## Scenario #1: Drowned in the Data Lake?

You are a new CDataO at Acme Software Solutions Corp., a firm that provides a multitude of online services that businesses of all shapes & sizes use to varying degrees (i.e. a given customer may not utilize all your services; some are huge, multinational organizations, others may only have 10 total employees). Some of these services ingest and process sensitive customer business operations data. Other services only work with “public” customer business information from data aggregators.

All these services were designed and rolled out over time by distinct business units. There is a new corporate strategy to unify the “customer experience” for those customers that utilize multiple online services of your firm which will require data access, processing, and interoperability between these business units and the online services.

Along with the customer data, there is metadata on how customers use the service that is collected by the individual business units and analyzed on an ad-hoc basis by some of them. Part of the new corporate mandate is the need for senior management to have a “big picture” view of what types of customers are using which components of each service, along with how much revenue each service makes as compared with the compute and storage costs.

Finally, the head of sales has been given the mandate to create more realistic and consolidated new sales and renewal forecasts from the disparate CRM, lead generation, and support ticket systems, spread across the existing business units.

You’re getting “pressure” from the matrixed “data team” to “make life easy” and funnel all this data into one, giant “data lake”.

This organization has an in-house General Council (+ small legal team) and a CISO that also runs the compliance functions.

We’ll walk through the following questions together:

Now that you’ve been introduced to the modern threats/risks/compliance areas:

* What high-level concerns would you have about building this giant lake of data?
* Walk through your answers for the four questions in the context of this scenario:
  + Which type of attacker might be interested in the C•I•A of this data/process?
  + What tactics and techniques might they use to achieve their anti-C•I•A goals?
  + What design elements can we build in to prevent attackers from achieving these goals and/or detect what they’re doing/have done?
  + What must we do to avoid violating XYZ compliance framework?

## Scenario #2: Classify This!

One of the services Acme provides is invoice reconciliation and anomaly detection. This is an API and supports interactive customer logins to gain access to a “portal” for the service.

Another service is customer list “clean up” — entity resolution, de-duplication, and enhancing with more individual metadata. All this processing happens in Amazon AWS. Customers upload data in batches and download results in batches via an online “portal”.

For the internal sales team, all customer metadata, transactions, locations, usage information will be consolidated.

Pick one of these situations and work through the eight areas:

1. Determine how much protection your information needs
2. Collect only what is necessary
3. Provide minimum necessary access
4. Disclose only the minimum information necessary
5. Safeguard information in transit
6. Secure physical equipment and resources
7. Safeguard information in storage
8. Dispose of information securely when no longer needed

and discuss your data storage, access, classification, processing, and retention “plan”. (NOTE: there is a key term used in the first scenario description that you *might* need to consider here).

Classification example (from the presentation):

* **Confidential** — High risk of significant financial loss, legal liability, public distrust or harm if this data is disclosed.
* **Sensitive** — Moderate requirement for Confidentiality and/or moderate or limited risk of financial loss, legal liability, public distrust, or harm if this data is disclosed.
* **Public** — Low or no requirement for Confidentiality; low or insignificant risk of financial loss, legal liability, public distrust or harm if this data is disclosed.

## Scenario #3: To the Cloud!

Senior management wants to save some $$$ and has put forth a mandate to “move operations to the cloud”. (Yes, you will hear mandates with that generality. Sigh.)

Use the situation you chose from Scenario #2, choose what type of “aaS” you’d recommend along with why you chose it:

* IaaS — Infrastructure
* PaaS — Platform
* SaaS — Software
* FaaS — Function

What would “keep you up at night” if the organization pressed “go” on your recommendation? What events would have to take place for you to pull back this service from a cloud deployment?