Data Minimisation Checklist

Data Minimisation for Risk and Compliance

The ultimate checklist

Data minimisation is a security and privacy principle that requires organisations to limit the amount of sensitive information they hold, knowing that data in their systems could be breached at any time. Castlepoint has assembled a complete checklist that can be used as a reference to implement data minimisation for organisations at risk of privacy or security breach. Following are recommendations that help design, implement, and maintain minimisation.

Introduction to the checklist

Designing

It is important to understand the threat environment, and the internal governance structures for minimization policy and process. You must know what rules apply to your information (including retention, secrecy provisions, and other regulatory obligations). Finally, you need a minimisation business case, with ROI.

Implementing

To achieve data minimisation, you need to collect only the required information for the purpose, keep it only as long as required, and ensure minimum access and privileges on the data. To balance the tension between the value of the information, and its risk, you need to implement records lifecycle controls.

To understand your data at scale, artificial intelligence is essential. But implementing autoclassification and automated decision making (ADM) requires consideration of risks (supportability, sustainability, and ethics).

Maintaining

Your data changes constantly, at high velocity. New risk is introduced faster than existing risk can be treated in traditional models. You must be able to monitor changes in your data estate, so that you can dispose of superfluous risk data as soon as possible. You need to ensure that your data minimization does not affect data that must be preserved. And you must be ready to respond in the event of a breach, by reporting on the impact.

1 Designing

Understand Minimisation

Understand Regulations

Make a Business Case

Understand Data Minimisation

- [ ] Review threat assessments, external and internal, and map them to your data types and assets. Understanding the types of bad actor, and what benefit they would get from taking your data, helps make a clear case for what data to focus on minimising first.
- [ ] Establish formal engagement with your governance teams on data. You have to work closely with the parts of the organisation responsible for your data obligations.

Understand Regulations

- [ ] Apply Privacy by Design in every sprint. Research your obligations, not just for your jurisdiction but also for the type of data you hold, and what types of people it is about.
- [ ] Build in records registration, classification, sentencing, and disposal into all your systems that hold sensitive and high value data (not just traditional EDRMS).
- [ ] Apply Defence-grade national security coding practices like OWASP when you develop your software, whoever you are. Whether you are a national security agency or not, you have an obligation to protect sensitive and high-value information from adversaries.

Make a Balanced Business Case

- [ ] Calculate the likely cost of a breach based on the number of records you hold, using established metrics. Price up your own database in terms of impact. The average *per-record* cost of a data breach was US$165 in 2023. Consider how many people you would have to compensate or insure, and any civil penalties that you would incur.
- [ ] Map the potential cross-functional benefits for your planned data minimisation investment or development, and quantify them where possible. Being able to understand and articulate how your data minimisation development can positively affect stakeholders across the business is vital for achieving Executive support for investment and innovation.

2 Implementing

Minimise Collection

- [ ] Develop an automated asset inventory of every document, email, database row, web page, and chat in every system, on every platform – including its location, attributes, and content. The only way to know how much data you have collected about individuals or high-risk topics, and where you are collecting and keeping it, is to audit your entire environment.

Minimise Access

- [ ] Implement monitoring of user activity on sensitive and high-value data. Ensure download, share, deletion, and downgrade of security markings, for example, are captured for later review. Logging who is accessing content helps detect early signs of breaches, and also helps plan for and enact credential revocation that may otherwise be missed.
- [ ] Ensure that systems you develop can proactively alert on data activity, specifically activity that may lead to data proliferation, precipitous destruction, or unauthorised access.

Minimise Retention

- [ ] Ensure records management capability is dynamic and manages records for their entire lifecycle. Information changes over time. Every time its context and content changes, and as users interact with it, both the applicable retention rules and the disposal trigger date will change. Your records management capability needs to detect and adapt to these changes, and, in the end, alert as soon as a record can be disposed of under law.

Implement Autoclassification and Automated Decision Making

- [ ] Plan for rollout of data minimisation capability to new, planned systems, and plan for retrofit to existing systems, prioritised on a risk basis. All your data can have risk, and all of it can have value. Autoclassification needs to run across structured and unstructured data, on prem and in the cloud, across any type of system (including legacy and bespoke systems).

- [ ] Implement data minimisation capability in a way that does not introduce tight coupling between technologies or affect the source data. Design and develop an interface model, rather than an integration model. Rollouts of technology that causes a detrimental impact on users, source systems, original data, or the governance team will not succeed.
- [ ] Review the ethics of your planned AI and ADM before deploying it, using tools like the ALTAI self-assessment. Ensure sufficient human oversight at all stages of development and use. The Assessment List for Trustworthy AI is a wizard-driven tool created by the European Commission that will give you a picture of how your systems and processes comply.

3 Maintaining

Monitor Data Risk

Preserve Data Value

Be Ready to Respond

Monitor Data Risk

- [ ] Implement capability to proactively destroy or deidentify personal and other attractive data on a daily basis, as soon as it comes due for disposal. Continually grooming the environment to remove data that carries risk, but no longer has business value, helps to minimise the success of a foreign state or sophisticated criminal operation.

Preserve Data Value

- [ ] Ensure that all types of retention policy are considered, and overlaid, to ensure that records are not deleted prematurely by triggers in your system as built. Ensure that appropriate subject matter experts are involved in the final decision to destroy data. While there is privacy legislation requiring you to destroy personal data as soon as possible, the exception (within the allowed uses) is where it has continuing value to the business. That be value can archival, social, or operational.

Be Ready to Respond

- [ ] Build in reporting to quantify the impact of a breach. Plan for a breach, and game out how you would be able, from the back end or the UI, to report quickly to the Executive on who is in your spill, what exactly was taken, and what laws were broken in the process.

Castlepoint's Ethical AI

Castlepoint manages over 286.5 million records for its clients in over 1.6 million separate systems. We have identified more than a quarter of a billion sensitive and high-risk records in enterprises that require protection.

With Castlepoint's ethical, explainable AI, organisations can sentence their records across dozens of systems with up to 98% accuracy. With our XAI, we have applied records retention rules at scale to these records so they can be appropriately preserved or lawfully destroyed. This has helped reduce the potential harm of any future data spill which can result in financial and reputational damage.

Castlepoint is explainable artificial intelligence (XAI) that reads, registers, and automatically classifies and sentences all the information in a network, in every format and system, to make sure you can find anything, anywhere, anytime. And it does this without any impact on users or source.

It provides a single view of all your information that enables:

- Enterprise search, eDiscovery, and eHold
- Manage-in-place compliant records management
- Cyber security and privacy risk management
- Auditing of data, users, and Generative AI

About Castlepoint

Castlepoint Systems' powerful AI-driven platform is the intersection between cybersecurity and data governance. It is the only solution using ethical AI to unlock information management, eDiscovery, governance, risk, and compliance for the fastest, easiest way to find, manage, protect, and de-risk all an organisation's information in a single interface. Founded in 2012, the multi-award-winning solution allows governance and risk teams to have a complete view across the whole environment and optimise investment in security, eDiscovery, and compliance processes. With offices in Australia and the United Kingdom, Castlepoint's scalable and transparent AI enables organisations to safeguard communities and people from data mismanagement and risk.

Follow Us:

@CastlepointSys2

castlepointsystems