Data Quality and Data Security

Network Security, Fall 2019

Seth James Nielson

Part IV
Security in a Data Context

- Data Governance
- Data Quality

Data Governance

Data Governance:
is the organization and implementation of policies, procedures, structure, roles, and responsibilities that outline and enforce rules of engagement, decision rights, and accountabilities for the effective management of information assets

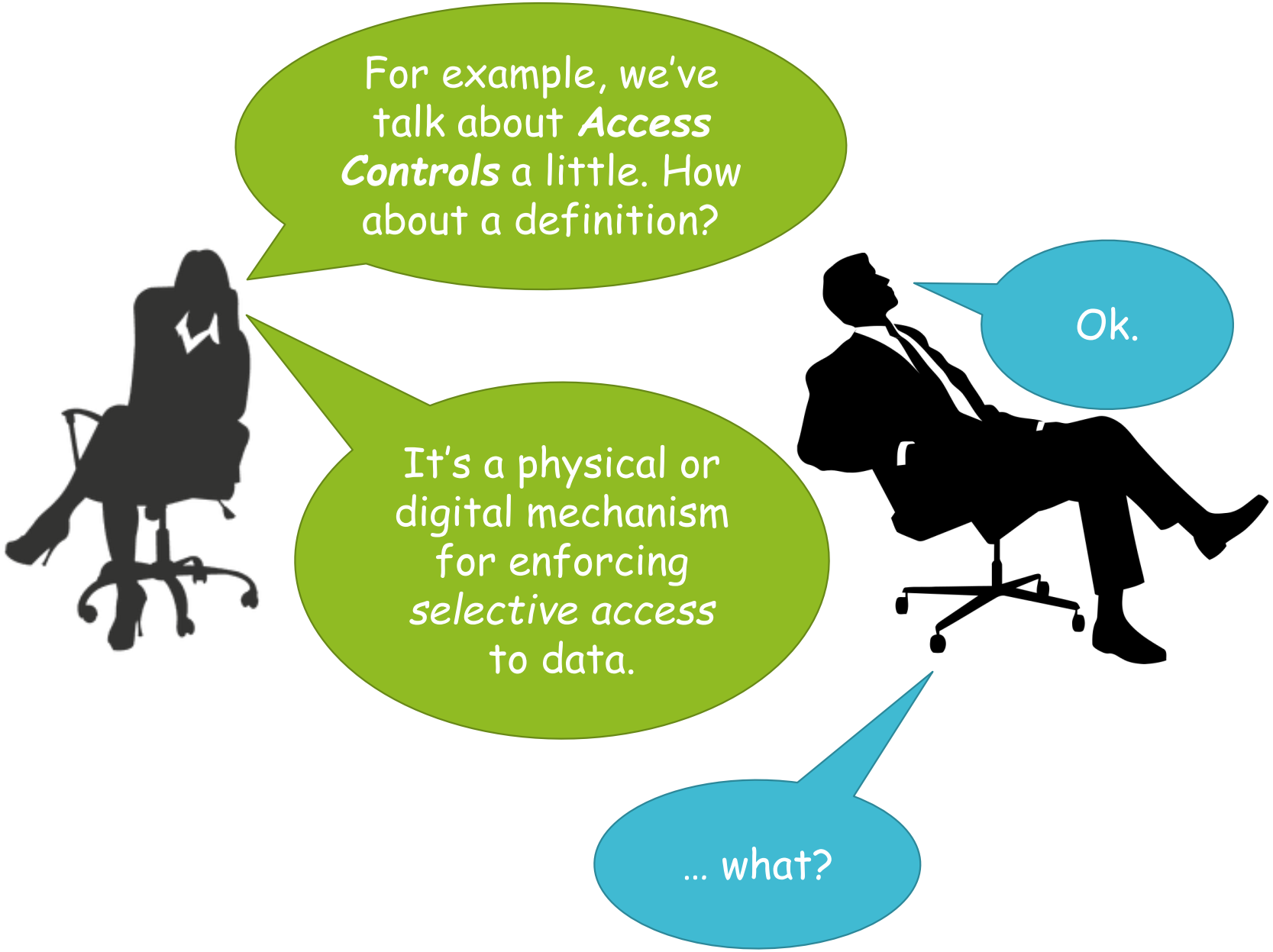
- John Ladley, Danette McGilvray, Anne-Marie Smith, Gwen Thomas (***McGilvray***, p. 52)

What does this have to do with security and privacy?

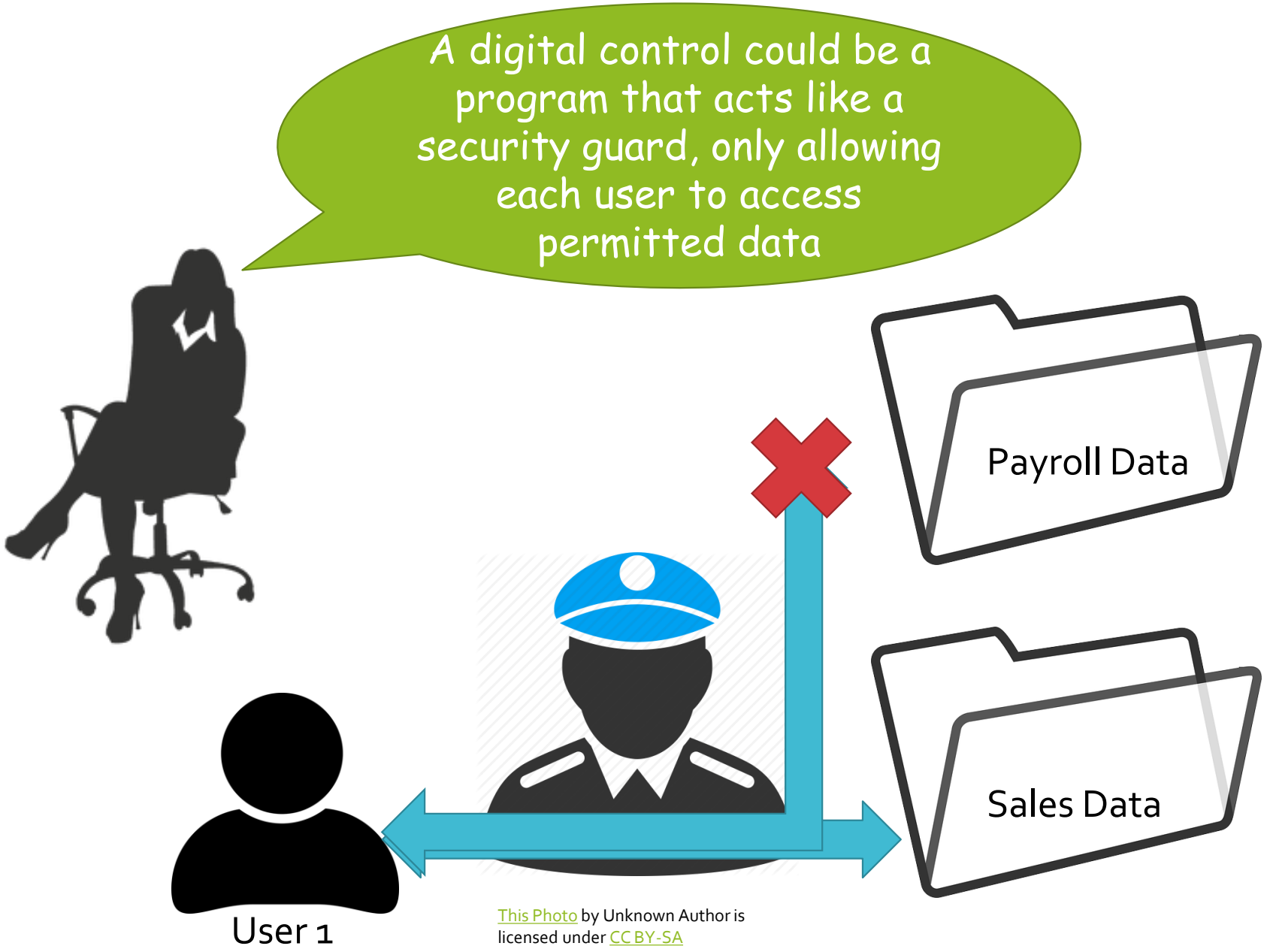
Data Governance and Data Security

Data Governance:
is the organization and implementation of policies, procedures, structure, roles, and responsibilities that outline and enforce rules of engagement, decision rights, and accountabilities for the effective management of information assets
- John Ladley, Danette McGilvray, Anne-Marie Smith, Gwen Thomas (***McGilvray***, p. 52)

Example: Access Controls

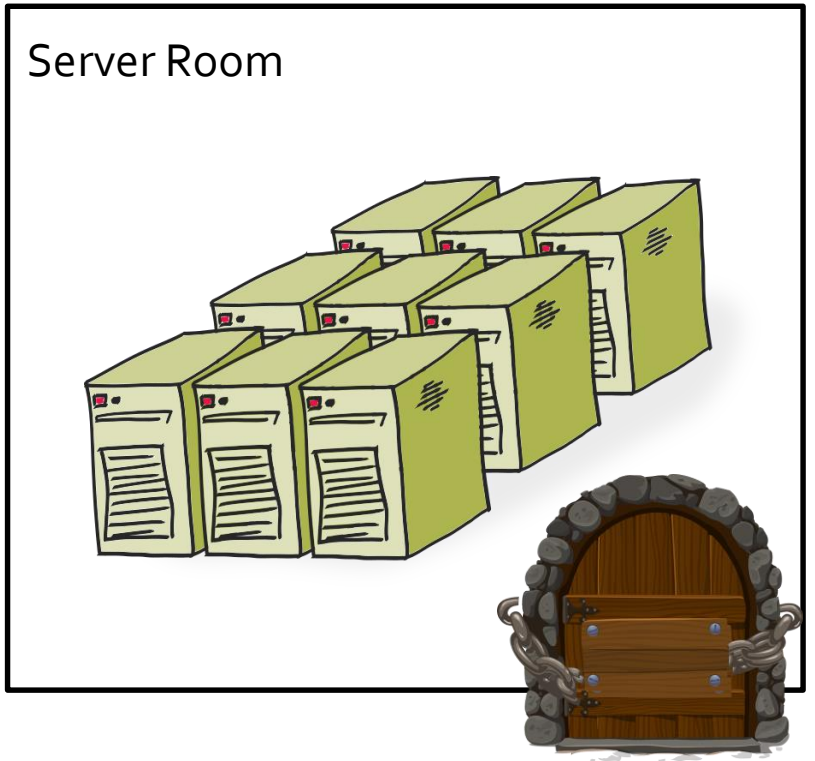


Digital Control Example

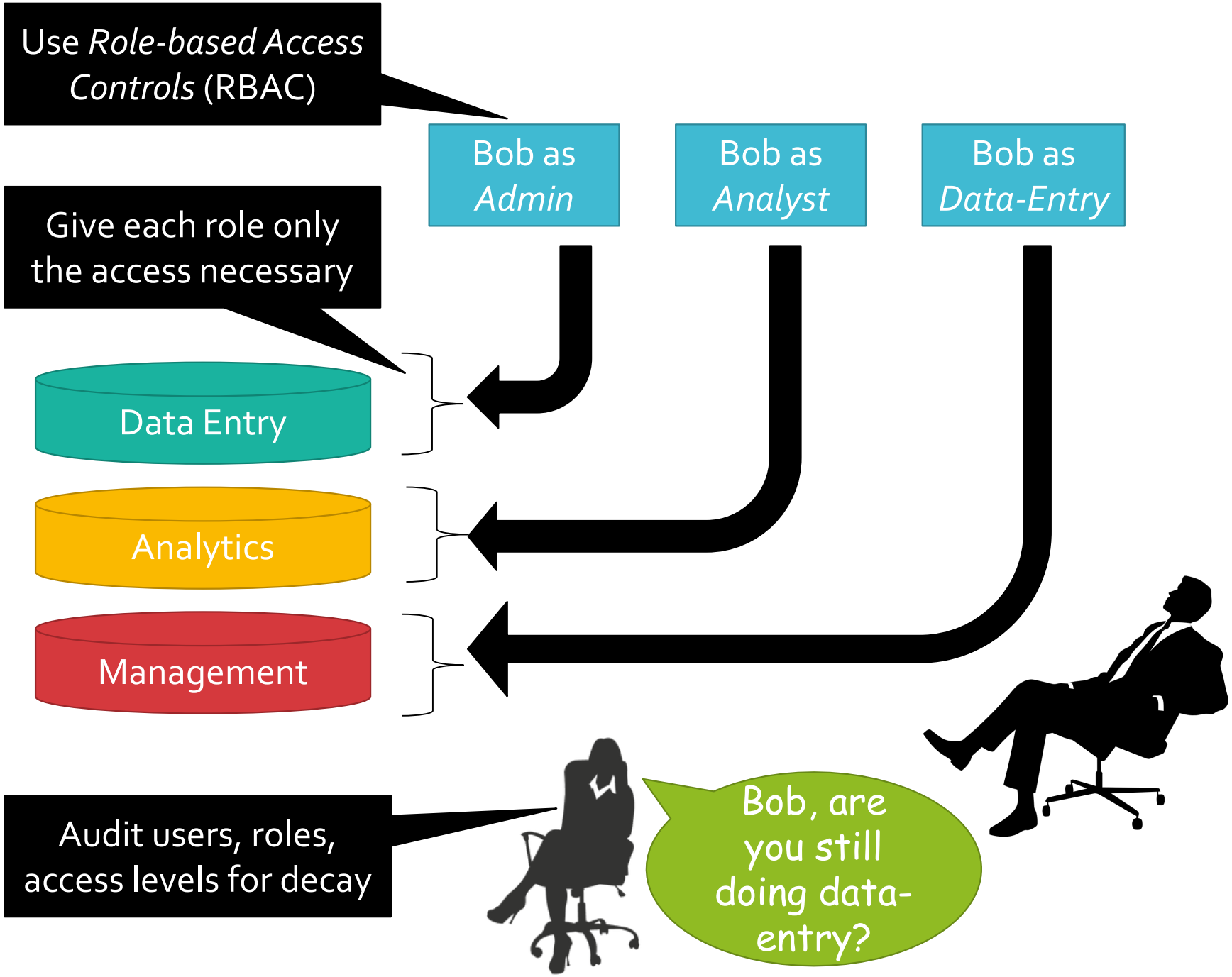


Physical Control Example

A physical control might be something as simple as locking a server room and limiting who has keys.



Access Control Best Practices



Rules of Engagement and Access Controls

Data Governance:
is the organization and implementation of policies, procedures, structure, roles, and responsibilities that ***outline and enforce rules of engagement****, decision rights, and accountabilities for the effective management of information assets*

- John Ladley, Danette McGilvray, Anne-Marie Smith, Gwen Thomas (***McGilvray***, p. 52)

But if you have weak ***business rules*** about who can access data and how…

Exactly.

… your access controls will be weak too!

Rules of Engagement and Access Controls

Data Governance:
is the organization and implementation of policies, procedures, structure, roles, and responsibilities that ***outline and enforce rules of engagement****, decision rights, and accountabilities for the effective management of information assets*
- John Ladley, Danette McGilvray, Anne-Marie Smith, Gwen Thomas

Or let's go back to privacy. Restricting PII to a single DB is no good if…

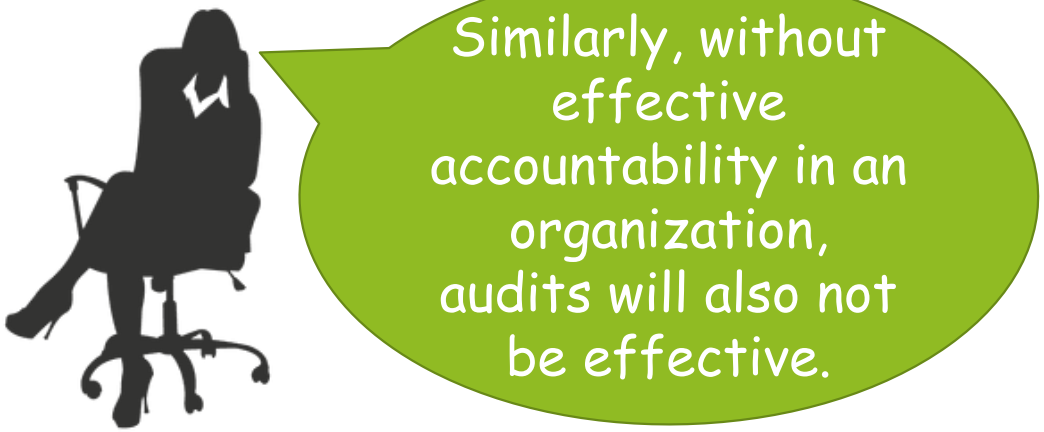
You're catching on!

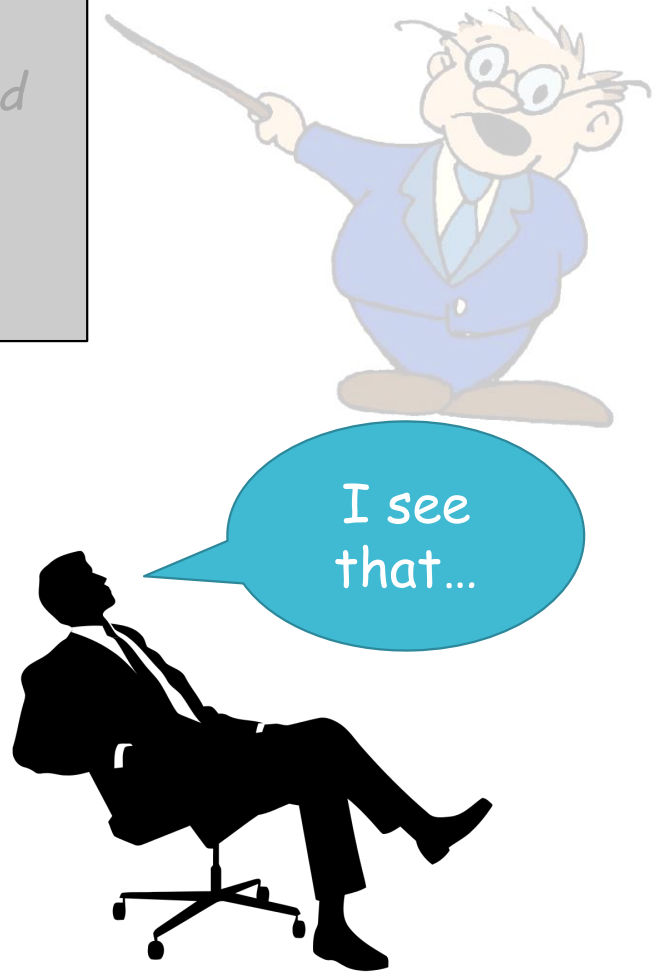
… you have no idea who can access the DB!

Accountability and Audit

Data Governance:
is the organization and implementation of policies, procedures, structure, roles, and responsibilities that outline and enforce rules of engagement, decision rights, and ***accountabilities*** *for the effective management of information assets*

- John Ladley, Danette McGilvray, Anne-Marie Smith, Gwen Thomas (***McGilvray***, p. 52)

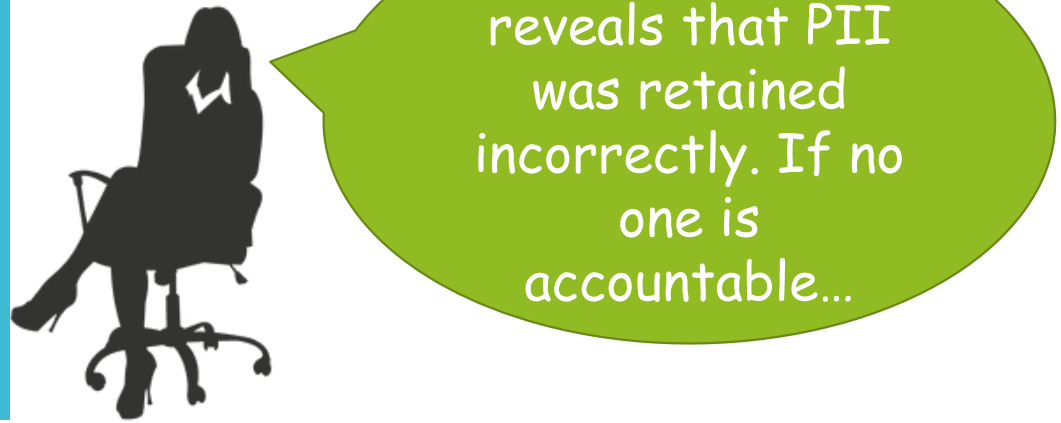




Accountability and Audit

Data Governance:
is the organization and implementation of policies, procedures, structure, roles, and responsibilities that outline and enforce rules of engagement, decision rights, and ***accountabilities*** *for the effective management of information assets*

- John Ladley, Danette McGilvray, Anne-Marie Smith, Gwen Thomas (***McGilvray***, p. 52)



Key Management

Data Governance:
is the organization and implementation of policies, procedures, structure, roles, and responsibilities that outline and enforce rules of engagement, decision rights, and accountabilities for the effective management of information assets
- John Ladley, Danette McGilvray, Anne-Marie Smith, Gwen Thomas (***McGilvray***, p. 52)

Data-at-Rest Key Management

- **Key management is always the hardest, weakest part of Crypto**
- Keys have a lifecycle
  - For at-rest, rotation required.
  - Should be audited

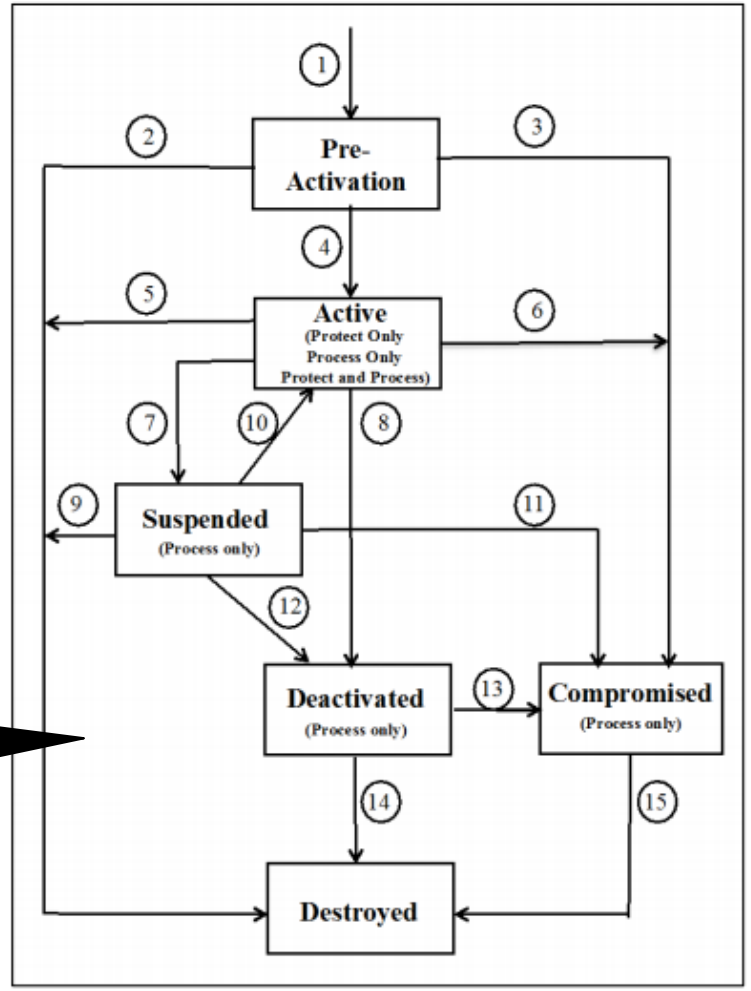


Figure 3: Key state and transition example.

From NIST SP 800-57 Pt. 1 Rev 4.

Key Management Best Practices

- Keep keys separated from the data
- Store master keys in a Hardware Security Module (HSM)
- Do not keep HSM's with master keys connected to the network
- Consider a complete, automated key management solution
- Consider splitting keys to sensitive data using m-of-n sharing

Decision Rights and Key Management

Data Governance:
*is the organization and implementation of policies, procedures, structure, roles, and responsibilities that outline and enforce rules of engagement, **decision rights**, and accountabilities for the effective management of information assets*
- John Ladley, Danette McGilvray, Anne-Marie Smith, Gwen Thomas (***McGilvray***, p. 52)

But if you don't know who should have a key, for how long, or for what purpose…

Indeed.

… then what good is key management ?

Not Done Yet!

Data Governance:
is the organization and implementation of policies, procedures, structure, roles, and responsibilities that outline and enforce rules of engagement, decision rights, and accountabilities for the effective management of information assets
- John Ladley, Danette McGilvray, Anne-Marie Smith, Gwen Thomas (***McGilvray***, p. 52)

Bob, much of our discussion on "governance" has focused on users. But what about the data itself?

I don't follow.

What's the problem here?

Remember in our PII discussion what we said about data inputs?

Yes, you said to only admit PII from authorized entry points.

Do you see what assumption that approach depends on?

Uh… no?

That you ***know all of your data entry points.***

Where is the Data?!

- Do you know where the data is in your system?
- By system, I mean the entire business "system".
- Do you know if you have any duplication?
- Do you know all the sources of a given piece of data?
- Do you know how a given piece of data is shared?
- Do you know how data is disposed of?

YOU CAN'T SECURE WHAT YOU DON'T KNOW

Information Quality

Information Quality:
*is the degree to which information and data can be a trusted source for… all required users· It is having the right set of **correct information**, at the **right time**, in the **right place**, for the right people…*
– ***McGilvray***, p.5, emphasis added.

An organization with low data quality will, almost certainly, have poor data security and privacy.

Got it!

Correlation from the Security World

"many systems fail because their designers protect the wrong things, or protect the right things but in the wrong way."

Ross Anderson, Security Engineering, 2nd Ed.,

Dr. Nielson's Version:

"many systems fail because their designers protect the wrong data, or protect the right data but in the wrong way…

… at least in part because ***they don't know what data they have, how correct it is, where it came from, and what it's used for****…"*

Data Quality Again.

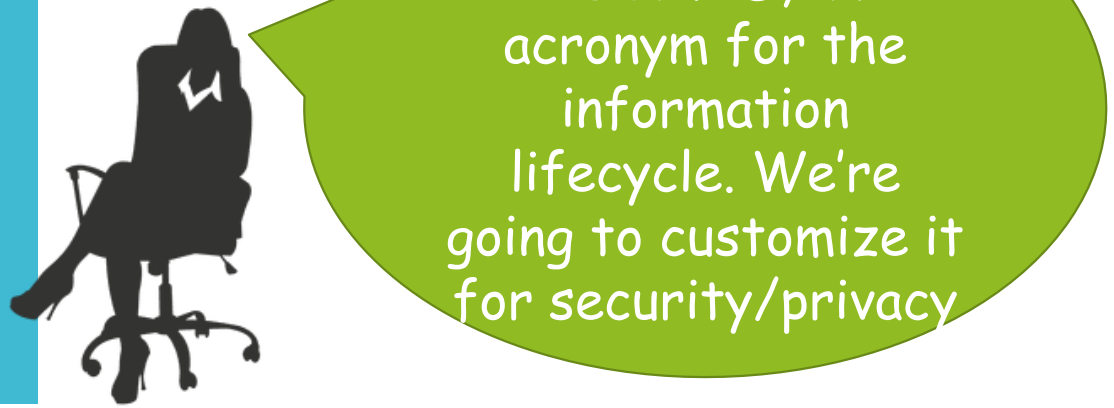
- I highly recommend *Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information*™ by Danette McGilvray
- The book describes Danette's Ten Step process for Quality Data
- To repeat, an organization without information quality is almost certainly going to be without information security
- In my opinion, a data quality analysis must be included in any kind of data security analysis
- Let's take a look at Danette's acronym "POSMAD"

POSMAD

POSMAD (Information Life Cycle):

P - Plan for data	M - Maintain data
O - Obtain data	A - Apply data
S - Store/share data	D - Dispose data

- ***McGilvray*** p.23 (adapted from ***English*** pp. 200-209).



Prepare Security & Privacy for the Resource

POSMAD (Information Life Cycle):

P - Plan for data	M - Maintain data
O - Obtain data	A - Apply data
S - Store/share data	D - Dispose data

- ***McGilvray*** p.23 (adapted from ***English*** pp. 200-209).

In the plan phase, you identify your data's security and privacy requirements. Then you design and implement controls.

So, what kind of encryption to use?

… no…

Plan – Measure Once, Cut Twice

- Encryption is probably the ***last*** thing to think about.
- What are the security/privacy requirements of the data?
  - Upon entry?
  - Upon exit?
  - While in custody of your organization?
  - Disposal?
- Who will have access to the data?
- Will access to the data require keys? How will keys be managed?
- What are the regulatory requirements?
- What are the ethical obligations?

Obtain the Data Securely

POSMAD (Information Life Cycle):

P - Plan for data	*M - Maintain data*
O - Obtain data	*A - Apply data*
S - Store/share data	*D - Dispose data*

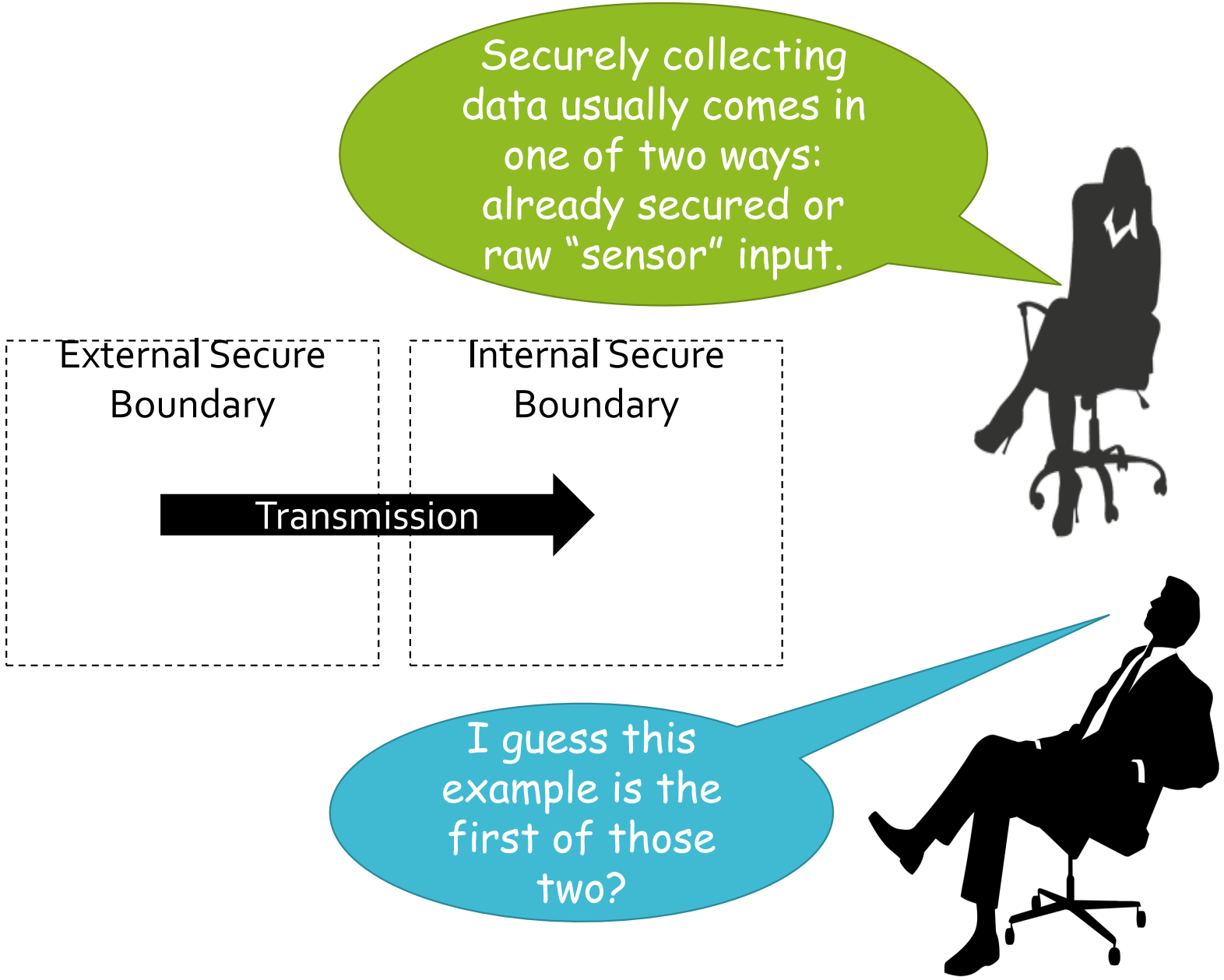
- ***McGilvray*** p.23 (adapted from ***English*** pp. 200-209).

Boundaries are often where things go wrong in security/privacy.

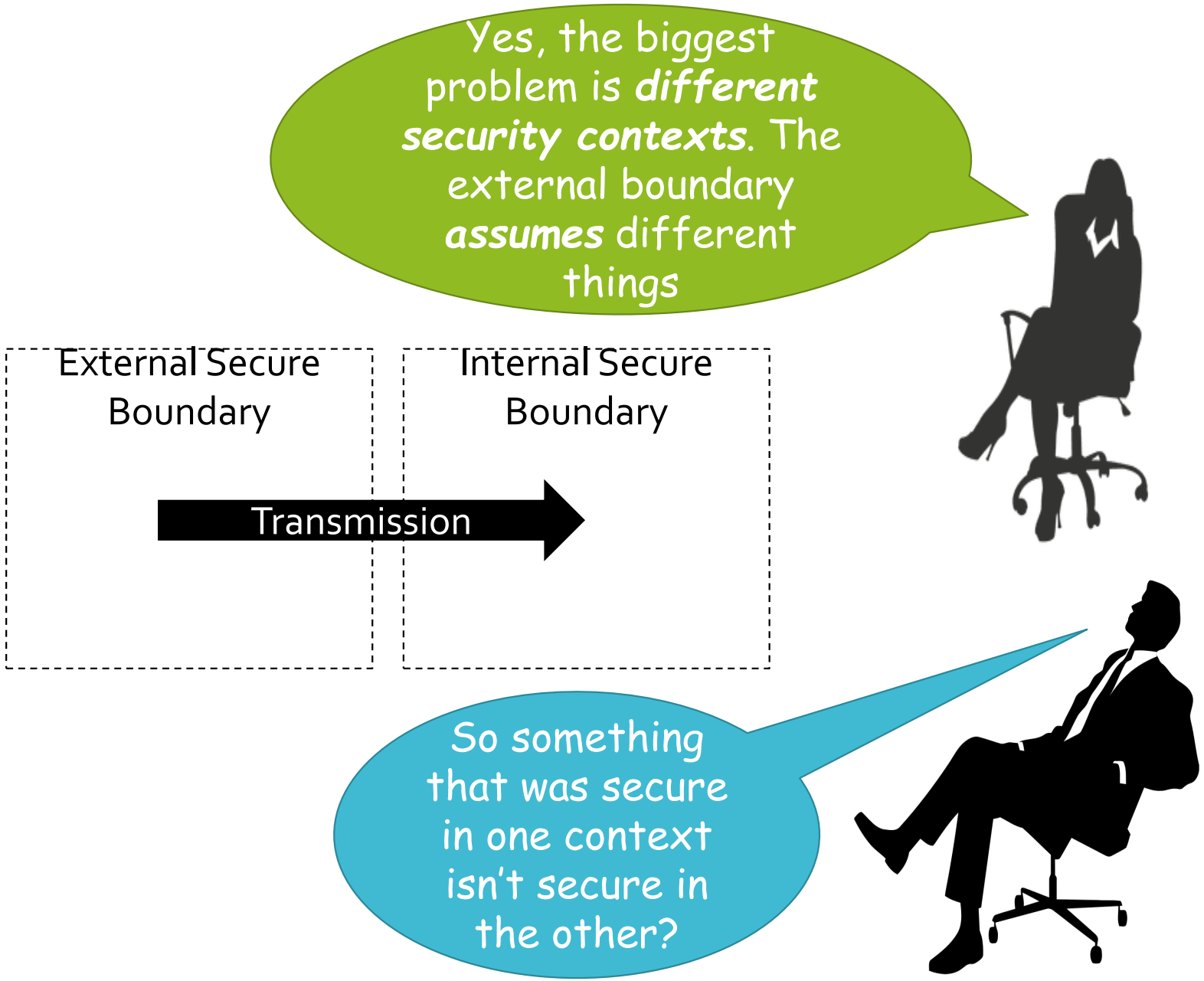
Doors and windows are always vulnerable

Right.

Data in Motion Again!



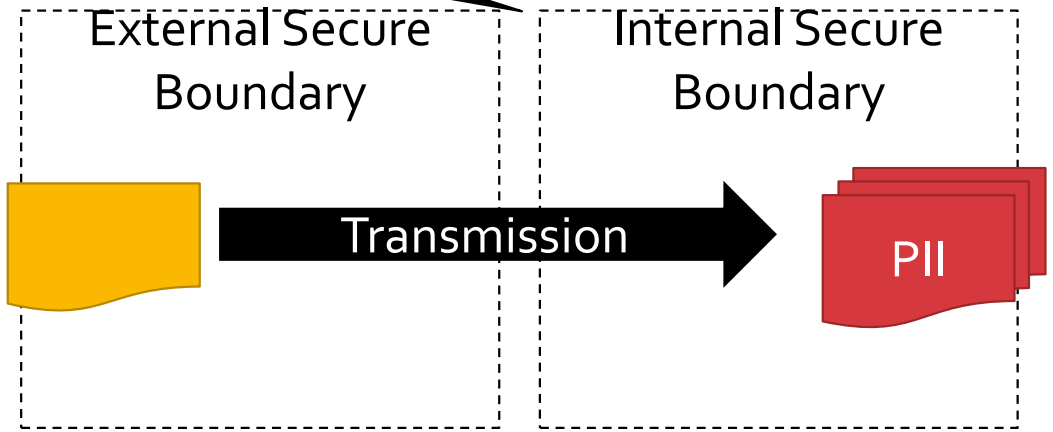
Security is Always About Context



PII Boundary Issue

Contexts can be different for many reasons including regulatory environment. When bringing in data, make sure the security assumptions for both sides are understood and documented

Here's a privacy example. Data is **Linkable** PII externally, but when it enters our system becomes ***Linked***



I see. We have more data than the external party. We have bigger privacy issues.

Store/Share Data Securely

POSMAD (Information Life Cycle):

P - Plan for data	M - Maintain data
O - Obtain data	A - Apply data
S - Store/share data	D - Dispose data

- ***McGilvray*** p.23 (adapted from ***English*** pp. 200-209).

Once you get data into the system, you have to control where it is…

… and who can access it. We've already talked about this

It won't hurt to repeat.

Security and Privacy in Sharing and Storing

- Ensure that ***all*** locations where the data is stored are known
  - Watch out for hidden storage locations such as cache/replication
  - Ideally, store it in just one location if at all possible
- Ensure that equivalent security controls are used in all locations
  - Don't require 2FA auth on the DB and leave a hard copy unsecured!
- Ensure that equivalent access controls are used in all locations
- Use Role-based Access Controls and minimize access
- For external access/sharing, review security context changes

Continuously Ensure the Security of the Data

POSMAD (Information Life Cycle):

P - Plan for data	**M - Maintain data**
O - Obtain data	A - Apply data
S - Store/share data	D - Dispose data

- ***McGilvray*** p.23 (adapted from ***English*** pp. 200-209).

Bob, have you heard of ***Data Decay?***

Unfortunately, security also "decays". Thus, security must be maintained.

Of course.

Security Decay

Sources of Security Decay:

- New OS/App vulnerabilities
- New crypto guidance
- Key lifetime expirations
- Certificate expirations
- Staff turnover
- New/updated applications
- New/updated regulations
- Corruption/mistakes in config
- Bugs
- Changes in requirements
- Changes in risk profile, tolerance
- Complacency over time
- Aggregation of minor issues
- Data decay
- Access controls decay

Security Decay Painful Example

Sources of Security Decay:

- ***New OS/App vulnerabilities***
- New crypto guidance
- Key lifetime expirations
- Certificate expirations
- Staff turnover
- New/updated applications
- New/updated regulations
- Corruption/mistakes in config
- Bugs
- Changes in requirements
- Changes in risk profile, tolerance
- Complacency over time
- Aggregation of minor issues
- Data decay
- Access controls decay
- Ignored audits
- Unmonitored logs/ignored alarms

Use Data Securely (Data-in-Use!)

POSMAD (Information Life Cycle):

P - Plan for data	*M - Maintain data*
O - Obtain data	***A - Apply data***
S - Store/share data	*D - Dispose data*

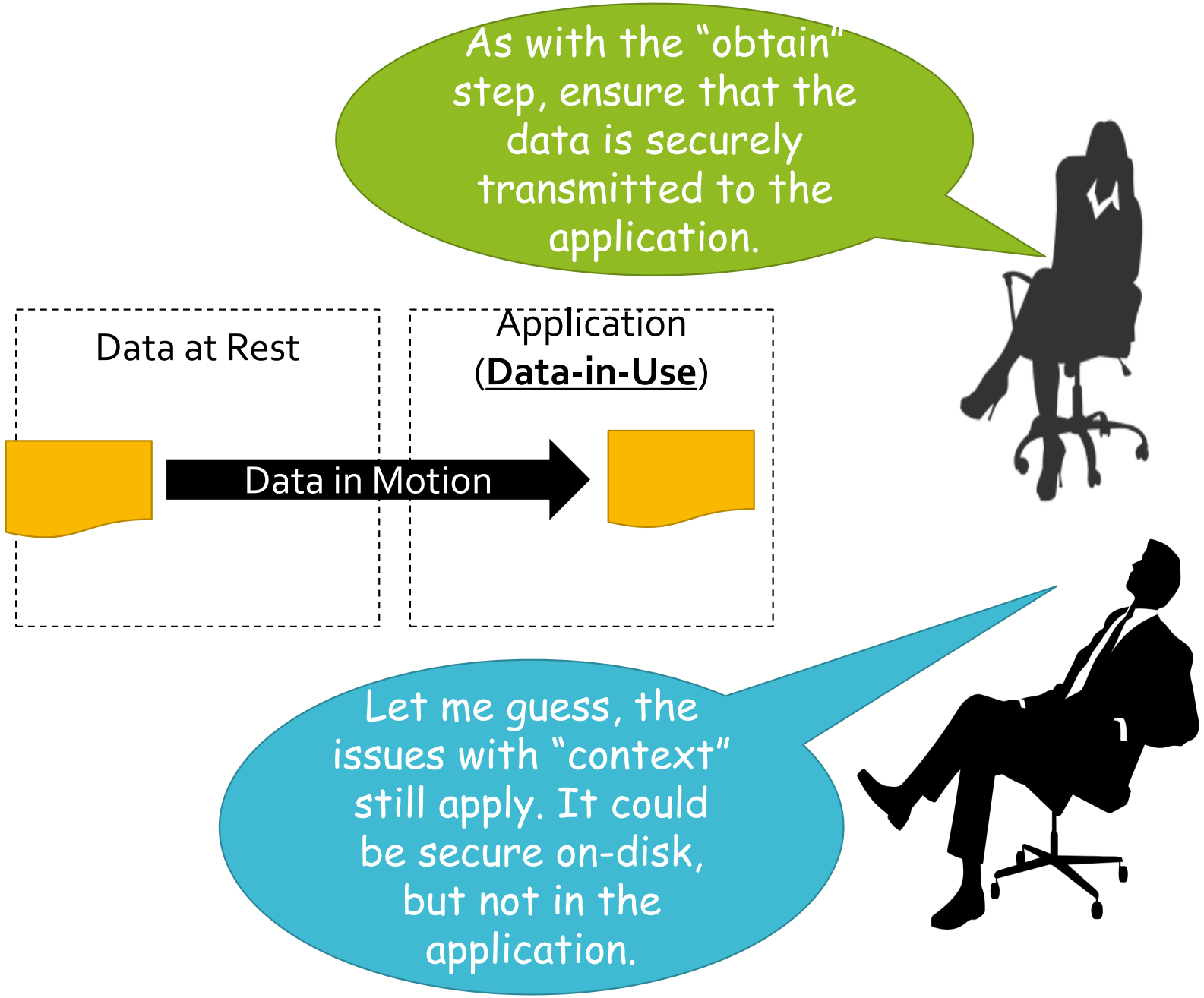
- ***McGilvray*** p.23 (adapted from ***English*** pp. 200-209).

And, of course, you must ***use*** data securely.

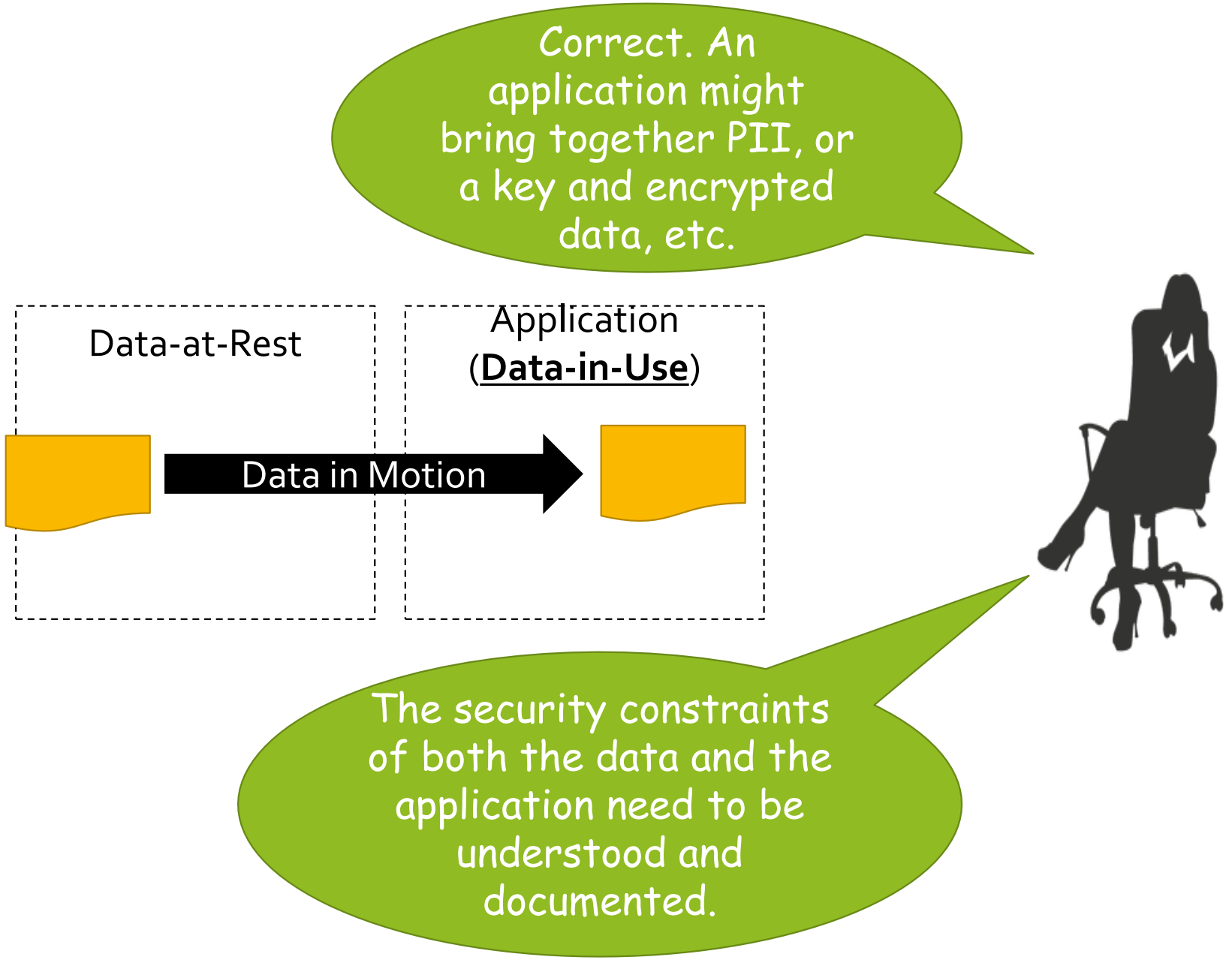
In many ways, it's a microcosm of all the other elements we've discussed.

How do you do that?

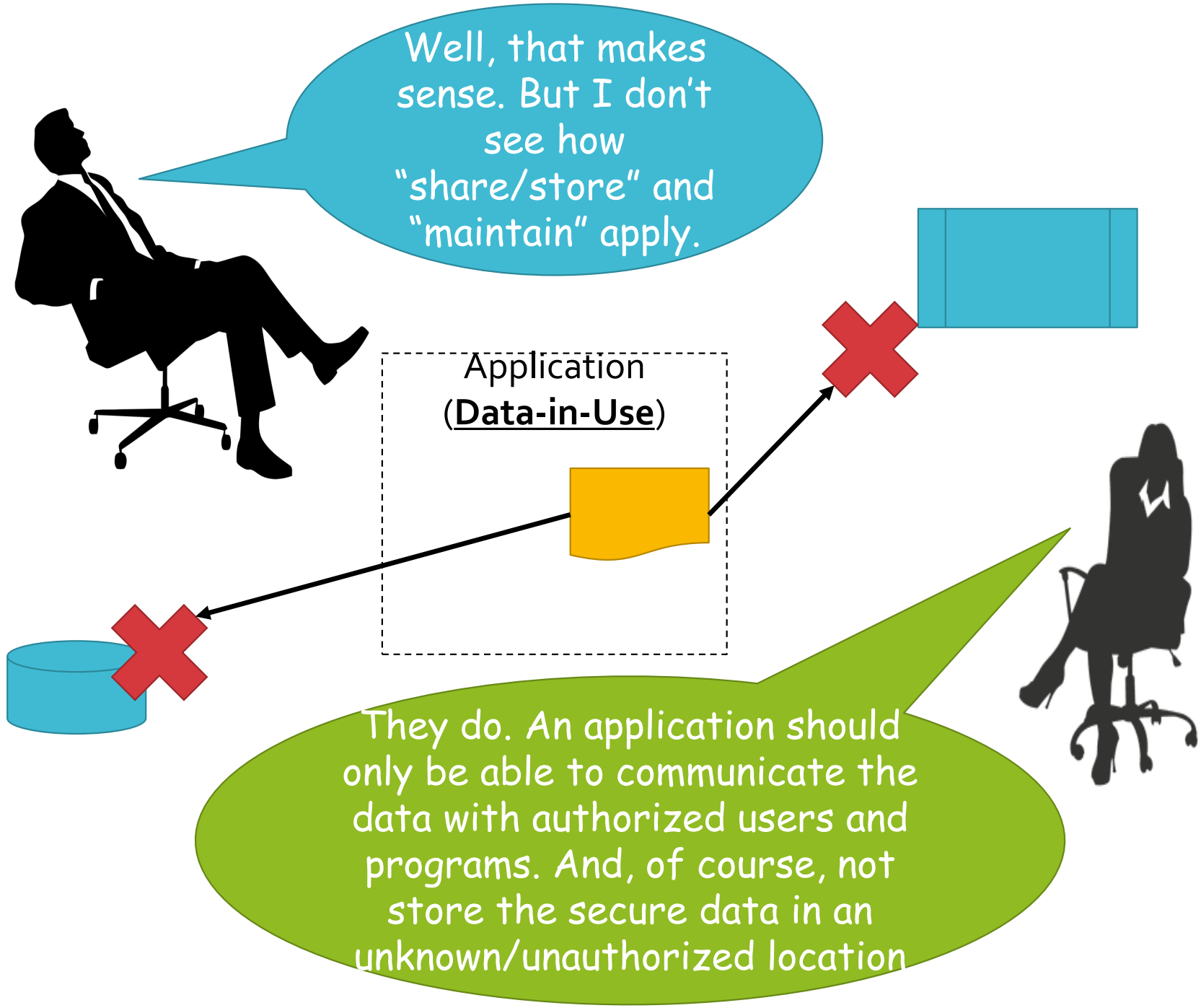
Visualizing Secure Data Use (1)



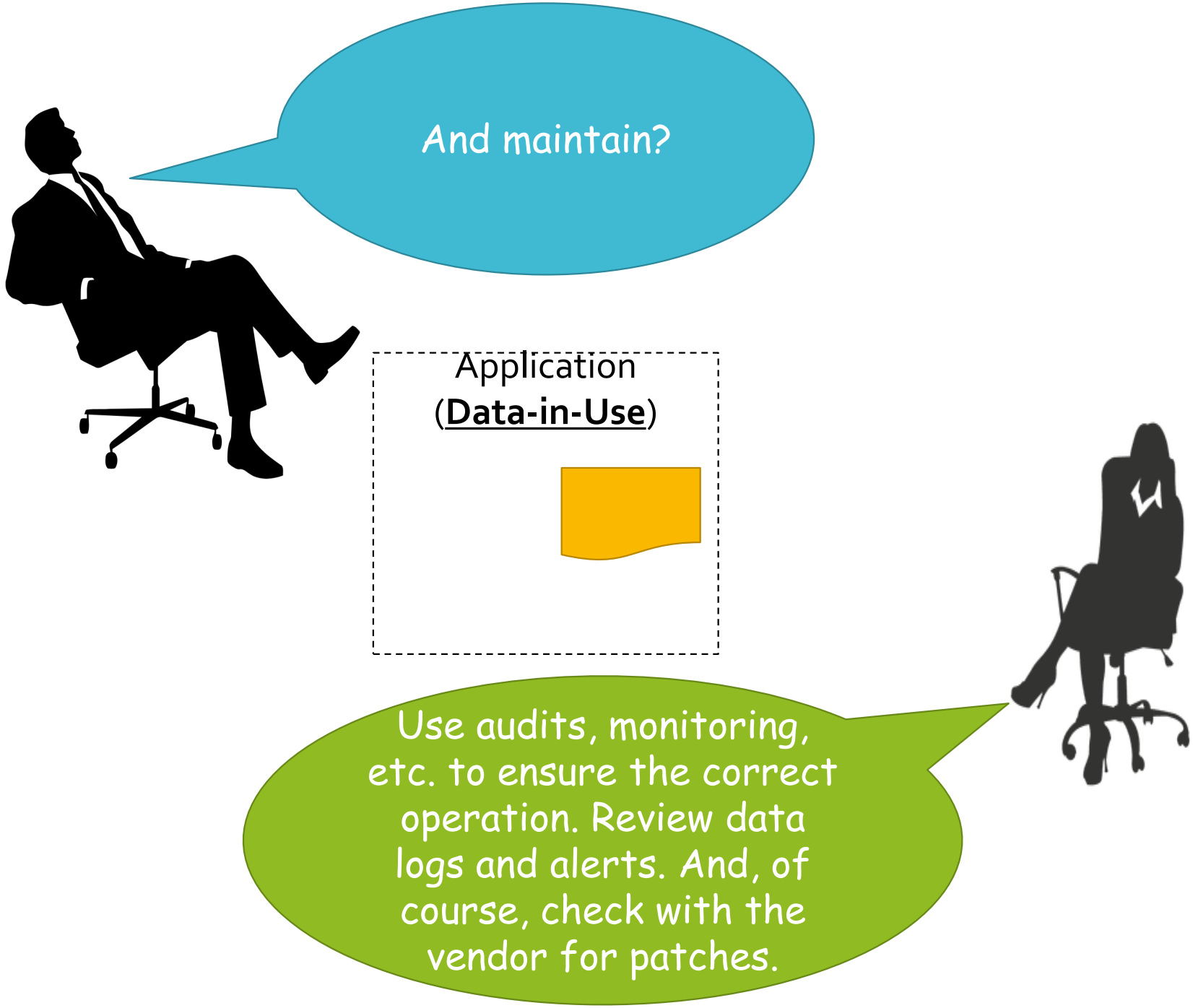
Visualizing Secure Data Use (2)



Visualizing Secure Data Use (3)



Visualizing Secure Data Use (4)



Secure Data Disposal

POSMAD (Information Life Cycle):

P - Plan for data
O - Obtain data
S - Store/share data
M - Maintain data
A - Apply data
D - Dispose data

- ***McGilvray*** p.23 (adapted from ***English*** pp. 200-209).

Last, but not least, secure data disposal.

That's ***media disposal***. This is a more holistic process.

Like with the hard drive?

Possible Elements of Secure Data Destruction

- Ensure that all copies of the data are accounted for
  - Revoke all access
  - Terminate any processes using the data
  - Ensure remote device usage is also accounted for
  - Include physical copies in the accounting
- Securely delete data; use cryptographic shredding or overwriting
- Document the destruction if required by policy or regulation
- Release keys, access controls, etc associated with the data
- Release metadata or other data no longer needed