



# An Introduction to Data Catalogs

# Josh Smith

Database Administrator

Friend of Redgate

SQLSaturday Spokane Organizer

Spokane SQL Users Group Leader

[josh@accitentionaldba.com](mailto:josh@accitentionaldba.com)

[@sqldeployhelmet](https://twitter.com/sqldeployhelmet)

[www.linkedin.com/in/joshsmithdba/](https://www.linkedin.com/in/joshsmithdba/)



This developer refuses to stop  
using nolock hints.

# Agenda

- ▶ Why Data Catalog?
- ▶ Who Data Catalogs?
- ▶ Where Data Catalog?
- ▶ How Data Catalog?
- ▶ Catalog Demo!
- ▶ Questions?

# Voyage of Discovery

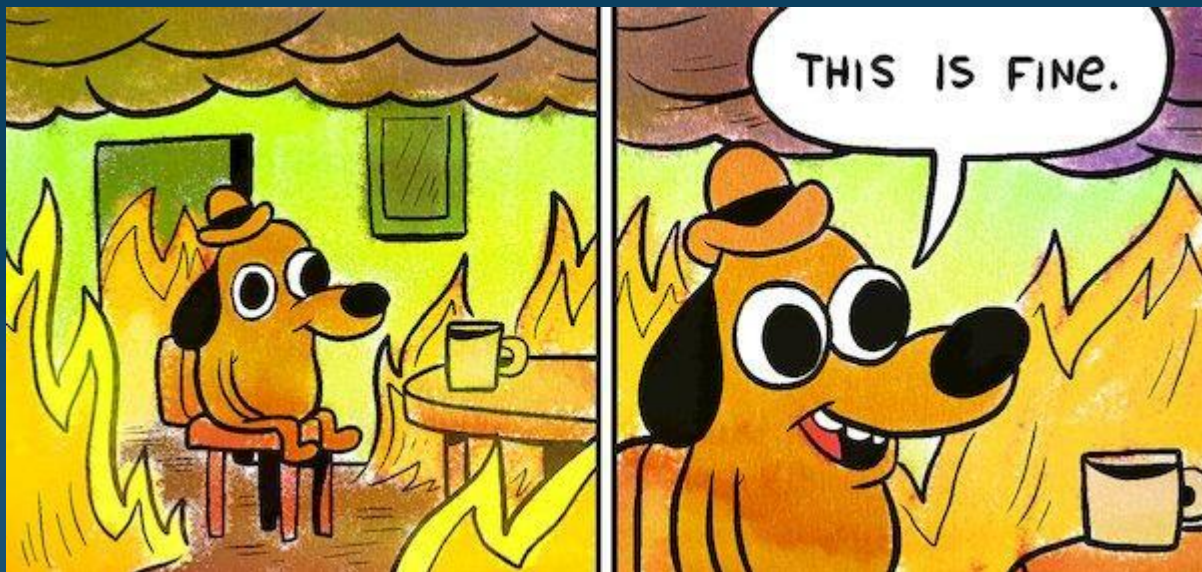
Where does your data live?

- ▶ Data that can't be found has no value.
- ▶ Susan in the back office has a wealth of knowledge about your core system but:
  - ▶ She's only one person,
  - ▶ Helping the data analysts is a low priority,
  - ▶ She's 2 years out from retirement.





# Data Breaches Are Just A Thing Now



- ▶ 14,717,618,286 Records lost or stolen since 2013<sup>1</sup>
- ▶ Only 4% of breaches were considered “secure” where the data stolen was encrypted.
- ▶ In 2018 only half of these were caused by an outside hack.
- ▶ 1/3 were the result of accidental loss or exposure.

# Let's Have More Meetings!



Increased focus on Data Governance has been the response to new legislation, the increased pace of data breaches as well as internal demands to sensitive data.

# Registered Assets

## Azure Data Catalog\*

- ▶ SQL Server (table or view)
  - ▶ Structural meta data (table, schema)
  - ▶ Type, row count, size, recent update, more.
- ▶ Cosmos DB Collection
- ▶ Many more:  
<https://docs.microsoft.com/en-us/azure/data-catalog/data-catalog-dsr>




## SQL Data Catalog (Redgate)

- ▶ SQL Server (instance, database, columns)
  - ▶ Structural meta data (table, schema)
  - ▶ Type, row count

# Death & Taxonomies

Azure Data Catalog/SQL Data Catalog allow you to create a collection of metadata tags to apply to your data.

- ▶ Information Type
- ▶ Data Sensitivity level
- ▶ Business terms
- ▶ Treatment intent
- ▶ Environment
- ▶ SLA
- ▶ Data Stewards
- ▶ Free text

Classification tags			
Categories <a href="#">+ Add category</a>	Tags	Multi-valued	Apply to
Environment	Development Test Production <a href="#">+</a>	<span>×</span>	
Data Residency	United States Europe Canada Brazil Asia Pacific Australia China India Japan Korea South Africa United Arab Emirates <a href="#">+</a>	<span>×</span>	
Applicable Regulation	GDPR HIPAA CaCPA PCI DSS SOX DPA Washington Privacy Act <a href="#">+</a>	<span>✓</span>	
Treatment Intent	Static Masking Encryption Dynamic data masking Row level security <a href="#">+</a>	<span>×</span>	





# Demo Time!

IMMA LET YOU FINISH...

# Questions?

A data catalog should start to answer common questions users have about your data:



**Jeff Albro, CDMP**  
@BeantownData

The WTF Guide to Metadata:

Data Dictionary - What TF is this data?

Data Catalog - Where TF is my data?

Data Lineage - Where TF did this data come from?

Data Steward - Who TF knows about this data?

Business Glossary - What TF should we call this data?

11:46 AM · May 11, 2019 · [Twitter for Android](#)