# Data Loss Prevention Fundamentals

Begins with CIA – Confidentiality, Integrity, and Availability

* Confidentiality: passwords, encryption, firewalls
* Integrity: hashes, checksums
* Availability: backups, failovers

DLP Policies are designed to protect the leakage of:

* Client’s sensitive information, financial data, user activity, system configurations, customer details, usernames/passwords, credit card details (PCI), and intellectual property.

There are two types of losses: Direct and Indirect. Indirect losses are intangible.

* Direct: violations of regulations (fines), customer compensation, investigation costs, litigation, reduced sales, restoration fees, etc
* Indirect: Share price falls, lower reputation, customer loss of faith, intellectual property loss, brand reputation hit.

DLP Systems are designed to detect and prevent data loss through 3 states:

* Data in motion (data moving around the system): Network traffic scanning
* Data in use (applications currently reading data into memory): application scanning/monitoring
* Data at rest (hard drives, static memory): storage scanning

4 Questions:

1. What state is the data in?
   1. DAR: local drives, remote drives
   2. DIU: what are people copy/pasting? Are they taking screen captures of protected docs? Printing or faxing sensitive data?
   3. DIM: understand well-known protocols (HTTP, FTP, Telnet) so you can identify suspicious protocols (Malware using P2P networks to create tunnels).
2. Where do we place the DLP system?
   1. Endpoint: place on the local host.
   2. Network: utilize firewall and IDS
3. What is our approach? Are we only detecting or do we also prevent?
   1. Preventing: encrypting, implementing access control
   2. Detection: content-based inspection, context based inspection (what state is the data being access in?), content tagging (add metatags to classify data).
4. What actions should we take?
   1. Audit
   2. Block
   3. Notify
   4. Modify (clean up data – delete cc info)
   5. Encrypt (automatically encrypt communications)
   6. Quarantine data for further inspection

Methods for searching large dumps of data: regular expressions!

I wrote the following regular expressions to find IP addresses, email addresses, credit card data, year, and telephone numbers.

Telephone: \d{3}[-.]?\d{3}[-.]?\d{4}

Year: [0-9]{4}

Email: [a-zA-Z0-9.\_%+-][+@[a-zA-Z0-9.\_%+-](mailto:+@[a-zA-Z0-9._%25+-)]+

Mastercard/Visa: [3|4|5]\d{3}(\s|-)?\d{4}(\s|-)?\d{4}(\s|-)?\d{4}

IP Addresses: [0-9]{1,3}\.[0-9]{1,3}\.[0-9]{1,3}\.[0-9]{1,3}