RSAConference2022
San Francisco & Digital | June 6 – 9

TRANSFORM

SESSION ID: **PDSC-W01**

DLP: An Implementation Story

Micah K Brown

Greater Cincinnati ISSA
@MicahKBrown

Disclaimer

Presentations are intended for educational purposes only and do not replace independent professional judgment. Statements of fact and opinions expressed are those of the presenters individually and, unless expressly stated to the contrary, are not the opinion or position of RSA Conference LLC or any other co-sponsors. RSA Conference does not endorse or approve, and assumes no responsibility for, the content, accuracy or completeness of the information presented.

Attendees should note that sessions may be audio- or video-recorded and may be published in various media, including print, audio and video formats without further notice. The presentation template and any media capture are subject to copyright protection.

Agenda:

- What is DLP
- DLP: The Cast
- DLP: Getting Started
- DLP: Architecture
- DLP: How to detect and bypass a DLP Solution
- DLP: Wrapping it up and Q&A

RSAConference2022

A DLP Story:

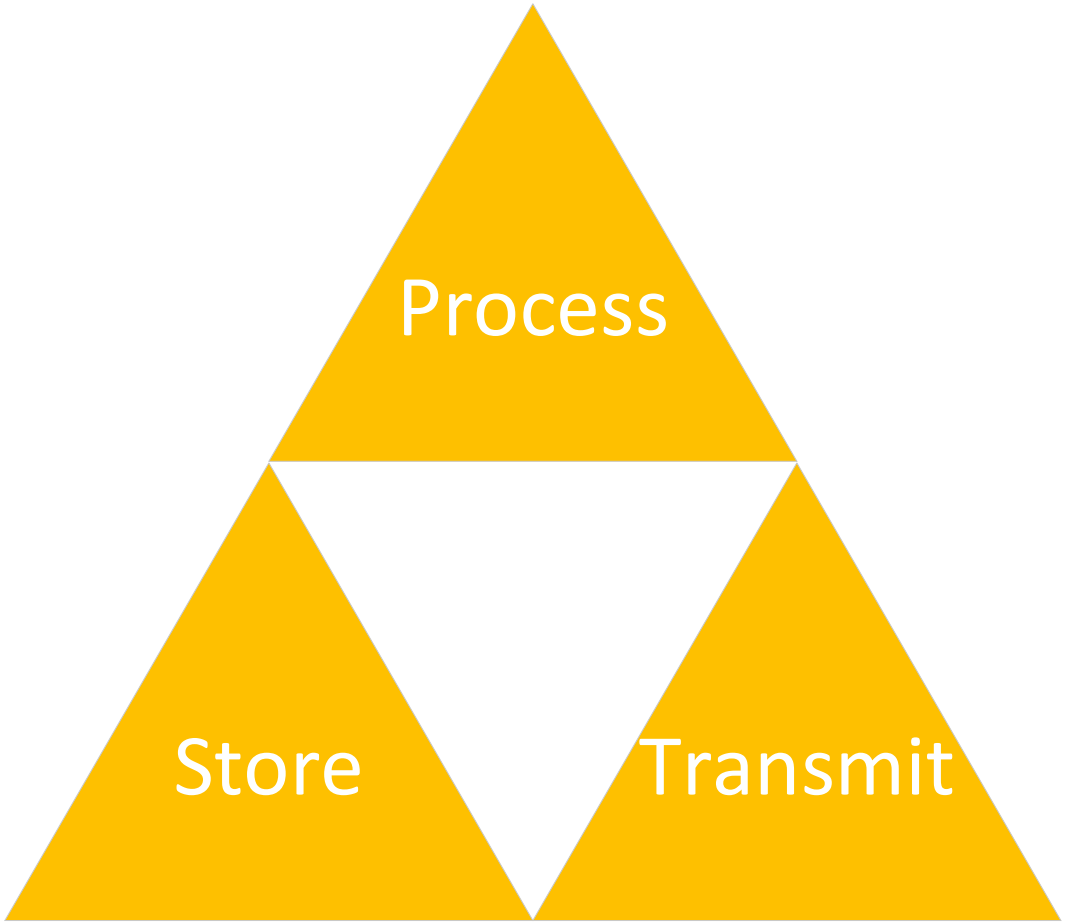
An Introduction

What is Data Loss Prevention (DLP)?

- **DLP is NOT . . .** a 'monolithic' stack of technology by one single vendor
- **DLP is NOT . . .** designed to stop a sufficiently advanced advisory [internal | external] by itself
- **DLP is NOT . . .** 'Digital Rights Management'
- **DLP is NOT . . .** 'User Behavioral Analytics'
- **DLP is NOT . . .** a 'Silver Bullet'

- **DLP IS** becoming a 'feature' in many different tools
- **DLP IS** designed to help establish proper data ownership / stewardship
- **DLP IS** more aligned with Business Risk Management than IT Security
- **DLP IS** a collection of tools that allows the business to define policies on how data is Stored, or Processed, the environment

What is DLP?



The State of Data in our Environment

DLP is made up of many different tools that work together. Each tool will inspect / interact with data in at least one or more of these states of data.

RSAConference2022

A DLP Story:

The Cast

The Cast:

- **DLP Client** – Software on the client system / server that monitors the actions of the user.

- **DLP Network** – A machine that inspects network traffic for potentially sensitive data as it traverses the network. Best to place at ‘choke zones’ such as internet egress or sensitive network.

- **DLP Repository Scanner** – A machine that scans data repositories (shares, SharePoint sites, databases, and more) for potentially sensitive information.

S Store P Process T Transmit

The Cast:

- **DLP Email –** A machine that partners with your email system to scan outgoing and / or incoming email for sensitive information.

- **DLP Web –** A machine that partners with your internet proxy to scan outgoing and / or incoming web requests for sensitive information.

- **DLP Management console –** A machine that manages all of the other DLP components. This also includes reporting, incident management, and evidence management.

S Store P Process T Transmit

DLP Concerns: 1 of 2

- **DLP Client can overlap functionality with other DLP Components**
  - DLP client can have high performance hit (CPU / Memory / Disk I/O).
  - You can run into incompatibles that cause other software not to work.
  - Incompatibilities with other applications can spike CPU / Memory/ Disk I/O.
  - Distribution of Client and Policy.

- **Guiding Principal:**
  - **Keep the DLP Client Policy as simple and lightweight as possible.**
  - **Lift as much functionality to other DLP components.**

DLP Concerns: 2 of 2

- **DLP Management console 'evidence' of violations of policy which could contain data governed by rules / laws / regulations / contractual agreements.**
- **Many different products are adding on "DLP" functionality. Interoperability might become a challenge between diverse systems.**
- **DLP Network can overlap with DLP Email and DLP Web.**
  - Prefer to use the 'best fit' DLP Tool for use case.

RSAConference2022

A DLP Story:

Getting Started

DLP: Setting Scope

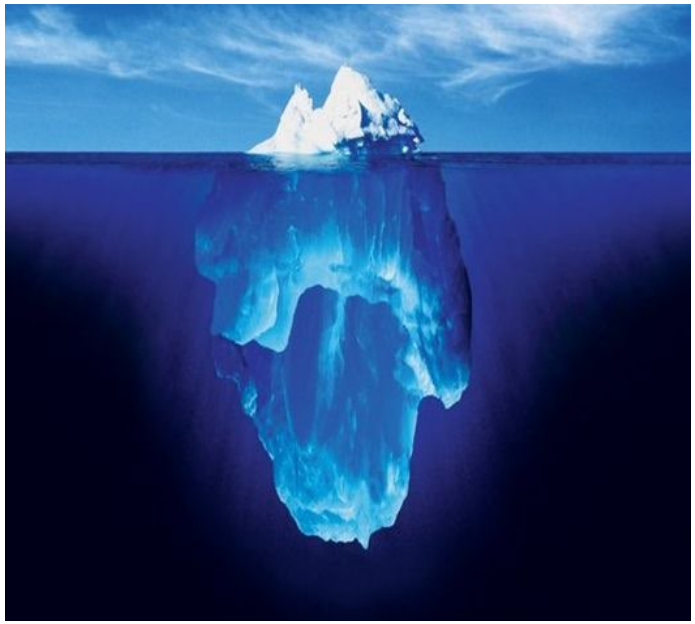
- **DLP has become a very popular concept at the Executive Level to protect data and prevent breaches.**
  - Clearly define what DLP is, what it can do, and what it can not do.
  - Clearly define what features / tools to turn on and what they do.
  - Is DLP part of the organization's "active defense toolset" or historical system of record?
  - Bring in Legal, HR, Compliance, and other appropriate business groups to understand any implications.
  - Define who responds to DLP incidents.
  - Define how DLP Incidents are handled.

DLP: The High Cost of fractured Policies

- **We implemented a three policy environment Test, Pilot, Production. This lead to some complications and lessons.**
  - We adopted the this architecture after a single REG-EX logic error was introduced into the client config that turned all the test systems into literal potatoes.
  - Each time you fragment your policy, it has HUGE implications for support, maintenance, testing, documentation.
  - You need to find an appropriate balance between a monolithic policy and micro fragmentation of your DLP policy. (This is challenging!)

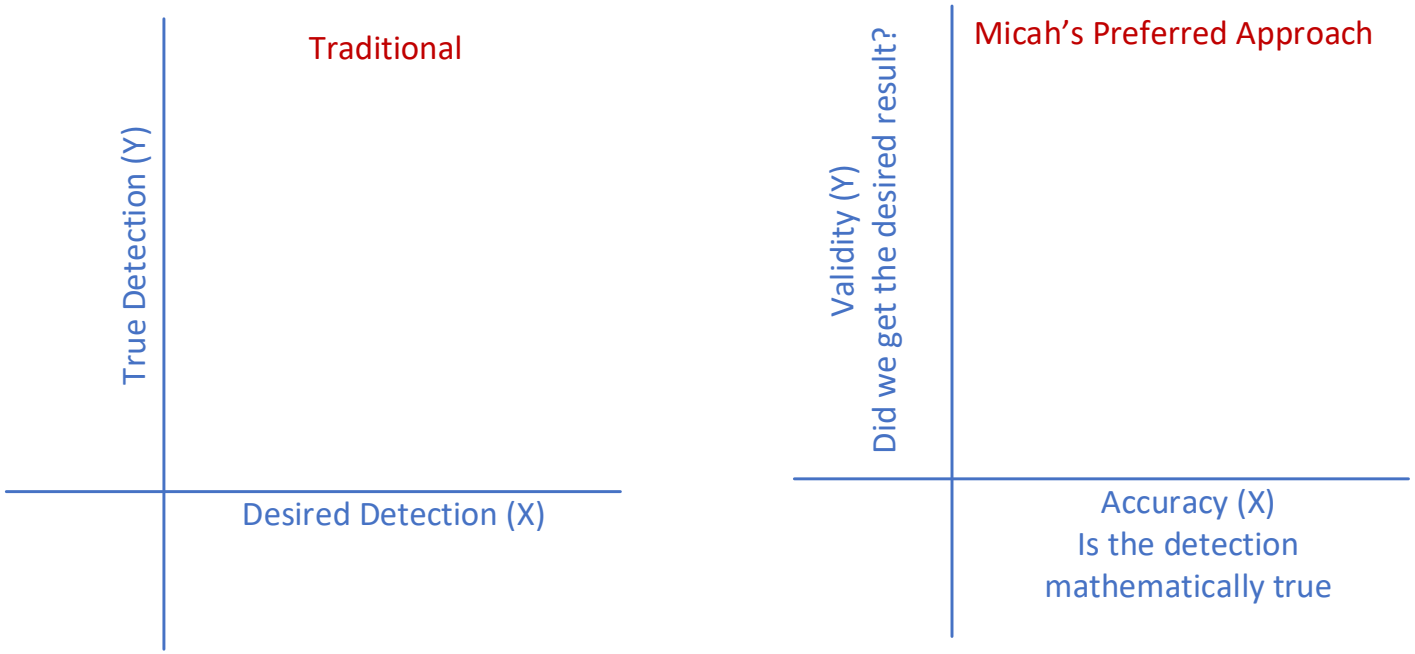
DLP: Tuning

- **Tuning is one of the hardest parts of a DLP Project**
  - Tuning involves both Functionality and Data Classification
    - If a Function adds more work that value, consider turning it off.
    - When tuning Data Classifications, ask how could this change be abused or if it introduces a chance for a **FALSE NEGITIVE**
  - The firs 80% of tuning is relatively easy and low risk.
  - We will discuss the challenges I had with US Driver License shortly.



There are no FALSE POSITIVES, Only poorly written rules*

On Tuning



- **Measuring data is not an exact science**
- **Traditional ([True | False], [Positive | Negative]) vs (Validity, Accuracy)**
- **Qualitative vs Quantitative**
- **@InfoSecBenjaminDisraeli: "lies, damned lies, and spreadsheets"**

RSAConference2022

A DLP Story:

DLP Architecture

DLP: Traditional DPL Policies

- **Traditional DLP is built upon REG-EX and word dictionaries. This can be problematic:**
  - 50 US states each have multiple definitions of driver's license.
  - Many websites use 15/16 digit numeric codes to reference content. This overlaps with traditional credit card definitions.
  - REG-EX and word dictionaries are not sufficient optimal.
- **When possible follow the main rule of improve "YES AND" in building traditional classification.**
- **"Proximity" rules are your friend!**

DLP: Object Orientated Rules

CLTv11DEF - Payment Card Industry Compliance	CREDIT-CARD-NUMBER DLPv9	Proximity	Patterns: Credit Card Number (Mastercard),Credit Card Number (China UnionPay),Credit Card Number (JCB),Credit Card Number (Visa),Credit Card Number (Diner's Club),Credit Card Number (Simple, dash delimited),Credit Card Number (Discover),Credit Card Number (American Express),Multiple Common PCI Cards and Dictionaries: CLTv11DEF - Credit-Report-TEXT is less than 90 characters and found at least 1 times
CLTv11DEF - Payment Card Industry Compliance	MAGSTRIPE-TRACK-NUMBER DLPv9	Advanced Pattern	CLTv11DEF - MAGSTRIPE-TRACK-NUMBER-EXPR
CLTv11DEF - Payment Card Industry Compliance	PCI multiple cards advanced pattern	Proximity	Patterns: Credit Card Number (Mastercard),Credit Card Number (China UnionPay),Credit Card Number (JCB),Credit Card Number (Visa),Credit Card Number (Diner's Club),Credit Card Number (Simple, dash delimited),Credit Card Number (Discover),Credit Card Number (American Express),Multiple Common PCI Cards and Dictionaries: CLTv11DEF - PCI GLBA-TEXT is less than 90 characters and found at least 1 times
CLTv11DEF - US-PII-Violations	DRIVERS-LICENSE-EXPR	Advanced Pattern	CLTv11PLT - DRIVERS-LICENSE-EXPR
CLTv11DEF - US-PII-Violations	SSN and (DOB or First Name or Last Name)	Dictionary & Advanced Pattern	CLTv11DEF - Social Security Number-EXPR AND (Date Of Birth, First Name, Last Name)
CLTv11DEF - US-PII-Violations	CLTv11DEF - DRIVERS-LICENSE-EXEMPT (applied in each rule where we have an exception)	Proximity	Patterns: CLTv11DEF - DRIVERS-LICENSE-EXPR and Dictionaries: CLTv11DEF - DRIVERS-LICENSE-EXEMPT-TXT is less than 100 characters and found at least 1 times

Use standardized naming conventions to flag the type of each “building block” and what policy each “building block” belongs to. We did this by using a “short code” at the start of custom dictionaries, and a dictionary “type flag”.

DLP: Tyranny of the Drivers License 1:3

Driver's License Positive match

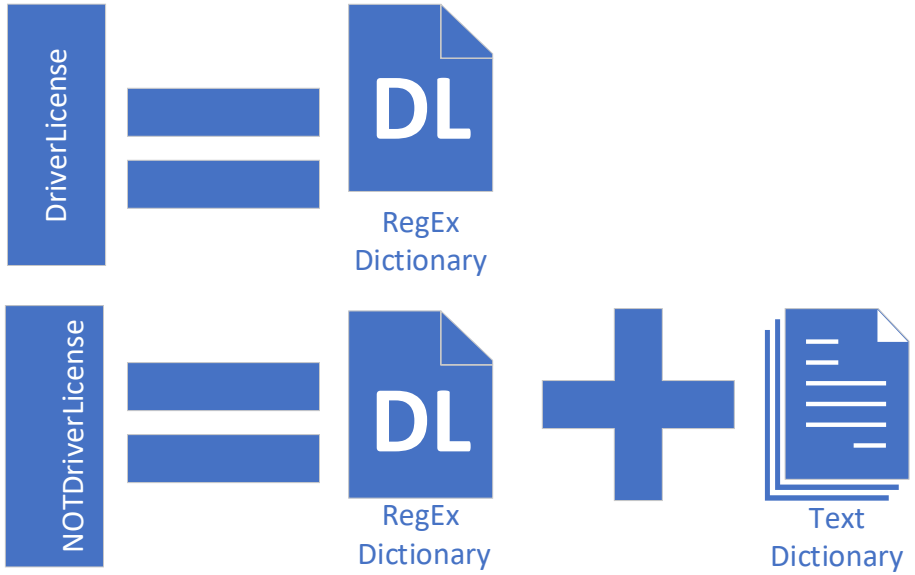
[.,-]00\d\d\d\d\d\d\d[.,-]
[.,-][xX]\d\d\d\d\d\d\d\d[.,-]
[.,-]\d\d\d\d\d\d\d[aA][.,-]
[.,-][rtRT]\d\d\d\d\d\d\d\d[.,-]
[.,-][a-zA-Z][a-zA-Z][a-zA-Z][a-zA-Z][a-zA-Z][a-zA-Z][a-zA-Z]\d\d\d\w\w[.,-]
[.,-][a-zA-Z]\d\d\d\d\d\d\d[.,-]
[.,-]\d\d\d\d\d\d\d\s
[.,-]\d\d\d\d\d\d\d\0x2C
[.,-][a-zA-Z]\d\d\d\d\d\d\d\d[.,-]
[.,-]\d\d\d\d\d\d\d\d\s
[.,-]\d\d\d\d\d\d\d\d\0x2C
[.,-]\d\d\d\d\d\d\d\d\d\s
[.,-]\d\d\d\d\d\d\d\d\d\0x2C
[.,-][a-zA-Z]\d\d\d\d\d\d\d\d[.,-]
[.,-][a-zA-Z]\d\d\d\d\d\d\d\d\d\d\d\d[.,-]
[.,-][a-zA-Z][a-zA-Z]\d\d\d\d\d\d[a-zA-Z][.,-]
[.,-][a-zA-Z]\d\d\d\d\d\d\d\d\d\d\d\d[.,-]
[.,-]\d\d\d\d\d\d\d\d\d\d\s
[.,-]\d\d\d\d\d\d\d\d\d\d\0x2c
[.,-]\d\d\d[a-zA-Z][a-zA-Z]\d\d\d\d[.,-]
[.,-]\d\d\d\d\d\d\d\d\d\d\d\d\s
[.,-]\d\d\d\d\d\d\d\d\d\d\d\d\0x2c
[.,-]\d\d[a-zA-Z][a-zA-Z][a-zA-Z]\d\d\d\d\d[.,-]
[.,-][a-zA-Z]\d\d\d\d\d\d\d\d\d\d\d\d\d\d[.,-]
[.,-][a-zA-Z][a-zA-Z]\d\d\d\d\d\d[.,-]
[.,-][a-zA-Z]\d\d\d\d\d\d\d\d\d\d\d\d\d\d\d[.,-]

Exceptions:

REDACTED: exempted pastern that matched user ID
[.,-][iI][mM]\d\d\d\d\d\d\d[.,-]
[.,-][pP][mM]\d\d\
d\d\d\d\d\d[.,-]
[.,-][sS][dD]\d\d\d\d\d\d\d\d[.,-]
[.,-][qQ][cC][ ]\d\d\d\d\d\d\d[.,-]
[.,-][qQ][cC]\d\d\d\d\d\d\d[.,-]
[.,-][cC][hH]\d\d\d\d\d\d\d[.,-]
20[012]\d[01]\d[0123]\d
[0123]\d[01]\d20[012]\d
[.,-][cC]\d\d\d\d\d\d\d[.,-]
[.,-][tT]\d\d\d\d\d\d\d[.,-]
[.,-][rR][fF]\d\d\d\d\d\d[.,-]

DLP: Tyranny of the Drivers License 2:3

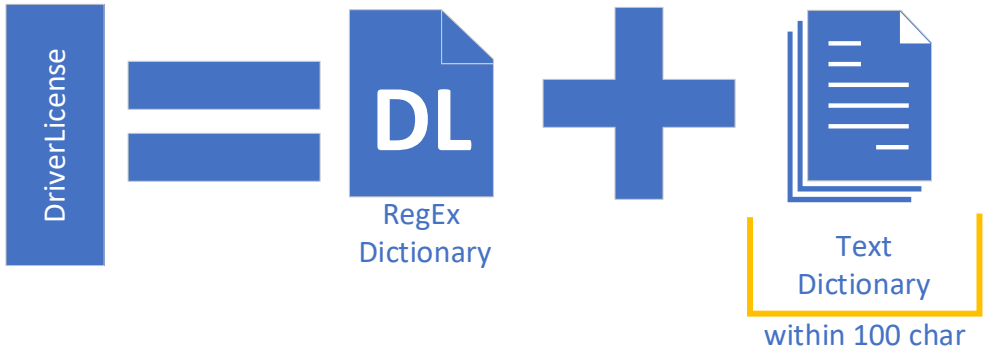
- **Our old strategy had two classifications around Driver's License**
  - Text dictionary: tele, mobile, meeting code
  - Under this config, every piece of data that matched NOTDriverLicense also matched the DriverLicense.
    - We had introduced a FALSE NEGATIVE where the existence of just one word from the text dictionary.



Alerts were configured to ignore NOT Drivers License!

DLP: Tyranny of the Drivers License 3:3

- **New Driver's License strategy**
  - Positive match dictionary sample: driver, license, vehicle, lic#, lic:, vin#, vin:, dl#, dl:
  - Text dictionary match needs to be within 100 char of the RegEx match. (reduces errors)
  - Significantly increased positive matches



DLP: Weaponize your users with Data classification

- Empower your users to tag your data with your organization's data classifications!
- This data is saved in the document meta-tag and is viewable by anyone or software that can look at the document or the document properties.
- This data tag is vendor agnostic.
- Very powerful when combined with traditional classification.
- Engender proper Data Stewardship in your company.
- Can create a custom data classification for a honey pot files.

RSAConference2022

A DLP Story:

Setting DLP Expectations

DLP: Setting DLP Expectations

- **DLP is a great tool but we can not set the expectations that a DLP Program will stop all data leaks.**
  - DLP helps the business take a policy driven approach to how they store / process / transmit data on a proactive level.
  - DLP can help build a culture of better Data Stewardship and help prevent accidental data leaks.
  - DLP can be used to meet certain GRC use cases.
  - DLP is not perfect and requires appropriate support.

DLP Detection and Bypass

- Once an attacker establishes a foothold on a client system, they can dump a list of processes to see if the system has a DLP Client.
- Almost all traditional DLP Solutions are overcome by simple data obfuscation or data encryption.
- **Take away: Communicate these points to management.**

RSAConference2022

A DLP Story:

Wrapping it up

“Apply” Slide

- In the next week :
  - Review / define the sensitive data in your environment and the implications if lost.
- In the next month:
  - Review / define your Data Classification Policy.
  - Update / define what data is stored / processed / transmitted in your applications.
  - Identify any applicable rules / laws / regulations / contractual agreements for each system.
- In the next quarter:
  - Either
    - Audit existing DLP systems to ensure meet your requirements.
    - Build business justification for a DLP implementation with solid and measurable use-cases.