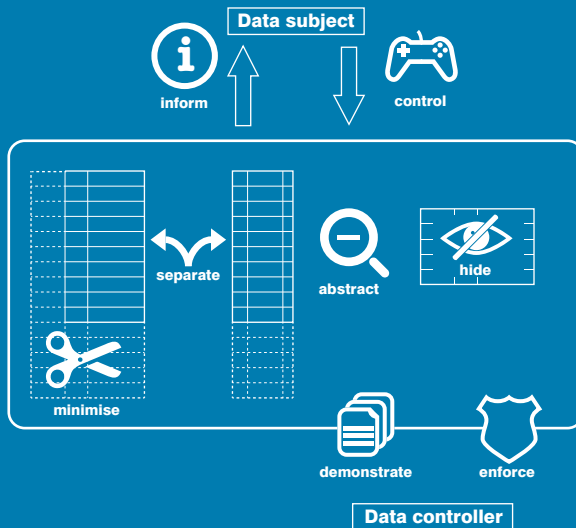


Privacy Design Strategies

(The Little Blue Book)



Jaap-Henk Hoepman

January 27, 2020

Copyright ©2018 – 2019 by Jaap-Henk Hoepman.



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0). To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

For commercial use, or to obtain hard copy versions of this book, please contact the author at info@deprivacycoach.nl.

1 Introduction

We want to determine by ourselves which personal details we share with others, and how they are used. We don't want everyone to know what we do or think. We don't want our boss to know what we discuss with our friends. We don't want our personal information to be misinterpreted outside the original context. In other words, we want some privacy please. But that is increasingly less self-evident in a world where digital innovations surround us and literally starts getting under our skin.

Therefore privacy is important. It protects us against intrusive companies and an omniscient government. The goal is to maintain a balance of power between the individual and society. In a democratic state this is not only of personal interest but also of interest to society at large. This is why privacy is a fundamental right. Strong European laws protect the privacy of all citizens on European soil.

Unfortunately, these laws are complex and vague. They offer little concrete guidance to designers and system developers. This is a problem if you want to design privacy-friendly systems. For example by applying the *privacy by design* philosophy, which demands that privacy requirements are taken into account right from the start and throughout the system development life cycle. This makes privacy, like security, a software quality attribute. Privacy by design is a legal requirement from 2018 onward. But you can also use it to go beyond the bare minimum required by law, and use it as an innovative force.

But how to make privacy by design concrete? And how to apply it in practice? Those questions are answered by this little book.

There are many privacy-enhancing technologies, but these are only relevant once you start developing your system. They are of little use when you are just starting to think about what the system you are about to build must be capable of, and how those capabilities can be realised in a privacy-friendly manner.

Where and how to begin?

Privacy design strategies aim to answer that question. They translate vague legal norms in concrete design requirements. They provide talking

points to explore the design of the system. They guide the initial design sketches into a privacy-friendly direction, forcing one to make fundamental design choices early on.

1.1 Audience

This book targets all organisations (businesses or government) that process personal information. It is in particular intended to be used by designers and engineers who build systems processing personal information, and the people responsible for these systems.

1.2 Reading guide

This book describes eight privacy design strategies. Every strategy is briefly explained and illustrated through several practical examples. A few concrete technologies that can be used to implement each strategy are also mentioned.

In the third to last chapter we will discuss how to use the privacy design strategies in practice, and how to integrate them into existing system development methodologies.

If you want to learn more about privacy in general, or about privacy by design in particular: the end of this book lists a few pointers to books, websites and other sources of additional information. There you can also read how to stay informed about further developments.

We conclude with a glossary.

1.3 Acknowledgements

I would like to thank Gergely Alpár for his comments.

2 The eight privacy design strategies

We distinguish eight different privacy design strategies, divided over two different categories: data oriented strategies and process oriented strategies.

The data oriented strategies focus on the privacy-friendly processing of the data themselves. They are more technical in nature. There are four of them.

Minimise

Limit as much as possible the processing of personal data.

Separate

Separate the processing of personal data as much as possible.

Abstract

Limit as much as possible the detail in which personal data is processed.

Hide

Protect personal data, or make it unlinkable or unobservable. Make sure it does not become public or known.

The process oriented strategies focus on the processes surrounding the responsible handling of personal data. They deal with the organisational aspects and the procedures that need to be in place. We distinguish the following four.

Inform

Inform data subjects about the processing of their personal data in a timely and adequate manner.

Control

Provide data subjects adequate control over the processing of their personal data.

Enforce

Commit to processing personal data in a privacy-friendly way, and adequately enforce this.

Demonstrate

Demonstrate you are processing personal data in a privacy-friendly way.

Within a strategy we distinguish several *tactics* that each provide a clearly different way in which the overarching strategy can be made more concrete. In the next eight chapters we will describe each of the strategies and their associated tactics, and provide examples of how they can be applied in practice.

3 Minimise



Limit as much as possible the processing of personal data.

The most obvious strategy to protect privacy is to minimise the collection of personal data. Nothing can go wrong with data you do not collect: they cannot be abused, they cannot be misused, or get leaked accidentally. Think carefully about the personal data you really need. Sometimes a radically different approach may lead to a design that requires significantly less data, or may require no personal data at all.

3.1 Tactics

Minimisation of personal data can be achieved by collecting data of less people, or collecting less data of people. Several different tactics are applicable.

Select Select only relevant people and relevant attributes. Determine beforehand which people and which attributes are relevant, and process only that data. Process only incoming data that satisfies the selection criteria. Be conservative when establishing the selection criteria: only select what is strictly necessary. Use a white-list.

Exclude Exclude people or attributes in advance. Determine beforehand which people or attributes are irrelevant. Do not process that data or immediately throw it away if you happen to receive it. Be liberal in grounds for exclusion: exclude as much as possible, unless you are certain, and can justify, that you need it. Use a black-list.

Strip Remove (partial) data as soon as it is no longer necessary. Determine beforehand the time you need a particular data item, and ensure it gets automatically deleted as soon as this time expires. If the data item is part of a larger data record, update the field to a default value indicating it is unspecified. Changes in organisation, processes or services may render certain data items irrelevant before their expiry time. Prune them.

Destroy Completely remove personal data as soon as they are no longer relevant. Ensure that the data cannot be recovered, even in unforeseen ways. Do not rely on logical removal strategies, that only block

the access to the data, but leave (traces of) the data still accessible to skilful attackers. Remove data from backups as well, and use secure ways to erase data from hard disks.

Considering minimisation forces you to think carefully about the fundamental values of your organisation and the core business you are in. “A cobbler should stick to his last”: if you are not in the primary business of profiling your customers for say advertising purposes, you should stay away from that. Minimisation forces you to be specific about your purposes; don’t be tempted to phrase them broadly.

In 2014 ING, a large Dutch bank, decided to offer third parties the opportunity to reach its bank account holders with targeted advertising, based on their transaction histories. This caused an uproar, and the bank quickly backtracked. ING customers considered this a huge breach of trust: financial transactions are quite sensitive, especially in the Netherlands where people do not easily share their financial situation or even their salaries.

3.2 Examples

Exclusion or selection of data is not only relevant when collecting information, or when personal data is obtained in another way, but also when using already collected data. Ensure that internal processes and applications only use the personal data that are truly relevant. Ensure that only relevant data is shared with third parties. And take care when the processing of personal data creates new personal data: select or exclude also in this case the new data that are truly (not) necessary.

Data mining, deep learning and Big Data create new insights. Select only the insights that are relevant. Throw the others away, and do not keep them because “it doesn’t hurt to try”: it sometimes does hurt.

In other words, also when using, sharing, analysing and enriching data one has to consider ways to minimise the final set of personal data retained.

The difference between stripping and destroying data is subtle: stripping happens in the application layer, while destroying focuses on the physical storage layer.

There are tried-and-tested methods to truly erase and destroy data from hard drives (for example by writing random data several times to each of the sectors involved). An efficient method of destroying backups is by encrypting the data before backing it up. By associating certain keys with particular retention periods, all data for a particular data can be destroyed by simply destroying the associated key.

A radically different system architecture may result in a much more privacy-friendly system. Consider for example a system for paying fees on toll roads. One approach would be to uniquely identify each passing car at each entrance and exit of the toll road. This would provide the operator with detailed information about everybody's driving habits: a huge breach of privacy. Another approach would be to install a small box in each car that automatically deducts the required fee from an associated account.

4 Separate



Separate the processing of personal data as much as possible.

Another important strategy is to (logically or physically) separate the processing of personal data. This makes it harder to combine or correlate data. By processing personal data from different contexts, the risk that data from one context becomes known in another context is reduced. Especially when the separation is physical. Hence separation implements contextual integrity.

4.1 Tactics

The following two tactics help implement the separate strategy.

Isolate Collect and process personal data in different databases or applications. These databases or applications are either logically separated, or actually run on different (yet still centrally located or controlled) hardware. Ensure that the logical separation is tightly enforced.

Distribute Distribute the collection and processing of personal data over different physical locations using databases and systems that are not under the control of a single entity. Use the equipment (PC, laptop, smartphone) of the data subject himself as much as possible, and use central components as little as possible. Use decentralised or even distributed system architectures instead of centralised ones.

4.2 Examples

A social network allows people to share status updates and pictures with friends and acquaintances. Current social networks (such as Twitter or Facebook) are centralised architectures: the platform sees everything that its users share with each other. That also determines the (tremendous) value of these social networks, and also underlies the business models of these networks. A privacy-friendly version of such a social network

would allow its users to store all updates and pictures locally on their own devices (e.g. their smartphones) and would share that information directly with their friends and acquaintances in a peer-to-peer fashion. In this case no data is stored or processed by a centralised component at all.

In general peer-to-peer networks or distributed algorithms could be used instead of centralised approaches, to improve privacy protection.

One small example to further illustrate this idea is the historical use of a passbook to record bank deposits or withdrawals, back in the days when people did not have bank accounts and wages were paid in cash. The passbook served as a local copy of all transactions, owned and controlled by the holder. Strictly speaking only the passbook itself could have served as the authentic record, alleviating the bank of the need to maintain any information about the 'account' and its holder. (In practice banks would have kept a record too to prevent fraud.)

Apple's iOS 10 allows users to group photos based on the people in them. This uses facial recognition software. In iOS 10 this software runs locally in the user's phone. The photos are therefore not sent to a central server for analysis.

An extremely privacy-friendly way of processing data in a distributed fashion is the use of *secure multiparty computation*. With this technology an arbitrary function over private inputs distributed over many different devices can be computed without these inputs ever leaving the devices. This technique has been used in Denmark to perform secure and privacy-friendly auctions. Every bid was secret, except the largest: that one could be determined.

5 Abstract



Limit as much as possible the detail in which personal data is processed.

While ‘minimise’ forces one to decide whether or not to process a particular piece of personal data, ‘abstract’ addresses the more subtle question of the level of detail in which to process personal data. The less detailed a personal data item is, the more we ‘zoom out’, the lower the privacy risk is.

5.1 Tactics

Processing personal data in less detail can be done both at the data subject level as well as the attribute level. The following three tactics apply.

Summarise Summarise detailed attributes into more coarse-grained, general attributes. For example, use an age category instead of a birth date, or a city of residence instead of a full address, whenever possible.

Group Aggregate information about a group of people instead of processing personal information for each person in the group separately. Compile group profiles with average information concerning the members of the group.

Perturb Do not process the exact value of a personal data item. Instead use an approximation of that value, or adjust the value with some random noise. For example, instead of reporting the exact current location of a person, report his location within some random distance from the real location.

Detailed information often becomes less relevant over time. Detailed logs are necessary to act swiftly in case of a hack or a disruption, but after some time only an aggregate overview of the total number of users of a service (perhaps with some information about which specific parts of the service were used) is still relevant. Be aware of this, and clean logs accordingly.

Note that even group profiles pose a privacy risk when individuals can easily be classified to belong to a particular group (such as people with a certain medical condition, or with a certain financial risk profile).

5.2 Examples

In many cases (like special discounts for seniors or juniors, or rules that require one to check whether someone is an adult) only age matters, and not the particular date of birth. In these cases it suffices to record the attribute “over eighteen” or “senior citizen”.

Smart meters are an example of a system that abstracts both over *space* as well as over *time*. For the stability of the electrical grid detailed information on energy use of a single household is not relevant. It suffices to monitor, in real time, energy use of a whole street or block of houses. To determine the electricity bill it is not necessary to record the energy use in real time. Instead it is sufficient to send the accumulated energy use of say three months at a time to the energy supplier.

Homomorphic encryption allows one to perform calculations on encrypted data without learning the real (unencrypted) data themselves. This allows one party to compute the sum of a set of encrypted measurements. Another party can then decrypt this sum. At the same time no one learns anything about the individual measurements.

Location-based services need the location of the user in order to show relevant information (like a nearby restaurant, for example). But depending on the service this location does not have to be exact. Precise GPS coordinates are usually not necessary. Sometimes a coarse location rounded to a square kilometer suffices, and sending all relevant data for that larger area to the device. The more detailed information is filtered out locally.

This type of clustering or cloaking implements the k -anonymity principle. This principle demands that the data is perturbed or cloaked in such a way that it could apply to at least k different people, that are also represented in the data set, or are using the service. In other words: apart from you, the data could belong to at least $k - 1$ other people. In location based services, k depends on the size of the area, and the average number of people there.

6 Hide



Protect personal data, or make it unlinkable or unobservable. Make sure it does not become public or known.

This important strategy targets the confidentiality, unlinkability and unobservability of personal data. This is in contrast to the minimise strategy that is aimed at explicitly deciding whether or not to process certain personal data at all. The adequate protection of personal data is a legal requirement.

6.1 Tactics

The hide strategy therefore contains the following tactics.

Restrict Restrict access to personal data. Ensure personal data is properly protected. Setup a strict access control policy. Only allow access to those who really need it (the ‘need to know’ principle). Make it difficult to accidentally share or leak personal data.

Obfuscate Prevent understandability of personal data to those without the ability to decipher it. Encrypt data so that it becomes unintelligible without the key. Hash personal data, e.g. to create a pseudonym.

Dissociate Break the link and remove the correlation between events, persons, and data. Remove directly identifying data.

Mix Mix personal data to hide the source or their interrelationships. Anonymise data. Hide data in a ‘cloud’ of other data. Break the correlation between two events, for example by not responding immediately. Collect a few events first, and then process them in bulk.

Hiding personal data can be achieved by protecting it (you know it is there but you cannot access it), making it unlinkable (you know the data, but not to which person it belongs), or making it unobservable (you are not even aware of the existence of the data). This latter aspect is only relevant to behavioural data (metadata), like location data, or information about who is communicating with whom.

Typically a combination of the above tactics is used to hide a particular data item.

6.2 Examples

Hashing and encryption are standard cryptographic techniques that one can use to protect personal data. Use them both for data transmitted over networks ('data in transit') as for data stored somewhere ('data at rest'), and also pay proper attention to key management.

Some communication services and cloud storage services use end-to-end encryption. In this case users agree, in a secure manner, on the keys to use for encryption in such a way that the service provider does not learn the key. This ensures that the service provider cannot decrypt and read the data it stores or forwards. The data is only available at the 'endpoints' (i.e. the smartphone or laptop) of the users themselves.

Attribute-based credentials (ABCs) allow a privacy-friendly form of identity management. Attributes are personal qualities, such as name, age, weight, income, etc. Using ABCs you can prove you possess such an attribute, for example that you are over eighteen, without revealing any other piece of information about yourself. Moreover, ABCs are unlinkable: reuse of a credential cannot be detected. If you prove a hundred times to the same service provider that you are over eighteen, then as far as the service provider is concerned, a hundred different people over eighteen used its service.

Tor, the onion router, makes web browsing anonymous. Your browser no longer connects to the web server directly. Instead the connection is established through three different Tor nodes. These intermediate connections are all encrypted. This way the web server, your internet service provider, any intermediate nodes, even the Tor nodes you use, cannot tell which websites you are visiting.

Note that in practice completely anonymous data sets do not exist: often one can use the retained data elements to deduce the identity of the person they pertain to. Therefore do not rely on anonymisation entirely.

7 Inform



Inform data subjects about the processing of their personal data in a timely and adequate manner.

Transparency about which personal data is being processed, how they are processed and for which purpose, is an essential prerequisite for better privacy protection. It allows users to take informed decisions about using a system and agreeing to the processing of their personal data (see also the control strategy). Moreover it allows society at large to verify whether organisations are processing our personal data responsibly. (“Sunlight is said to be the best of disinfectants.”)

7.1 Tactics

Transparency can be achieved following these tactics.

Supply Supply information about *which* personal data is processed, *how* they are processed, and *why*. Clearly specify how long personal data is retained, and how it is deleted. List all third parties with which you share this personal data, be clear about the conditions that cover each third party data exchange, and specify how these conditions are enforced. Put a link to your privacy policy on your homepage, and in your app. Clearly indicate how people can get in touch with questions about their privacy.

Explain Explain which personal data you process, and why. Argue why this is necessary. Do this in a clear and easy to understand manner, even for a layperson. Target this information to different user groups: novices, experts, the authorities. Consider using a layered approach: first provide an overview, and provide links to more detailed information.

Notify Notify users (in real time) the moment you process their personal data, share it with third parties, or as soon as you become aware of a data leak. Prepare clear procedures for this. Make notifications short but informative. Be sure not to notify too often. Allow users to control for which events they wish to receive a notification.

Informing users about the processing of their personal data (through a privacy statement) presupposes that there is a privacy policy in place (see the enforce strategy) on which the processing is based. Moreover, this information must be complete and up-to-date. This is harder than it first appears.

Several years ago we, together with a few students, performed an experiment to test data subject access rights. The results were astounding. We received literally screenshots from data-bases. One student even got a phone call from the help desk of his mobile phone company. The poor help desk person asked him whether he really wanted to pursue his data subject access request, as it would take hours if not days to complete. This is what happens when you do not develop your system with such access requests in mind.

7.2 Examples

The set of icons created by the **Creative Commons** can be used to summarise the copyrights associated with a particular document. Similarly, **privacy pictograms** can communicate the essence of a privacy policy at a glance. For example these icons could indicate which type of personal data is processed, where they are processed, and whether they are shared with third parties (and if so, with whom).

A personal privacy dashboard clearly shows to the users which data are collected, how they are processed, for which purpose, and with whom they have been shared. Companies like **Google** have implemented such privacy dashboards. Make sure the access to this dashboard is tightly secured!

Apple's iOS shows a notification icon in the status bar whenever an application accesses the location services. This is an example of an 'ambient notification': the user is informed in a subtle, non-invasive, way about the use of his personal data.

8 Control



Provide data subjects adequate control over the processing of their personal data.

Control is a fundamental principle to protect the privacy of users. The main goal of privacy is not to totally prevent the processing and sharing of personal data. Not at all! But users want to have control and have a say in how their personal data is processed and shared.

8.1 Tactics

Users get control over the processing of their personal data through one of the following tactics.

Consent Ask users explicit consent to the processing of their personal data. Inform them beforehand exactly about which personal data will be processed, how they will be processed and for which purpose ('informed consent', see also the inform strategy). It should be possible to withdraw consent.

Choose Offer users a real choice: basic functionality should be accessible for people who do not consent to the processing of their personal data. Offer a (paid) alternative.

Update Offer users the means to review and update the personal data collected about them. A logical approach is to combine this with the approach that allows users to view the personal data collected (e.g. through a dashboard).

Retract Offer users a means to retract (or to ask for the deletion of) their personal information. Again, this can be part of a privacydashboard.

Consent is not always required when processing personal data, for example when you have a legitimate interest to process it. Always consult a lawyer to make sure.

It is not always possible or even required to allow users to correct their data or to delete their data when they ask you to. Sometimes personal data are simply required to be retained. In medical records it is undesirable to allow patients to edit entries with medical significance.

8.2 Examples

In many cases processing personal data is simply allowed, for example because they are strictly necessary to execute a contract (e.g. the postal address is required when ordering something on-line) or because there is a legal requirement to do so (e.g. banks verifying your identity). In other cases consent is required. Inform users clearly about the purposes. And offer them a real choice (so that they do get access even when they do not agree, perhaps to a limited part of the service). Use *opt-in* (no processing without prior consent) instead of *opt-out* (processing takes places, unless consent is withdrawn at a later stage): the default choice does *not* constitute consent. So don't use pre-checked checkboxes to subscribe people to newsletters, for example...

Websites have to ask permissions to place cookies. Many of the 'accept cookies' notifications do not comply: they do not offer a real choice as the website cannot be visited unless the cookie is accepted. A proper cookie statement offers a real choice to accept cookies, and offers the option to choose which cookies to accept or not (e.g. those needed for anonymous website statistics, or those used to connect with social networks).

A radically different approach shifts the control of the personal data completely to the users themselves. Instead of organisations storing the personal data for all their customers, they ask their customers to each store that data themselves. The organisations later access that information through a standard interface, under the control of the user. This is sometimes called 'customer managed relations' as opposed to 'customer relations management'.

9 Enforce



Commit to processing personal data in a privacy-friendly way, and adequately enforce this.

Privacy should not only be guaranteed through technical means, but also through organisational means. It should be part of the organisational culture and be propagated by higher management. Otherwise nobody will feel responsible. A clear privacy policy will provide scope and guidance. The enforce strategy is internally oriented, towards the organisation itself. The strategy ensures that the externally communicated privacy statement (see the inform strategy) is also enforced internally through a privacy policy.

9.1 Tactics

Privacy-friendly processing is enforced by applying the following tactics.

Create The organisation should commit to privacy. Take responsibility. Create a privacy policy. Assign resources to execute this policy. Determine for each process the goal and the (legal) ground: is there a legitimate interest, or is consent required? Be clear about the business model.

Maintain Maintain the policy with all the necessary technical and organisational controls. Implement these controls. Assign responsibilities. Create an awareness campaign and train all personnel. Make sure third parties (processors) also comply with the policy.

Uphold Circumstances change. Verify the privacy policy regularly, and adjust its implementation whenever necessary. Establish prior criteria, and evaluate your policy against them.

The privacy policy should be aligned with the overall business plan and mission statement of the organisation. Make sure it is consistent with all other organisational policies.

9.2 Examples

One approach is to implement a privacy management system similar to the plan-do-check-act cycle from the **information security management standard (ISO 27001)**. This could be integrated with the data protection impact assessments that have to be performed anyway (and that are further discussed under the demonstrate strategy).

Another technical approach is the use of so called 'sticky policies' that are attached a data item to indicate the original purpose for which the data item was collected, or to specify the processing that is allowed to be performed on the data item. Processes in the organisation are set up in such a way that for each data item the sticky policy associated with it is automatically verified.

One could also consider the use of techniques that register 'unusual' activity and that block access to data items when necessary, for example when one person tries to access an unusual amount of data, or if significantly many people access a particular data item (for example the medical records of a VIP).

Consideration should also be given to the system development processes within your organisation. Adhere to the privacy by design philosophy, and address privacy from the start. This little book helps you with that! ;-) If you do not develop your own systems, privacy by design thinking, and in particular the use of the privacy design strategies, also helps setting the requirements for products or services to be procured.

10 Demonstrate



Demonstrate you are processing personal data in a privacy-friendly way.

This strategy addresses the new requirement that organisations need to *demonstrate* compliance to privacy regulations. The demonstrate strategy is externally oriented, towards the data protection authorities (possibly through the internal data protection officer).

10.1 Tactics

The following tactics help organisations to demonstrate compliance.

Record Document all (important) steps taken. Record decisions, and motivate them. Collect system logs (and respond to anomalies)¹.

Audit Audit the logs regularly, but audit also the organisational processes in general, and the way personal data is processed within the organisation.

Report Report the results of such audits to the Data Protection Authority (DPA), or keep them for later reference. Consult, when possible, the DPA regularly.

Document extensively yet efficiently the ways in which the organisation protects personal information. Do so in a surveyable and clarifying manner. Make sure that the documentation corresponds with reality. Get certified.

10.2 Examples

Performing a data protection impact assessment (DPIA), and especially recording the findings and properly documenting the decisions made based on it, is a good starting point. A DPIA has to be performed when personal will be processed. Sometimes a limited DPIA suffices. When this indicates that significant privacy risks may arise, a full-fledged DPIA

¹This tactic was called 'log' in earlier work.

must be performed. A DPIA must be repeated every once in a while as circumstances may have changed.

Another approach is to get certified against an (internationally) recognised standard for privacy friendliness (like TRUSTe or EuroPriSe). Alternatively, a benchmark with other organisations that operate within the same (business)sector could provide additional support for the privacy maturity of your organisation.

Using formal methods and structured development environments when developing new systems is an advantage.

11 Applying the privacy design strategies

Traditionally, system development and deployment is a cyclic process. This system life cycle proceeds through several phases: ideation¹, definition, design, development², deployment, operation, evaluation and de-commissioning (see figure 11.1). The privacy design strategies were developed because existing tools (design patterns and privacy-enhancing technologies) apply mostly to the design and development phase. While during the first two phases (ideation and definition) important decisions are made that have a significant impact on the overall privacy properties of the system under development.

The fact that the privacy design strategies have been developed in the context of the classical waterfall development methodology does not mean that they cannot be applied in other, more modern, approaches like agile software development. Also in these approaches concepts are formulated and defined. It's just that these phases are embedded differently in these modern approaches.

The privacy design strategies impose certain (technical) goals that, when reached, improve the privacy of the overall system. One could also view them as questions one can ask during the software development process. How can I separate the processing of personal data? How can I properly inform my users about the processing of their personal data?

This can be done in a structured manner. Make sure all stakeholders involved in the project are represented, including the process owner and the technical expert. Also ensure that the end users (whose personal information will be processed) are represented as well. This guarantees that the design process (and hence the analysis of the risks) not only takes the perspective of the data controller into account, but also that of the data subjects.

The privacy design strategies are not only relevant when you develop systems yourself. They can also be used in procurement when determining the specifications for a product to be bought elsewhere.

¹Also known as the concept formulation phase.

²Includes testing and evaluation.

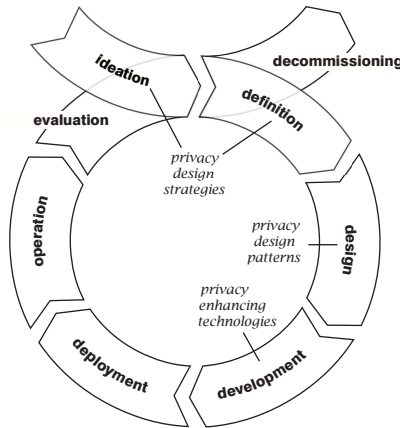


Figure 11.1: System life cycle

Don't focus on just one strategy. It's not an either-or decision. They are all potentially useful. So apply them all, one by one, to make your system as privacy-friendly as possible. Depending on the context, though, you may find certain strategies more applicable than others.

Consider all aspects related to the processing of personal data. That is to say, consider operating on, storing, retaining, collecting, sharing, changing, breaching and deleting personal data. Investigate for each of the strategies (and associated tactics) whether they apply to each of these aspects.

Don't limit your attention to ordinary personal data, but also consider metadata (that may inadvertently be collected).

Finally, consider applying the design strategies to components of the overall system (instead of the system as a whole). In fact one can apply the privacy design strategies iteratively: an initial design based on a first analysis using the privacy design strategies can be refined by applying the strategies again at a lower level of detail.

12 Concluding remarks

The latest version of this book can be found at <http://www.deprivacycoach.nl>.

The icons for each of the strategies are available in [pdf](#) and [SVG](#) format.

Want to stay informed? Follow [@deprivacycoach](#) on Twitter, or subscribe to the email notifications at <http://www.deprivacycoach.nl>.

Please contact me by email at info@deprivacycoach.nl for questions or comments. I love to hear from you, especially if you used the privacy design strategies, and if you wish to share your experiences!

And why is this called ‘The Little Blue Book’? Because blue is the colour [associated with trust](#).

May 2018,
Jaap-Henk Hoepman.

12.1 Sources

- M. Colesky, J.-H. Hoepman, and C. Hillen. A Critical Analysis of Privacy Design Strategies. In 2016 International Workshop on Privacy Engineering (IWPE’16), pages 33-40, San Jose, CA, USA, May 26 2016.
- The [General Data Protection Regulation](#) (GDPR).
- G. Danezis et. al., [Privacy and Data Protection by Design](#), ENISA Report, december 2014.
- The [Privacy design patterns](https://privacypatterns.org) database (<https://privacypatterns.org>).

12.2 Institutes

- The [European Data Protection Board \(EDPB\)](#) (Formerly the [Article 29 Working Party](#)).
- The [European Data Protection Supervisor \(EDPS\)](#).
- The [Federal Trade Commission \(FTC\)](#).
- The [Electronic Frontier Foundation \(EFF\)](#), [Privacy International](#) or

European Digital Rights (EDRI).

- The [Privacy & Identity Lab](#) (PI.lab).

12.3 Learn more

- D. Solove, “Understanding Privacy”, Harvard University Press, 2010.
- P. E. Agre and M. Rotenberg (eds.), “Technology and privacy: The new landscape”, MIT Press, Cambridge, MA, 2001.
- B. Schneier, “Data and Goliath”, W. W. Norton & Company, 2016.
- The [Privacy Wiki](#).

12.4 Commercial use

This book is released under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

In practice this means I expect some form of remuneration for certain forms of commercial use. The idea is that I receive a fair compensation for the time and effort I put into writing this booklet, proportional to the *direct* revenue you generate using this work.

Not all use of this work in a commercial context is restricted. For example, if you use this work to implement an internal privacy by design approach, or use this to train your own employees to increase the level of privacy protection, then this is totally fine. My reward in this case is that your products and services become more privacy friendly.

However, if you use this material as a significant part of a privacy by design training that you offer as a commercial service to others, or if you use this as a significant part of a privacy by design methodology that you apply commercially for external clients, then the use of my work creates *direct* revenue for you. This applies in particular when you distribute (a link to) this book as part of the training or support material. The amount of remuneration I expect depends on the significance of my work for you, and the amount of revenue it generates for you. But as I cannot determine this at all, I simply depend on your honesty to fairly assess what my work is worth to you.

For further information, or payment options, please contact me at info@deprivacycoach.nl.

13 Glossary

Data controller Entity that determines the means and purposes of the processing of personal data. In general this is the entity that offers a product or service to the end user.

Data subject Natural person whose personal data are processed. Often a user of a service or a product.

Design pattern Describes a common recurring structure of interrelated components, that solves a generic problem in a particular context.

Data protection authority (DPA) The national authority responsible for enforcing data protection laws.

Personal data A piece of data that can either directly or indirectly be linked to a (natural) person. A name or social security number is personal data. But so is a license plate, or an IP address. In other words: many data are considered personal.

Privacy by design Design philosophy that demands that privacy requirements are addressed from the start when designing and developing a new system.

Process The act of collecting, storing, retaining, using, processing, sharing, changing or deleting data.

(Data) processor Entity that processes personal data under the authority of another entity (typically the data controller).

How to make privacy by design concrete? How to apply it in practice? Those questions are answered in this book.

This book was written by Jaap-Henk Hoepman. He is a privacy expert, associate professor in the department of computer science of the Radboud University, Nijmegen and the faculty of law of the University of Groningen. He is principal scientist of the Privacy & Identity Lab.