A practical guide to implementing a data retention policy

drs. J. Blaauw en Y. Ajibade Msc*

Trefwoorden: data management, data retention, data retentie, record retention, GDPR, AVG, data, dataverwerking, privacy

Every organization processes[1] data for different reasons and in different ways. In a data driven world, an organization largely depends on the data it has and uses.[2] Part of fruitful processing data is making sure that you know when data must be kept and when it must be removed.

Data[3] is subject to different data retention periods, which may vary per country and/or industry. As a result, thousands of retention rules require you to either keep or destroy data and it is challenging to get advice on this topic. Implementing a data retention policy will help organizations to be in control of their data and it will reduce the risk of being non-compliant with laws and regulation, including the General Data Protection Regulation (GDPR).[4] Properly implementing such a policy takes effort, commitment and management support. As a bonus, up-to-date and relevant data increases the value of your organization.

This article offers a helping hand to organizations that are in various stages of implementing a data retention policy. In the first part of this article, we will focus on the legal framework. Here we will zoom in on the complexity of various rules and regulations. In the second part, we outline eight steps to build a solid data retention policy. We will conclude by looking at the creation and implementation of a data retention policy through real-life examples and insights.

Chapter 1: Legal Framework - GDPR and... all the other laws in the world

The GDPR and record retention closely relate to each other. Privacy laws often require organizations to justify their processing of personal data and to set an end date for the data life cycle. After the end date lapses personal data needs to be deleted.

There are also data retention rules relating to non-personal data. These rules often tell organizations to store such data for a minimum period (meaning that is allowed to store non-personal data for a longer period than the legally required minimum).

The GDPR tells organizations never to process personal data longer than necessary for the purpose for which the data have been collected or used.[5] Based on this very generic GDPR rule some Data Protection Authorities have issued guidelines with regard to certain categories of data (e.g. recruitment data, CCTV or health data). Where Data Protection Authorities have not given specific guidance, it is up to organizations to make their own risk assessment. This can be quite a daunting task. It means that organizations should implement a personal data retention period per purpose for which it processes personal data. So how to determine what is 'necessary'? Virtually every minimum retention period relating to a piece of information containing

* Joris Blaauw is Senior Compliance Officer, Group Data Protection Officer and Group Privacy Officer within SHV Energy. Yomi Ajibade is legal analyst at filerskeepers.co and focuses on global data retention.

1 Processing means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

2 Nicolaus Henke et al. (McKinsey Global Institute), *The Age of Analytics: Competing in a Data-Driven World (December 2016).*

3 Data retention was historically called 'record retention'. In a digitalized and post-GDPR we prefer the term 'data' as it more adequately describes information as a production factor of immense value.

4 Rita Heimes, 'Top 10 Operational Responses to the GDPR – Part 5: Preparing and implementing data-retention and record-keeping policies and systems', IAPP (February 26, 2018), https://iapp.org/news/a/top-10-operational-responses-to-the-gdpr-part-5-preparing-and-implementing-data-retention-and-record-keeping-policies-and-systems/ accessed on September 18, 2018.

personal data becomes a maximum retention period if the GDPR applies: when the retention period is fulfilled, there is no legal requirement to store the associated personal information anymore (if there is also no legitimate interest to keep the data for potential litigations). Organizations will need to ask themselves per data category: what is the shortest personal data retention period my organization could implement before it gets into trouble? Beware, the European Data Protection Authorities may ask to provide examples to illustrate! An illustration can be either the law that requires you to keep data for a specific amount of time or case examples that justify the chosen retention period. The latter can be specific requests from a data subject that require you to keep data for x amount of time. The examples should be reasonable and proportionate.

Then again, record retention can be very difficult and 100% compliance is often impossible. Per country, hundreds of record retention rules could apply to your business. Only China is good for over 500 retention periods and Russia counts over 1200 retention periods for both personal and not personal data. Retention rules are often conflicting. What would you do if you have a central system and your payroll data should be stored at least 50 years in Poland and Romania while in France those same payroll data should be deleted after 6 years? Compliance with record retention rules in one country can lead to non-compliance or loss of litigation position in another[6], if the payroll system cannot handle different retention schedules for different countries.

There is at least one commonality in most countries: keeping data forever is generally not allowed. Multinational organizations will have to make sense of all those retention periods applying to their data from around the world. Here comes the tricky part: a granular approach to record retention is probably not technically possible, as most IT-systems or Cloud applications where data resides do not allow differentiation per country, data category or data point. The choice you have comes down to keeping data forever or deleting it after one generic retention period has expired. As a result, many organizations implement simple custom tailored retention periods to ensure compliance with most record retention requirements instead of all. We call these optimal retention choices **'golden standards'**. These golden standards are often written down in a neatly structured global data retention policy. Below you will find the eight steps to create your perfect retention policy.

Chapter 2: 8 steps to create a perfect global record retention policy

A good global retention policy contains all retention periods applicable to data held in your organization. It tells you who is accountable for compliance with the policy, who owns the policy, who should keep what data for what purpose, for which time period, starting when, and if it is a maximum or minimum period, and preferably all with a link to the legal reference. Moreover, it tells you how to destroy data. Let us get started:

1. **Determine your retention strategy**
   First decision: is your organization a data hoarder or a strategic litigator? Data retention is not an exact science and it very much depends on your geographic scope, the industry you are in and the choices that your organization makes with regard to the use of data. Geography can matter. For example, if you are in China where the laws tell you to keep some data forever while in the US or Brazil you may risk eDiscovery procedures, which often inspire organizations to keep their data a bit shorter.[7] Your industry also matters. For example, telecom operators are often regulated to keep data for a shorter period compared to organizations dealing with nuclear waste. Oftentimes it just depends on the business model of a certain organization and the use cases for its data. In that case, a retention policy focused on data use maximization is often the preference.

   In other word, do you want to have your information destroyed before the eDiscovery subpoena comes in? Or are you just too attached to your data because your business is data driven? Whatever it is, plan your strategy. It will help you to choose the most optimal golden standard (see step 4 below).

[5] Article 5 (1)(e) GDPR.

[6] https://www.quora.com/Why-is-records-management-important.

[7] Electronic discovery (also eDiscovery) refers to discovery in legal proceedings in the US such as litigation and government investigations, where the information sought is in electronic format (often referred to as electronically stored information). Electronic discovery often entails far reaching data mining and is subject to rules of civil procedure often involving review for privilege and relevance before data is turned over to the requesting party. Organizations often experience eDiscovery as impactful and burdensome.

2. **Determine your governance**
Who is accountable for your retention policy? Will it be the CIO or Chief Privacy Officer (CPO), or would you like to go higher up in the tree, say CEO, COO or CFO? The more you care about your data, the higher to place accountability. Yes, it is just that important. Even if you are not data driven in the 2018 kind of sense, record retention rules have impact on your core business processes. Moreover, who better than the CEO or CFO to make risk decisions about your core business operations, right?

   Which other people or functions will need to be involved? Legal and finance of course, they will need to be able to advise on changing rules and issue legal hold or tax hold notices which are aimed at ensuring that data is kept in the event of an investigation or litigation. Moreover, do not forget about your privacy, compliance and information security functions who are there to ensure that the business does not end up in the arms of the Data Protection Authorities.

3. **Determine what deletion means**
What happens after the sign 'delete' has been given? Deletion means destroying in the purest sense: a record has been disposed of when it is really gone and you cannot access and recover it. No copies, no cache and no backup tapes are available and you have no means to reconstruct the record. EU data protection regulators call it 'irreversible deletion'.[8] Dust off your shredder, burner and eraser and destroy those documents. It is not painful; it is part of data retention (and life).

   However, please do not start stressing out right away; there are also soft-deletion options or restricted archives which you could consider. Soft-deletion means that data is 'fake' deleted first.[9] In case someone misses the data dearly, you can still retrieve it. The same goes for restricted archives, with the exception that an archive often serves a specific purpose (e.g. historical purposes or litigation). This means that you should not use or access the data outside of the archive's purpose. In any case, soft-deletion and restricted archives are not the same as deleting data. However, it does improve your story line when facing a regulator or judge. In any case, your global data retention policy should include an instruction on deleting data.

4. **Don't miss out on any important categories**
When choosing your most important retention categories, take a close look at your data. What data do you have? What are your core business processes, and what data is used in these processes? We have a few ideas to get you started. Every organization has a few must have business data:

   - Accounting data;
   - Tax data;
   - HR data;
   - Health and safety data;
   - Environmental data;
   - Personal data.

   In addition, do not forget about litigation. Statutory limitation periods (the maximum time within which someone can file a claim) are often very good guidance on what and how long you need to store data. When you have included these, it is all about the industry related storage terms that matter. Relevant industries include:

   - Financial service;
   - Telecoms;
   - Pharma;
   - Construction;
   - Transport;
   - Aviation;
   - Critical infrastructure;
   - Manufacturing;
   - Real estate;
   - Hospitality;
   - Professional services.

[8] Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland and Seitlinger and Others v. Ireland*. 8 April 2014, http://curia.europa.eu/juris/document/document.jsf?docid=150642&doclang=EN accessed on September 19, 2018.

[9] Eugene Gilburg, 'Soft-deletion is actually pretty hard', (March 23, 2018) https://medium.com/build-acl/soft-deletion-is-actually-pretty-hard-cb434e24825c accessed on September 19, 2018.

It is all about keeping it simple. There are so many retention periods in the world that the minimum retention period is often actually good enough and you do not need to store data longer. Limit the number of categories and per category limit the number of retention periods (if possible).

5. **Establish your golden standards**
How to determine your golden standard? Now comes the difficult part from a more political and international relations perspective. What value do you attach to all those different retention laws applicable to your data? It may be simpler than you think; determine what the important countries are, then choose a retention trigger and a strategy to find your optimal retention period. Make sure to achieve compliance in as many countries as feasible – within your organization's technical limits.

How this could work:
Geography: First, identify the countries in which your company is based. Then attach a value to these countries. How do they rank in terms of:

- main establishment and regional headquarters;
- number of employees;
- turnover;
- location of data and data centers.

Interesting thought: some organizations will deem all laws equally important and do not want to rank countries at all.

Choose a trigger: to be able to compare retention periods, you will need to understand the different triggers that mark the start of the retention period, think of:
- Moment of creation of a document or entering information into a database;
- Close of calendar year in which a document was created;
- Close of tax year in which a document was created;
- Moment of termination of a contract;
- Date of last activity.

Once you know when the retention period first starts running in a certain country it will help you to determine exactly how long these retention terms are in practice (5 years after creation of a document will mean a much shorter retention period than 5 years after the termination of a contract).

Select your desired retention length: now that you know which laws matter and what the different triggers are, you can calculate the optimal retention period, by looking at:
- The various minimum retention periods applicable to the multinational;
- The various maximum retention periods applicable to the multinational.

Practical note: some multinationals will give precedence to minimum retention periods instead of maximum retention periods. Now mix and match, and you will have your desired trigger and retention period based on the laws that matter to you!

6. **Create actionable retention periods**
The number one problem with retention policies is that IT cannot implement an unlimited number of retention periods. That is why we insist on keeping it simple and actionable. This will allow you to actually implement the retention period. No vague language such as 'current +10 years' but '10 years from the date on which the book year ended'. Also, be clear on what should be stored. Do not use 'user data' if the law only requires you to store an 'IP-address and login timestamp'. This will allow your IT team to implement the retention period.

7. **Ask for feedback from data users**
Although successful record retention policies are driven top down and not bottom up, it is important that you test your golden standards with the business. Stakeholder management is key. There may very well be a good reason why some departments keep data longer or delete it a bit sooner. **Remember, it is the business that is responsible for its own compliance. The business will have to implement the retention policy.** Collaborating with department heads to gain their support for a global retention policy, and ensuring that their own efforts are leveraged as part of the broader policy, is essen-

tial. In a world of agile and scrum, feedback just makes processes better. Even data retention.

8. **Never forget about the law (deal with residual risks)**
   The bad news in this story is that 100% compliance with record retention rules is probably not possible due to conflicting laws and limitations set by technology. If your golden standard is 7 years while French law has a maximum retention period of 6 years and Russia wants you to store forever, you simply cannot be compliant with all laws except for France and Russia. For these countries, you will need to make a risk assessment and fix it (or not but that is up to you). This goes for all outliers from the golden standard. Think of:
   - A minimum retention period that is longer than the golden standard;
   - A maximum retention period that is shorter than the golden standard.

   After this initial risk assessment, you will need to assess how these risks can be classified in terms of non-compliance and what the chances of materialization of such risk are. Also, please accept the simple fact that the laws are constantly changing. What applies today may not apply tomorrow.

   Finally, never forget about the GDPR (and all other non-EU data protection laws). Data protection laws tell organizations never to store personal data longer than necessary for the purpose for which the data have been collected or used.[10] As a result, virtually every minimum retention period relating to a piece of information containing personal data becomes a maximum retention period.

If you use these principles, you are well on your way to build a future proof and compliant retention schedule.

Chapter 3: Theory into practice

When you start applying the eight steps, you will encounter some challenges. In this chapter, we share examples and best practices that we have seen when creating and implementing a global record retention policy.

Creating a policy
When drafting the data retention policy, one of the major tasks is actually finding the relevant and applicable data retention periods that apply to your organization. There are several ways to collect the information about relevant data retention periods. As it is extremely labor-intensive to do it yourself, we recommend collecting this information from a central source such as country retention schedules provided by the filerskeepers.[11]

Before deciding on what a global data retention policy looks like, make sure that you do not only have management buy-in, but also organizational buy-in. Involving key stakeholders will be a mandatory requirement as we are talking about data 'owned' by the business. There should be a common agreement who in the organization has the ownership of the data and who has ownership of the data retention policy and related activities. Within an organization, individuals might feel hesitant to call themselves data owner of a particular set of data. It is common that individuals will refer you to the highest level of senior management. This level of management might be ultimately responsible, yet they are generally not involved with day-to-day management of data at lower levels and therefore not aware what data is present and needed on a micro-basis. Properly determining the right level of ownership is essential. Data ownership can be delegated by the highest management level to the appropriate level and this delegated data ownership must be supported by the policy and guidelines.[12]

Data ownership and ownership of the data retention policy is also important when weighing the need to delete data from a privacy/compliance perspective versus the business need to retain the data. Various functions and departments play a role in this. Legal, Information Security and Compliance take part in deciding golden standards and implementing these in systems. HR, Marketing/Sales, Finance etc. provide input on how long they need data for their business processes. This should be covered in the policy and tasks should be assigned to the project team that creates and implements the policy.

[10] Ibid (n. 3).
[11] www.filerskeepers.co.
[12] Michael Gregg, 'Logical Asset Security: Data Management- Determine and Maintain Ownership' (CISSP Exam Cram, 4th Edition) (March 10, 2017). http://www.pearsonitcertification.com/articles/article.aspx?p=2731933&seqNum=3 accessed on September 19, 2018.

For organizations that operate in multiple countries, it might be worthwhile to create various data retention periods. Especially organizations with a decentralized character, local operations with independent processes and systems allow for a decentralized approach to data retention. The organization's headquarter should write an overall data retention policy that clarifies the data retention choices made and provide direction concerning the overall data retention strategy. This guidance combined with information about the retention periods applicable to the country, provides the local organization with the tools to create their own local retention policies.

Whether an organization operates in one country or worldwide, whether it is small or large, template data retention policies are available online. filerskeepers has recently posted one on their website. A solid template can be used as a starting point for your organization's very own data retention policy.

Implementing a policy
Implementing a data retention policy may prove to be an even bigger challenge than creating one as most organizations have collected vast amounts of data, dating back to various moments in time, processes and (former) data owners. It is important to understand that properly implementing a data retention policy may take years (but no reason to postpone!). Here are a few tips to get started:

- **Set up a data retention taskforce.** Ideally, the record owners take on this responsibility. However, in some cases the data have become orphans in the course of time. In our experience, it pays off to take a pragmatic approach. A small project team consisting of experts with regards to privacy, data retention and other legal specialists such as labor law and litigation can take data owners by the hand and make decisions when no data owner can be identified. Identified data owners might know what the data was meant for, but will need assistance with determining if the data can be kept in line with the applicable data retention period. Additionally, in many cases it is required to keep a record of data sources that have been deleted. This record should contain, yet not limited nor restricted to, the name of the files, former location of the file(s), date of destruction and the authority of the one(s) who gave instruction to destroy the data.

- **Create a long-term implementation plan.** With vast amounts of data stored in different places, and new data being collected continuously, planning and assigning responsibilities is vital.

- **Find the right place where data is stored.** Data can be structured or unstructured, part of a current process or merely a souvenir of former times. Finally yet importantly, data can be stored in a digital environment or stacked away in a file cabinet. Finding the right place to start implementing the data retention policy will help you prioritize and plan your efforts.

- **Find quick wins together with stakeholders.** Old systems that do not allow the deletion of data might be replaced by a new system that allows you to build in data retention periods on short term. Finding quick wins makes implementation and maintenance easier yet depends on the circumstances.

- **Always look ahead.** Ideally, data retention should be covered in any new system before that is implemented. In order to get to such a situation, we advise you to apply the privacy-by-design principles. Data retention is an important factor in the privacy-by-design cycle. When this concept is also used for the rare processes and systems that do not contain personal data, data retention will be completely covered. When an organization takes this approach for some years, many issues related to data retention will be solved.

- **Organize special clean desk days.** Many employees store data in their personal file cabinets or on personal drives on their devices. Some organizations have organized special clean desk days on which employees are asked to remove data that should no longer be kept. During these days, experts are available for questions and guidance. These days also provide momentum for raising awareness within the organization regarding data management.

- **Use anonymization and pseudonymization.** Alternatives that are sometimes used instead of deletion are anonymization and pseudonymization. Pseudonymized data might

still allow for some re-identification, while anonymous data cannot be re-identified. Because of the chance of re-identifying someone, the data retention period for pseudonymized data is no longer than the original data. Pseudonymization can be an alternative to keep data for a longer period of time. However, the usability of this data is in most cases limited.

The challenge of unstructured data: email

Another source of unstructured data that necessitates a thorough assessment is email. Email contains different content and has various kinds of documents attached to it. In theory, it is triggering all data retention periods. It is important to understand that **email is not an archive,** contrary to actual practice in most organizations. Instead of archiving important documents in a tool designated for data management, many employees store data in their email folders. This poses various risks including the following:

- Data stored in email boxes are subject to data retention requirements. Keeping them longer than required can lead to a violation of applicable laws.
- Current and future delegates may have access to information that they should not have access to, merely because it is stored in the mailbox of the person who gave them access to their folder.
- During employment, the documentation is only accessible to employees that have received this document directly. Other employees with a need to know cannot access it, reducing the learning curve of others and reducing knowledge sharing.
- When an employee leaves the organization, the data most likely will be lost when his or her account is deleted.

Given that email will contain different content and various kinds of documents, it is quite a beast to be tamed. However, we see two options:

a) Technological tools are available to categorize emails and documents, for example, Titus or certain corporate email suppliers. These kinds of tools require employees to indicate for every email what kind of data is included in that email. Based on various predefined categories or archives, the employee will make a selection. The applicable data retention period corresponds to the category or the selected archive. The advantage is that this option allows for differentiation in data retention periods associated with a category or archive. The downside is that it requires a manual action from the employee for every email that is sent or received. Something that most employees do not appreciate on top of their busy work schedule. And this also increases the risk that employees will select the wrong category or archive.

b) By setting generic data retention periods for all email within an organization, you can implement two time frames (x and y period). After x amount of years all emails are 'soft deleted' and placed in an archive that is not accessible to employees. In order to get access to emails in this archive a request needs to be filed with the IT department. This request will be judged on relevant variables that include necessity and the data retention period for the requested data. This might mean that the request of an HR officer to pull up a resume of a job candidate that was not hired two years earlier will not be granted whereas a request for a contract that is still in effect will be granted. After y amount of years, where y is larger than x, the emails will be deleted altogether. In practice, this could mean that email is placed in an archive after two years and deleted after five. The exact number of years are based on legal retention periods and should be defined by the organization.

   This approach might disturb employees. Therefore, it is important to explain the logic behind this and offer employees alternatives to accommodate their way of working. A relevant question to ask is how often the employee has recently looked into an email that was sent or received more than x years ago.
   As with most things, there will be exceptions to the rules. When introducing these kinds of generic rules, the organization has to identify the exceptions. Commonly found exceptions are legal functions, which need documentation for litigation purposes or health and safety functions that need to comply with long retention terms. When identifying the exceptions to the rule, one must be critical and evaluate if an exception is appropriate in view of applicable law. As there is a bigger risk for non-compliance with data retention rules when an exception is applicable, additional mitigating measures should be considered.

Final remark

In this article, we have shared a legal framework, eight steps to creating a solid data retention policy and shared practices. It will allow you to gain control over data management and generate value from data. As the saying goes, 'you better use it, before you lose it'. ■