME MOVER TYPE
 - Data Type: Controlled Value
 - Value: DIESEL ENGINE
 - Value: ELECTRIC MOTOR
 - Value: GASOLINE ENGINE
 - Value: HYDRAULIC MOTOR
 - Value: PNEUMATIC MOTOR
 - Value: STEAM ENGINE
- Property: PRIME MOVER QUANTITY
- Property: CURRENT TYPE
- Property: VOLTAGE IN VOLTS
- Property: FREQUENCY IN HERTZ
- Property: PHASE
- Property: DUTY CYCLE
- Property: INCLOSURE TYPE

Attributes of Data Type: Controlled Value

Attribute Name	Attribute Value
Controlled Value Type Name (eOTD Ref)	0161-1#04-000006#1

Selection Hierarchy

Data Requirement Specification:

Class: WINCH,DRUM,VEHICLE MOUNTING

Property: PRIME MOVER TYPE

Data Type: Controlled Value

PiLog™ MDOM
Master Data Ontology Manager
Data Requirement Manager

Record: 9 of 10 (Queried Records: 10)

Add Data Requirement to Corporate Registry View Data Requirement

- Ontology Manager
 - OM1, Dictionary Manager
 - Concepts
 - Terms
 - Definitions
 - Dictionary Manager
 - OM2, Data Requirement Manager
 - Data Requirement Manager
 - OM3, Classification Manager
 - Classification Manager

Corporate Classification Structures

- Corporate Classification Structures
 - Fin 3
 - Financials 1
 - Financials 2
 - Spend Analysis
 - 0161-1#01-086022#1 - BEARING
 - 0161-1#01-086023#1 - BEARING, ASSEMBLY
 - 0161-1#01-086024#1 - BEARING, BALL
 - 0161-1#01-086025#1 - BEARING, CAM FOLLOWER
 - 0161-1#01-086029#1 - BEARING, PLAIN
 - 0161-1#01-086031#1 - BEARING, ROD END
 - 0161-1#01-086032#1 - BEARING, ROLLER
 - 0161-1#01-086036#1 - BEARING, THRUST
 - 0161-1#01-067473#1 - GASKET
 - 0161-1#01-014416#1 - GASKET AND PREFORMED PACKING SET
 - 0161-1#01-014417#1 - GASKET AND SEAL SET
 - 0161-1#01-087270#1 - GASKET SET
 - 0161-1#01-087271#1 - GASKET, CAMPROFILE
 - 0161-1#01-087272#1 - GASKET, CORRUGATED METAL
 - 0161-1#01-087273#1 - GASKET, JACKETED
 - 0161-1#01-087274#1 - GASKET, PRE CUT
 - 0161-1#01-087275#1 - GASKET, RING JOINT
 - 0161-1#01-087276#1 - GASKET, SPIRAL WOUND

Building Elements

eCLR - Full		eCLR - PiLog Preferred	
eOTD Classes - Full		eOTD Classes - PiLog Pref	
Concept ID	Term	Term ID	Origin
0161-1#01-087267#1	GAS, NITROGEN	0161-1#TM-1100153#1	Starno
0161-1#01-087268#1	GAS, OXYGEN	0161-1#TM-1100154#1	Starno
0161-1#01-087270#1	GASKET SET	0161-1#TM-1100156#1	Starno
0161-1#01-087271#1	GASKET, CAMPROFILE	0161-1#TM-1100157#1	Starno
0161-1#01-087272#1	GASKET, CORRUGATED METAL	0161-1#TM-1100158#1	Starno
0161-1#01-087273#1	GASKET, JACKETED	0161-1#TM-1100159#1	Starno
0161-1#01-087275#1	GASKET, RING JOINT	0161-1#TM-1100161#1	Starno
0161-1#01-087277#1	GASSIFIER	0161-1#TM-1100163#1	Starno
0161-1#01-087278#1	GATE	0161-1#TM-1100164#1	Starno
0161-1#01-087281#1	GATE, VALVE	0161-1#TM-1100167#1	Starno
0161-1#01-087282#1	GALIGE	0161-1#TM-1100168#1	Starno

Record: 1 of 4136

Definition

Definition

A printed, typed page with spaces for filling in data E.g. questionnaires, job applications, etc. and may have carbon paper attached for duplication.

Build a Spend Analysis Classification System using eOTD, PiLog Preferred or Corporate Classes

PiLog™ MDOM Master Data Ontology Manager Classification Manager

Language Example

PiLog Data Quality Solutions ECCMA
Electronic Commerce Code Management Association

PiLog - Master Data Ontology Manager

Username : USER1 Logout

Home About Status Help Profile

Ontology Manager

- OM1, Dictionary Manager
 - Concepts
 - Terms
 - Definitions
 - Dictionary Manager**
- OM2, Data Requirement Manager
- OM3, Classification Manager

eOTD Terminology - Full | eOTD Terminology - PiLog Preferred | eOTD Terminology - Heidelberg Preferred

Disable paging | Page of 1

Language	Concept Type	Concept ID	Term
English en_US	Class	0161-1#01-853020#1	ADHESIVE
English en_US	Class	0161-1#01-045277#1	ADHESIVE
English en_US	Class	0161-1#01-085775#1	ADHESIVE
English en_US	Class	0161-1#01-067206#1	ADHESIVE

Record: 2 of 4 (Queried Records: 4)

Language	Concept Type	Concept ID	Term
French fr_FR	Class	0161-1#01-045277#1	ADHESIF
Hungarian hu_HU	Class	0161-1#01-045277#1	RAGASZTÓ
Ukrainian uk_UA	Class	0161-1#01-045277#1	ADHESIVE
English en_US	Class	0161-1#01-045277#1	ADHESIVE
Romanian ro_RO	Class	0161-1#01-045277#1	ADEZIV
Bulgarian bg_BG	Class	0161-1#01-045277#1	ADHESIVE
Czech cs_CZ	Class	0161-1#01-045277#1	LEPIDLO
German de_DE	Class	0161-1#01-045277#1	KLEBSTOFF
Spanish es_ES	Class	0161-1#01-045277#1	PEGAMENTO
Polish pl_PL	Class	0161-1#01-045277#1	KLEJ
Dutch nl_NL	Class	0161-1#01-045277#1	KLEEFMIDDEL
Italian it_IT	Class	0161-1#01-045277#1	ADESIVO
Korean ko_KR	Class	0161-1#01-045277#1	접착제
Slovenian sl_SI	Class	0161-1#01-045277#1	ADHESIVE
Slovak sk_SK	Class	0161-1#01-045277#1	LEPIDLO
Finnish fi_FI	Class	0161-1#01-045277#1	LIIMA-AINE
Greek el_GR	Class	0161-1#01-045277#1	ADHESIVE

DR Viewer (0161-1#MG-009511#1)

XML View | Validate | Setup

Attributes of Class: ADHESIVE

Attribute Name	Attribute Value
Class Name (eOTD Class ...)	0161-1#01-045277#1

Data Requirement Specification:

- Class: ADHESIVE
- Property: PHYSICAL FORM - [M]
- Property: APPLICATION TYPE - [M]
- Property: COLOR - [M]
- Property: CURING METHOD - [M]
- Property: HEAT VULCANIZABLE FEATURE - [M]
- Property: RATED DRYING TIME
- Property: SPECIFIC USAGE DESIGN
- Property: SEPARATE CATALYST
- Property: TEMPERATURE RATING
- Property: UNIT PACKAGE TYPE
- Property: QUANTITY WITHIN EACH UNIT PACKAGE
- Property: FURNISHED ITEMS
- Property: SPECIFICATION/STANDARD DATA

Selection Hierarchy

Data Requirement Specification:

Class: ADHESIVE

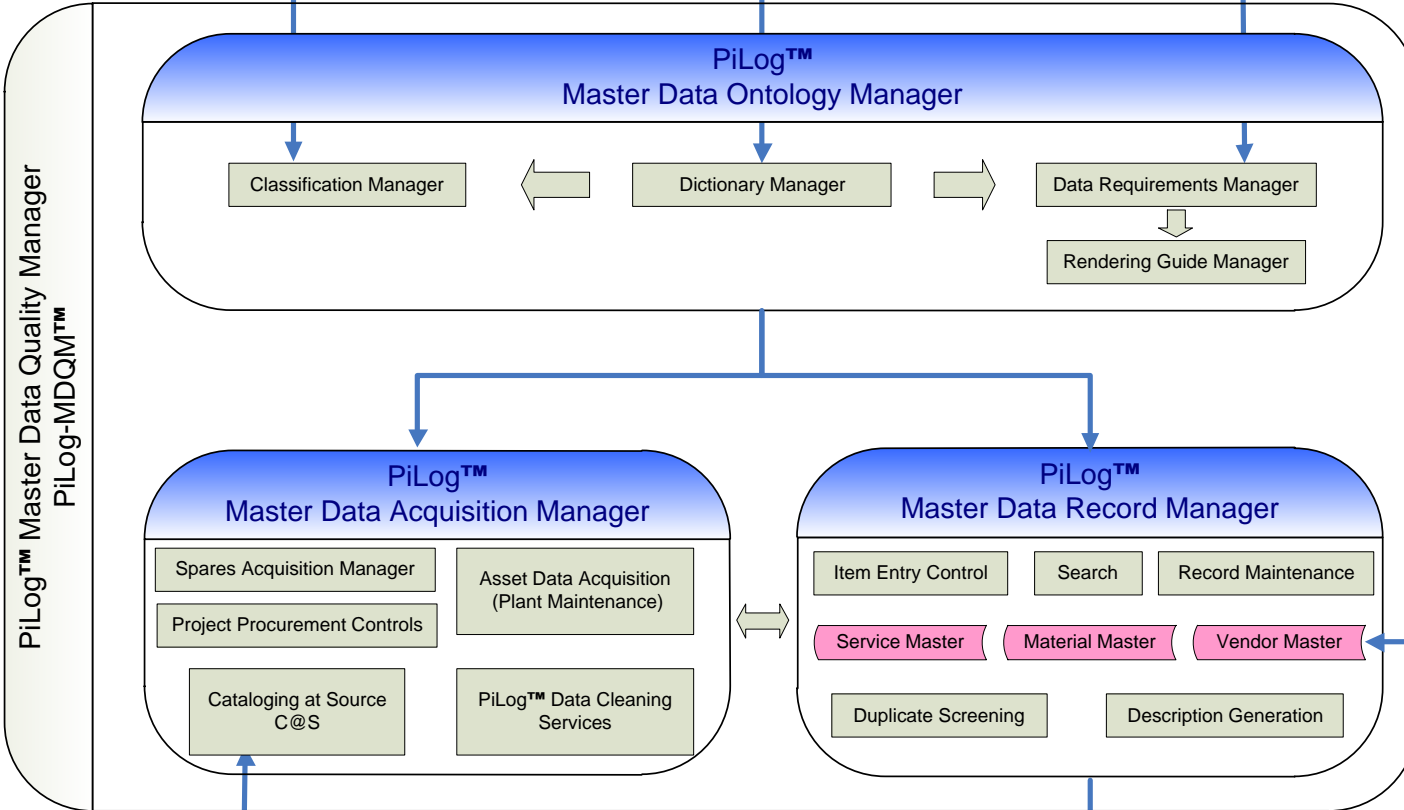
0161-1#TM-404573#1	NCB Poland	FIIG=A535P0;INC=11297
0161-1#TM-455746#1	NCB Netherlands	FIIG=A535P0;INC=11297
0161-1#TM-534194#1	NCB Italy	FIIG=A535P0;INC=11297
0161-1#TM-578306#1	NCB Korea	FIIG=A535P0;INC=11297
0161-1#TM-910206#1	NCB SLOVENIA	FIIG=A535P0;INC=11297
0161-1#TM-864997#1	NCB SLOVAKIA	FIIG=A535P0;INC=11297
0161-1#TM-1046274#1	NCB FINLAND	FIIG=A535P0;INC=11297
0161-1#TM-1091629#1	NCB GREECE	FIIG=A535P0;INC=11297



ECCMA
Classification Registry
eCLR™

ECCMA
Open Technical Dictionary
eOTD®
ISO 22745

ECCMA
Data Requirement Registry
eDRR™
ISO 22745



ISO 8000
eOTD-q.xml
ISO 22745-35
eOTD-r.xml
ISO 22745-40

Vendors and Manufacturers

Integration
Interface

EAM/CMMS/ERP

eOTD-r.xml
ISO 22745-40
Master Data
and
Structured
Descriptions

ECCMA
Global Organization
Registry
eGOR™

