



Hacking for Diplomacy

DS 6 – Data Architecture for
Cybersecurity

Meet the Team

Jaime Campanelli, A.J. Musacchio, Jenelle Salazar, Randall Weber





E.D.A.C Consulting

Enhanced
Data
Architecture
for
Cybersecurity



Initial Problem Statement

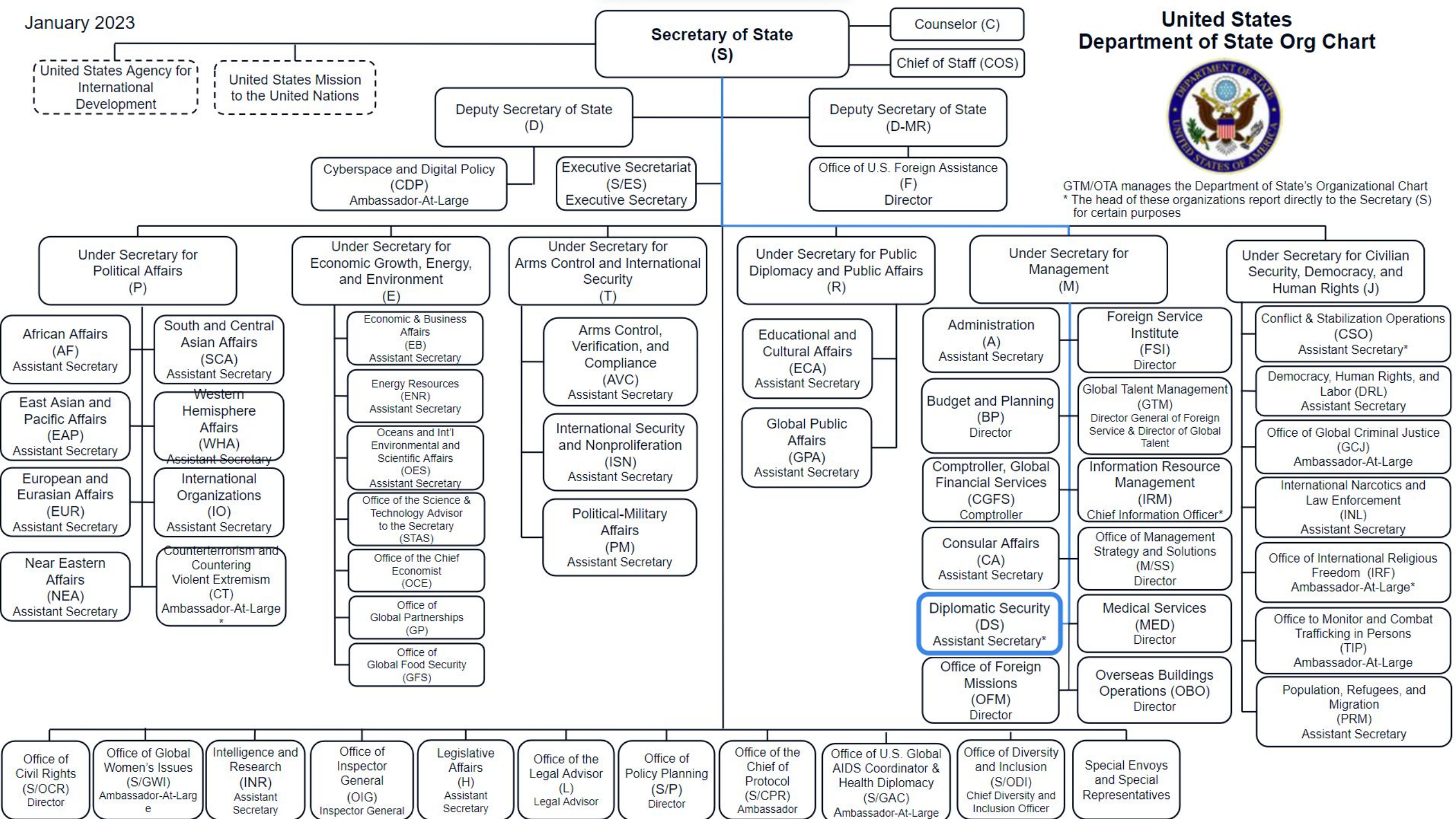
Network defenders in the Office of Cyber Monitoring and Operations need a better way to query and correlate data in a hybrid and multi-cloud data ecosystem in order to develop analytics capability at the network defender level and inform insight-driven decisions on cybersecurity incident response at the senior leadership level.

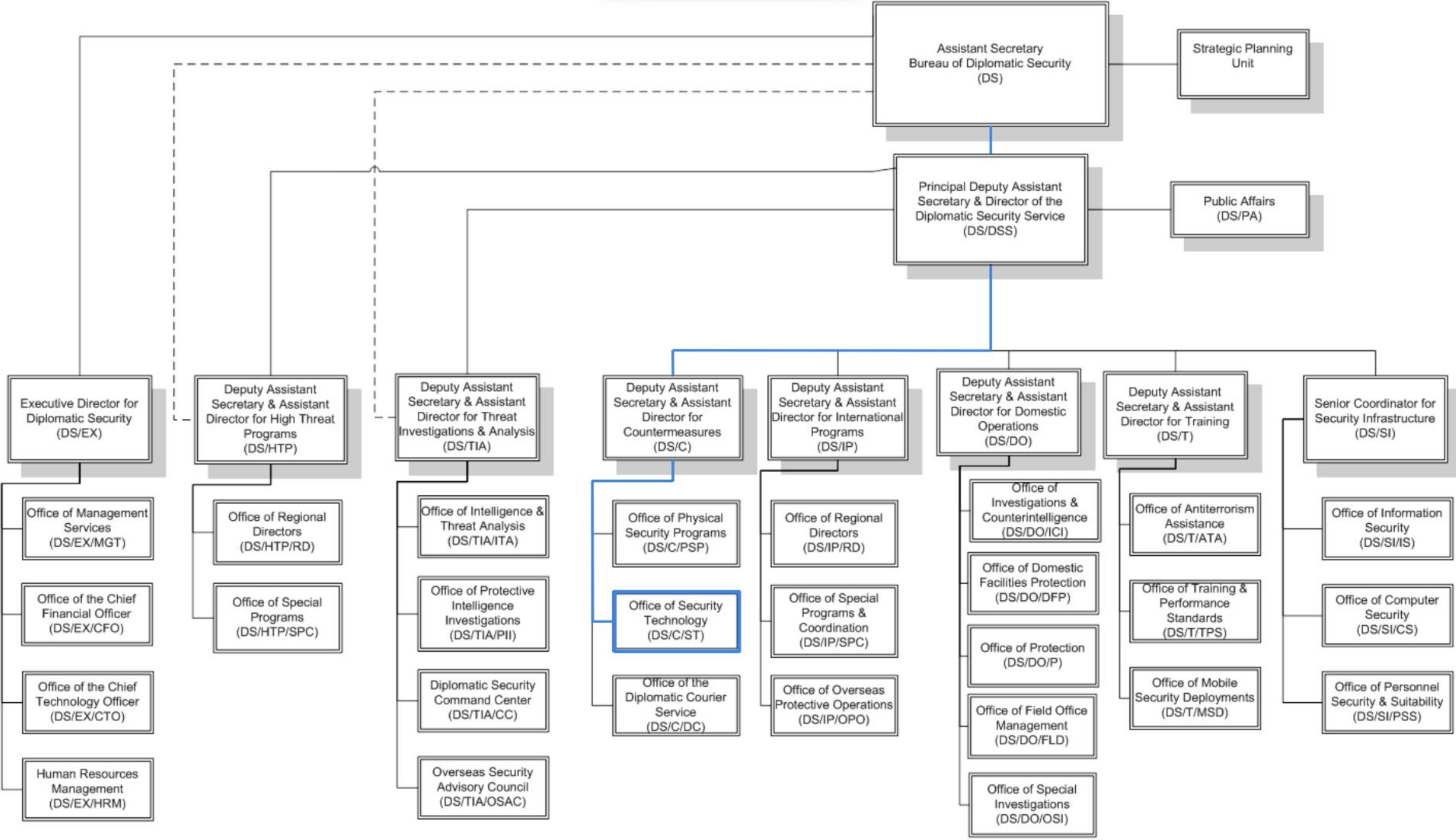
Revised Problem Statement

Network defenders in the Office of Cyber Monitoring and Operations need a better **agnostic way to collect, store, and analyze logs. This system will be used to inform cybersecurity related decisions** on the network defender and incident response level. **To pair with this, policy will be required to help Network Defenders implement new changes and become more aware.**

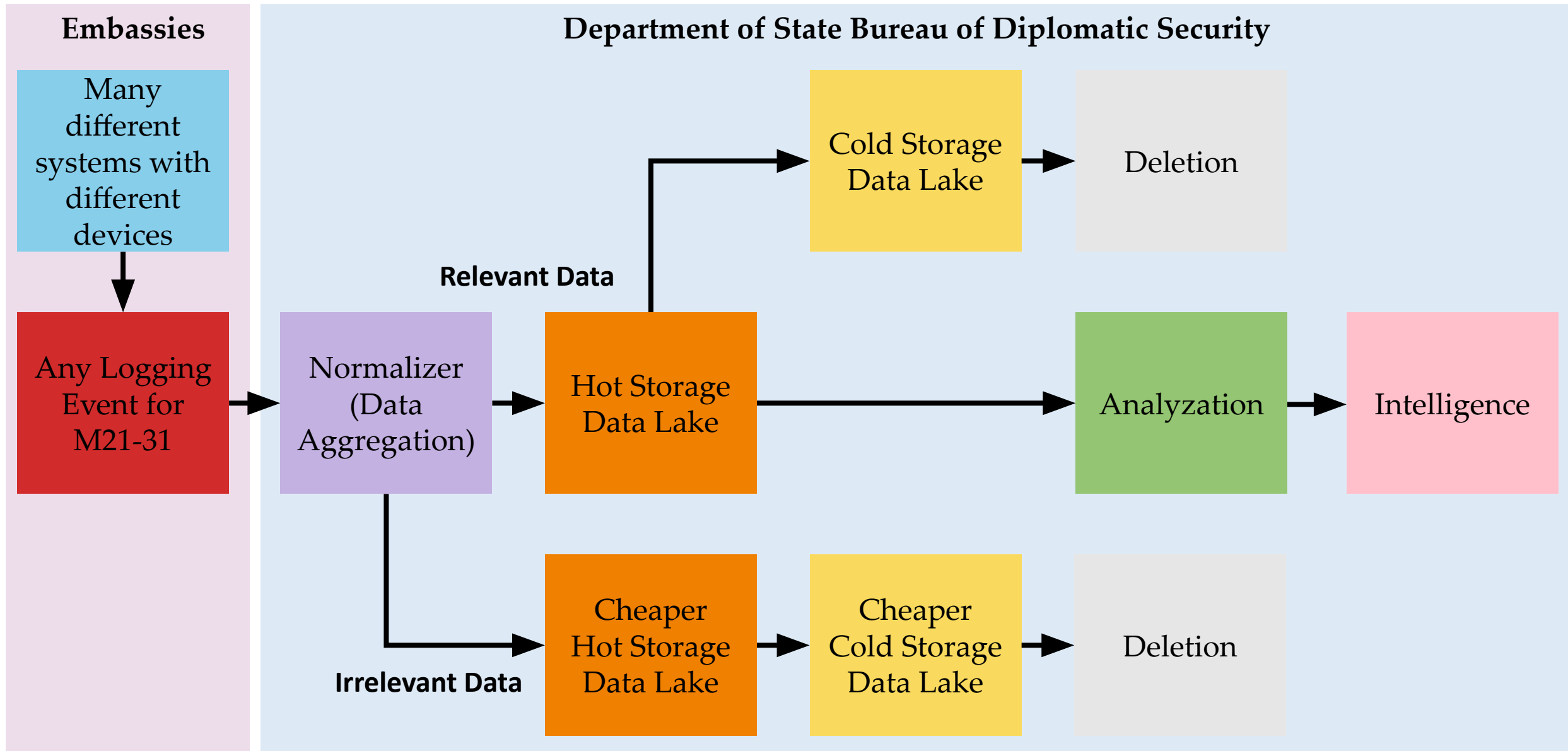


GTM/OTA manages the Department of State's Organizational Chart
* The head of these organizations report directly to the Secretary (S) for certain purposes





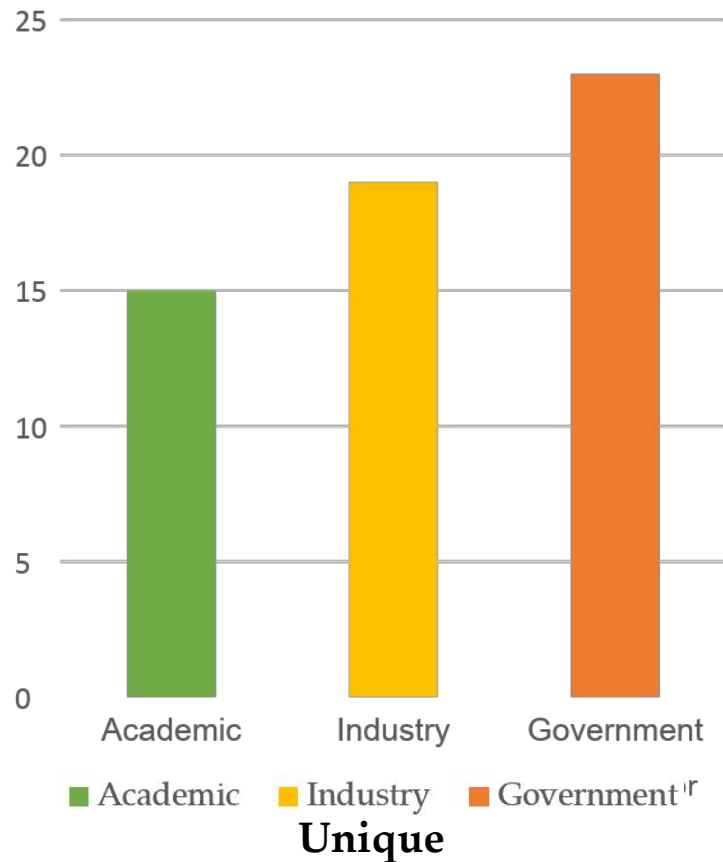
Final MVP



Interview Breakdown By Sector

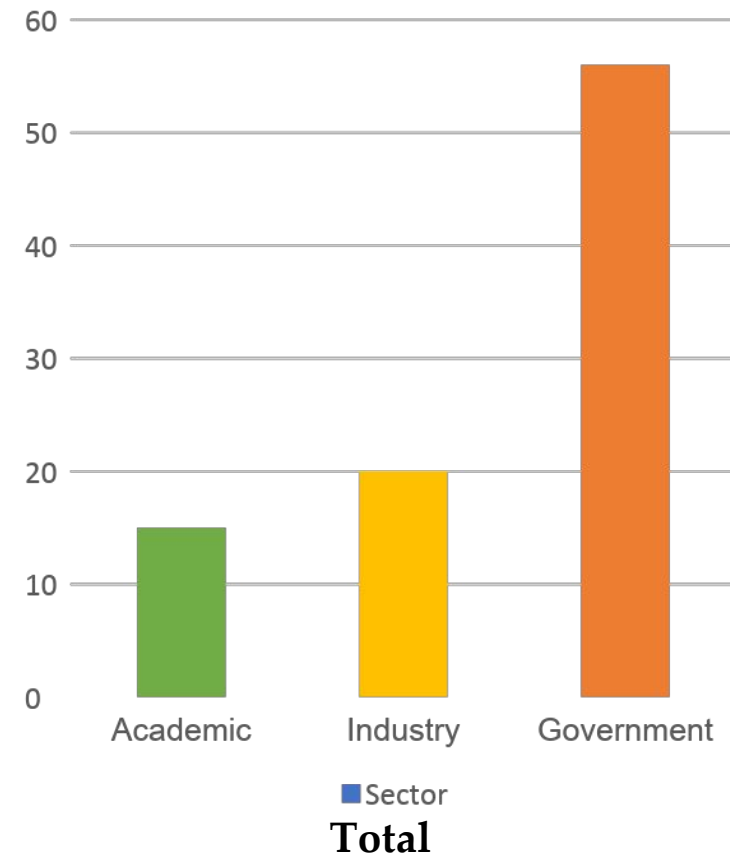
❖ Across all 15 weeks we had a total of **59 unique** interviews across 3 sectors:

- ❖ **Academic** – 15
- ❖ **Industry** – 19
- ❖ **Government** – 23

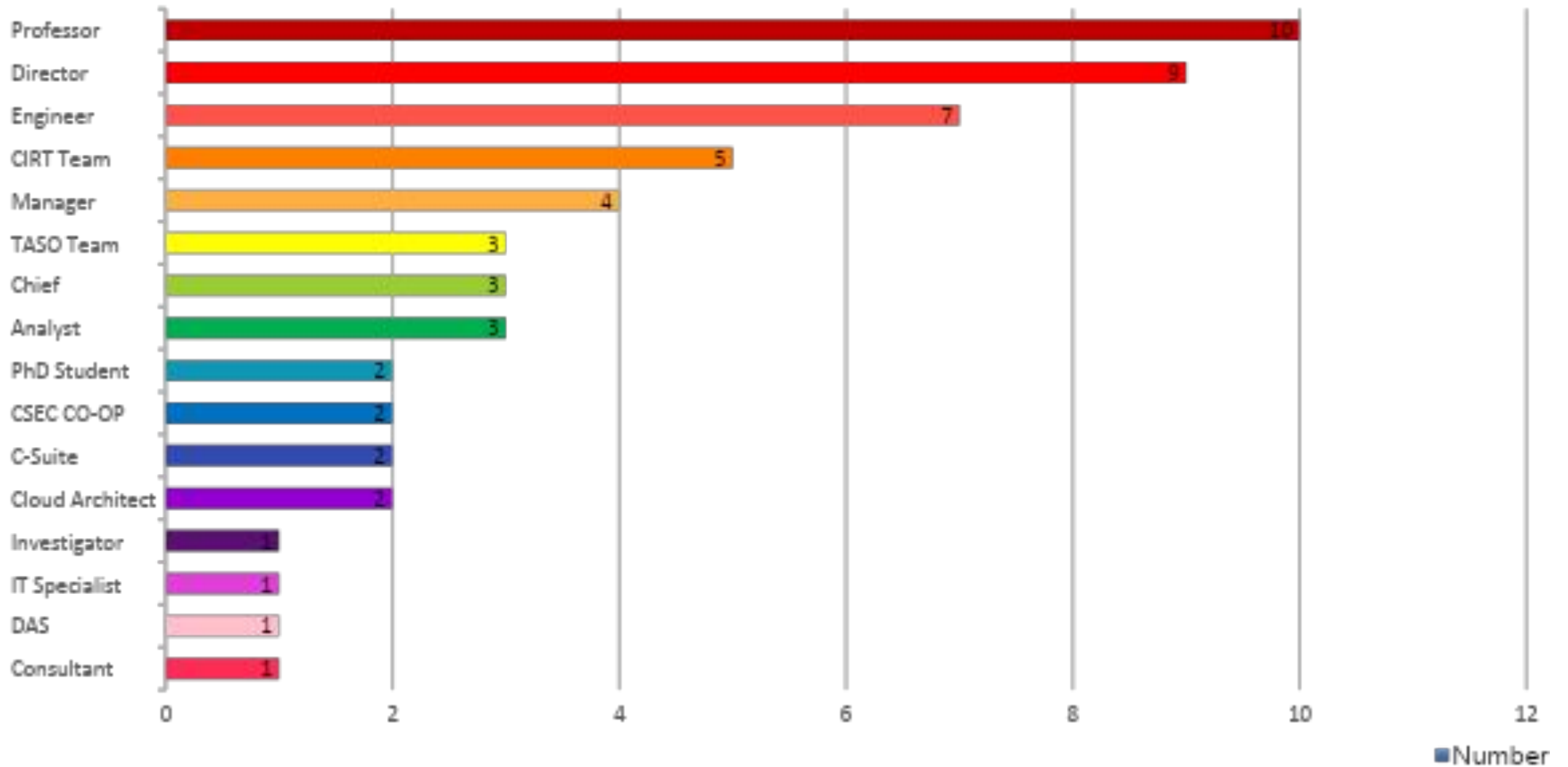


❖ Across all 15 weeks we had a total of **91 total** interviews across 3 sectors:

- ❖ **Academic** – 15
- ❖ **Industry** – 20
- ❖ **Government** – 56

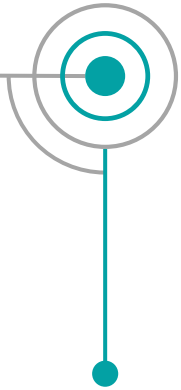


Interview Archetype Breakdown



Project Journey

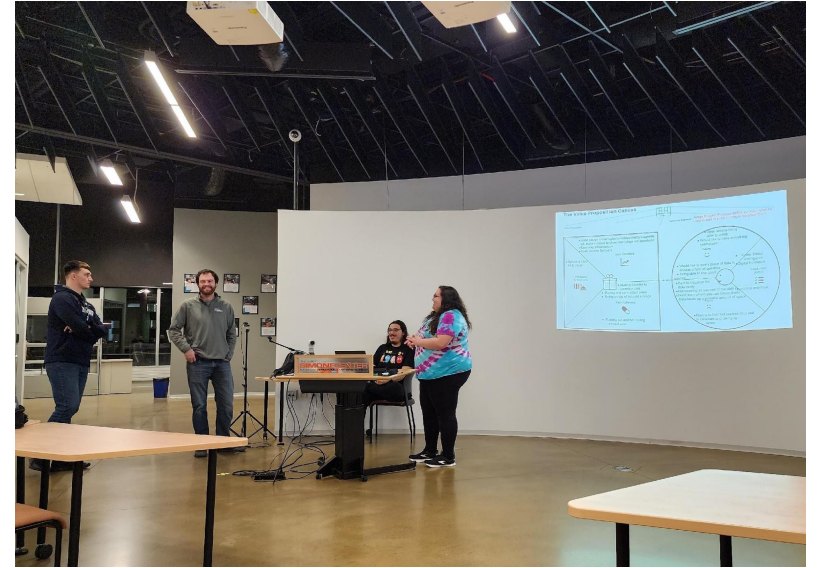
Weeks 1 - 5



Total Interviews: 17
Research Papers: 20
Days Left: 70

Weeks 1-5

- ❖ Meetings with sponsors to focus on understanding the problem
- ❖ Discovery Interviews
 - ❖ Interviews with consultants, security engineers, and educators
 - ❖ Understanding technologies involved in the Problem Statement and their capabilities



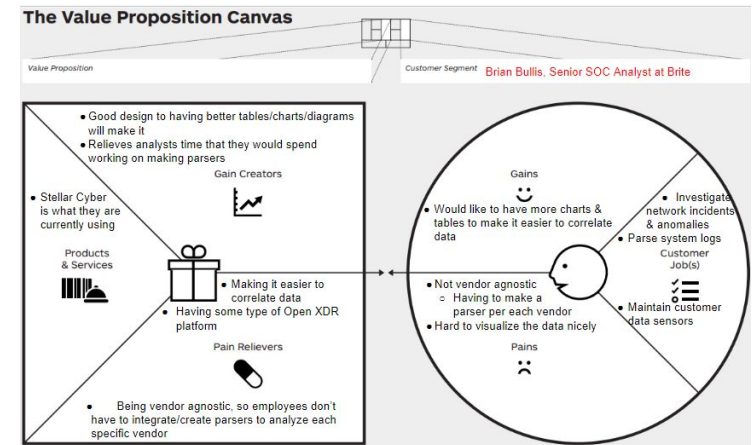
Weeks 1-5: Interviews

David Hagan – Cloud Data Architect, Office of Cyber Monitoring and Operations

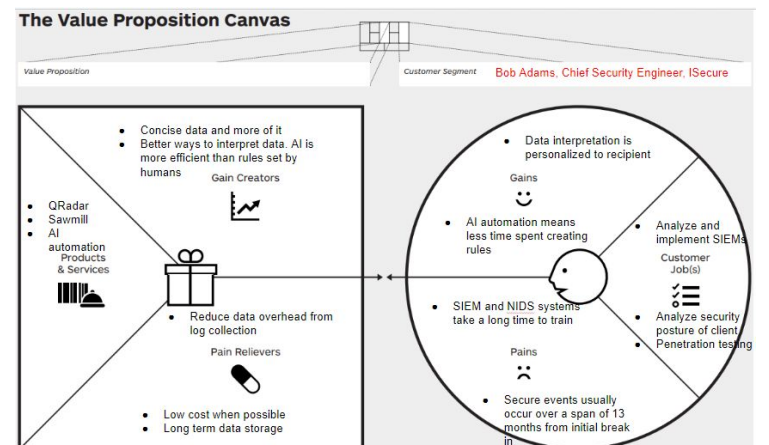
- ❖ Data storage and log aggregation
- ❖ Big takeaway: Data Lakes for storage
 - ❖ Aggregate events and then place in cloud storage

Other Important VPCs

Brian Bullis



Bob Adams



Weeks 1-5: Mission Model Canvas










The Mission Model Canvas

Mission/Problem Description:

Designed by:

Date: Week 4

Version:

<p>Key Partners </p> <ul style="list-style-type: none"> •Cyber and Technology Security (CTS) Directorate •Cybersecurity Consultants and Managers •SOC Analysts who work on SIEM •Professors and Researchers in fields of Information Technology, Cybersecurity, and Database Management 	<p>Key Activities </p> <ul style="list-style-type: none"> •Create ways for data uniformity across many teams •Provide solutions for availability of data to CTS <p>Key Resources </p> <ul style="list-style-type: none"> •Cloud Systems •Security Information Event Management (SIEM) Tools •Enterprise Resource Planning tools (ERP) •Data analytics/visualization software 	<p>Value Propositions </p> <ul style="list-style-type: none"> •CTS will have a better ability to secure the DOS infrastructure and visualize security data •Better data visualization will lead to faster and more effective incident response 	<p>Buy-in & Support </p> <ul style="list-style-type: none"> •Emphasize the importance of security to other teams in DOS. •Demonstrate how data visualization improves response times and methods <p>Deployment </p> <ul style="list-style-type: none"> •Data uniformity and availability •Deploying a Beta Version •Data visualization charts/deliverables 	<p>Beneficiaries </p> <ul style="list-style-type: none"> •Incident Response personnel •Cyber Security Engineers •Network Defenders and Threat Hunters •Employees who utilize SIEM platforms •Senior IT Leadership •Office of Cyber Monitoring and Operations •Pen-testing teams
<p>Mission Budget/Cost </p> <ul style="list-style-type: none"> •Cost implementation of a new or updated system 		<p>Mission Achievement/Impact Factors </p> <ul style="list-style-type: none"> •CTS has widespread access to necessary company data in order to properly secure the network infrastructure. 		

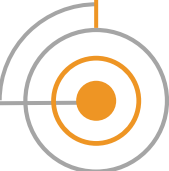
Project Timeline

Weeks 1 - 5



Total Interviews: 17
Research Papers: 20
Days Left: 70

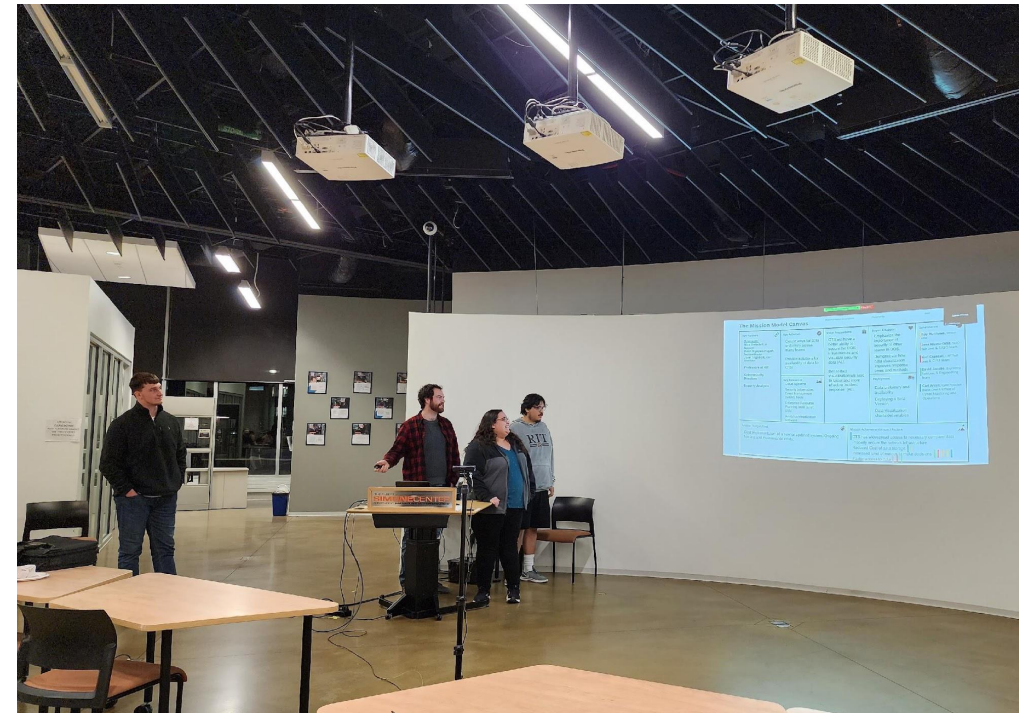
Total Interviews: 32
Research Papers: 28
Days Left: 56



Weeks 6 - 7

Weeks 6-7

- ❖ Shifted focus to more specific issues and how to solve them
 - ❖ How is the data collected?
 - ❖ How is data stored and retrieved?
- ❖ Ways to filter out data and sort it by relevance to security
- ❖ Centralization of data into Data Lakes



Weeks 6-7: Interviews

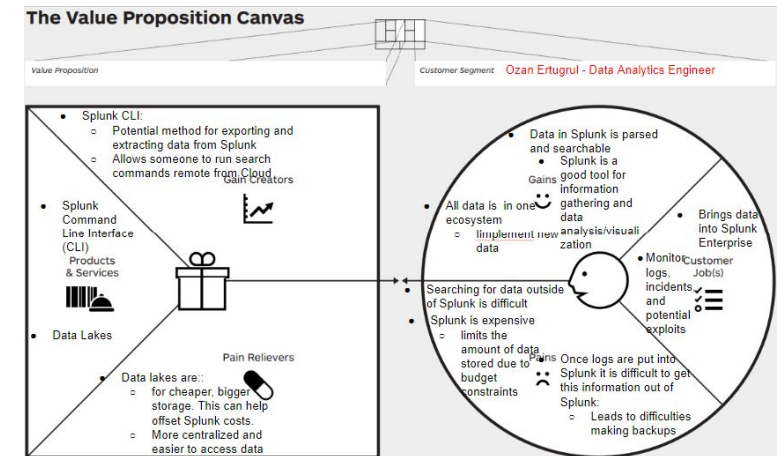
Mike Pinch – Director, Security Risk Advisors

- ❖ Presented the idea of a data pipeline and fusion center
- ❖ Cribl allowed easy traversal of unsorted logs within a data lake
- ❖ Importance of filtering incoming information as "useful" or less-useful.

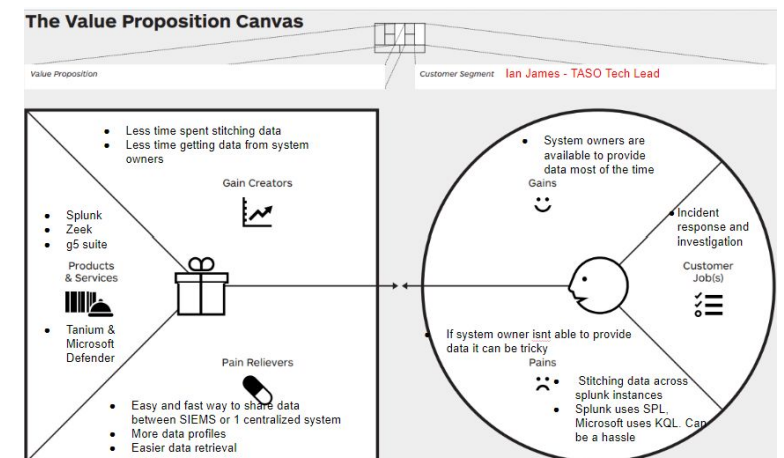
**“The problem with SOCs across the industry view and use SIEM as the center of their data universe... which it shouldn't be”
– Mike Pinch**

Other Important VPCs

Ozan Ertugrul



Ian James



Weeks 6-7: Mission Model Canvas

The Mission Model Canvas

Mission/Problem Description:

Designed by:

Date: Week 7

Version:

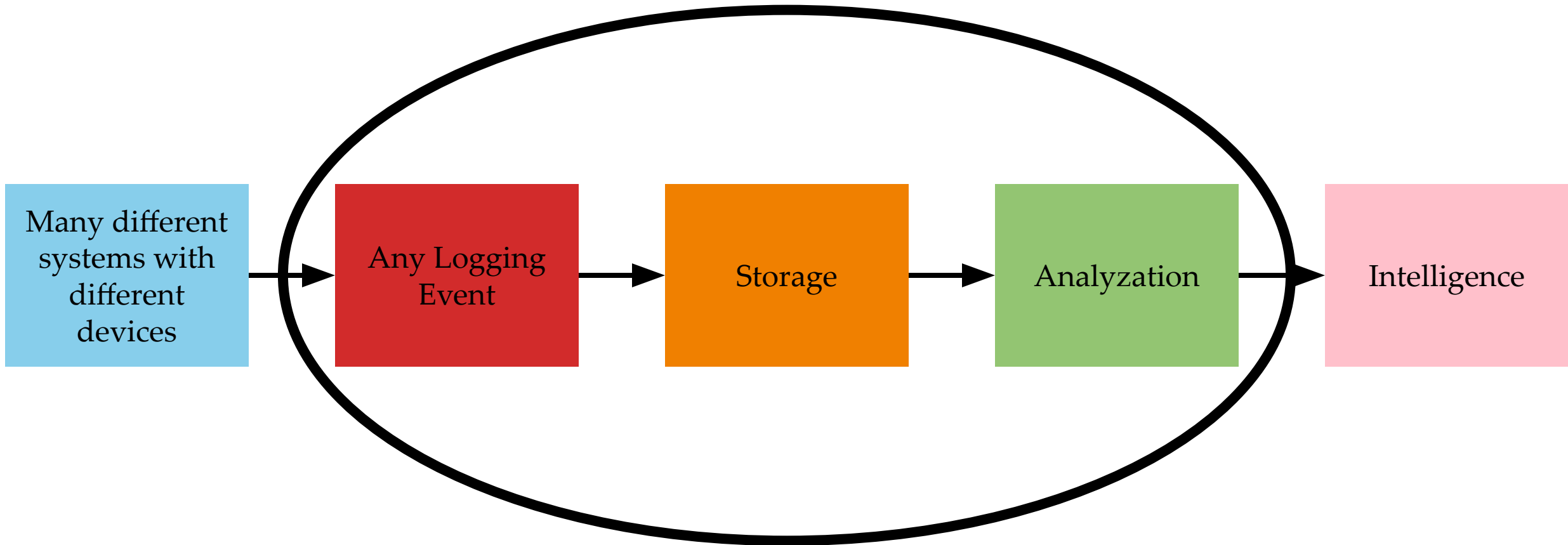
<p>Key Partners </p> <p>Sponsors: Nick Swindell, IT Specialist Danh Nguyen-Huynh, Technical Director Jake Trigoboff, CIRT Cloud Lead</p> <p>Professors at RIT</p> <p>Cybersecurity Directors</p> <p>Security Analysts</p>	<p>Key Activities </p> <p>Create ways for data uniformity across many teams</p> <p>Provide solutions for availability of data to CTS</p> <hr/> <p>Key Resources </p> <p>Cloud Systems</p> <p>Security Information Event Management (SIEM) Tools</p> <p>Enterprise Resource Planning tools (ERP) Data</p> <p>Analytics/visualization software</p>	<p>Value Propositions </p> <p>CTS will have a better ability to secure the DOS infrastructure and visualize security data (ALL)</p> <p>Better data visualization will lead to faster and more effective incident response (ALL)</p>	<p>Buy-in & Support </p> <p>Emphasize the importance of security to other teams in DOS.</p> <p>Demonstrate how data visualization improves response times and methods</p> <hr/> <p>Deployment </p> <p>Data uniformity and availability</p> <p>Deploying a Beta Version</p> <p>Data visualization charts/deliverables</p>	<p>Beneficiaries </p> <p>Roy Matthews, Division Chief</p> <p>Jose Rivera-Ortiz, TASO Tech Lead, & TASO team</p> <p>Karl Crandall, CIRT Tech Lead, & CIRT team</p> <p>David Jacobs, Engineering Tech Lead, & Engineering team</p> <p>Carl Wyatt, Cyber Protection Branch Chief & Office of Cyber Monitoring and Operations</p>
<p>Mission Budget/Cost </p> <p>Cost implementation of a new or updated system. Ongoing hosting and maintenance costs.</p>		<p>Mission Achievement/Impact Factors </p> <p>CTS has widespread access to necessary company data Properly secure the network infrastructure.</p> <p>Reduced Cost of data storage </p> <p>Increased level of metrics to make decisions </p> <p>Faster access to data </p>		



This work is licensed under the Creative Commons Attribution-ShareAlike 3.0 Unported License. To view a copy of this license, visit: <http://creativecommons.org/licenses/by-sa/3.0/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

DESIGNED BY: Strategyzer AG & Steve Blank
 The makers of Business Model Generation and Strategyzer

Week 6 – First Problem Flowchart



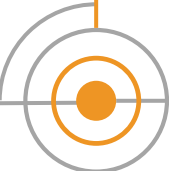
Project Timeline

Weeks 1 - 5



Total Interviews: 17
Research Papers: 20
Days Left: 70

Total Interviews: 32
Research Papers: 28
Days Left: 56



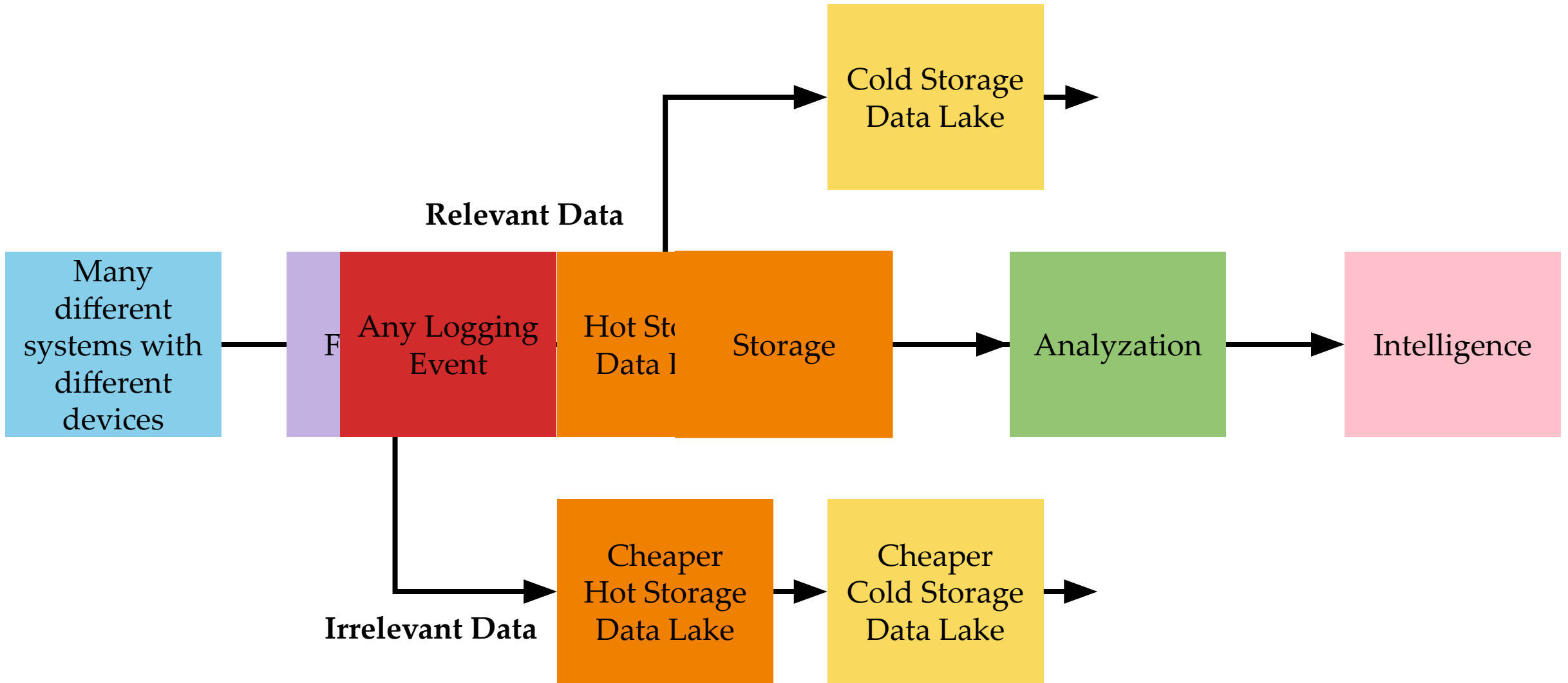
Weeks 6 - 7

Weeks 8 - 9



Total Interviews: 77
Research Papers: 36
Days Left: 42

Week 8 – Second MVP



Weeks 8-9

- ❖ During Weeks 8 & 9 we visited our sponsors in Washington D.C., Maryland & Virginia
 - ❖ Toured DoS SA-20 location & off-site data center
 - ❖ Met with 15+ people including, but not limited to:
 - ❖ Senior management
 - ❖ Incident response personnel



Weeks 8-9: Interviews

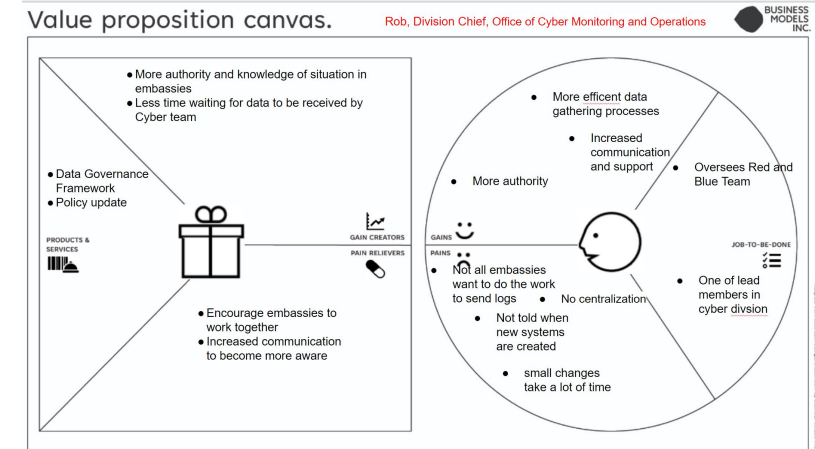
Roy Matthews - Division Chief, Office of Cyber Monitoring and Operations

- ❖ Discussed change management and onboarding processes
- ❖ New software goes through engineering management process to ensure it complies with standards/"meets baseline"

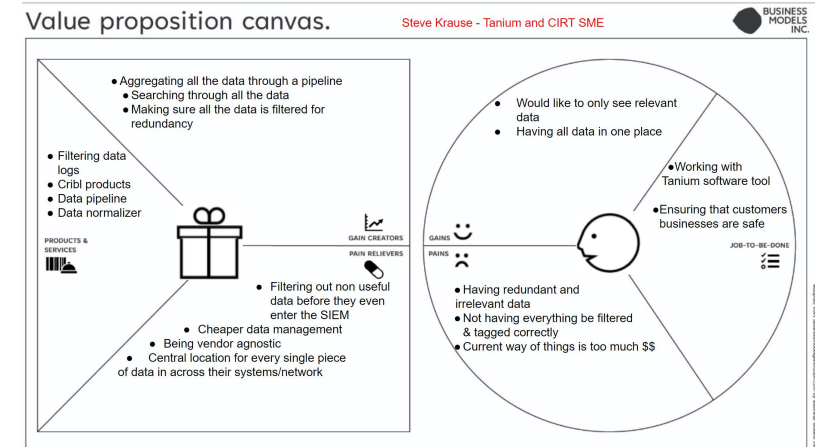
**"Walk through the critical path with milestones."
- Roy Matthews**

Other Important VPCs

Rob



Steve Krause



Weeks 8-9: Mission Model Canvas

The Mission Model Canvas

Mission/Problem Description:

Designed by:

Date: Week 8

Version:

<p>Key Partners </p> <p>Sponsors: Nick Swindell, IT Specialist Danh Nguyen-Huynh, Technical Director Jake Trigoboff, CIRT Cloud Lead</p> <p>Professors at RIT</p> <p>Cybersecurity Directors</p> <p>Security Analysts</p>	<p>Key Activities </p> <p>Create ways for data uniformity across many teams</p> <p>Provide solutions for availability of data to CTS</p> <hr/> <p>Key Resources </p> <p>Cloud Systems</p> <p>Security Information Event Management (SIEM) Tools</p> <p>Enterprise Resource Planning tools (ERP) Data</p> <p>Analytics/visualization software</p>	<p>Value Propositions </p> <p>CTS will have a better ability to secure the DOS infrastructure and visualize security data (ALL)</p> <p>Better data visualization will lead to faster and more effective incident response (ALL)</p>	<p>Buy-in & Support </p> <p>Emphasize the importance of security to other teams in DOS.</p> <p>Demonstrate how data visualization improves response times and methods</p> <hr/> <p>Deployment </p> <p>Data uniformity and availability</p> <p>Deploying a Beta Version</p> <p>Data visualization charts/deliverables</p>	<p>Beneficiaries </p> <p>Roy Matthews, Division Chief</p> <p>Jose Rivera-Ortiz, TASO Tech Lead, & TASO team</p> <p>Karl Crandall, CIRT Tech Lead, & CIRT team</p> <p>David Jacobs, Engineering Tech Lead, & Engineering team</p> <p>Carl Wyatt, Cyber Protection Branch Chief & Office of Cyber Monitoring and Operations</p>
<p>Mission Budget/Cost </p> <p>Cost implementation of a new or updated system. Ongoing hosting and maintenance costs.</p>		<p>Mission Achievement/Impact Factors </p> <p>CTS has widespread access to necessary company data Properly secure the network infrastructure.</p> <p>Reduced Cost of data storage </p> <p>Increased level of metrics to make decisions </p> <p>Faster access to data </p>		

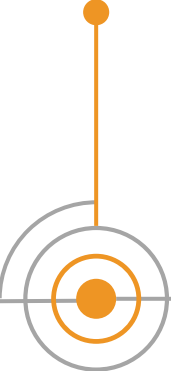
Project Timeline

Weeks 1 - 5



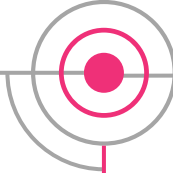
Total Interviews: 17
Research Papers: 20
Days Left: 70

Total Interviews: 32
Research Papers: 28
Days Left: 56



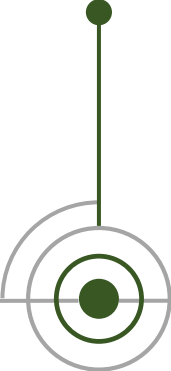
Weeks 6 - 7

Weeks 8 - 9



Total Interviews: 77
Research Papers: 36
Days Left: 42

Total Interviews: 91
Research Papers: 50
Days Left: 0



Weeks 10 - 15

Weeks 10-15

- ❖ Created proper Gantt Deployment Chart
- ❖ Finalized our MVP with sponsors
- ❖ Interviews focused on following areas:
 - ❖ Disaster Recovery & Business Continuity
 - ❖ Onboarding Procedures
 - ❖ Risk Management



Weeks 10-15: Interviews

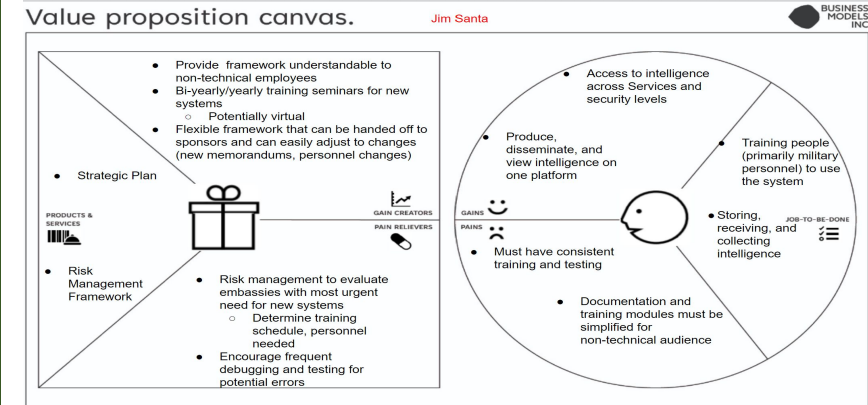
Jake Trigoboff - CIRT Cloud Lead, Office of Cyber Monitoring and Operations

- ❖ Onboarding and introduction phases
- ❖ Collaboration with Technical and Management
- ❖ KPIs are focused on incident statistics (types, frequency, logging requirements)

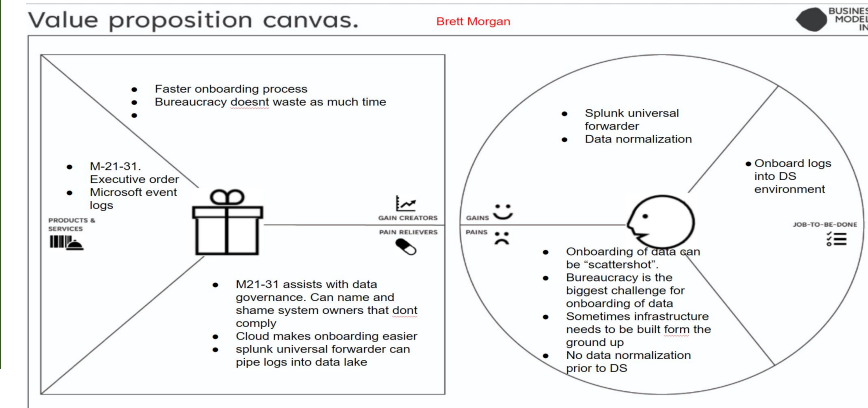
"Once we figure out flow of change management having some integrations between management and technical side will be important."
- Jake Trigoboff

Other Important VPCs

Dr. Jim Santa



Brett Morgan



Weeks 10-15: Mission Model Canvas
















The Mission Model Canvas

Mission/Problem Description:

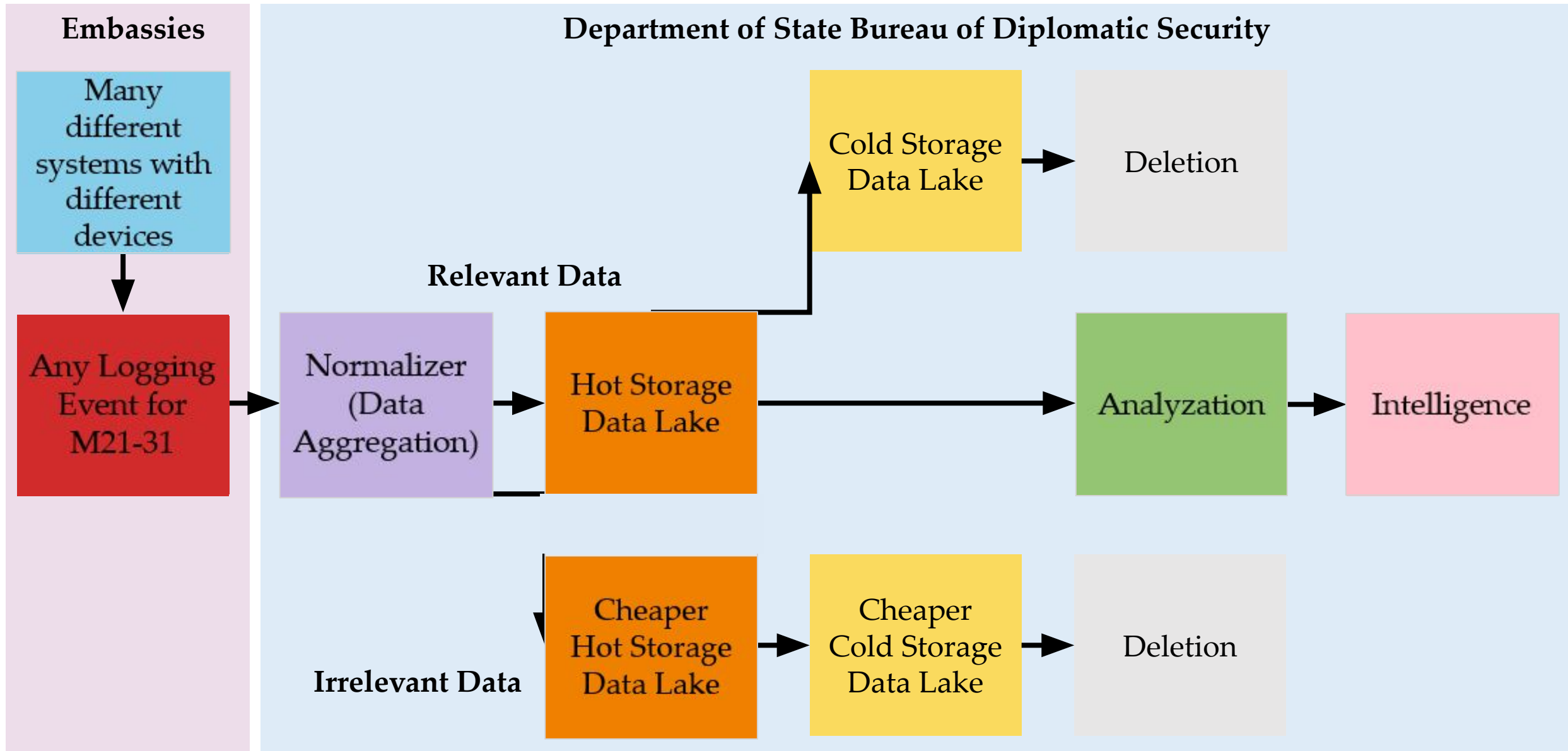
Designed by:

Date: Week 13

Version:

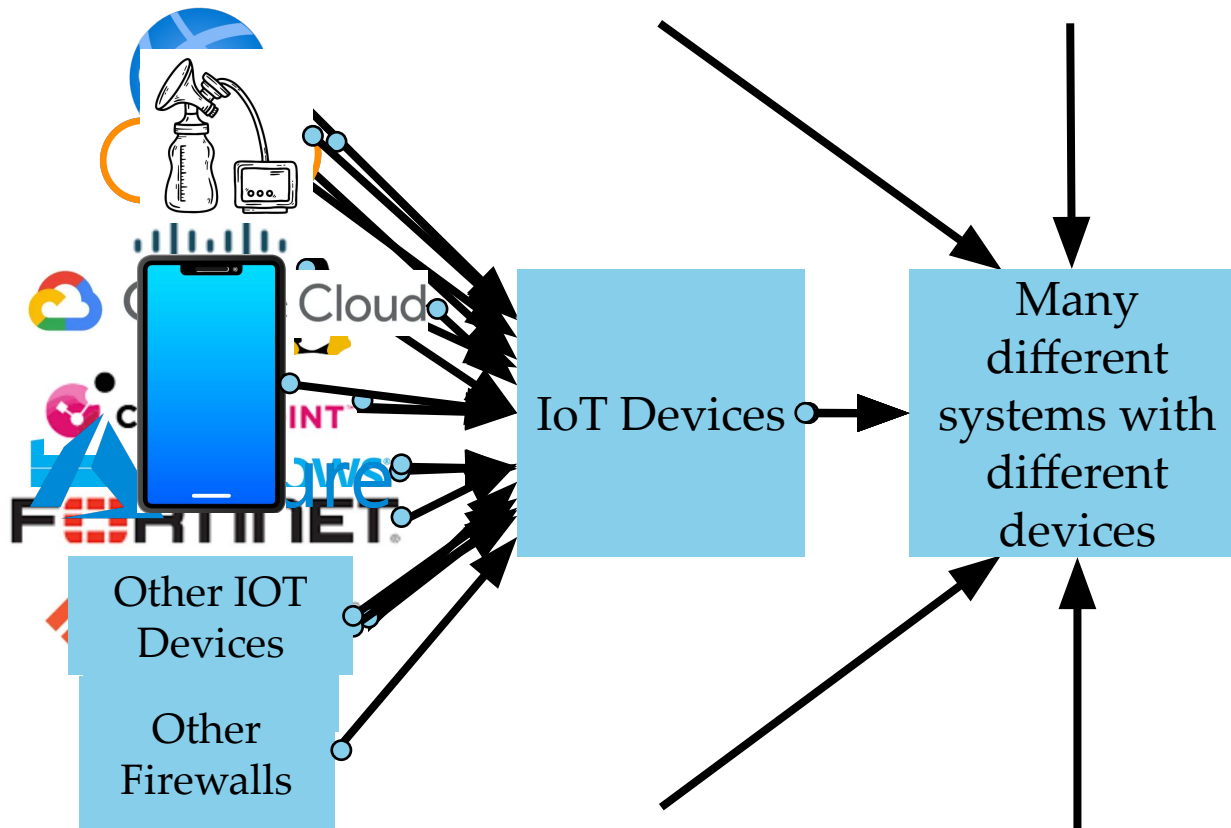
<p>Key Partners </p> <p>Sponsors: Nick Swindell, IT Specialist Danh Nguyen-Huynh, Technical Director Jake Trigoboff, CIRT Cloud Lead</p> <p>Professors at RIT</p> <p>Cybersecurity Directors & Managers</p> <p>Security Analysts</p> <p>Consultants</p> <p>Engineers</p>	<p>Key Activities </p> <p>Create ways for data uniformity across many teams</p> <p>Provide solutions for availability of data to CTS</p> <hr/> <p>Key Resources </p> <p>Cloud Systems</p> <p>Security Information Event Management (SIEM) Tools</p> <p>Enterprise Resource Planning tools (ERP) Data</p> <p>Analytics/visualization software</p>	<p>Value Propositions </p> <p>CTS will have a better ability to secure the DOS infrastructure and visualize security data </p> <p>Better data visualization will lead to faster and more effective incident response </p>	<p>Buy-in & Support </p> <p>Emphasize the importance of security to other teams in DOS.</p> <p>Demonstrate how data visualization improves response times and methods</p> <hr/> <p>Deployment </p> <p>Data uniformity and availability</p> <p>Deploying a Beta Version</p> <p>Data visualization charts/deliverables</p>	<p>Beneficiaries </p> <p>Roy Matthews, Division Chief</p> <p>Jose Rivera-Ortiz, TASO Tech Lead, & TASO team</p> <p>Karl Crandall, CIRT Tech Lead, & CIRT team</p> <p>David Jacobs, Engineering Tech Lead, & Engineering team</p> <p>Carl Wyatt, Cyber Protection Branch Chief & Office of Cyber Monitoring and Operations</p>
<p>Mission Budget/Cost </p> <p>OMA: The allocation of 20hrs/wk of 1 current FTE engineer's time for system maintenance. OPA: Cost implementation of a new or updated system. Ongoing hosting and maintenance costs. Storage: \$500,000+ monthly SIEM tool: \$10,000 monthly Data aggregation/Normalizer tool: \$34,000 monthly Management Solution: \$200,000 monthly</p>		<p>Mission Achievement/Impact Factors </p> <p>Access to all Memorandum 21-31 logs </p> <p>Capable storage for all logs </p> <p>Vendor Agnosticity </p> <p>Data Governance support and stakeholder buy-in </p>		

Week 10 - Final MVP



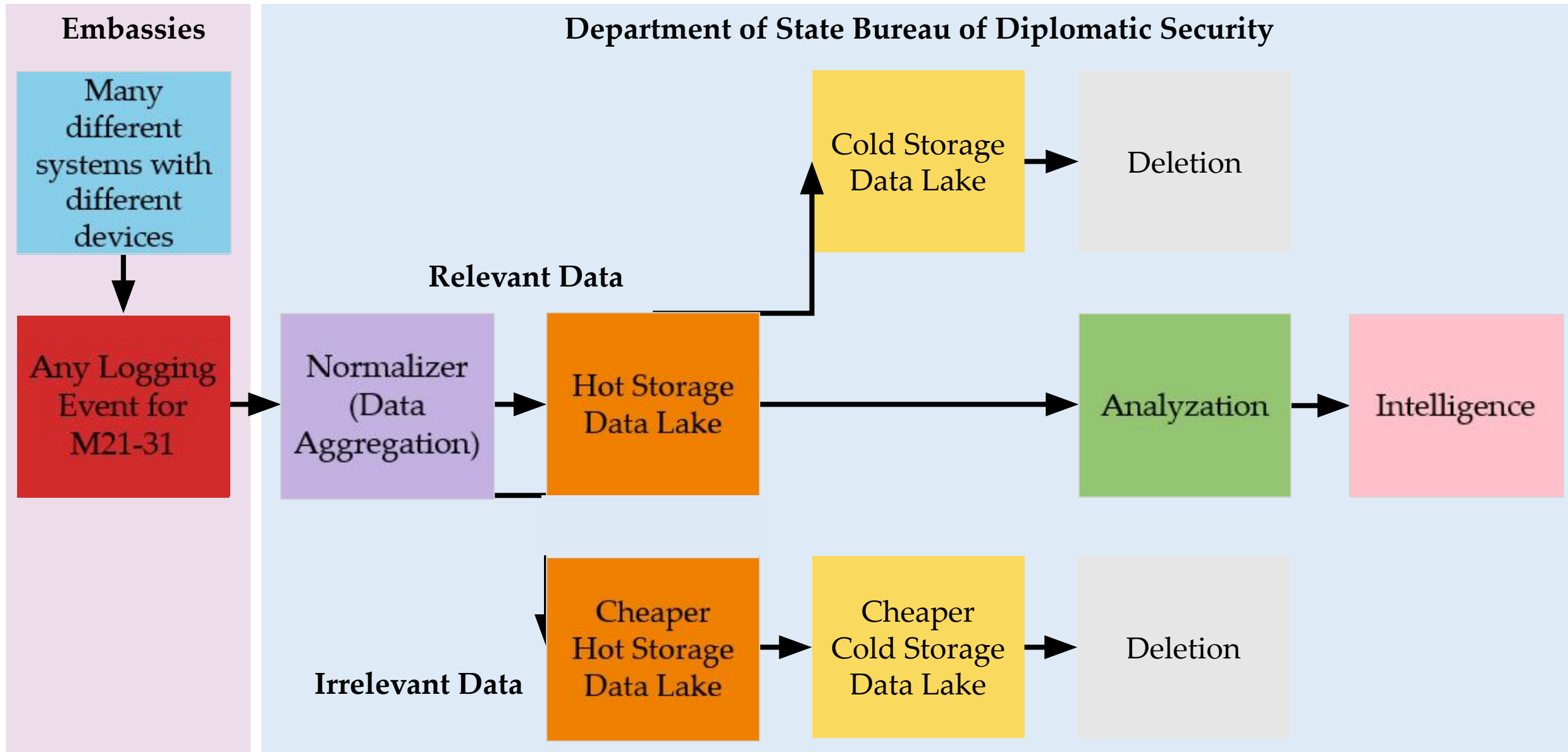
Step 1 - Sources

Embassies Side



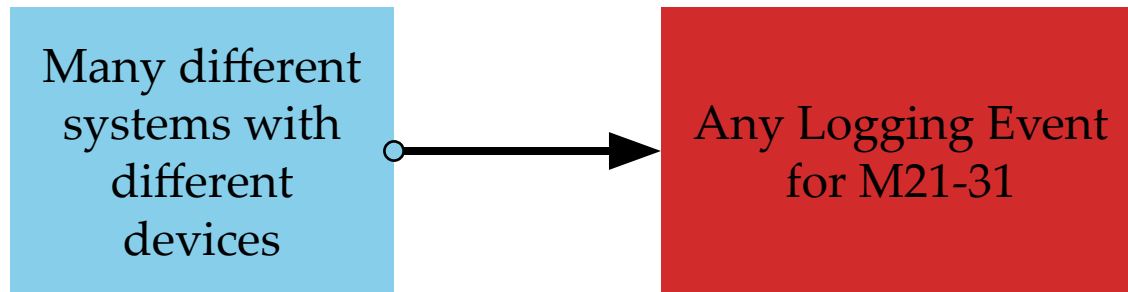
- ❖ Each embassy will collect all of their source logs
- ❖ Every data source will be categorized by its type
- ❖ Including all
 - ❖ Operating Systems
 - ❖ Networking Devices
 - ❖ Firewalls
 - ❖ Cloud Devices
 - ❖ IoT Devices

Week 10 - Final MVP



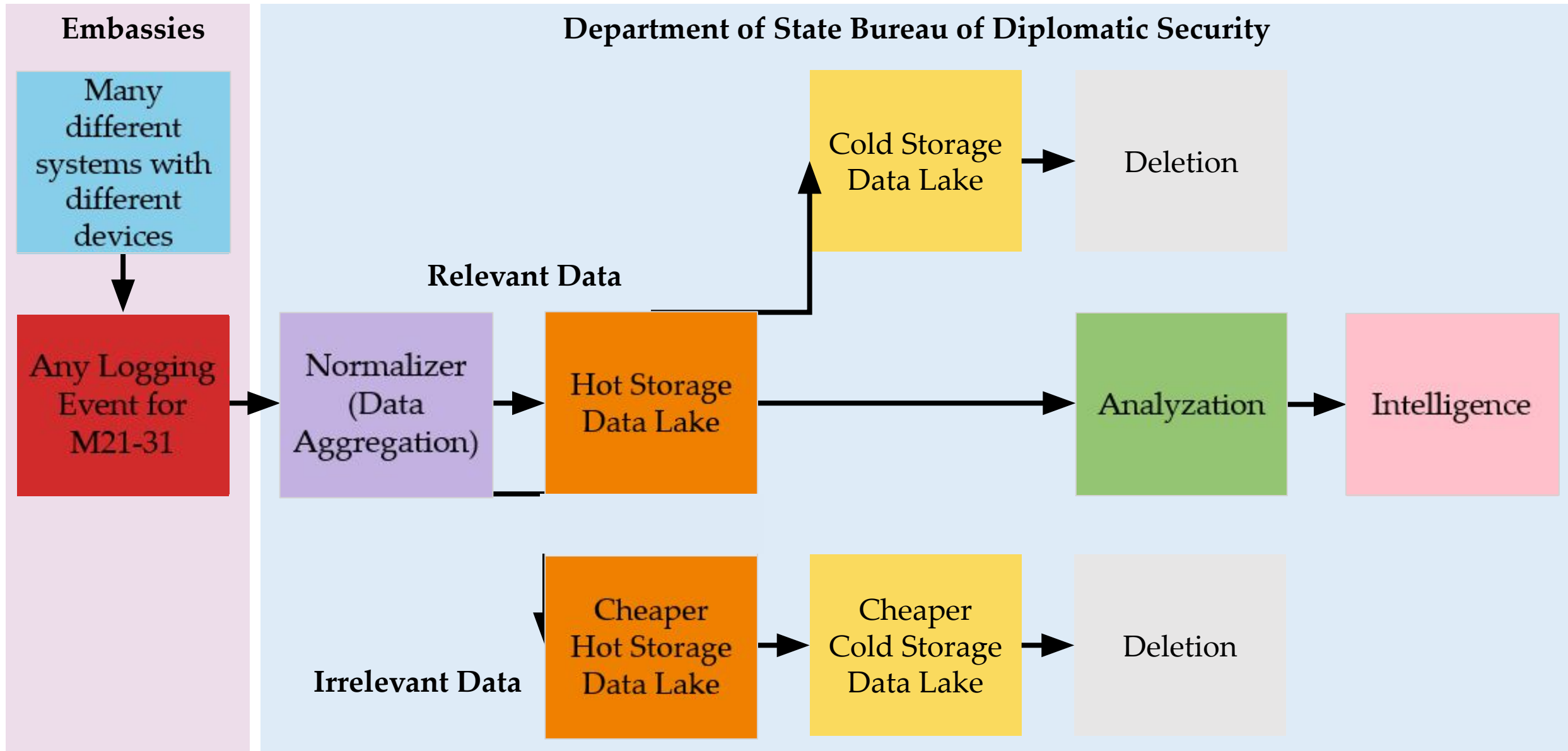
Step 2 - M21-31 Logging Event

Embassies Side

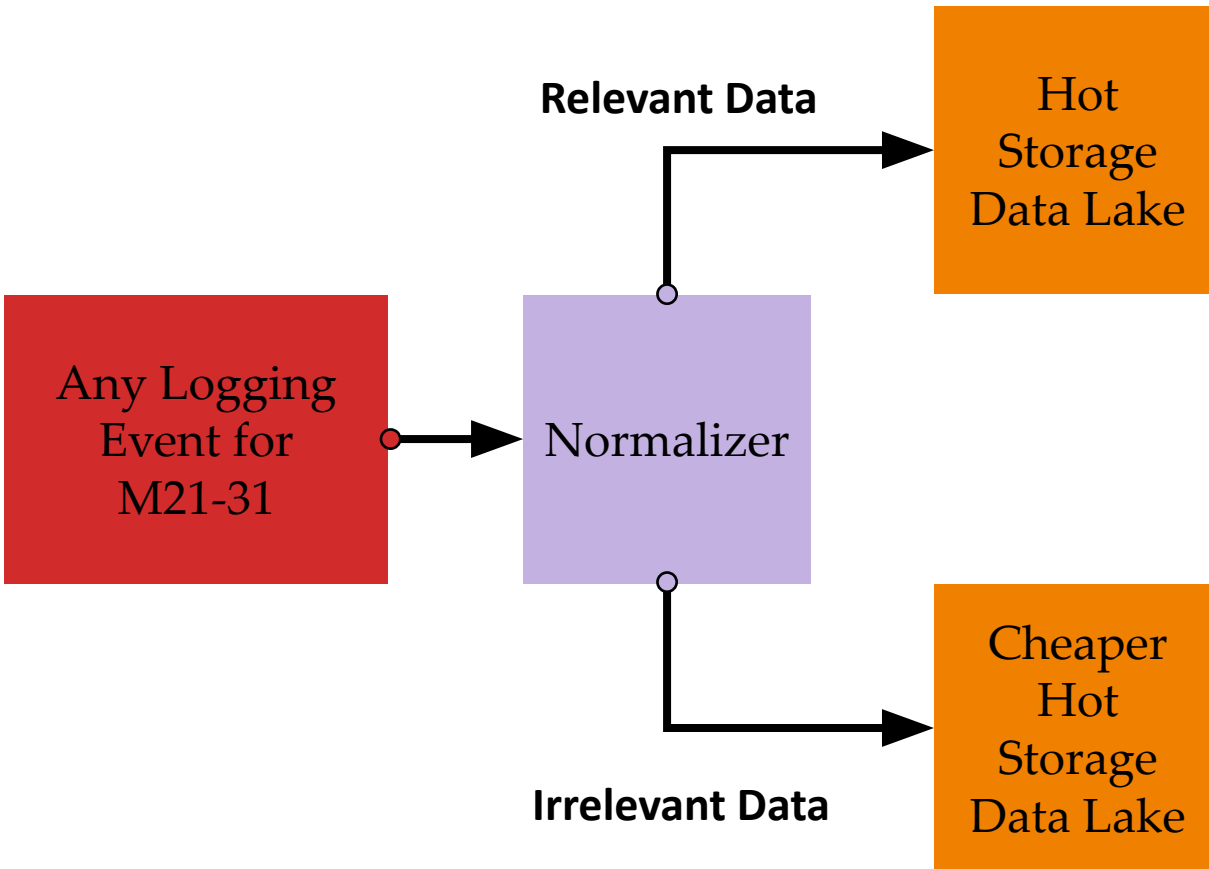


- ❖ All gathering of data & logs must be compliant to Memorandum 21-31
- ❖ Developed to ensure centralized visibility for Security Operations Center (SOC) of federal agencies.
- ❖ It addresses:
 - ❖ Logging
 - ❖ Log retention
 - ❖ Log management

Week 10 - Final MVP



Step 3 - Normalizer



- ❖ Categorized logs are sent to the normalizer
- ❖ Makes sure logs will be tagged with its source
- ❖ Normalizer aggregates all the data and logs
- ❖ Analyze and classify each log as relevant or irrelevant
 - ❖ **Relevant:** This log will help in investigation and is useful
 - ❖ **Irrelevant:** This log may not help, not have useful info and potentially wastes space

Recommend Software Solutions – Normalizer



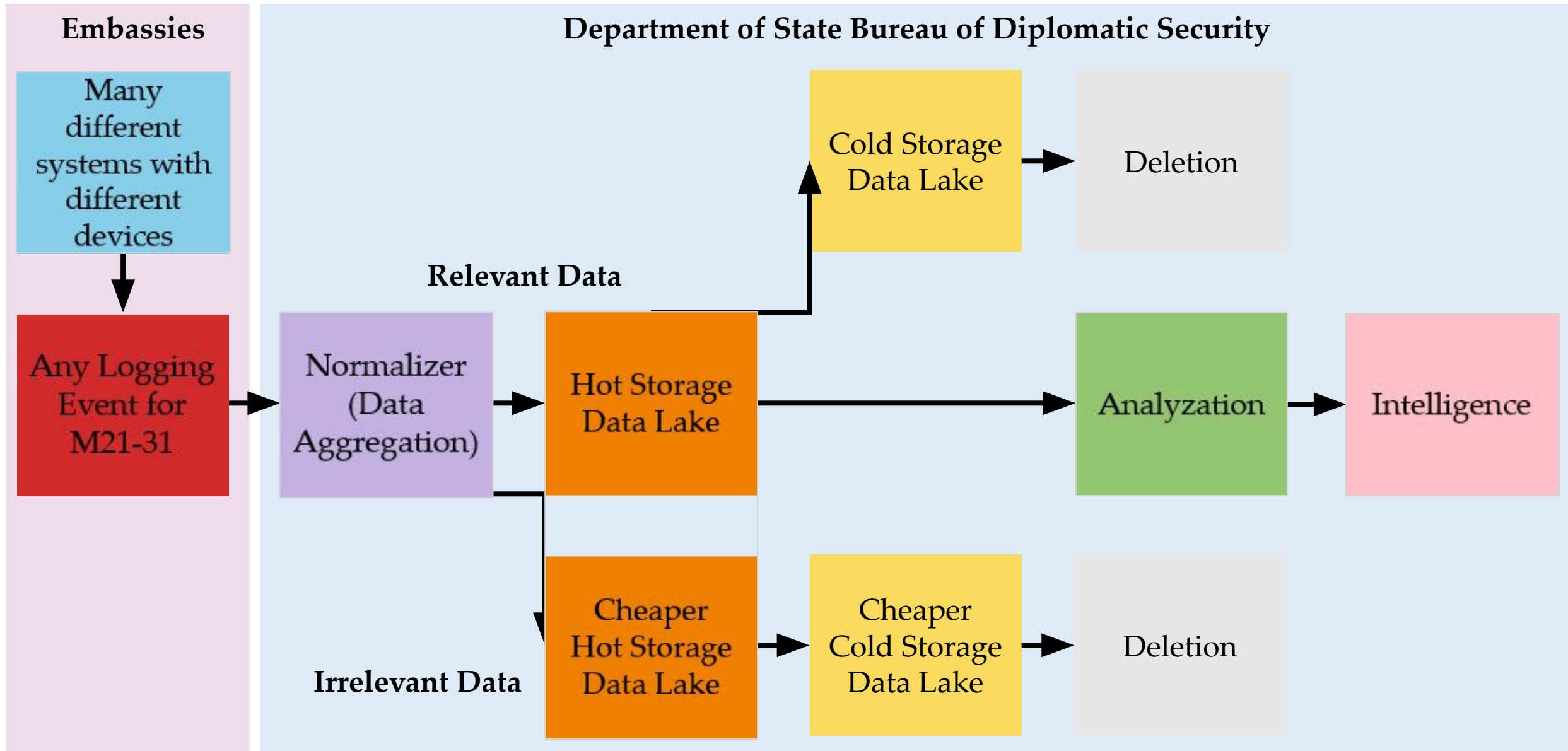
Pros:

- ❖ Universal Receiver
- ❖ Dashboard – Easy to understand/use, can click and drag sources to destinations
- ❖ Integration with other software including other recommended software.

Cons:

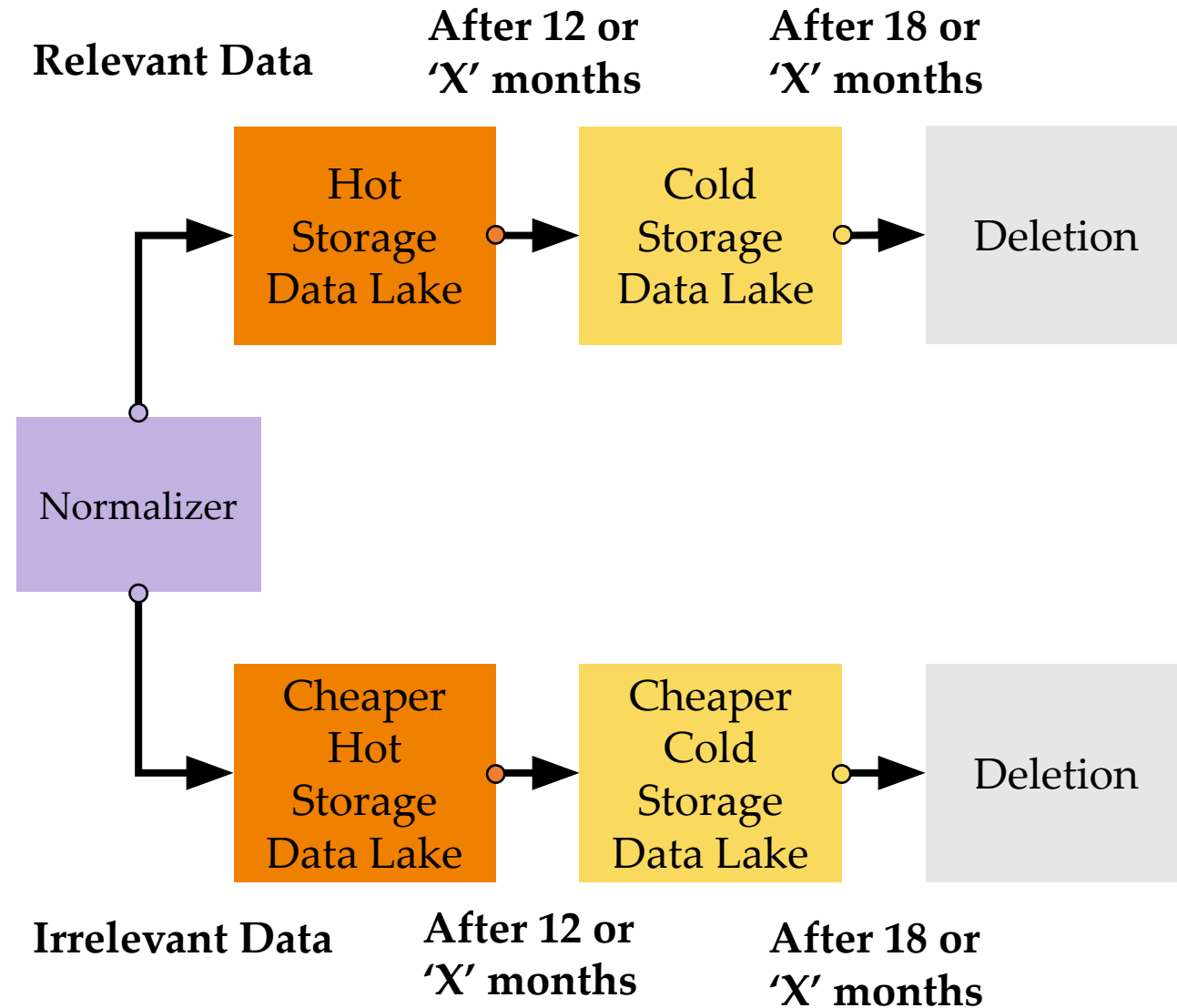
- ❖ No Artificial Intelligence or Machine Learning functionality

Week 10 - Final MVP



Steps 4, 5, & 6 - Data Storage

DoS Side



- ❖ Data will be stored in a Data Warehouse in the Department of State
- ❖ Relevant data and irrelevant data will be split up into different data lakes for cost and relevance
- ❖ Mandated in M21-31:
 - ❖ After 12-'X' months in hot storage, data will be moved to cold storage
 - ❖ After 18-'X' months in cold storage, data will be deleted



Recommend Software Solutions – Data Storage



Pros:

- ❖ Pricing based on computing usage
- ❖ Auto-scaling
- ❖ Real-time data lineage

Cons:

- ❖ Only at-rest encryption on data



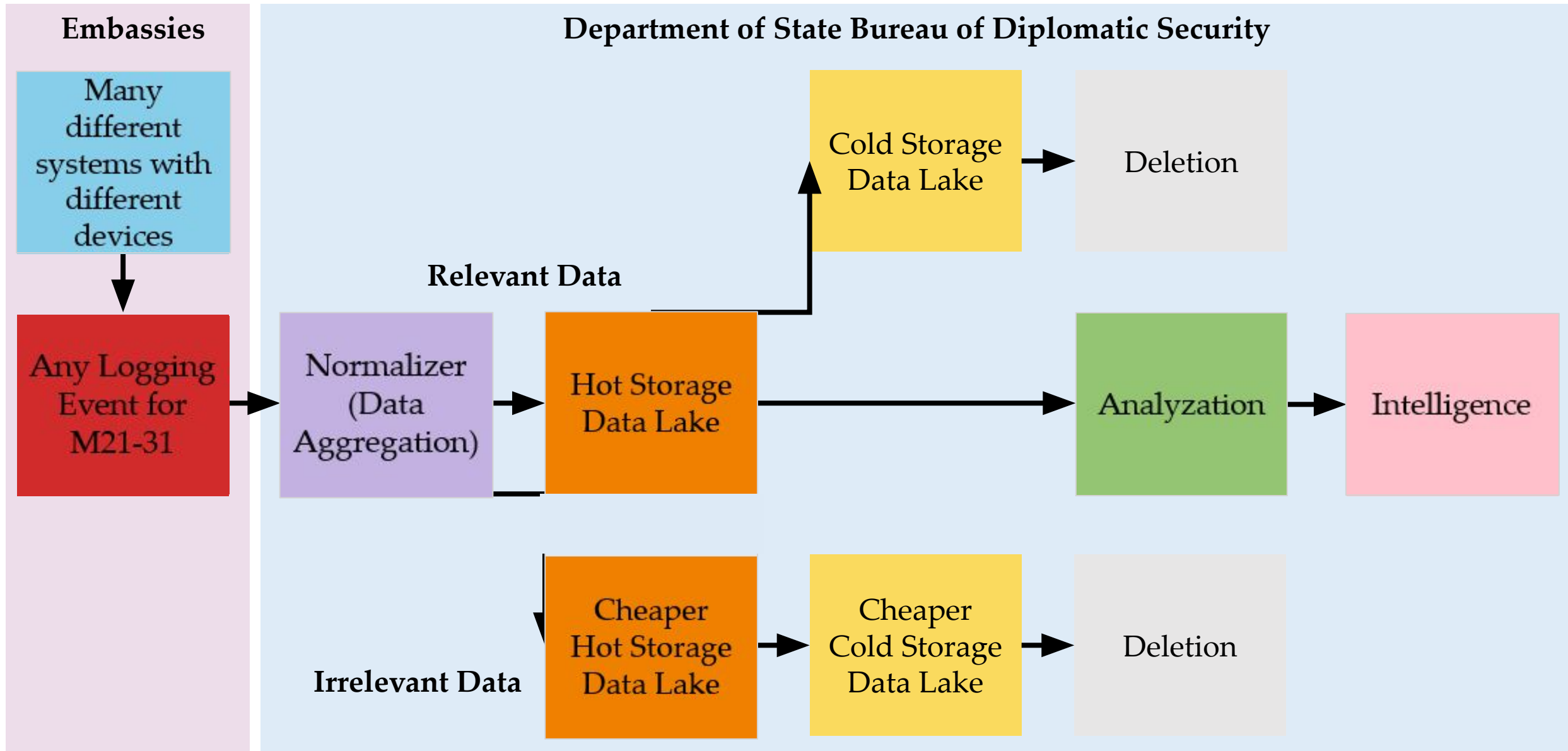
Pros:

- ❖ At-rest and in-transit encryption on data
- ❖ Good scaling capabilities
 - ❖ Easy to increase and decrease size of data warehouse based on needs

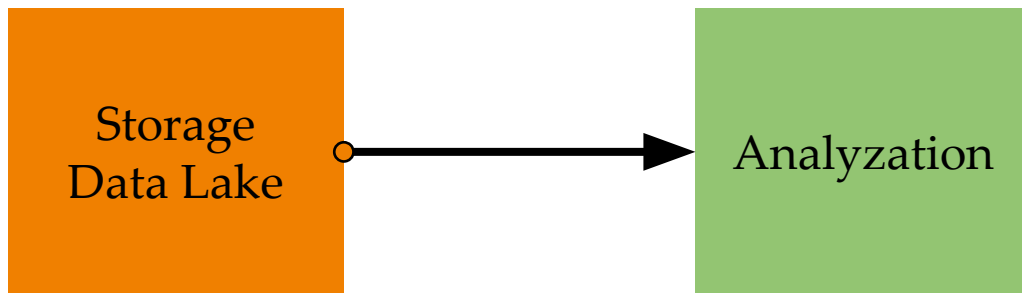
Cons:

- ❖ Pricing based on storage volume
- ❖ No auto-scaling – not as fast

Week 10 - Final MVP



Step 7 - Analyze



- ❖ The SIEM tool will pull the data from the data lakes
- ❖ The SIEM tool will display all the data to showcase what is happening on the network & sources

Recommend Software Solutions – Analysis



Pros:

- ❖ Current SIEM tool
- ❖ Well integrated data collection and analysis

Cons:

- ❖ Current contract is expensive
- ❖ Not easy to get data out of Splunk once it is put in (indexed)



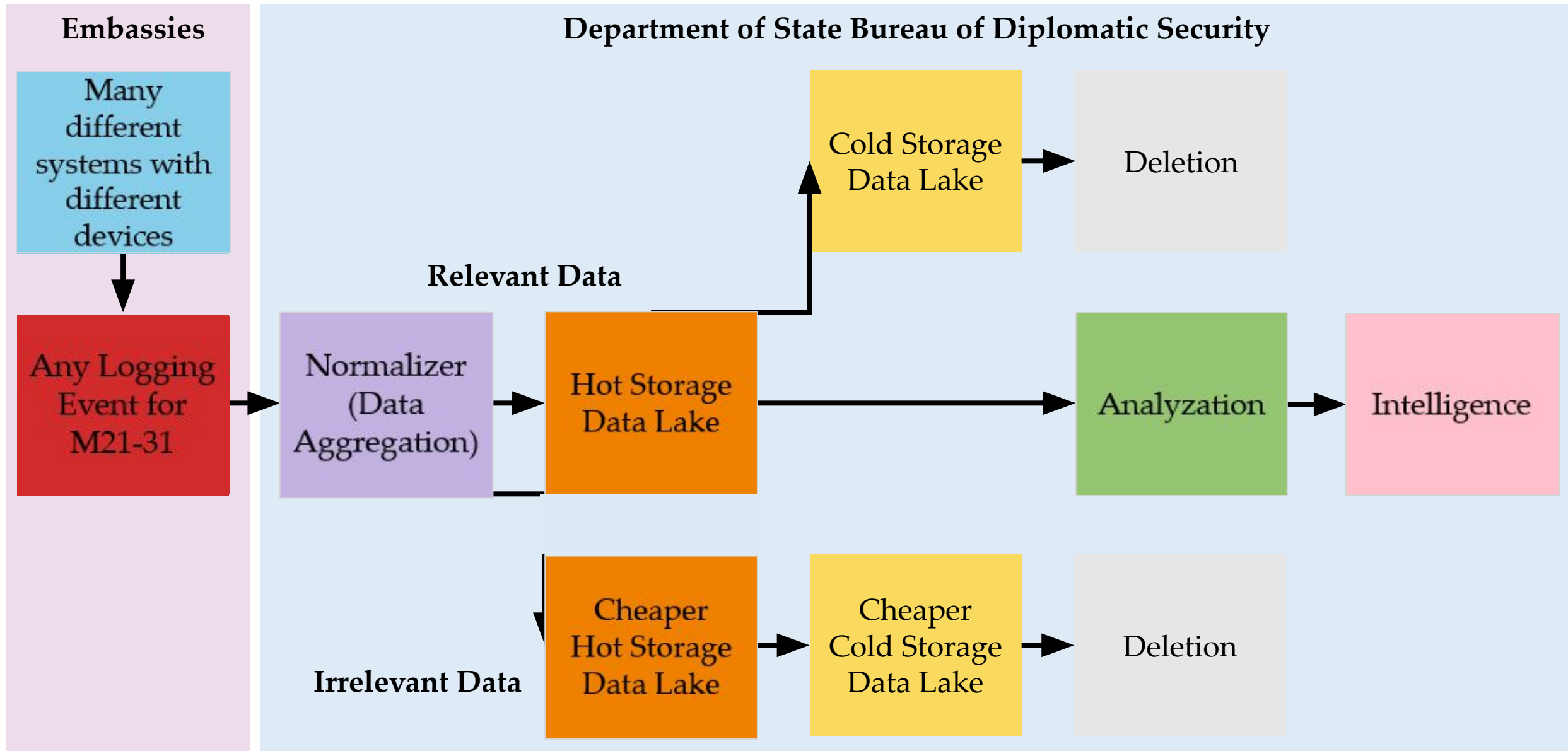
Pros:

- ❖ Use of historical data (can bring from cold to hot)
- ❖ Artificial Intelligence and Machine Learning
 - ❖ Includes behavioral analysis data and task automation

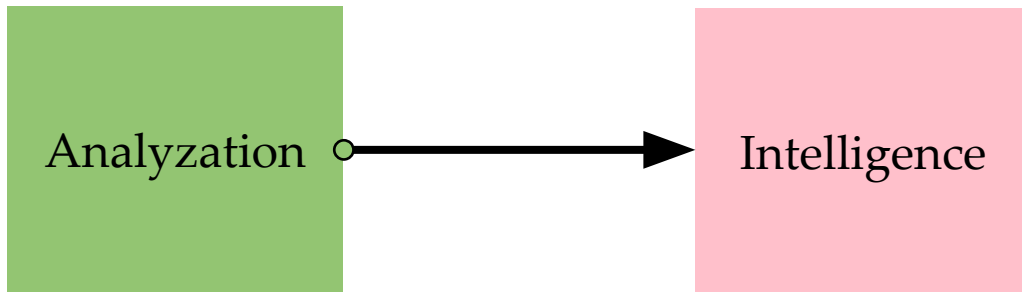
Cons:

- ❖ Complex pricing system

Week 10 - Final MVP

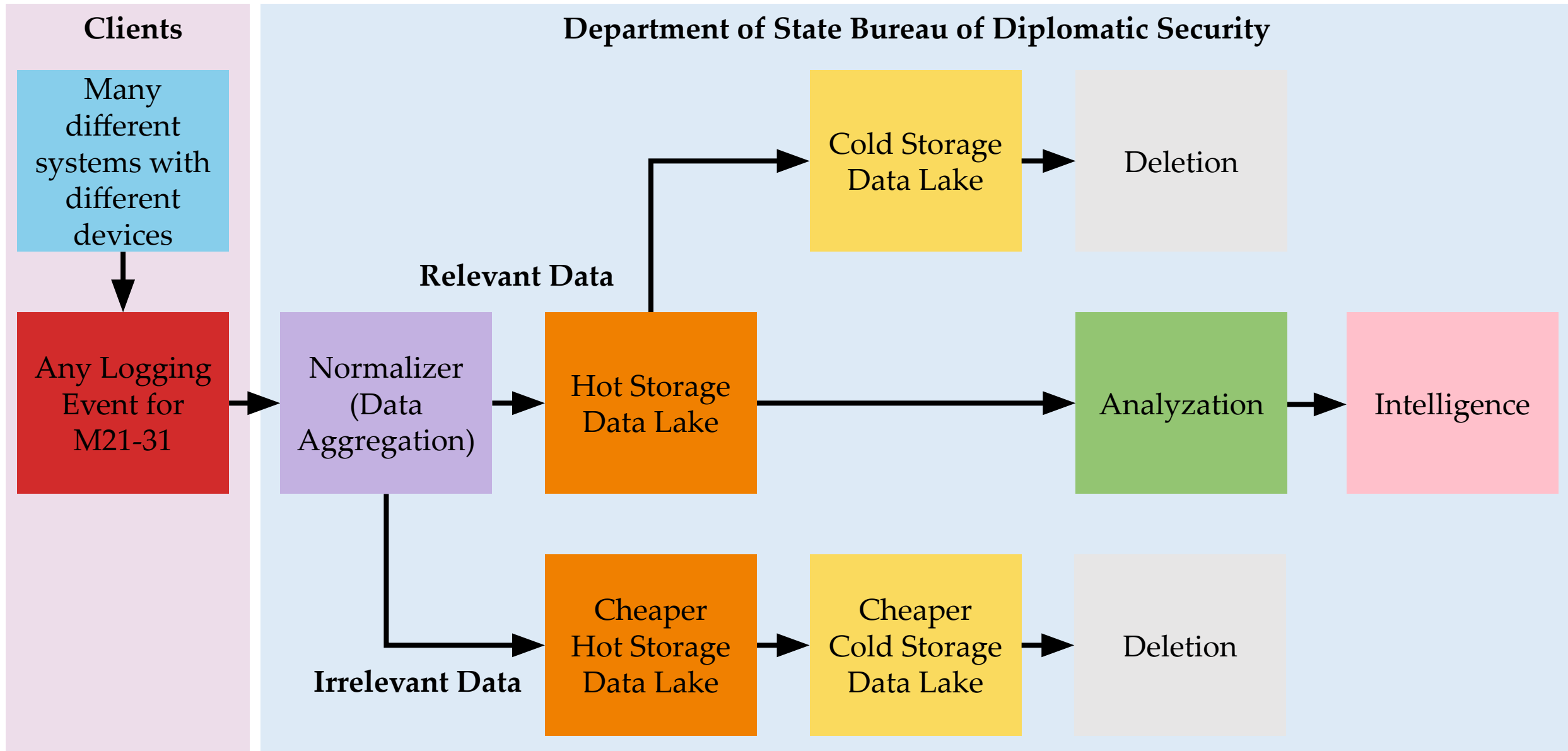


Step 8 - Intelligence



- ❖ Analysts at DoS will analyze & examine the data for any malicious incidents.
- ❖ DoS Analysts will work with the embassies to remediate the issue.
- ❖ Once the issue is resolved, embassies and the DoS Analysts will make sure to prevent similar events from happening in the future

Week 10 - Final MVP



Onboarding Notes

- ❖ Annual training highlighting new features, software updates, other changes
- ❖ Make sure current equipment can handle project
- ❖ Determine implementation on specific embassies with key metrics

**"Training on software should be a continuous, long-term process."
- Bryan Reinicke, MIS Capstone Professor**

Disaster Recovery Notes

- ❖ Ensure any new software additions meet security standards and do not increase vulnerability.
- ❖ How long are servers active? 24 hours?
- ❖ What is risk tolerance?

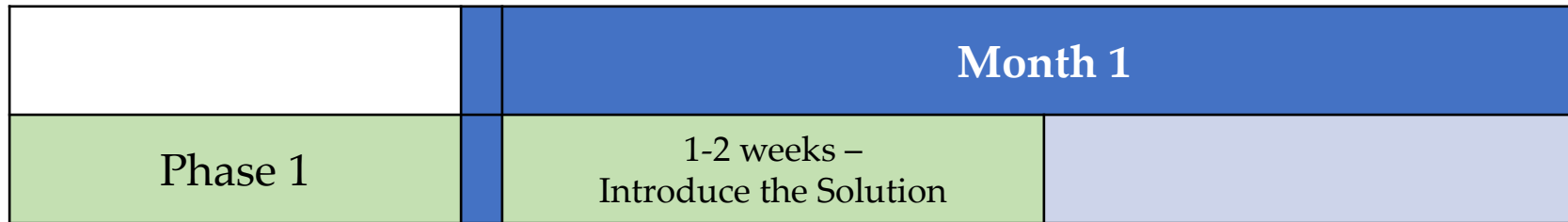
“Backups, backups, backups!!!”
– Paul Centanni, CISO at Acture Solutions

Deployment – Overview

Approximate Total Time Length: 10.5 Months – 14 Months



Deployment – Phase 1



-
- ❖ The focus of Phase 1 be introducing the solution to sponsors and senior management.
 - ❖ Share research on software
 - ❖ Develop estimates of implementation time
 - ❖ Outlining risk management process, key performance indicators, and goals.

Deployment – Overview

Approximate Total Time Length: 10.5 Months – 14 Months



Deployment – Phase 2

	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10	
Phase 2			9 Months: Approximate Full Approval of Project								
		8 – 9 Weeks	← Meet with Sponsors and Get Initial Approval								
			24 – 26 Weeks: Get Solution Budgeted								
			Resource Allocations for Implementation →						8 – 9 Weeks		
			24 – 35 Weeks: Get Approval from Senior Management								

-
- ❖ The focus of Phase 2 will be approval and resource allocation.
 - ❖ Official project approval from senior management
 - ❖ Budgeting and cost projection
 - ❖ Resource allocation
 - ❖ Hardware, software, personnel

Deployment – Overview

Approximate Total Time Length: 10.5 Months – 14 Months



Deployment – Phase 3

	Month 1	Month 2 – Month 9	Month 10	Month 11	Month 12	Month 13	Month 14
Phase 3			2 – 3 Weeks	← Build Out Infrastructure			
				2 – 3 Weeks	← Coordination and Discussion with Customers		
	Employee Training and Onboarding of New Systems & Tools →			2 – 3 Weeks			
	Onboard Log Sources, Collaborating with Customer’s Engineers, Pushing Customer’s Data Down, Incorporate the Data Aggregator →				4 – 10 Weeks		

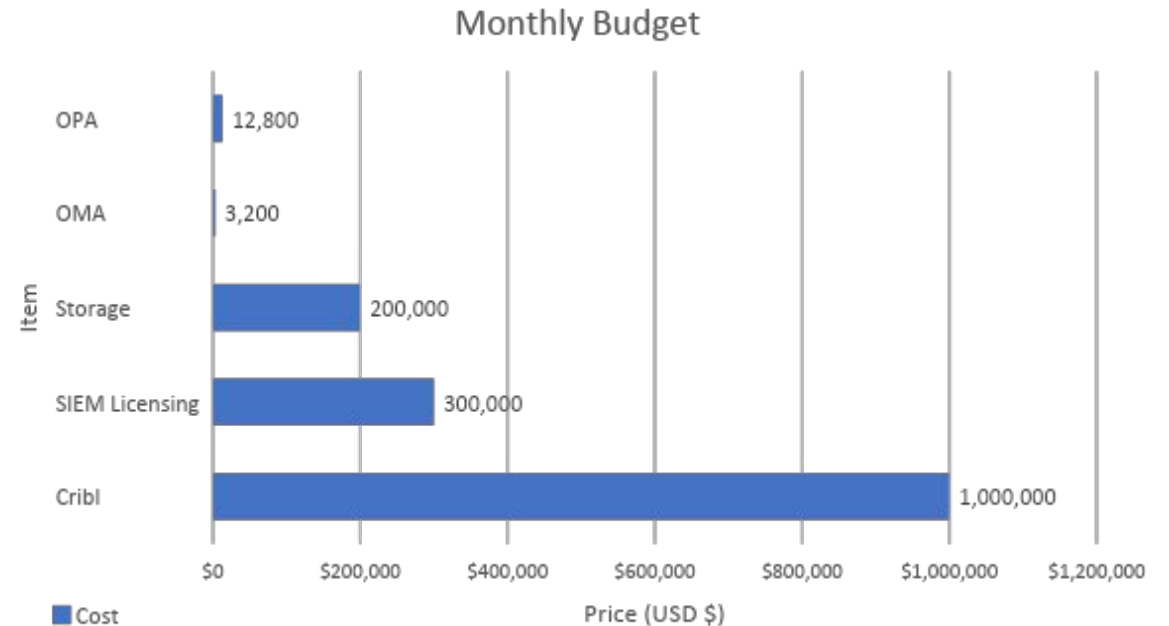
-
- ❖ The focus of Phase 3 will be onboarding and deployment.
 - ❖ Onboarding logs and data
 - ❖ Implementing new software
 - ❖ Training employees

Monthly Budget

Implementation:

- ❖ Data Normalizer software (Cribl): \$1,000,000
- ❖ Cost of licensing and use of Data Lake Storage: \$200,000
- ❖ SIEM Licensing and use: \$250,000
- ❖ OMA cost of system maintenance per month: \$3,200
- ❖ OPA cost for system implementation: \$12,800

Total: \$1,516,000



Special Thanks

Thank you, Nick Swindell, Danh Nguyen-Huynh, and Jake Trigoboff, and our other Sponsors!



Special Thanks

Thank you to all the people we interviewed!

Manny Medrano
Ozan Ertugrul
Carlos R. Rivero
Christine Shely
Alex McPherson
Gharun Lacy
Andy Meneely
David Hagan
Danh Nguyen-Huynh
Bob LaBanz
Ken Miller
Brian Bullis
Tom Kopchak
David Kirk
Sara Kastner
Bryan Reinicke
Dave Ballard
Roy Matthews
Ray Romano
Mohammed Saidur
Quang Bui
Justin Balroop
James Bridgen
Ali Tosyali
David Loshin
Steve Krause
John Topp
Nick Swindell
Chad Rooney
Michael Wofford
Jim Santa
Bill Stackpole
Myra Rowell
Rob
Anthony Henry
Paul Centanni
Sean Doran
Jon-Michael Lacek
Carl Randall
Mark Johnson
Mehdi Mirakhorli
Jose Rivera-Ortiz
Nick Ortiz
Jake Trigboff
Rob Naik
Michael Kelly
Demetrius J Gooden
Rob Mennell
Ian James
Mike Pinch
Nate Matthews
Saikat Biswas
Bob Adams
Maxwell Baron
James H. Moore
Brett Morgan

Special Thanks

Thank you to our mentor Rob Leonard!



Special Thanks

Thank you, Jim Santa and Suvam Barui!





Questions?