



10 April 2014

The Redesigned WSDOT Data Catalog

Andy Everett, Metadata Repository Librarian, Washington State DOT

The New WSDOT Data Catalog



Andy Everett
DOTS Librarian
Washington State Department of Transportation
April 10 2014



Agenda:

- **The Data Problem at WSDOT**
- **Goals for Data Management**
- **What is Metadata?**
- **What is a Metadata Repository?**
- **A Typical Problem**
- **WSDOT Metadata Repository Architecture**



The Data Problem

- Ambiguous
- Invisible
- Redundant
- Disparate
- Poorly Named
- Poorly or not Defined
- Poorly Structured
- Difficult to Find
- Unknown Quality



Data As A Corporate Asset

- Shift in the perception of data as a component of IT to the Business
- Data is no longer “owned” by IT
- Active Stewardship of the Data by the Business



Goals

Provide a tool for data management:

- Harvest the Physical Database Inventory (What data do we have?)
- Understand data—Discover the actual meaning of data. (What is it?)
- Find data—Discover the data that is available for a subject. (Where is it?)
- Find people—Discover who is the data steward is (Who is responsible?)

What is Metadata?

- *Metadata* are data that describe the content, quality, condition and other characteristics of data.
- Structured, descriptive information about a **resource**

What is a Metadata Repository?

- A metadata repository is a collection of information that documents databases. It tells what data we have, where they are located, what they mean, who collects the data, who is responsible for the data quality and how to get access to the data. A good metadata repository works like a card catalog system for a library—helping you to inventory, organize, research, and easily find your information assets and application components.

A Typical Problem

- A search for data elements that have “**EMPL**” and “**NAME**” as a part of the Element name returns a list of **66** separate data elements.
 - How many are duplicates of the same data?
 - Are they all storing the data in the same format?
- Within the **66** Elements there are **21** distinct variants on the element names.
 - Does “**CRSE-ORG-EMPL-NAME**” store the name of the Organization or the Employee?



Physical Data Inventory

Database

- Mainframe Database
- WSF_CORP_DV

File/Table/View

- DOT-EMPLOYEE-MASTER-FILE
- tbl_WSF_Employee

Element

- EMPL-NAME
- Empl_Name



Associating Business Concepts with Physical Data

Business Concept

Has one or more...

Business Terms

with at least 1...

Preferred Term

that may have at least one ...

Data Object

Data Object can represent a Database, table/view, Data Element



What Metadata is Captured

- **Terms, Synonyms and Definitions**
- **Data Stewards**
- **Business Rules**
- **Technical Metadata (about the database)**
- **Metadata for IT Portfolio Management**
- **Metadata for Data Security Classification**



Data Sets

- Data Set Defined:

A finite grouping of related data that has context based on a data variable such as date range or geographical extent.

- Example:

Geospatial Data Sets, Reference Data Sets, Data Warehouse Reports, 911 Calls during a time period



Data Sets and DOTS

DOTS does not currently contain descriptive Metadata Sets such as the North American Profile of ISO19115:2003

DOTS focus is describing the schema of the database, the definitions of the data and the business stewards of the business terms and data objects. What? Where? Who?

Enhancements are on the horizon to accommodate metadata application profiles and standards such as NAP to describe data sets that are either open and accessible, or limited access and secured.



Usability Changes

1. Developed Search with a set of Search rules.
 1. Begins With, Contains Ends With
 2. Use Synonyms in Search
 3. Search Definitions
2. Dropped display taxonomy
3. Developed a thesaurus management tool
 1. Broader Terms (parent)
 2. Narrower Term (child)
 3. Related Terms
4. Concept based with Terms related to Concept
 1. Allows for easy promotion of Synonym to Preferred Term



Usability Changes

1. Developed easy to navigate Detail pages
2. Left hand navigation for Data Objects
3. Improved Term to Data Object Association

Data Catalog

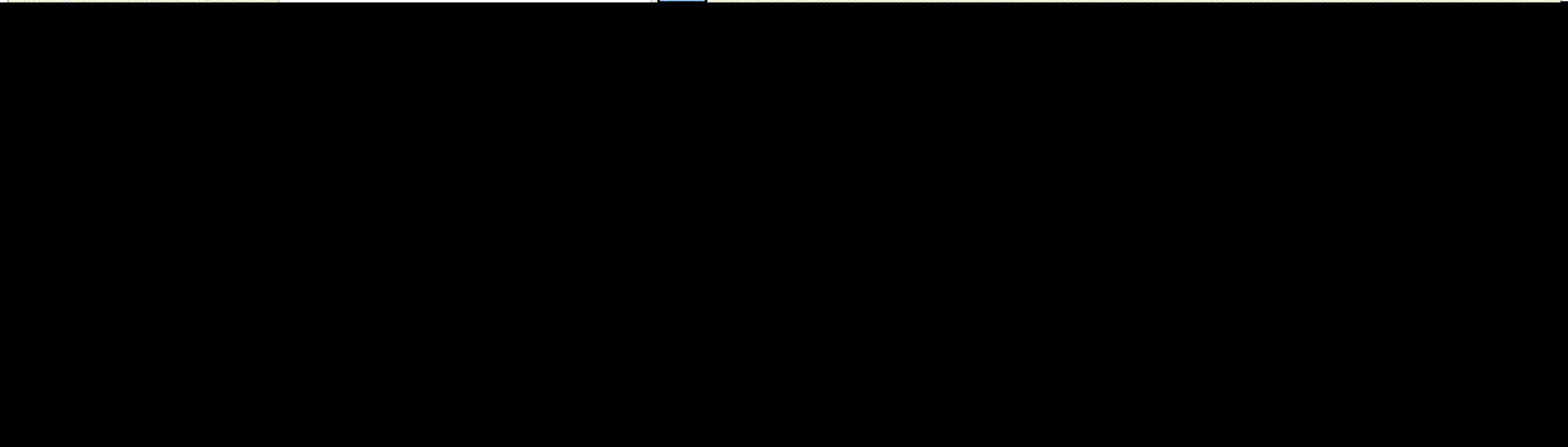


Basic Search

Advanced Search

Enter Keyword(s)

Find



DOTS

[Search](#) [Manage](#) [Reports](#) [Help](#) [Admin](#)



[?](#)

[Data](#) [Business Term](#) [Application](#)

Send Comments to: [DOTS Support](#)

Version: 6.0.0

DOTS

Search Manage Reports Help Admin



Collision 

Data **Business Term** **Application**

Showing items 1 through 100 of 2,695.

T Collision
Definition: An occurrence of injury or damage involving at least one Motor Vehicle or Pedalcycle for which a Police Traffic Collision Report (PTCR), Ve...
Business Area: WSDOT Agency Wide **Definition Status:** Approved

D Collision
Type: Table
An occurrence of injury or damage involving at least one Motor Vehicle or Pedalcycle for which a Police Traffic Collision Report (PTCR), Vehicle Collision Report (VCR) or a report via SECTOR is filed.
Full Path: HQOLYMSQLDS02\MSSQLSERVER.CollisionDataMart.dbo.Collision

D Collision
Type: Table
An occurrence of injury or damage involving at least one Motor Vehicle or Pedalcycle for which a Police Traffic Collision Report (PTCR), Vehicle Collision Report (VCR) or a report via SECTOR is filed.
Full Path: HQOLYMSQLDS02\MSSQLSERVER.CollisionDataMartFrozen.dbo.Collision

DOTS

T Collision [ADD](#) [EDIT](#)

Definition An occurrence of injury or damage involving at least one Motor Vehicle or Pedalcycle for wh (VCR) or a report via SECTOR is filed. [FOLLOW](#) [DELETE](#)

Business Area: WSDOT Agency Wide [EDIT](#) **Definition Status:** Approved [EDIT](#)

Preferred Term: Collision **Synonyms:** crash, wreck, accident, SECTOR [EDIT](#)

Abbreviations: None [EDIT](#) **Acronyms:** None [EDIT](#)

[+ Business Term Relationships](#)

[+ Additional Notes](#)

[+ Associated Data Objects \(18\)](#) [ADD](#)

[+ Attachments \(0\)](#) [ADD](#)

[+ Discussion \(0\)](#) [ADD](#)

[+ Stewards](#) [EDIT](#)

Send Comments to: DOTS Support

Version: 6.0.0

DOTS

Business Term Relationships

Parent: Highway Safety [EDIT](#)

Child: Alcohol Test Results, At Angle Collision Indicator, Badge Number, Boundary Trafficway, Cable Median Barrier Collision Indicator, Case Number, City Street Collision, CLAS Entry Date, Closed to Public Traffic Indicator, Collision 12 Hour Time, Collision Barrier Collision Indicator, Collision Count, Collision Date, Collision Diagram, Collision Factors, Collision Identifier, Collision Occupant, Collision Quality Review Type, Collision Record Status Type, Collision Report Source Type, Collision Report Type, Collision Severity Type, Collision Time 24 Hour, Collision Type, County Location Coding Form Sent Datetime, Damage Threshold Met Indicator, Diagram Data Collision Type, Direction Movement From, Direction Movement To, Driver Passenger Indicator, Evident Injury Collision Indicator, Government Owned Indicator, History Indicator, Hit and Run Indicator, Injury Count, Injury Type, Large Truck Collision Indicator, Lighting Condition Type, Location Character Type, Mainline Milepost, Milepost Ahead/Back Indicator, Milepost Direction Type, Motor Vehicle Alcohol Impaired Collision Indicator, Motor Vehicle Alcohol Related Collision Indicator, Motor Vehicle And Pedacyclist Collision Indicator, Motor Vehicle And Pedestrian Collision Indicator, Motor Vehicle Driver Cell Phone Distraction Collision Indicator, Multiple Motor Vehicles Collision Indicator, MV Speed And Or Alcohol Impaired Collision Ind, Non Fatal Injury Collision Indication, Non Motorized Vehicle Indicator, On Duty Emergency Worker Indicator, On Off Roadway Indicator, Passenger Involved Count, Pedalcycle, Pedalcycle Driver Action Type, Pedalcycle Involved Count, Pedalcyclist Action Type Description, Pedestrian, Pedestrian Pedalcycle Type, People Involved Count, Person Partially Totally Ejected Indicator, Police Arrival Time, Police Dispatch Time, Posted Speed, Private Trafficway Name, Property Owner Involved Count, Quality Assurance Last Action Date, Rear End Collision Indicator, Reference Point Miles/Feet Indicator, Roadway Character Type, Roadway Surface Condition Type, Roadway Surface Type Code, Serious Injury Collision Indicator, Standard Day Part, State Reportable Indicator, Stolen Vehicle Indicator, Total Possible Injuries, Traffic Control Device Type, Trafficway Reference Name, Tribal Reservation Code, Tribal Reservation Name, TRIPS Last Maintenance Datetime, TRIPS Report Number, Vehicles Involved Count, Weather Name, Work Zone Related Collision Indicator, Working Vehicle Ownership Code [EDIT](#) [DISPLAY LIST](#)

Related: Police Traffic Collision Report [EDIT](#)

DOTS

Associated Data Objects (18) ADD

REMOVE MULTIPLE

Showing items 1 through 18 of 18. 1

Name	Type	Full Path	Action
CDMCollisions	Table	HQOLYMSQLDS02\MSSQLSERVER.TARIS.dbo.CDMCollisions	REMOVE
Collision	Table	HQOLYMSQLDS02\MSSQLSERVER.CollisionDataMartFrozen.dbo.Collision	REMOVE
Collision	Table	HQOLYMSQLDSD\MSSQLSERVER.CollisionDataMart.dbo.Collision	REMOVE
Collision	Table	HQOLYMSQLDS02\MSSQLSERVER.CollisionDataMart.dbo.Collision	REMOVE
Collision	Table	HQOLYMSQLDS02\MSSQLSERVER.SafetyAnalyst2009.dbo.Collision	REMOVE
Collision	View	HQOLYMSQLDS02\MSSQLSERVER.RoadwayDataMart.dbo.Collision	REMOVE
COLLISIONS	Table	HQOLYMSDE06P\MSSQLSERVER.Collisions.dbo.COLLISIONS	REMOVE
Collisions	Table	HQOLYMSQLDS02\MSSQLSERVER.TARIS.dbo.Collisions	REMOVE
CollisionsFlatFile	Table	HQOLYMSQLDS02\MSSQLSERVER.CollisionDataMart.dbo.CollisionsFlatFile	REMOVE
CollisionsFlatFile	Table	HQOLYMSQLDS02\MSSQLSERVER.SafetyAnalyst2009.dbo.CollisionsFlatFile	REMOVE
CollisionsFlatFile	Table	HQOLYMSQLDS02\MSSQLSERVER.StagingSafetyAnalyst.dbo.CollisionsFlatFile	REMOVE
CollisionsFlatFile	Table	HQOLYMSQLDS02\MSSQLSERVER.CollisionDataMartFrozen.dbo.CollisionsFlatFile	REMOVE
CollisionsFlatFileView	Table	HQOLYMSQLDS02\MSSQLSERVER.StagingSafetyAnalyst.dbo.CollisionsFlatFileView	REMOVE
CollisionsFlatFileView	Table	HQOLYMSQLDS02\MSSQLSERVER.CollisionDataMartFrozen.dbo.CollisionsFlatFileView	REMOVE
CollisionsFlatFileView	Table	HQOLYMSQLDS02\MSSQLSERVER.CollisionDataMart.dbo.CollisionsFlatFileView	REMOVE
CollisionsFlatFileView	Table	HQOLYMSQLDS02\MSSQLSERVER.SafetyAnalyst2009.dbo.CollisionsFlatFileView	REMOVE
CollisionsGIS500K2008	Table	HQOLYMSQLDS02\MSSQLSERVER.TARIS.dbo.CollisionsGIS500K2008	REMOVE
CollisionsGIS500KTest	Table	HQOLYMSQLDSD\MSSQLSERVER.TARIS.dbo.CollisionsGIS500KTest	REMOVE

DOTS

Stewards [EDIT](#)

Planning - Collision Data and Reporting

Activity	Name	Phone	Email
Executive Steward	Finch, Mark	360-570-2369	FinchM@wsdot.wa.gov
Managerial Steward	Jobe, Nadine	360-570-2398	JobeN@wsdot.wa.gov
Operational Steward	Bernard, Mike	360-570-2454	BernarM@wsdot.wa.gov
Operational Steward	Stanley, Warren	360-570-2497	StanleW@wsdot.wa.gov

DOTS

D Collision

Advanced Search

Host/Server:

Keyword:

Find it!

Clear Filter

Any Lifecycle: Yes No Any Platform: Yes No Any Group: Yes No No Assigned Term Definition

SQLSERVER

CollisionDataMart

dbo

+ BASECollisionMainlineFatalInc

+ CityCounty

+ CityCountyPopulation

+ CitySeg

+ CityStreet

+ Collision

+ CollisionDiagram

+ CollisionEndOfYear2004

+ CollisionEndOfYear2005

+ CollisionEndOfYearAADT

+ CollisionFactors

+ CollisionOccupant

+ CollisionReportType

+ Collisions2004

+ Collisions2005

+ CollisionsFlatFile

+ CollisionsFlatFileView

+ CollisionsGIS500K

+ CollisionSummary

+ CommercialCarrier

Save

Expand All

Collapse All

Add Column

Table Information

Name: Collision

Schema Name: dbo

Business Term: Collision [EDIT](#)

Definition: An occurrence of injury or damage involving at least one Motor Vehicle or Pedalcycle for which a Police Traffic Collision Report (PTCR), Vehicle Collision Report (VCR) or a report via SECTOR is filed.

Contains State Data Classification Level: Category 1 - Public

Data Category:

Requires Special Handling: No **Data Sharing Agreement:** No

Notes:

+ **Business Stewards**

+ **Technical Stewards (0)** [EDIT](#)

+ **Special Handling**

+ **Administrative Information**

+ **Business Intelligence Metadata**

+ **Attachments (0)** [ADD](#)

+ **Update Information**

Save

DOTS

Business Stewards

Planning - Collision Data and Reporting

Activity	Name	Phone	Email
Executive Steward	Finch, Mark	360-570-2369	FinchM@wsdot.wa.gov
Managerial Steward	Jobe, Nadine	360-570-2398	JobeN@wsdot.wa.gov
Operational Steward	Bernard, Mike	360-570-2454	BernarM@wsdot.wa.gov
Operational Steward	Stanley, Warren	360-570-2497	StanleW@wsdot.wa.gov

Technical Stewards (1) [EDIT](#)

Business Area	Activity	Steward	Phone	Email	Group
WSDOT Agency Wide	Database Admin				Data Resource Management

Special Handling

More Information

Andy Everett andy.everett@wsdot.wa.gov

Data or Term Search Librarian

Member TRB LIST Subcommittee on Transportation Research Thesaurus

