G3O Consultants Ltd

GDPR

What it means & what needs to be done

GDPR What and When

- General Data Protection Regulation
- Individuals will have greater control of who has their data, and how it will be used
- Organisations must report on data breaches within 72 hours
- Organisations will be bound by more stringent rules for obtaining consent from individuals on how their data can be used
- Comes into operation May 25th 2018
  - There is no grace period

Responsibilities of the Organisation

- It is made transparently clear that any Organisation, no matter its size, is responsible for all personal and personally identifiable data throughout the organisation no matter how collected, stored and accessed.
- This requires Data Governance

Cost of non-compliance

- Infringement of Articles 5, 6, 7 and 9 carries a penalty fine of up to **€20M or up to 4%** of total global revenue of the preceding year, whichever is greater.
- Infringement of Articles 8,11, 25-39, 42 and 43 carries a penalty fine of up to **€10M or up to 2%** of total global revenue of the preceding year, whichever is greater.
- These penalties are for each audited breach.

Data Governance

- The penalties force any Organisation to pin the overall responsibility for Data Governance to the Executive Board or its equivalent.
- The reporting on Data Governance may be carried out by a member of the Board as Data Controller, such as a CDO or CIO, but the overall responsibility is the Board's
- Data Governance must not be treated as a matter of Technology alone. Controls and Policies may be implemented using Technical Solutions but it must be remembered that the Data covered belongs to the individual and not to the Organisation that holds it. This means that the Policies adopted must make that clear in terms of Data Governance.

Principles

- Personal Data, or Personally Identifiable Information belongs to the Individual
- Data must be categorised, at a minimum as:
  - Personal
  - PII
  - Other
- Access must be authorised
  - Including the context or purpose of the access
- Where data is and its state must be known and recorded.
  - All data
- All movement, change and use of data must be auditable.

Personal Rights

- An individual's right to access their data
- An individual's right to port their data to a new service provider
- An individual's right to erasure to-be-forgotten
- An individual's right to notification within 72 hours in the event of a breach

Personal Data & PII

- ‘personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;
- This GDPR data covers both traditional Personal Data, things and facts we know relate to an identity, and Personally Identifiable Information (PII)
- **Personally identifiable information** (**PII**) is any **data** that could potentially identify a specific individual. Any information that can be used to distinguish one person from another and can be used for de-anonymizing anonymous **data** can be considered **PII**.
- The word ‘potentially’ is critical here.

Protecting against inadvertent leakage of identity

- PII includes data that was collected anonymously and later other collected data or behaviour may become associated with it to the point where it identifies an individual.
- This can become an inadvertent leakage at any time.
- The safest course is to treat all anonymous data and events as unique. This uses the concept of pseudonymous identifiers.
- Convert the identifiers in any event (IP address, cookies or any other identifier) to use a new pseudonymous identifier.
- The same pseudonymous identifier cannot be used in any other context or use, otherwise the same possible identity leakage could occur.
- The most common example of this is Access logs, though any event log may have the same issue.

Categorising Data

- Before any system changes or new logging, auditing or monitoring systems are implemented the following must take place:
  - Auditing and identifying all of the current data in the current system, including all of the Access policies, locations, processes, monitoring and alerts etc
  - Identify where Data is not currently controlled or where there are no understood Access policies
  - Categorise each piece of data as being Personal/PII and any other category that makes sense to the organisation, such as Confidential, Financial, etc
- Don't forget the processes that acquire data and categorise those in the same way.

Data Policy

- For each Category and context of use and retention define the access and usage Policy.
- The same Category (and individual sets of data within a Category) may have multiple Policies depending upon the context of who or what is accessing and for what purpose.
- The resulting access policies will be a matrix mapped to the data. Individual data stores and application systems will have different implementations. Where a Data Store cannot provide the necessary policy the Application or Operating System will have to provide it.

Security Policy

- Quite apart from any Access Security rules there is a fundamental policy that must be adhered to for all Personal/PII Data.
- Encryption
- All such data, or potential data, must be encrypted in flight and at rest.
- This means every transmission of data whether across public or private networks must be encrypted, unencrypted on receipt and reencrypted for the next hop.
- Data stored in any cache, file, database or other storage system must be separately encrypted.
- And most importantly the Key Management Policy and Procedures must be defined and implemented rigorously. Leakage of Keys leaks all data they protect.

Managing Data

- Principally for the purposes of GDPR audit the following needs to be known
- What data has been acquired, modified, stored, archived, deleted etc
- What was the source
- Who required it (including processes)
- When did it happen
- What happened
- Where is the destination(s)

Data Provenance

- Knowing and maintaining the record of these events is called Data Provenance
- The chain of events that happened to a piece or set of data from the time it entered the system, including the original source provenance.
- Often this is thought of as relating to Data Sets rather than data collected interactively from an individual. There is really no difference, individually collected data becomes part of the collective data set.
- This is the Big Data/Data Automation method.
- Systems like BigID https://bigid.com/gdpr-data-subject-automation/ claim this

SIEM

- An alternative way of Data Management is to use the logging of all events together with the categorisation of data, Policy rules and so on with a fairly standard SIEM system.
- Security Information and Event Management (SIEM pronounced SIM) systems are used in general to log all access and identify anomalous or threat like behaviour and then alert upon it.
- For the identification of threat patterns, access methods and so on SIEM systems do work, though they tend to take some time to train to avoid false positives.
- Whether they will be successful for the detection and alerting of nuanced and subtle breaches as in the correlation of sets of anonymous behaviour is unclear.
- SIEM can be implemented in a home grown way using Splunk and other Log shipping open source solutions. Regardless the volume of log data can overwhelm the unwary.

Baseline and incremental testing

- Testing whether personal data can be leaked or subverted requires more than external penetration testing.
- Penetration testing is basic to knowing how safe you are but it doesn't cover all the situations where personal data may leak.
- Internal systems, including any which generate or interface with email systems can leak data both inadvertently and due to ignorance on behalf of the internal user. Internal systems should be penetration tested as well.
- A great many tests can be automated but the critical tests are the social tests.
- Compliance becomes easier if it is part of every procedure, every process and frequently tested. This includes penetration testing, Phishing campaigns, etc.

Summary of Steps

- Take ownership fo Data Governance at the highest level
  - Appoint a Data Contrioller and provide Board Level approval for resourcing.
- Categorise all of your data.
- Update or create all of the policies relating to Data.
- Plan and implement systems that can report on the provenance of all Data entering, whilst its within and leaving your systems.
  - Do not forget paper records and systems
- Plan, schedule and implement external and internal penetration tests as well as social testing such as Phishing etc.
  - Implement as a matter of priority test findings.