Securing Content Management Systems

The Next Frontier in Data Leakage Prevention

Tamer Abuelsaad - Senior Software Engineer - IBM

Terms and Definitions

Data Leakage/Loss Prevention (DLP)

- Technology, products, or services that prevent sensitive information from being exposed

- Types of DLP Systems:
  - Network DLP – Typically installed on network egress points to monitor and enforce (if needed) on traffic leaving a perimeter
  - Endpoint DLP - Scanning and enforcement on end-user devices, such as, Laptops, servers, mobile devices, …
  - SMTP DLP – Scanning and enforcement on Mail Gateways targeting inbound/outbound mail inspection

Unstructured Data

- Information that either does not have a pre-defined data model and/or does not fit well into relational tables. Typically text-heavy, but may contain data such as dates, numbers, and facts as well. (*source: Wikipedia*)

- Examples: Excel, Power Point, PDF, Open Document Format (ODF), plain text, …

Content Management Systems (CMS)

- Computer program that allows publishing, editing and modifying content as well as maintenance from a central interface. Such systems of content management provide procedures to manage workflow in a collaborative environment. (*source: Wikipedia*)

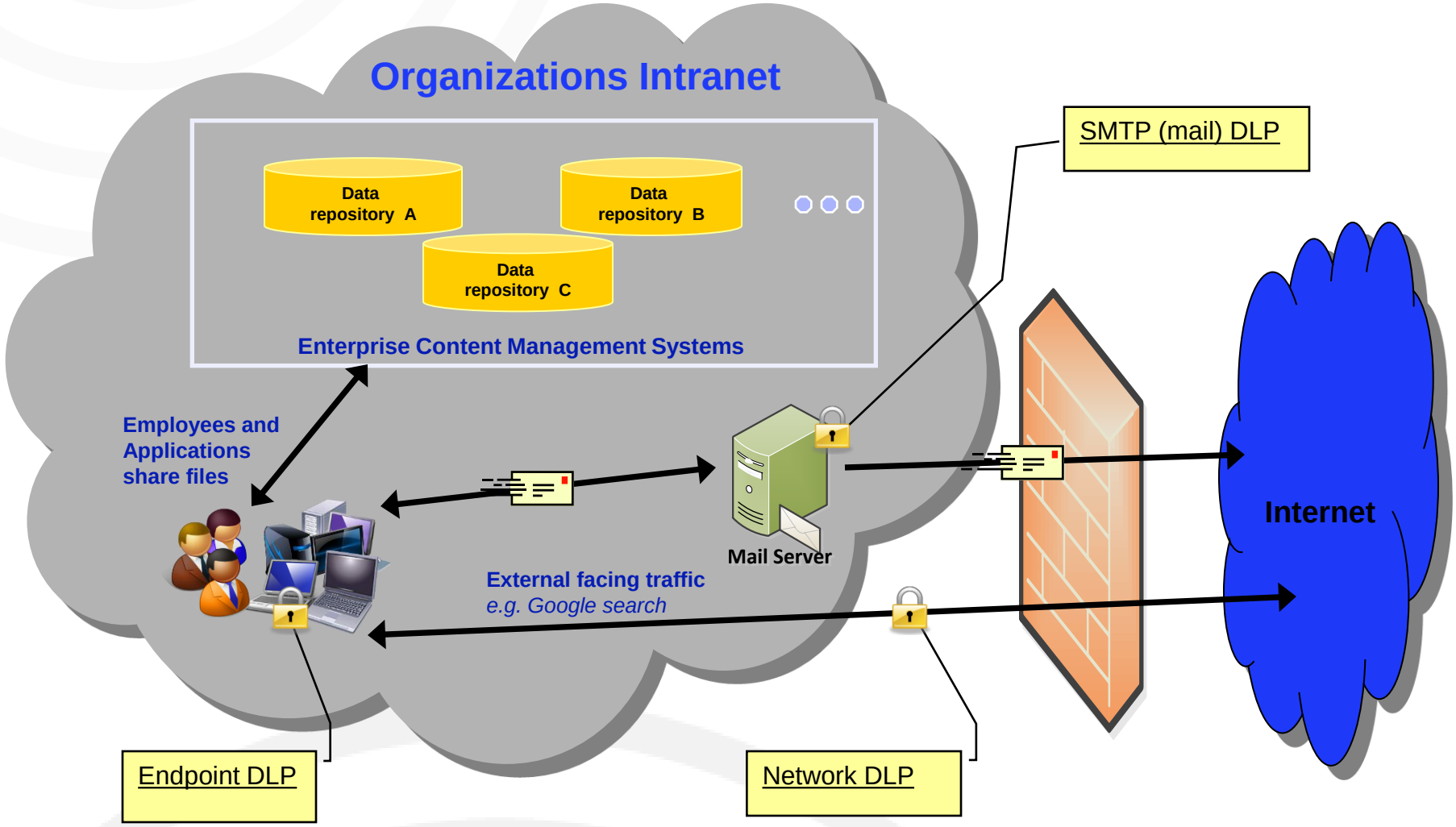
- Can also be referred to as Data Repositories

- Examples: Wikis, web based file storage (documentation, …), etc…

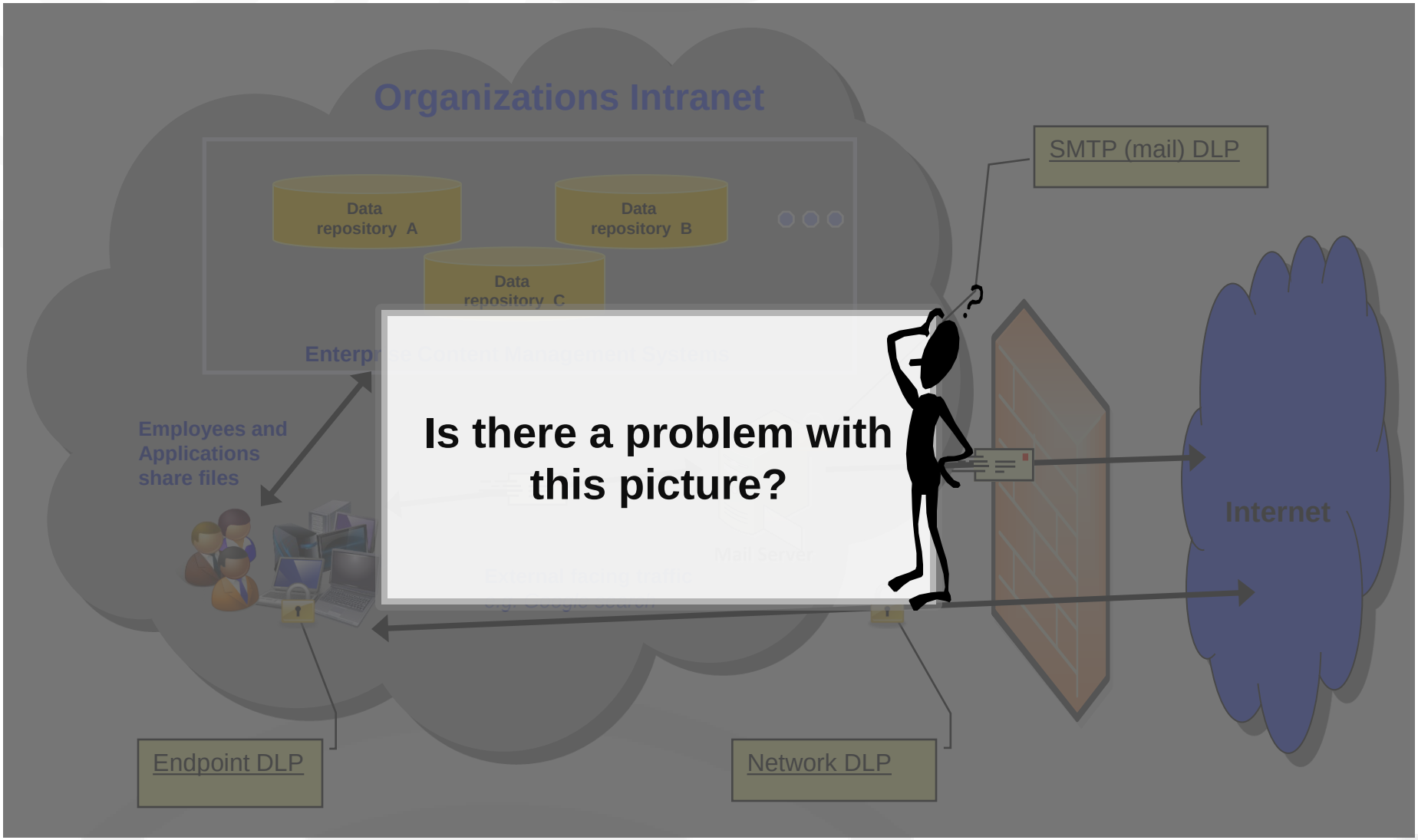
- Very important for Social Collaboration

End of Definitions

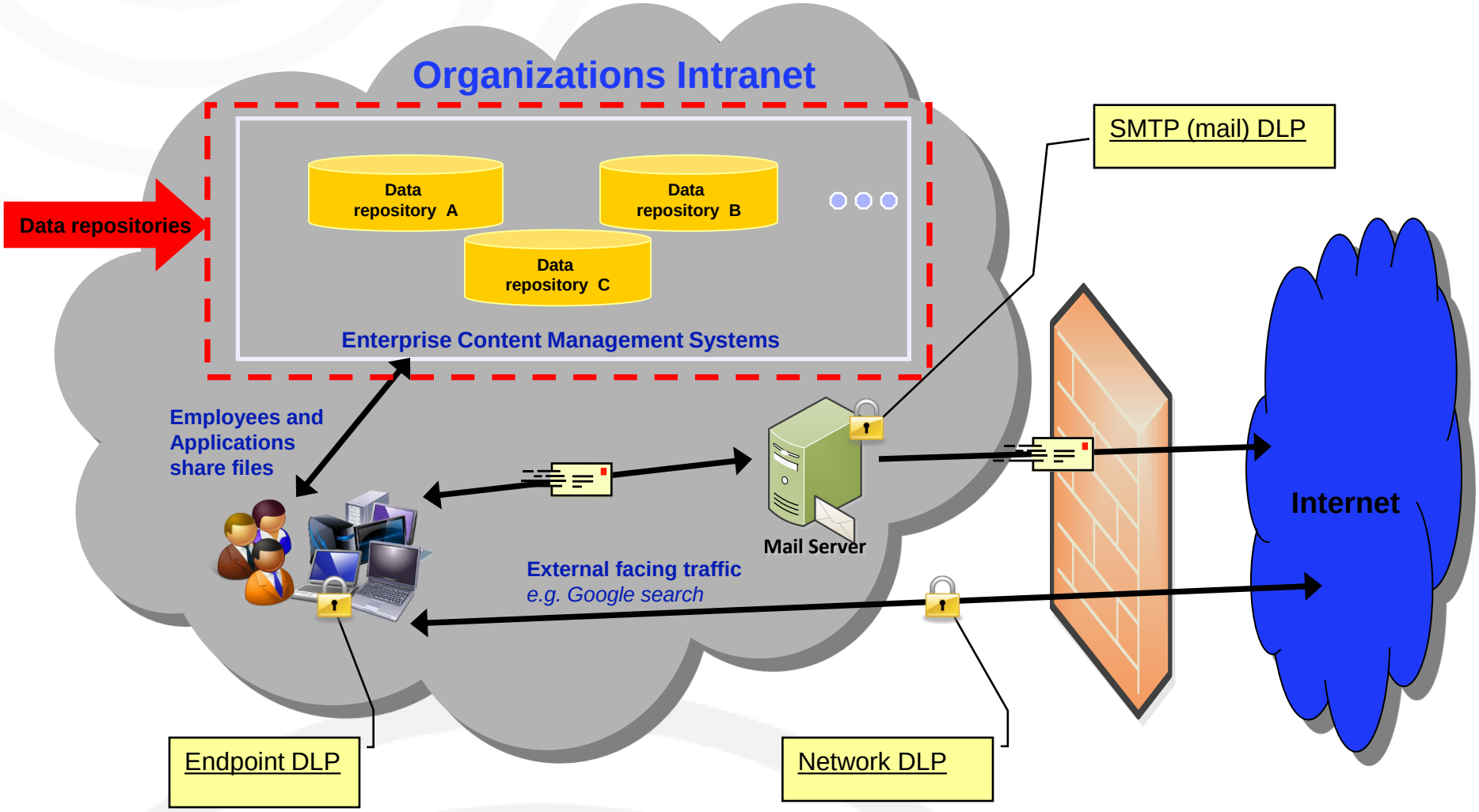
Typical DLP Coverage



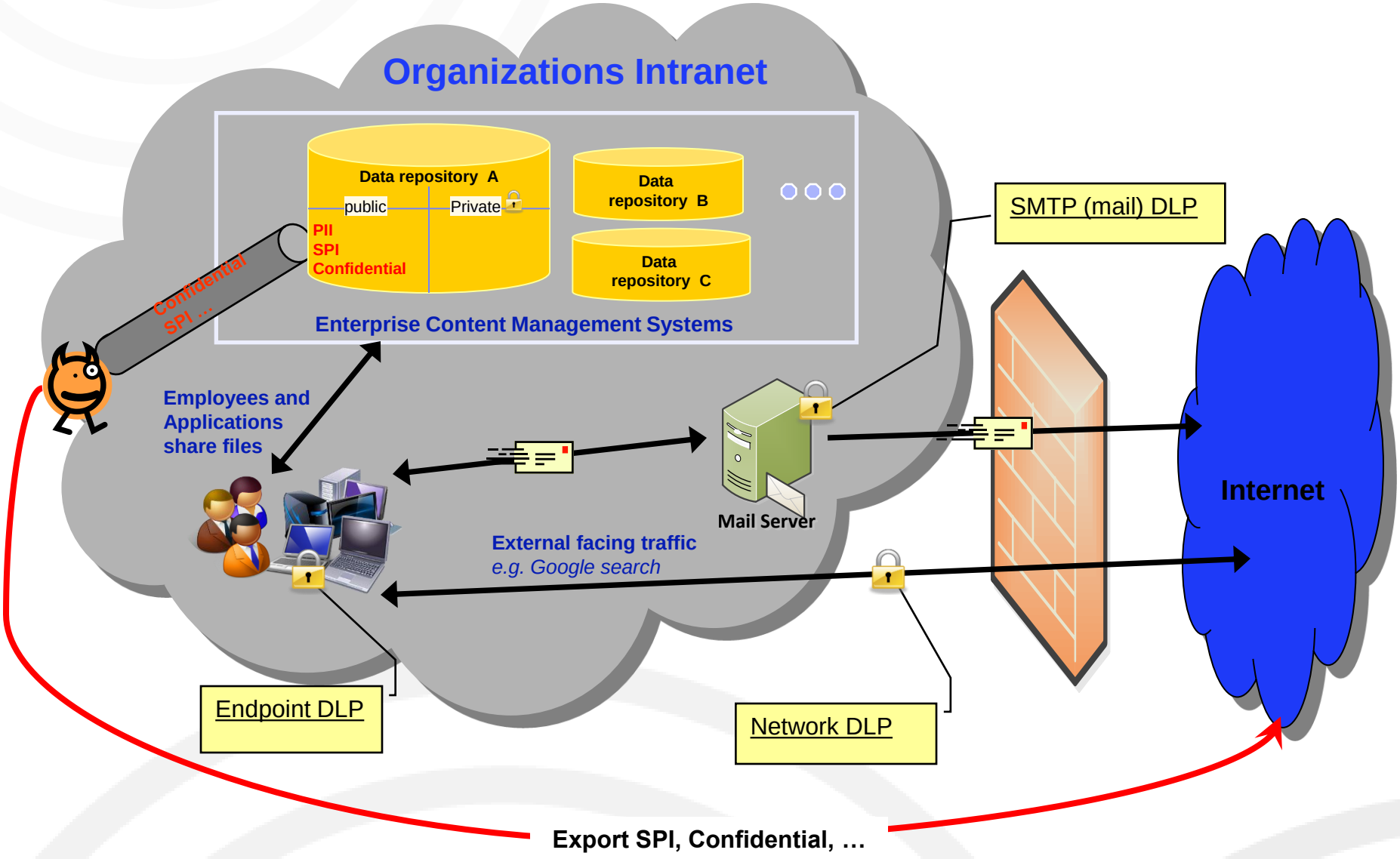
Typical DLP Coverage



Is There Something Still Exposed?



What Would a Hacker or Malicious Person Do?



Latest Themes in Data Protection

- Move security closer to data (*source: IDC DLP Report 12/2011*)

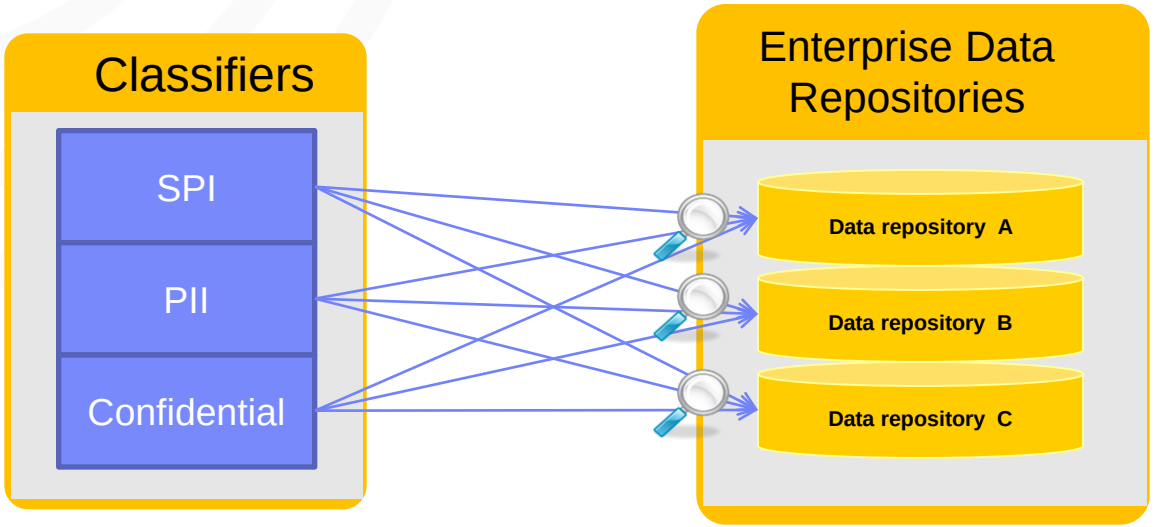
- Involve data owners in DLP Strategy (*source: Aberdeen Group 05/2012*)

Properties of Content Management System

- Owner centric – each file has one or more owner(s)
- Each file has Access Controls List (ACL), such as, Public, Private, or Controlled share
- APIs to *CRUD* file meta data, access controls, and file content (REST, SOAP, …)

Can Known CMS Properties Help Secure it?

1. Automated centralized file content scanning/classification

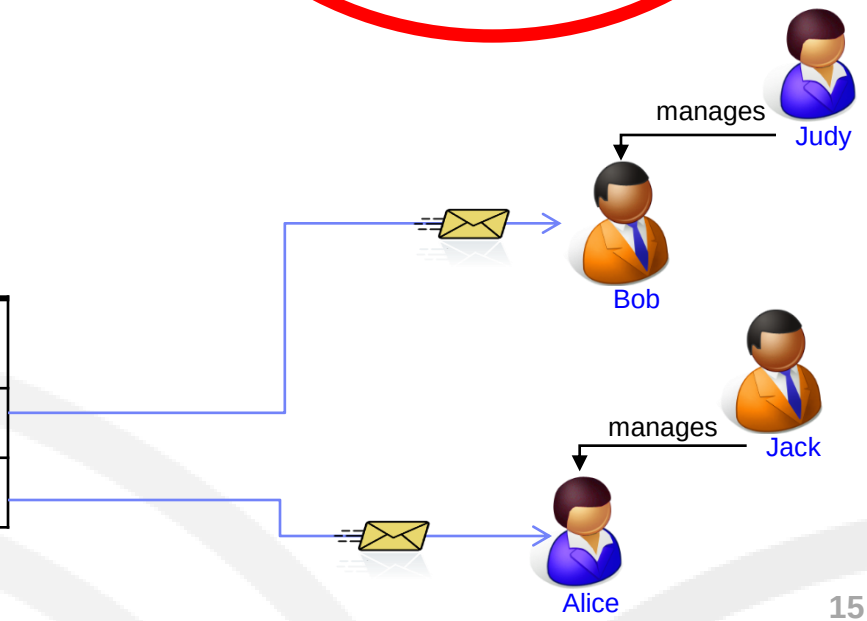


- Multiple scanners per data repository
- Use Data Repository APIs to access data/files

2. Engage Content Owners in Violation Remediation

- Engage content owner(s) to handle any content found in violation
- *e.g. email owner if Business Sensitive file is found without Access Controls*; avoid IT security guy chasing them

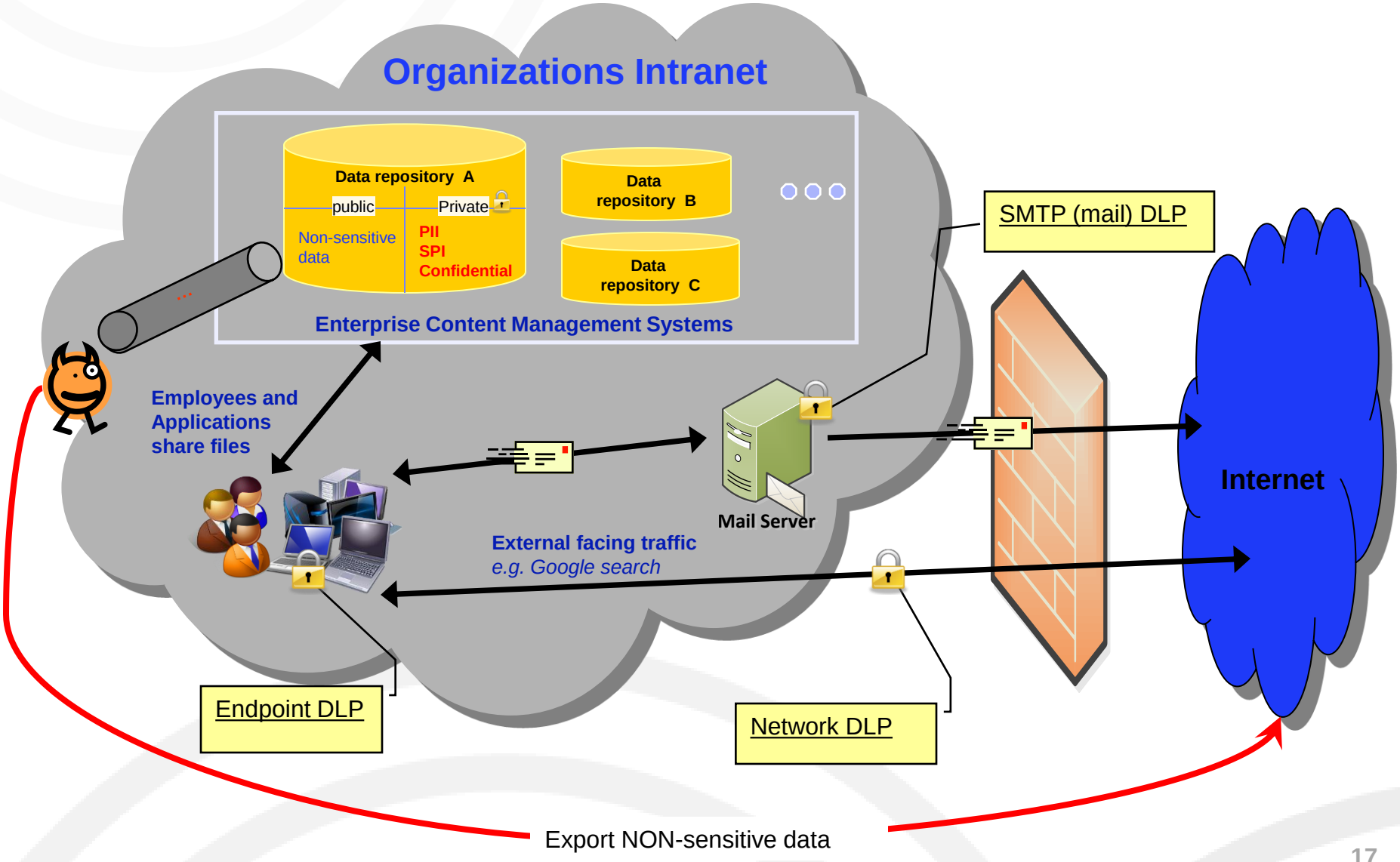
File Name	File Owner(s)	Classification	Access Controls
MyFile1.ppt	Bob	Sensitive	Public
MyFile2.ppt	Alice	Sensitive	Public



3. Automated Risk Remediation

- There will always be delinquent users
- Let the system close the gap using Data Repository APIs
- *e.g. public file with violation(s), the system can make it private via API usage*

The Result: Secured Content Management Systems

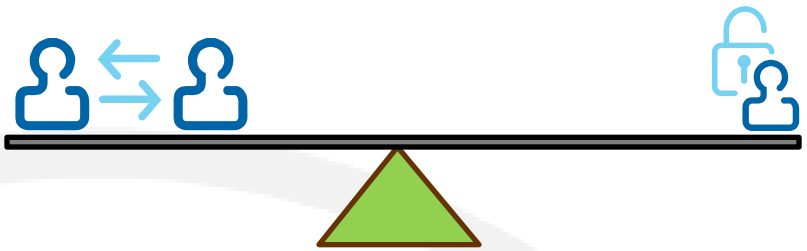


Proper Ways to Secure Unstructured Data Repositories

- Tell employees about data repository scanning and why you are implementing it
- Central content classification to be used for all data repositories
- Automated remediation
- Involving content owners in the remediation process, even as simple as email notification or escalation
- Empower content owner
  - Provide violation evidence (e.g. you put "*Company Confidential*" in this file)
  - Provide evidence location (e.g. pages 2, 4, 5, 6)

It is important in social collaborative environment not to threaten collaborators (users) and cause them to stop sharing.

A delicate balance between user sharing and sensitive assets protection is needed.



Thank You!

Tamer E. Abuelsaad
Software Engineer
CIO Lab
Innov Security & Compliance

294 Route 100
Somers, NY 10589

Tel 914 766 1368
IP Phone 720 342 0851
tamera@us.ibm.com