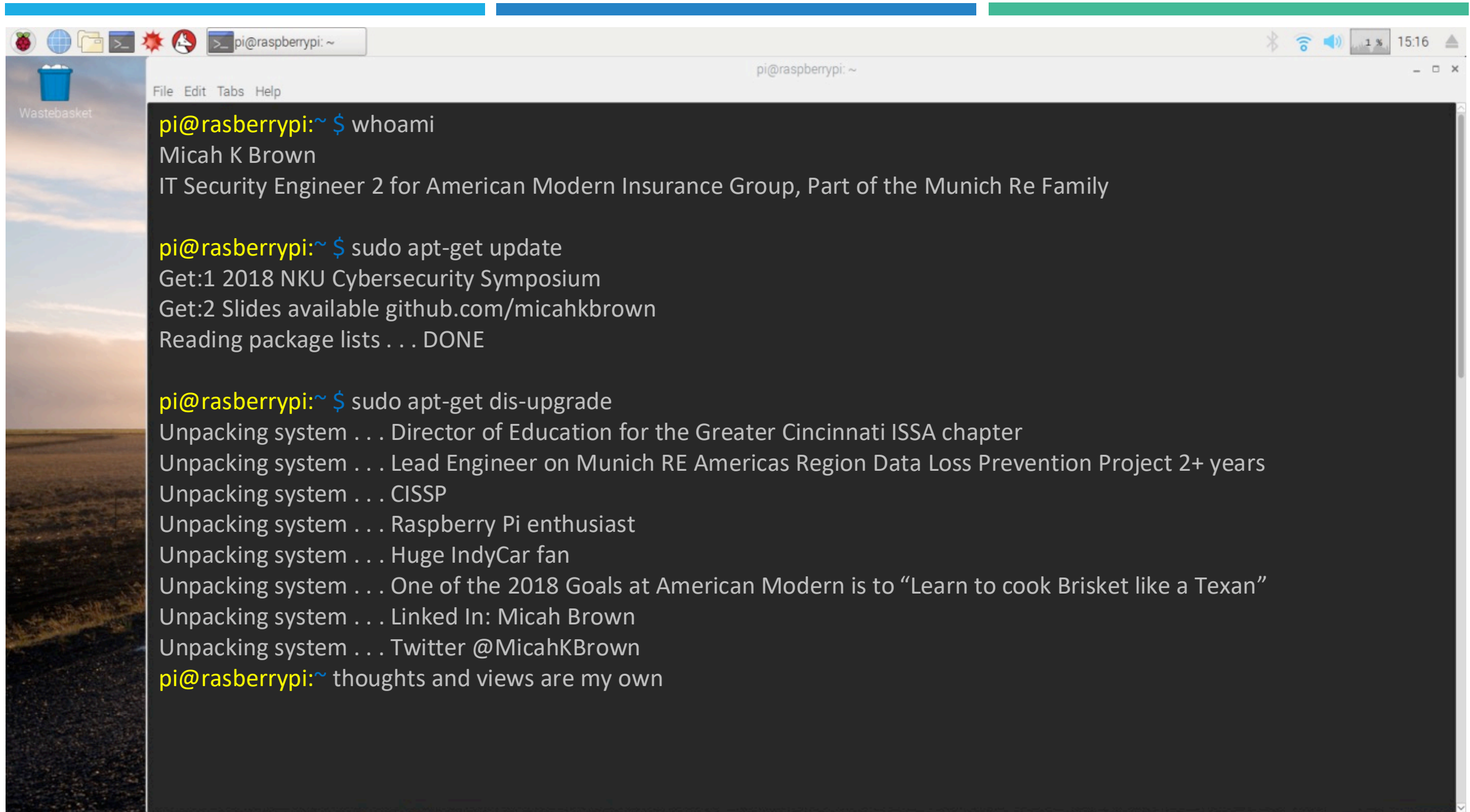




# DLP DEMYSTIFIED

HOW I LEARNED TO STOP WORRYING AND EMBRACE MY BLUE TEAM ROOTS





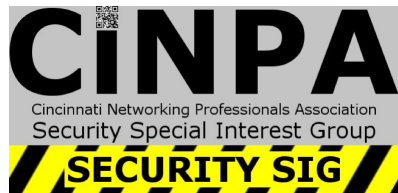
## Greater Cincinnati IT Security groups



<http://www.cincy-issa.org/>




<https://sites.google.com/view/cincysmba>



<https://www.meetup.com/TechLife-Cincinnati/events/>



<https://www.infosecincy.org/>



**“A process cannot be understood by stopping it.  
Understanding must move with the flow of the  
process, must join it and flow with it.”  
— Frank Herbert**

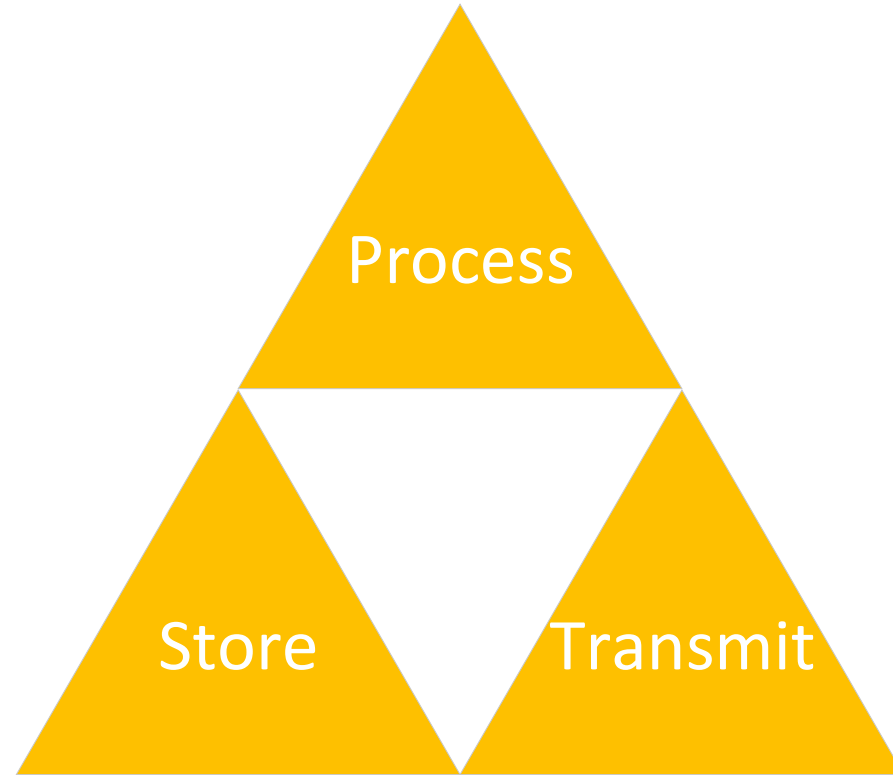
# DATA LOSS PREVENTION

## OVERVIEW



# WHAT IS DLP?

- **DLP is NOT** ... a 'monolithic' stack of technology by one single vendor
- **DLP is NOT** ... designed to stop a sufficiently advanced adversary [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
- **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



## THE STATES 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.

# DLP:THE MAIN PLAYERS



- **DLP Client** – a software that sits on the client system / server. Can have overlap with other parts of the DLP environment if not careful. Can impact performance.



- **DLP Network** – Think of an IDP / IPS Sitting on your network only instead of looking for attack patterns on a network segment, it looks for potentially sensitive information. Can be put inline or out of line.



- **DLP Repository Scanner** – a physical or virtual appliance that sits on your networks and scans data repositories (shares, sharepoint sites, databases, and more) for potentially sensitive information.



- **DLP Email** – a physical or virtual appliance that partners with your email subsystem to scan outgoing and / or incoming email for sensitive information. Similar AV for email.



- **DLP Web** - a physical or virtual appliance that partners with your internet proxy to scan outgoing and / or incoming web requests for sensitive information. Similar AV for web proxy.



- **DLP Management console** – the physical or virtual appliance that manages all of the individual DLP parts. This also includes reporting, incident management, and evidence management.

# DLP: CONCERNS

- **Most DLP Client implementations have a lot of functionality overlaps with the rest of the environment (DLP network, DLP Repository Scanner, DLP Email, DLP Web). Why not do it all on the client?**
  - DLP client can have high performance hit (CPU / Memory / Disk I/O) on the user given a sufficiently advance policy.
  - You can run into incompatibles with local software such as Chrome and Firefox that can break client web inspection.
  - Incompatibilities with other applications can spike CPU / Memory/ Disk I/O.
  - Client must be connected to network to get updates.
  - The client might not be on every user system / server.
- **DLP Network can overlap with DLP Email and DLP Web.**
  - This traditionally happens with unencrypted conversations.
- **DLP Management console can be a high value target due to access to evidence and may be governed by rules / laws / regulations / contractual agreements.**
- **Many different products are adding on “DLP” functionality. Interoperability might be a challenge between diverse systems.**





**“The beginning of knowledge is the discovery of  
something we do not understand.”  
— Frank Herbert**

# DATA LOSS PREVENTION

GETTING STARTED



# DLP: DEFINE YOUR GOALS

- **DLP has become a very popular concept at the executive level to protect data and prevent breaches. It is our job to help educate what the tool(s) can do and shape the conversation that will ultimately create a policy that provides the organization value and that can be supported.**
  - Lead the project from Engineering / Analyst up.
  - Clearly define what features / tools to turn on.
  - Is DLP part of the organization's "active defense toolset" or historical system of record?
  - DLP generally gathers a lot of NPI, PCI, HIPAA, GLBA and other sensitive data as part of evidence collection. Any splash damage due to evidence collection? (Rules / Laws / Regulations / contractual agreements)
  - DLP requires the cooperation of the entire organization. Bring them in early. Key players (HR, Compliance, Legal, IT, individual business units).
  - Who responds to incidents and how are true issues escalated?
  - What is the SLA for a DLP incident?

# 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: BUILD WHAT YOU CAN SUPPORT

- **Our original policy that we went to pilot had each user creating 100 incidents per day. @ 5min per incident we would need to hire 4235 DLP analysts just to keep up!**
  - Present operational costs to management in terms they can understand.
  - As a rule of thumb default rules stink!
  - Learn to love REG-EX!
    - Matt Scheurer - Regular Expressions (Regex) Overview 2017 Derby Con
    - <https://regex101.com/>
  - Acknowledge the risk of over tuning
    - Driver's License and Credit Cards are typically very challenging patterns to tune!
    - Is the juice worth the squeeze.
    - Does an exemption make it super easy for users to bypass the DLP environment.
    - When tuning, use positive inclusive matching.
    - The Dangers of FALSE NEGATIVE

---

There are no FALSE POSITIVES, only poorly written rules\*





**“To know a thing well, know it's limits; Only when pushed beyond it's tolerance will it's true nature be seen.” — The Amtal Rule  
Frank Herbert, Children of Dune**

# DATA LOSS PREVENTION

DLP ARCHITECTURE



## DLP:TRADITIONAL DLP POLICY

- 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.
  - We have to do better ...
- 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: THE TYRANNY OF THE DRIVER'S LICENSE 1:2

## Driver's License Positive match

```
[.-]00\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]\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\d0x2C  
[.-][a-zA-Z]\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\d0x2C  
[.-]\d\d\d\d\d\d\d\d\d\d\s  
[.-]\d\d\d\d\d\d\d\d\d\d0x2C  
[.-][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[.-]  
[.-][a-zA-Z][a-zA-Z]\d\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\d0x2c  
[.-]\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\d0x2c  
[.-]\d[a-zA-Z][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[.-]  
[.-][a-zA-Z][a-zA-Z]\d\d\d\d\d\d\d[.-]  
[.-][a-zA-Z]\d\d\d\d\d\d\d\d\d\d\d\d\d\d[.-]
```

## Exceptions:

REDACTED: exempted pastern that matched user ID

```
[.-][il][mM]\d\d\d\d\d\d\d[.-]  
[.-][pP][mM]\d\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\d[.-]  
[.-][qQ][cC]\d\d\d\d\d\d\d\d[.-]  
[.-][cC][hH]\d\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\d[.-]  
[.-][tT]\d\d\d\d\d\d\d\d[.-]  
[.-][rR][fF]\d\d\d\d\d\d\d\d[.-]
```

# DLP: THE TYRANNY OF THE DRIVER'S LICENSE 2:2

## Our old strategy had two classifications around Driver's License

- DriverLicense = [Reg-ex dictionary match]
- NOTDriverLicense = [Reg-ex dictionary match] + [text dictionary that would indicate REG-EX match is not DL]
  - Negative match dictionary sample: tele, mobile, meeting code

Under this setup every document that matched the NOTDriverLicense also matched the DriverLicense. Thus, in our incident rules we exempted any document that had the NOTDriverLicense classification. This had adverse effects in that if a document had multiple detection [HIPAA, GLBA, PCI] but also matched NOTDriverLicense, we would not get an incident. NOT COOL

## New Driver's License strategy

- DriverLicense = [Reg-ex dictionary match] + [text dictionary that would indicate a positive DL match] within 100 char
  - Negative match dictionary sample: driver, license, vehicle, lic#, lic:, vin#, vin:, dl#, dl:

# DLP: WEAPONIZE YOUR USERS WITH DATA CLASSIFICATION

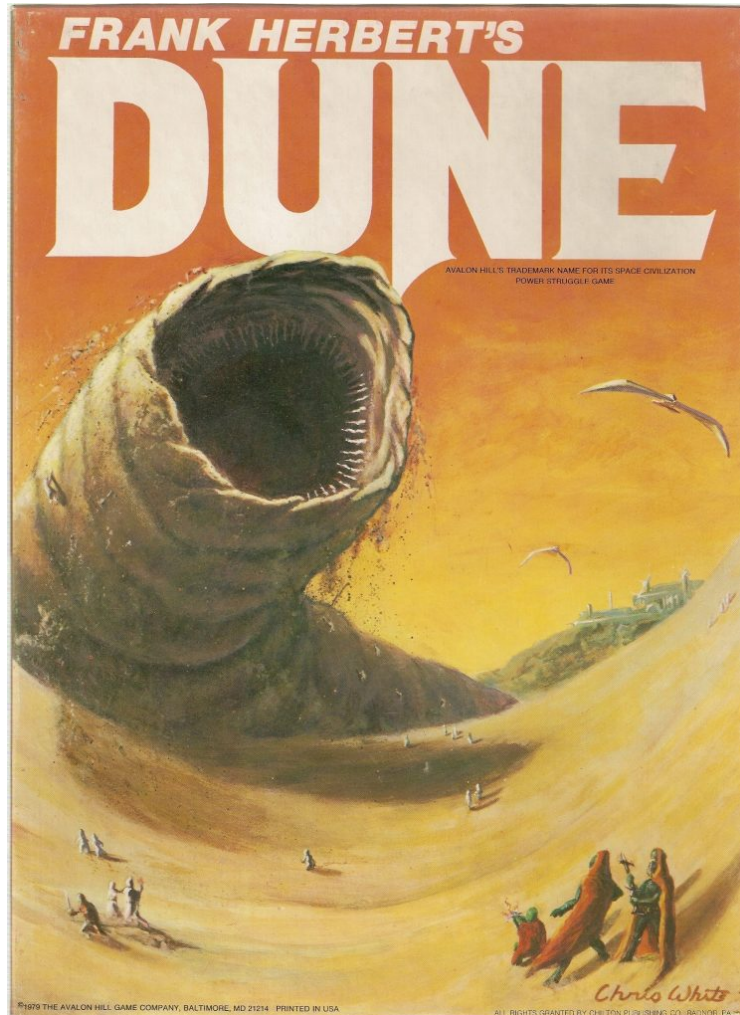
- **Recently many popular office software suites provide a way in which you can empower your users to tag your data with your organization's data classifications!**
  - This data is saved in the document meta-tag and is viewable by anyone that can look at the document or the document properties.
  - This data tag is vendor agnostic and can be leveraged by tools outside of DLP.
  - Classification plus traditional classification rules (REG-EX and word dictionaries) is powerful!
  - Engender proper data ownership culture by including your workforce.
  - If you create a custom data classification for a honey pot, make sure the meta-data resembles the other metadata tag.  
#verygroovey

**“Arrakis teaches the attitude of the knife - chopping off what's incomplete and saying: 'Now, it's complete because it's ended here.’” — from "Collected Sayings of Maud'Dib" by the Princess Irulan”  
— Frank Herbert, Dune**

# DATA LOSS PREVENTION

LET’S BUILD A POLICY

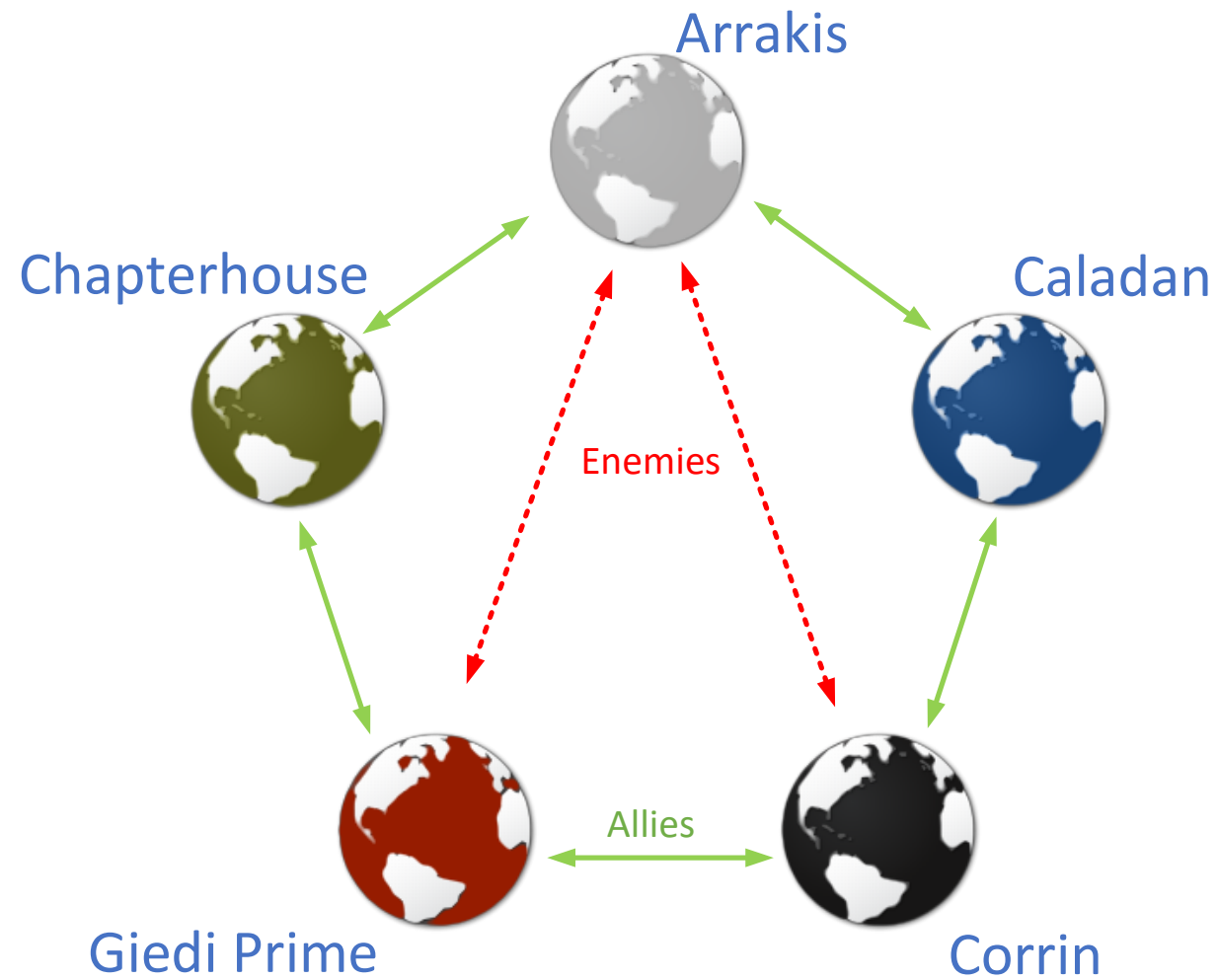
# DLP: LET'S BUILD A POLICY!



We will be modeling a DLP Policy loosely based on Frank Herbert's Dune universe. We will be working for Paul Atreides to build a DLP Policy for Planet Arrakis as it interacts with four other planets. The spice data must flow.

- **Caladan – Arrakis is a long time ally of the Caladan.**
- **Corrin – Relationships between Corrin and Arrakis has been troubled for several generations.**
- **Giedi Prime – For generations the leaders of Giedi Prime and Arrakis have been in open conflict.**
- **Chapterhouse – Chapterhouse and Arrakis have a successful partnership**

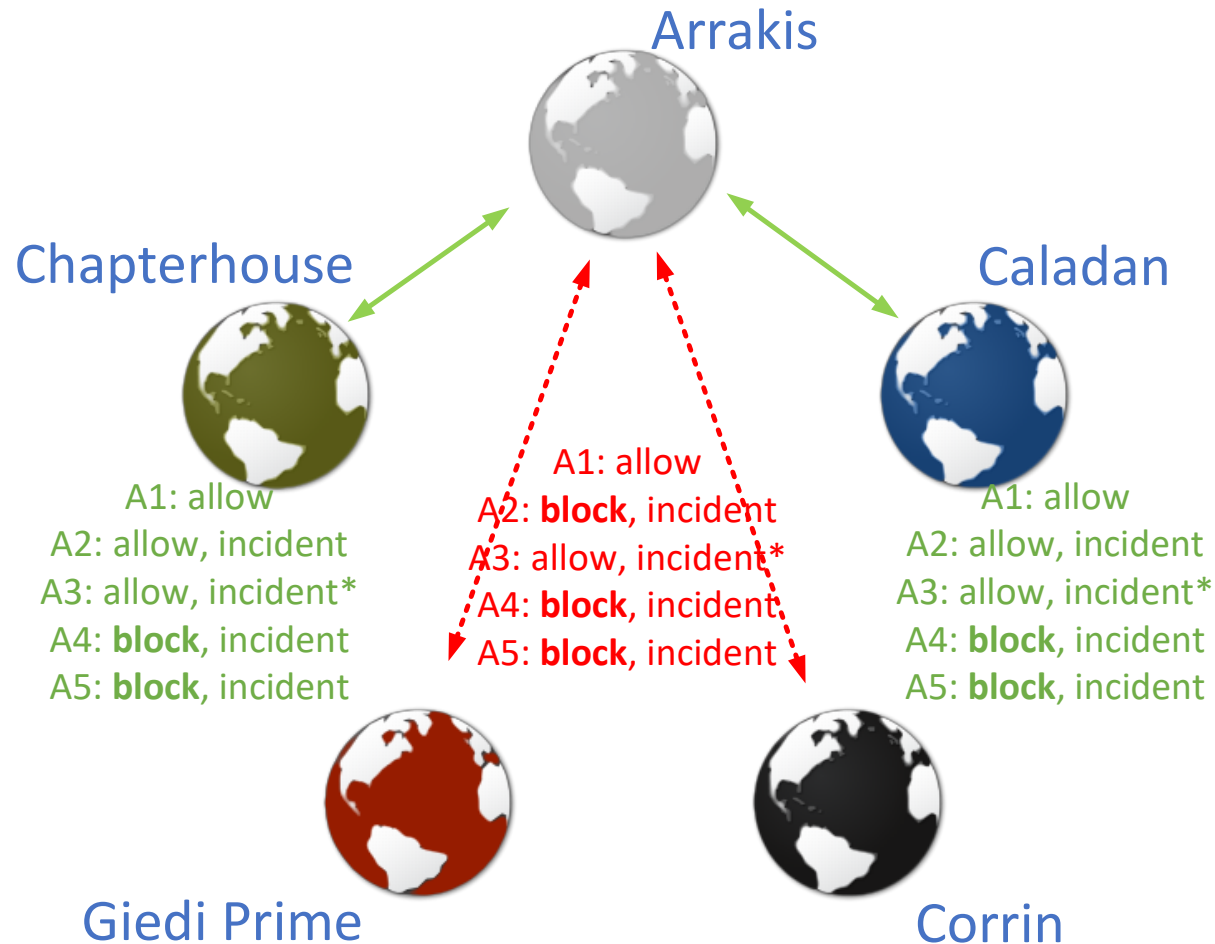
# ARRAKIS:ALLIANCES



# ARRAKIS: DATA CLASSIFICATION

- Planet Arrakis's data classification policy.
  - **A1 – Public:** suitable to be shared with anyone.
  - **A2 – Proprietary:** to be used by members of Arrakis's governments or allies.
  - **A3 – Diplomatic:** free to be shared with other authorized planetary governments and only from members of the diplomat department.
  - **A4 – Internal:** can not be transmitted to non-Arrakis government.
  - **A5 – Top Secret:** restricted to the highest levels with Arrakis government.

# ARRAKIS: DATA FLOW VISUALIZATION



- **A1 – Public:** suitable to be shared with anyone.
- **A2 – Proprietary:** to be used by members of Arrakis's governments or allies.
- **A3 – Diplomatic:** free to be shared with other authorized planetary governments and only from members of the diplomat department.
- **A4 – Internal:** can not be transmitted to non-Arrakis government.
- **A5 – Top Secret:** restricted to the highest levels with Arrakis government.



# ARRAKIS: DOCUMENTING THE POLICY

Arrakis Data Classification				
Classification	A2 – Proprietary			
Target Group	All			
	Caladan	Corrin	Giedi Prime	Chapterhouse
DLP Client	allow, incident	block, incident	block, incident	allow, incident
DLP Network	allow, incident	block, incident	block, incident	allow, incident
DLP Email	allow, incident	block, incident	block, incident	allow, incident
DLP Web	allow, incident	block, incident	block, incident	allow, incident
Classification	A3 – Diplomatic			
Target Group	Diplomatic Corp			
	Caladan	Corrin	Giedi Prime	Chapterhouse
DLP Client	allow, incident	allow, incident	allow, incident	allow, incident
DLP Network	allow, incident	allow, incident	allow, incident	allow, incident
DLP Email	allow, incident	allow, incident	allow, incident	allow, incident
DLP Web	allow, incident	allow, incident	allow, incident	allow, incident
Classification	A3 – Diplomatic			
Target Group	everyone else			
	Caladan	Corrin	Giedi Prime	Chapterhouse
DLP Client	block, incident	block, incident	block, incident	block, incident
DLP Network	block, incident	block, incident	block, incident	block, incident
DLP Email	block, incident	block, incident	block, incident	block, incident
DLP Web	block, incident	block, incident	block, incident	block, incident

- **A1 – Public:** suitable to be shared with anyone.
- **A2 – Proprietary:** to be used by members of Arrakis’s governments or allies.
- **A3 – Diplomatic:** free to be shared with other authorized planetary governments and only from members of the diplomat department.
- **A4 – Internal:** can not be transmitted to non-Arrakis government.
- **A5 – Top Secret:** restricted to the highest levels with Arrakis government.



**"The power to destroy a thing is the absolute control over it."**

**— Frank Herbert, Dune**

# DATA LOSS PREVENTION

HOW TO DETECT / BYPASS A DLP SOLUTION



# DLP: DETECT, BYPASS, AND ABUSE DLP

- On clients, look at processes and services to see if DLP is running.
  - Most DLP client solutions can create incidents on a read, write, execute. However some can not detect a file encrypt / file obfuscation.
- Without the help of an external device to break encryption, DLP Web and DLP Network can be easily subverted by encryption / obfuscation.
- DLP Repository scanner is subverted by encryption / obfuscation.
- Breaking encryption can introduce more vulnerabilities such as crypto downgrade attacks and / or another juicy target for a malicious attacker.
- DLP incidents can provide a lot of interesting information!