LEARNING ANALYTICS IN 5 STEPS

A GUIDE TO THE GDPR

CONTENTS

INTRODUCTION

Learning analytics is the "measuring, collecting, analysing and reporting of data from students and their environment to understand and optimise learning and the learning environment" (Siemens, 2011). The analyses of the data can provide various target groups (students, teachers, institutions) with valuable insights into various aspects of the education process. Collecting and analysing data often means processing personal data - data which, directly or indirectly, contains information about individuals such as students and teachers. When an educational institution processes personal data, the General Data Protection Regulation (GDPR) applies.

The GDPR places strict criteria on the basis for using learning analytics. The use of learning analytics requires careful consideration of the risks for individuals, and suitable measures to be taken. It is not as simple as asking for consent once and including a privacy statement on the website.

Within this this roadmap SURF aims to assist educational institutions to shape their learning analytics in accordance with the GDPR. This roadmap is only intended as a guide, each educational institution will need to make its own considerations in regard to privacy. It supports institutions in safeguarding the legal aspects of their learning analytics projects. This contributes to creating a secure learning environment which respects the privacy of individuals.

Important

This plan offers institutions a framework for introducing learning analytics at their institution. However, it does not constitute exhaustive advice on how to comply with legislation and regulations as an institution.

Getting started

Does your institution want to start with learning analytics and launch a project to this end? Please follow the plans' different steps with your learning analytics team to immediately survey the impact of learning analytics on privacy. Don't skip any steps. It is best to focus on the steps that are most relevant to your project. With each step, you can read about the relevant legal terms and what you have to do. Also make sure that your institution's Data Protection Officer (DPO) is involved in carrying out this plan.

Tip

For more context around the legal terms, you can consult:

- **the GDPR guide from the Government (in Dutch)**
- **the SURFwiki explaining the GDPR and how to interpret it (in Dutch)**
- **SURF's 'Privacy in research' online module (in Dutch)**

STEP 1: PREPARATION

When preparing the learning analytics process, it is important to start by putting together a project team to jointly define the purposes of the process. It is also important to identify all of the project's stakeholders. This step describes the preparatory actions you need to take.

1.1 Define the specific goals you wish to achieve with learning analytics

It is important to define the specific goals of the project as early on as possible. Use the first step to consider and describe what your institution actually wants to achieve with learning analytics. Start by defining the aimed result of the project followed by a description of how learning analytics can help you to achieve it. You will also use the goals you formulate in this phase during the design phase (step 2) and in assessing legitimacy using a Data Protection Impact Assessment (DPIA) in step 3.

Describing your specific objectives is essential, as it makes the processing of personal data fair, transparent, and understandable. Important principles are:

- **Purpose limitation:** you may only use data for specific, clearly described and legitimate purposes. Learning analytics is a way to achieve your goals. Under specific circumstances, collected data may also be used at a later date for another purpose they were originally obtained. However, this must be done compatible with the purpose for which the data was collected.
- **Data minimisation:** data must be adequate, relevant, and restricted only to what is necessary for the purposes for which it is processed.
- **Storage restriction:** set deadlines for deleting personal data based on the purposes for which they are required.
- **Responsibility:** you must record the use of personal data and explain the use of this data to individuals in a transparent, understandable, and easily-accessible format. For more information see 'Inform', paragraph 4.3.

Examples

Define and describe the purposes as specifically as possible. General descriptions such as 'improve the learning process' or 'storing answers when creating a pilot test' are not sufficient. The context and expectations of individuals determine the amount of detail required. You should therefore elaborate further on the overall purpose below, and how the data are used for that purpose:

- **Purpose:** 'To put more focus on topics where the most mistakes are made during lessons'. The full description of the purpose, including how learning analytics will be carried out, would then be, for example: 'The education coordinator of the 'basis of privacy law' module analyses the answers in the digital learning environment at a group level throughout the academic year in order to extrapolate the subjects in which most mistakes are made, so that more attention can be given to those subjects during lessons.'
- **Purpose:** 'To be able to offer excellent students honours programmes that are relevant to them'. The full purpose description, including how learning analytics is carried out, would in this case be: 'The honours programme coordinators want to recognise excellent students based on the completion speed and results of practice tests in the online learning environment on relevant subjects at the end of an academic year, so these students can be offered relevant programmes.'

A good, complete description of a purpose is structured as follows: 'who' wants to do 'what' using 'which' personal data and 'when', to achieve 'result'. Keep the SMART principle in mind: Specific, Measurable, Acceptable, Realistic, and Time-bound. Well-described purposes will make it easier to complete the step-by-step plan.

1.2 Set up a multidisciplinary project team

The next step in carrying out a learning analytics process is to set up a project team. Consider who in your organisation(s) has an interest in joining the project. Has a similar project team been set up in the past? If so, find out who was part of that group.

Involve people with the right expertise in your project team. For example with the following background: client, project leader, designer, functional manager, lawyer, and teachers.

In addition to the project team, you should also consider which stakeholders must be involved in the project now and in the future. A particularly important role is that of the Data Protection Officer (DPO). You must obtain advice from the DPO when assessing the processing and carrying out a Data Protection Impact Assessment (DPIA). This will be further discussed in step 3.

Define how you will create support within the institution and at an individual level. How can you convince students that learning analytics is important? Discuss the project with them. This is also an important part of carrying out the DPIA.

1.3 Describe stakeholders

Describe which specific groups of stakeholders will play a role in deploying learning analytics, define their positions, and what their expectations or wishes are (stakeholder analysis). This includes groups within the institution, those who will be working with and managing the learning analytics system, and also groups helping to maintain the system, subcontractors, and groups from whom data will be collected. Examples include students, teachers, managers, project leaders, and management. The institution may already have procedures in place for submitting a project to the participation council or student council, or to have it tested by a testing commission or Ethics Board. If there are experience experts in or outside the institution who have already applied learning analytics, involve them.

A stakeholder analysis will help you to learn from the experience of others, pick up on focus areas and risks you had not previously considered, and create support among stakeholders.

STEP 2: DESIGN

Once the framework of the learning analytics process has been defined and described in the initial version of your action plan, work out how you will implement it. The outcome of this step is a description of the systems required to facilitate learning analytics. This design makes it clear both technically and legally how the institution will carry out learning analytics.

The following design steps already took the relevant legal principles into account. However, in this phase, you don't consider yet whether the entire learning analytics system is legitimate and which basis of assessment can be applied. Before making this consideration, the design must be clear. Therefore, you will not assess the legitimacy until later on. This will be further discussed in step 3.

2.1 Include privacy protection in your design

Learning analytics will only work if you manage to process the data you have collected in the right way. When taking the subsequent steps, consider which data you need to collect to achieve your goals and include this outcome in the design plan.

Define which data you need as a minimum to achieve your purposes
Could you also achieve your objective if you collect data from fewer individuals? Or by collecting less data from each individual? When designing the learning analytics system, you are legally obliged to include privacy protection in your design from the start.

The concept of privacy by design means that from a product or service's design phase onwards, privacy-sensitive features are taken into account and sufficient privacy safeguards are integrated to protect and secure personal data.

The less data you collect from individuals, the easier it is to protect this data and the less of a risk there is of a data breach. You can also protect data by splitting it with the aim of separating data and processing it separately, so that data is only combined with a clear goal. You can also sometimes protect personal data by using it in an abstract way: aggregated instead of detailed. Statistics are the best known way of showing research data without disclosing personal data.

Tip

Read more about privacy by design and the applicable strategies here:
- **Privacy Designer (in Dutch)**
- **the Blauwe Boekje (pdf) privacy design strategies (in Dutch)**

2.2 Define the scope of your processing

It is important to already know in advance the scope of your project from a privacy perspective. It is important to determine in advance:
1. what sort of personal data you are using
2. the sensitivity of this data
3. the amount of data you are processing, and
4. whether there may be other relevant features that influence the scope of your processing, such as a geographical area.

1. Define which parts of your dataset contains personal data

The GDPR specifies that personal data is all information about an identified or identifiable natural person. This refers to information that either relates to somebody directly or can be traced back to him. Personal data therefore includes names and addresses, for instance, but also exam results or information about behaviour. By further combining this data (possibly with other data), it is then possible to trace the data back to individuals. This concerns indirectly identifiable data.

Describe what personal data will be collected as part of this project. For more information see the GDPR guide from the Government (in Dutch) and the SURFwiki explaining the GDPR and how to interpret it (in Dutch).

2. Determine whether you are processing sensitive personal data

Sensitive personal data is data about a person relating to their health, ethnicity, sexuality, political preferences or religion. Processing this data is prohibited unless the institution has the individual's consent, or if its use is regulated by law. Consent means asking separately for the relevant data and explaining why, while giving consent is optional.

Obtaining explicit consent is the most obvious exception for processing sensitive data using learning analytics. For the basic elements, see 'Consent' in paragraph 3.1. If asking for consent is impossible, there is a potential exception for scientific research (see article 24 GDPR Implementing Law).

Tip

For more information about the difference between explicit consent to process sensitive personal data and consent as a basis of assessment, see the directives on consent from the European Data Protection Board (pdf in Dutch).

Please note

The citizen service number (BSN) - the same number the government uses as an education number - is also a special kind of personal data. Processing the BSN is only permitted if ratified by law, which must also state the purpose for doing so. For example, when copying a proof of identity: when doing so, the BSN must be made illegible.

As an institution you may also process sensitive personal data unintentionally, for example, if the institution carries out a detailed analysis of study behaviours in online learning environments and the findings are specifically linked to medical (sleep patterns and lifestyles) or religious (behaviour on holy days) data. In this step, you should therefore try to make a note of any consequences learning analytics analysis may have.

3. Define the amount of data you are processing

Once you have determined what type of (sensitive) personal data you need at least to achieve your purpose, it is also important to define how many different types of data you will process within the learning analytics process. In doing so, it is important to at least answer the following questions:

- How many types of data do we plan to process?
- How often will we process this data?
- How long will the processing of data take (from collection to destruction/anonymising)?
- From how many data subjects (those who are the subject of the learning analytics process) will data be processed?

4. Define whether other elements are relevant to the scope of the process.

In some cases, there are additional factors that influence the definition of the scope of processing. One example could be the geographical area within which data is processed. It may be the case that if the personal data of individuals is processed from other (European) countries, some rules may be different. If you reach the conclusion that this is indeed the case, we recommend you consult the DPO to determine whether specific measures are required.

2.3 Describe how data will be processed

1. Define which actions you will take: processing personal data

Describe the different actions you will perform with the personal data. Processing means: all actions that an educational institution can perform with personal data, from collecting and storing it up to destroying the data. You must therefore describe each phase of processing for learning analytics like collecting, subsequently analysing, and then using the results. Tracking what someone does in a learning environment is for example a type of personal data processing. Using synonyms and making data anonymous are good measures you can take to protect individuals' privacy, but these actions are also considered processing.

2. Describe the data sources you will use

Describe the data sources (input) you will use, and the origin of these sources. Define what data you require that the institution already tracks, and where data can be found.

Educational institutions have a range of systems which may contain relevant data. These include learning management systems, student information systems, test systems, attendance monitoring systems, and even mobile applications with sensors such as GPS tracking. The power of learning analytics is that you can combine those data by using technical standards.

Important

Do not immediately use all of the existing data for your learning analytics project. Instead, first work through the roadmap and always use the principle of data minimisation.

3. Describe the systems used

Outline the systems and services in and outside the institution that are relevant to carry out the process. How will you set up the learning analytics system? Will you be using a separate website, or perhaps an app? Or will an existing system be used to apply learning analytics? Does the system include new technologies whose risks are not yet determined?

Keep in mind that when collecting data through cookies (and comparable technologies such as scripts), conditions apply under the Telecommunications Act. To this end, see the rules under 'Consent' in paragraph 3.1.

4. Describe the service providers involved

Educational institutions often turn to service providers such as suppliers for learning analytics. Define who will have access to the data and with whom the institution will share its data. It is possible that service providers will be able to access the data through learning analytics software. This could happen for example with cloud services for students or external systems which measure performance during tests and generate reports on this.

If you are using third-party software or services, there are two things you must pay attention to:

1. From a legal perspective, the institution is always liable for the quality of, and problems with service provision. This will also be the case when the software supplier does not wish to accept any liability. The institution cannot avoid this liability by, for instance, including a liability limitation in the learning analytic system's acceptance statement or a disclaimer on the software start screen.
2. If the external service provider also receives personal data, like with cloud services, the institution must make separate agreements on what the service provider may do with this data. Here, it is important to determine the service provider's role. To define which privacy role a service provider has, read the explanation in paragraph 2.5.

Selecting service providers

When selecting service providers, include GDPR requirements such as the security and information obligations and measures when processing data outside the European Economic Area in the selection process. If you entrust processing activities to a service provider, you need to rely on service providers that offer sufficient guarantees, particularly with regard to expertise, reliability, and resources. This allows you to ensure that the technical and organisational measures comply with privacy regulations, including with regard to secure processing.

2.4 Describe the (security) measures

Institutions using personal data must protect this data in accordance with the GDPR. This reduces the risk of a data breach and other potential breaches to the protection of personal data. Under the GDPR, you must take suitable technical measures to this end, such as using modern technology to secure personal data. You must not just look at the technology, but also at how you approach personal data as an organisation. For example, who can access which data?

Take technical and organisational measures to guarantee data confidentiality, integrity, and availability

If you store personal data, you must protect it adequately. This means that you, as an institution, must reasonably protect all of the personal data you obtain against unauthorised or illegal access or use. This applies not only to data which you have asked for, but also personal data you received inadvertently. We therefore recommend you to create a security and data breach policy.

There is still no generally applicable norm or standard that specifically focuses on personal data protection. In some branches, specific norms such as the NEN 7510 apply. SURF has created a norm and test framework for information security based on the NEN 27001.

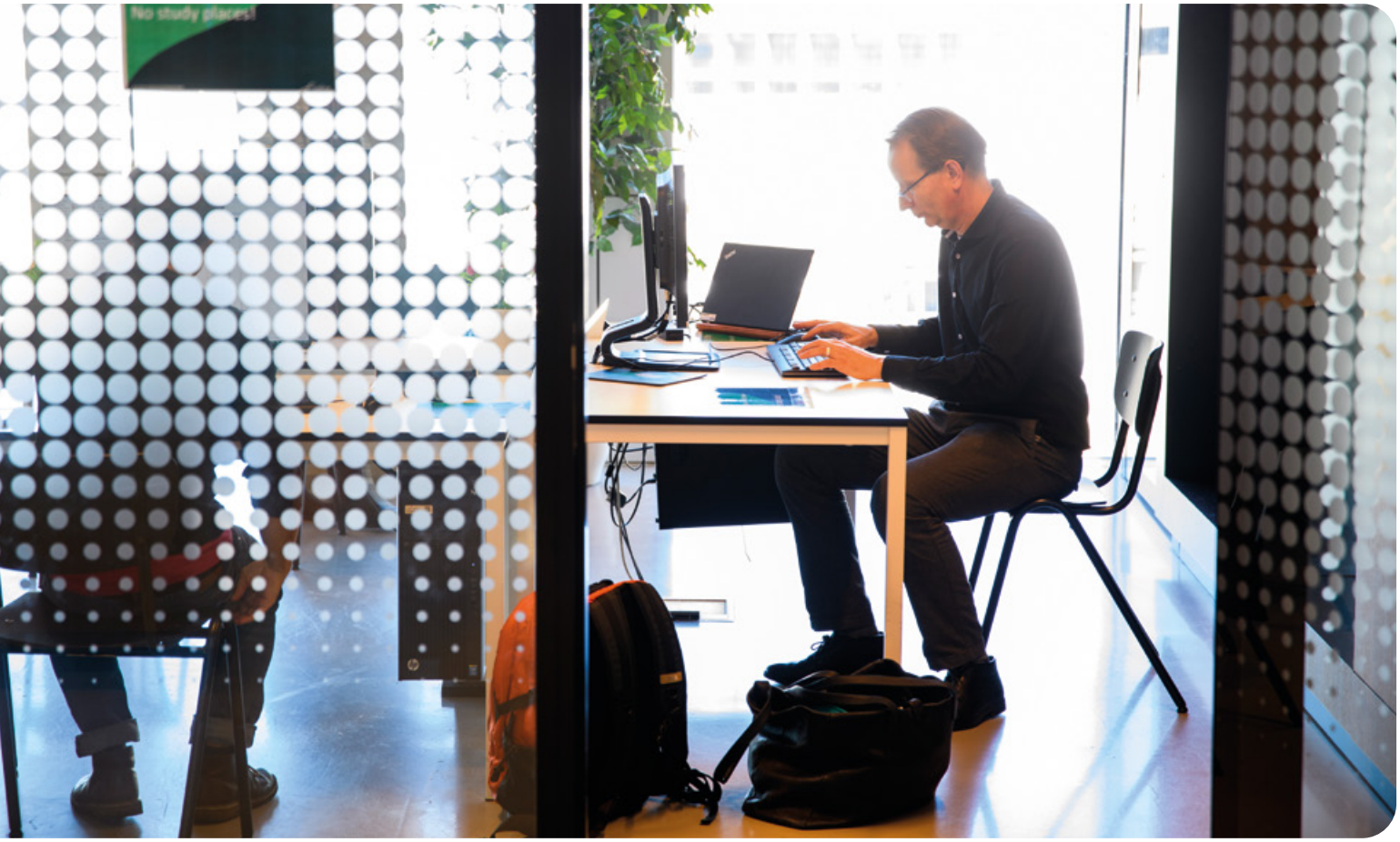
The data protection authority has published guidelines on how institutions can meet security requirements. These guidelines state that security must be an integral part of the development and improvement of services, and that you must regularly check whether security remains adequate (PlanDoCheckAct).

In addition, data must be kept accurate and available. Therefore, take measures that prevent data adaptation and reduce and remove threats that can make data no longer available. For example, take measures to fend off ransomware and DDoS attacks.

Tip

Consult additional resources on security:

- **the website of the Personal Data Authority**
- **The useful reports from the French privacy regulator and ENISA/Teletrust**



Use Aggregates and pseudonyms

In some cases, measures are taken to better protect or secure data. Hashing - a practice which processes data cryptically - is an example of this. The supervisory authority believes that even these practices constitute processing personal data. Just hashing names or identifying different data, without additional measures, does not mean that personal data is not being processed. In these cases it is therefore still possible to trace back the data to individuals. This is also known as pseudonimisation. This data is still covered by the GDPR.

Data is only anonymised if there is no longer any reasonable way to trace the data back to a person. For instance, by assigning random unique numbers to data after which the list that links the names to the numbers is destroyed. Even then, however, a person can frequently still be identified by using other data, even if there is no name or traceable unique number linked to it. A collection of a student's test results and modules is unique, for example; no two students in the same year get exactly the same grades for the same curriculum. This collection therefore constitutes of a set of personal data, even if the student's name or unique number is not included.

Data may lose its status of personal data if it is sufficiently aggregated, i.e. when it is merged with statements regarding several people. 'Eighty percent of students failed this exam' is not personal data, for example. There are no legal requirements around statistics such as these.

Important

Sufficiently aggregated data gives you more options, but the source data continues to be personal data governed by the GDPR. Only processing actions that start with aggregated data will give you more options. For example, if a teacher measures how quickly individual students work through the course material at an online learning platform, a basis is required. This continues to be the case if the teacher only makes statistical statements about their findings. For more information, please see the guidelines on research and anonymising data published by the predecessor of the European Data Protection Board (EDPB).

Define the organisation's deletion period for data sets

Personal data may not be retained longer than necessary to achieve the purpose of the processing. If data is no longer required, it must be destroyed or deleted. You must therefore define in advance a deletion period for the different data sets. This period depends entirely on the purposes for which you are retaining the different data sets and how the system processes them.

- Estimate how long you will require the data sets to achieve your goals.
- Keep in mind, as much as you can, additional processing purposes such as substantiating investigations in reports, archiving, academic accountability, and the Personal Data Authority's directives.

Tip

Also consult the handbook of the National Coordination Point Research Data Management (LCRDM).

2.5 Define stakeholders' privacy roles

Define which privacy role the involved parties have for each instance of data being processed. The roles defined by the GDPR are:

- Controller: the party who defines the purpose and means of processing
- Joint controllers: those who jointly define the purpose and means of inextricably linked processing activities
- Processor: the party who processes the personal data on behalf of the controller, and
- Third party: other parties.

Example: *A university of applied sciences decides to analyse student data to measure which modules cause students to fail more often. To analyse this, the university uses a cloud tool into which they put all of the students' grades for a year. In this case, the university defines which data it analyses for what purpose. The cloud solution supplier only carries out analysis on behalf of the university. In this case, the university of applied sciences is a controller, responsible for the processing, while the supplier of the tool is a processor.*

Tip

For more context around the legal concepts, please consult the GDPR guide from the Government.

STEP 3: ASSESSMENT

3.1 Legality of the learning analytics process

It is important to investigate whether the learning analytics purposes are proportionate to the infringement on the individuals whose data you process (proportionality), and whether you can achieve your purpose in a way which is less disadvantageous for the individuals involved (subsidiarity). If everything has been done properly, you will already have considered these points during the project design phase (see step 2).

Determine whether the application is proportionate
Proportionality depends on whether the use of learning analytics is proportionate. You should therefore consider what your institution is seeking to do with learning analytics and define how this purpose relates to the potential impact on individuals. Although the ends do not justify the means, generally speaking, more will be permitted for a more difficult purpose. The data must be necessary to achieve the purpose. This is stricter than 'desirable' or 'useful'.

Define whether less intrusive means can achieve the same purpose
Learning analytics must be assessed not only for the proportionality, but also for subsidiarity. This means that there must not be a less invasive way of achieving the same purpose. In the context of learning analytics, this means inter alia that you must examine the overarching purpose and consider whether this can be achieved without using learning analytics. Take into consideration the need to process data, the complexity, and the consequences of using learning analytics.

Once the learning analytics design has been created (outcome of step 2), and you have defined the proportionality and subsidiarity, you can identify the relevant basis or bases for processing data. Each instance of personal data processing requires an underpinning basis or justification. The basis justifies the institution's processing of personal data for the purpose of learning analytics. In principle you require one of the following six bases for each phase (collection, analysis, use):

- consent
- execution of an agreement
- legal obligation
- protection of vital interests
- fulfilment of public service
- legitimate interest

Using data for a purpose other than that for which it was originally collected
An institution may only collect and process data based on one of the aforementioned bases. As an institution you can use this data later for scientific research, statistical purposes or other compatible purposes, if the institution has taken the measures necessary to ensure that further processing will only be carried out for these specific purposes. If you have someone's e-mail address because they've asked a question, you can respond to this person but cannot subscribe them to your newsletter, because this is a different purpose. A different purpose is permitted if it is consistent with the original one.

An example of this would be sending out a survey about the quality of the helpdesk to someone who has used the service. Because further processing (in this case learning analytics) must correspond to the specific purposes, purpose limitation is a difficult requirement. The idea is to gain new insights, to be able to ask new questions, and to look at data from new perspectives.

This is by definition not a 'specific purpose'. We therefore recommend listing as many concrete purposes as possible in the information provided and revising these regularly. When revising the purposes for which your institution has obtained consent, you must ask for consent again for these revised purposes.

Example: *A teacher uses a monitoring tool during an exam in order to measure whether people who are retaking the exam perform differently than those taking it for the first time. The basis applied here is 'legitimate interest'. The results must then be processed through statistical research into success factors for that exam. This is deemed further processing for research.*

Tip

For the framework terms for scientific research, also see the handbook of the National Coordination Point Research Data Management (LCRDM).

Is the institution applying learning analytics to data which was legally collected earlier, and is the purpose of learning analytics compatible with the purpose for which the personal data was originally collected? If so, there is no need for a separate legal basis different from the one on which the processing was originally based.

The following factors are important when determining whether the purposes are compatible:
- in particular, **the reasonable expectation of the individuals** based on their relationship with the institution regarding further use
- any links between the original purposes and the purposes of further processing
- the context in which data was collected: what is the relationship between the institution and the individual?
- the type and nature of the data: is it sensitive?
- the possible consequences of additional processing: what are the consequences for the individual?
- whether there are sufficient safeguards such as encryption, aggregation, or use of pseudonyms.

This 'compatibility' exception is appropriate if the individuals can expect that only data gathered earlier will be used for the specific learning analytics process, the purposes of which are similar to the original purposes and that the data used will only be non-sensitive data which will then not have any great impact on the individuals in the learning analytics process.

If the learning analytics depend in whole or partly on new sources, or if the purposes are not compatible with the original ones, the institution must examine whether one of the six bases can be used, as stated at the beginning of this paragraph.

Basis: Determine whether 'consent' is required for (parts of) the processing

In some instances, individuals must give consent to the institution that tracks them with learning analytics. Due to the conditions that the consent must meet, this is a challenge from a legal and technical perspective. Giving consent is not merely the formality of ticking a box or clicking on a pop-up to make it go away; consent must be a freely-given approval from individuals and give them control over (part of) the learning analytics processing.

As a rule of thumb, in all cases you can assume that individuals' consent is required if learning analytics:

- processes sensitive personal data (see paragraph 2.2);
- processes sensitive persona data such as location data. In some cases, this can present a serious risk to data protection, meaning high levels of individual control is appropriate;
- places or reads cookies or other data such as learning analytics scripts on the computers, laptops or telephones which individuals use ; and
- takes decisions based solely on an automated processing of personal data if those decisions have legal consequences or if those decisions otherwise affect the individuals to a significant extent.
- after weighing the interests and measures taken on the basis of 'legitimate interest', the individuals' interests are not sufficiently respected. This basis is described at the end of this paragraph.

The law does not precisely describe how you should ask for consent. But, how the institution asks for consent must meet the following specific requirements:

1. **Freely given:** consent must be given freely. Individuals must be able to say 'no'. This may not entail them being refused entry to a mandatory module, for example, or prevent them from doing exams.
2. **Unambiguous:** the act of giving consent must be clear and active. This means that boxes may not be pre-ticked.
3. **Informed:** as an institution, you must provide individuals with information on the most important elements (such as the purpose of each processing action for which you are asking consent, which data will be processed, and how to revoke consent) directly and in advance. It must be just as easy to revoke consent as it is to give it.
4. **Specific:** consent must always apply to a specific processing action and a specific purpose. 'I consent to learning analytics' is not specific. Make clear who will carry out monitoring, what data will be collected, and what is done with that data. An example: 'I consent to my academic performance being tracked and recorded in this module's online learning environment. It will be used to offer me personalised study advice. The tutor will be given this data to proactively be able to address risks of delays with me.'

If you ask for consent for several purposes during processing, you as an institution must inform those affected and ask for consent for each purpose separately. The purpose also may not change over time.

In short, clarify to which modules the consent relates, how far monitoring in each module extends, and what the impact is for each module. You can also use a few sentences to give a brief explanation, with a link to a clearer and more informative privacy statement with more specific information. Obtaining consent through terms of use, general terms and conditions, or a privacy statement is not permitted. However, you may refer to a privacy statement to offer further information about consent. For more information see the step 'Inform', paragraph 4.3.

If consent does not meet the requirements above, the consent is not valid. You may therefore not process personal data (for this part). In addition, during processing and for up to 5 years afterwards, you as an institution must be able to demonstrate that valid consent was obtained. This makes the basis for consent weak and difficult to apply.

[1] *This does not apply to cookies which are required for the technical functioning of the (educational) service provision or for analytical cookies which only infringe on individual privacy in a minor way (see also Article 11.7a of the Telecommunications act and the future ePrivacy Regulation and the step-by-step plan of the Personal Data Authority. (pdf in Dutch)*

Tip

For more information please see the Dutch translation of the GDPR guidelines on consent from the European Data Protection Board (EDPB).

Basis: Execution of an agreement

Personal data may be processed if doing so is necessary for the execution of an agreement. If an institution has entered into an agreement with someone, the institution may process this person's personal data to the extent that doing so is necessary to be able to execute the agreement. This must be an agreement to which the individual is an actor. Processing must be a necessary outcome of the agreement.

The mere fact that learning analytics is included in an agreement with an individual does not mean that the requirement has been met if the processing is necessary to be able to execute the agreement. When considering whether data processing is necessary for the input on the educational performance, the content and fundamental purpose of the agreement determine whether data processing is actually necessary for the (educational) performance. If this is the case, the institution may use this as a basis. However, the agreement itself may not focus solely on personal data processing, but must have a different (educational) purpose. Generally speaking, educational institutions do not have agreements with students. Some consider registering with the institution to be an agreement, but from a strictly legal standpoint this is not the case.

As an educational institution, for this basis you must demonstrate that the use of learning analytics is necessary. As learning analytics are reasonably new, it is easy for them to be deemed unnecessary for education. The general perception is that education can be provided just fine without learning analytics. It can only be presented as necessary if it is clear that education without learning analytics falls short. However, this can only be done if learning analytics proves its value over a period of several years. This basis can therefore be applied, though the probable link to specific educational purposes is less strong.

Basis: Legal obligation

The third basis listed in the GDPR is legal obligation. Data must be processed if mandated by legislation. To be able to base personal data processing on this basis, it must not be possible to meet this obligation without processing personal data.

An institution may take the position that it is required to implement the purposes of the legislation. This requires a good understanding of educational performance. Learning analytics is a way of gaining this understanding. That argumentation may justify the use of this tool. The same comments about necessity apply here as can be found under the 'Execution of an agreement' basis.

Basis: Public interest or public authority

This basis is relevant if the institution performs a public duty in the public interest or for a public authority, such as under the Higher Education and Scientific Research Act (WHW). This includes tasks that are based on the WHW Act and are relevant to the institution. Just as for the other bases, 'necessity' is the criterion you must be able to substantiate. The data cannot only be 'relevant' or 'useful'.

An educational institution, when substantiating the necessity, can take the position that it is broadly fulfilling its public duty. This requires a good understanding of educational performance. Learning analytics is a way of gaining this understanding. That reasoning may justify the use of this tool. The institution itself must be able to substantiate that learning analytics is truly necessary for the public duty and that there is no real other alternative.

Tip

For learning analytics processing which does not fall under the institution's public duty, please review the conditions of the 'legitimate interest' basis below. As a semi-public organisation, institutions may not base the performance of public duties on 'legitimate interest'.

Basis: Justifiable interest

This basis means that the processing of the personal data is necessary for the (educational) interests of the institution, an external party, or individuals and that thereby the individuals' privacy has been taken into consideration to the fullest extent.
Camera surveillance is a good example: it is difficult to ask for consent from everybody who enters a building, but the need for the specific interest of "surveillance and security" is clear. In this example, a notice and a regulation stating what happens with the recorded images are sufficient. In addition, filming is not permitted in areas where privacy weighs more heavily, such as in changing rooms or bathrooms.

The institution can use 'legitimate interest' as a basis if:

- the institution, an external party, or the individuals have a clearly defined legitimate interest or benefit;
- the learning analytics processing is necessary to defend this interest. So, as described earlier, this means that as an institution, you must examine whether the interest of the processing is proportionate to the infringement on individuals' privacy, and whether it would be possible for the institution to achieve its purposes in a different way which is less detrimental to the persons involved.
- the institution has balanced the interests and educational interests of those of the individuals, and has documented this. The institution may also have to take measures to ensure that individuals' rights and freedoms are given a greater importance than the legitimate interest. The reasonable expectation of individuals, based on their relationship with the institution, must be taken into account when doing so. Using this basis is subject to strict requirements: institutions must, for example, publish their considerations (e.g. on their websites), and in principle offer individuals the right to object (an opt-out) at any time.

Tip

Use a 'Legitimate Interest Assessment' or the explanations from the Personal Data Authority to weigh up interests for the basis of 'legitimate interest'.

3.2 Data Protection Impact Assessment (DPIA) for learning analytics

The GDPR specifies that in addition to applying a basis, in some situations, organisations must also carry out a *Data Protection Impact Assessment* (DPIA). This is a tool to outline and assess in advance the privacy risks of data processing in a structured, standardised way, so that measures can be taken later to minimise these risks. Your documentation also demonstrates that you as an institution have met the legal requirements for processing.

Even though carrying out a DPIA is not always mandatory, it is recommended for learning analytics projects. A DPIA not only helps to raise awareness for the consequences of using data, there is often also a *profiling* or a high privacy risk for individuals that make it mandatory to carry one out. SURF recommends that all institutions using learning analytics carry out a DPIA. See also the Personal Data Authority's explanation of DPIAs.

Please note

In many cases, using learning analytics will be categorised as profiling. Profiling is one of the mandatory processing actions listed by the Personal Data Authority, and it is described as follows: "*Systematically and extensively assessing personal aspects of natural persons based on automated processing (profiling). Examples include assessing work performance, students' performance, economic situation, health, personal preferences or interests, reliability or behaviour.*"

- Read more about Data Protection Impact Assessments on the website of the Personal Data Authority
- See also the Dutch translation of guidelines on profiling

3.3 How do you carry out a DPIA?

There are various methods for carrying out a DPIA. You can choose one of them as long as it meets the basic requirements described in the GDPR. The basic requirements are a systematic description of the data processing you will be carrying out, an assessment of the necessity for and proportionality of learning analytics vis-à-vis the stated purpose, an assessment of the privacy risks for individuals and measures to address these risks. It is then up to the institution to actually implement these measures. You are also required to obtain advice from the Data Protection Officer (DPO). This offers additional certainty that the DPIA sufficiently considers the risks and that sufficient measures will be taken to cover these risks.

If the DPIA reveals that learning analytics present a high level of risk and you as an institution are unable to find (sufficient) measures to mitigate these risks, then you must consult with the national Personal Data Authority (DPA) before you can start using learning analytics.

Tip

- (Data) Protection Impact Assessment templates and risk forms have been drawn up by SURF (in Dutch)
- There are also example DPIA templates from supervisory authorities which can be used or adapted:
  - Commission Nationale de l'Informatique et des Libertés (CNIL)
  - Information Commisioner's Office (ICO)

3.4 Define the impact of learning analytics (risk assessment)

Applying learning analytics must be in line with the purposes and social guidelines of the institution itself. The values your institution pursues, the relationship with students, sustainability, diversity, and so on, must all be reflected in the learning analytics system. Risk assessment is an important part of the DPIA.

When using learning analytics to achieve a certain purpose, it is done with the aim of realising benefits for the institution and the individual. Benefits may include freedom, wellbeing, sustainability, inclusivity and diversity, equality, suitable service provision, fairness, efficiency, and cost reduction.

Describe the benefits for the institution and for individuals

The benefits of learning analytics may exist on different levels and for different parties. An institution may apply learning analytics with the aim on reducing costs, in conjunction with achieving policy, educational, or research purposes.

Describe the risks for the institution and for the individual

As an institution, you must take suitable measures to mitigate the risks of learning analytics. In practice, this means assessing the risks presented by learning analytics on an ongoing basis to be able to detect when privacy risks for individuals arise. A "risk" is a scenario describing an event and its consequences, estimated according to severity and likelihood. As an institution, you must manage these risks by coordinating activities to guide the organisation and by implementing measures.

As part of the DPIA, you as an institution must investigate whether planned measures, safeguards and mechanisms to protect individuals' interests are sufficient or whether improvements can be made to further mitigate risks for the parties involved. You must therefore bear in mind the possible impact on individuals and any damage that processing data may cause, be it physical, emotional, or material damage. Look in particular at whether learning analytics constitutes a risk of:

- being unable to exercise rights (such as privacy rights)
- refusal of services or denial of educational opportunities
- loss of control over the use of personal data
- discrimination
- identity theft or fraud
- financial loss
- damage to reputation
- physical damage
- loss of confidentiality
- re-identification of pseudonymised data, or
- other significant economic or social disadvantage

3.5 Define risk mitigation measures

Once all risks have been outlined, your institution must describe the measures it intends to take to handle these risks, including safeguards, security measures, and mechanisms to guarantee personal data protection and to demonstrate compliance with the GDPR.

Record the source of each risk or potential disadvantage identified. Consider the options available to mitigate that risk. For example:

- decide not to gather certain types of data
- restrict the processing scope
- shorten retention periods
- take additional technological security measures
- train staff to ensure that risks are anticipated and managed
- anonymise or pseudonymise data where possible
- write internal guidelines or processes to avoid risks
- using different technology
- make clear agreements regarding sharing data (see 'Processor agreement')
- make changes to privacy declarations (see 'Inform')
- give individuals a reasonable measure of control over the data the institution processes. Perhaps you will be able to achieve the specified purposes merely by giving individuals access to the findings of learning analytics
- offer individuals the option to opt out of participating in the project, where necessary, or
- introduce new systems to help people exercise their rights.

This is not an exhaustive list; you can find more information in the official Dutch translation of the DPIA guidelines. Also ask your Data Protection Officer (DPO) for advice. Determine whether the measures reduce or eliminate the risk. Take the costs and benefits of each measure into account when deciding whether it is appropriate.

3.6 DPIA outcome

Once all of the information has been included in the DPIA, it is important to obtain advice from the institution's DPO. This offers additional certainty that the DPIA sufficiently considers the risks and that sufficient measures will be taken to cover these risks. This advice must be documented alongside the institution's decision of whether or not to proceed with the learning analytics process. It is also recommended to ask individuals (e.g. student representatives) what their opinion on the project is.

Tip

Submit the outcomes of the DPO to the Data Protection Officer (DPO) and/or internal legal department for a final review.

If the DPIA shows that processing constitutes a high level of risk for individuals and that this risk is not eliminated by taking risk mitigation measures, the institution must consult with the national Personal Data Authority before processing begins. This is known as a 'preliminary consultation'. For more information please visit the website of the Personal Data Authority.

STEP 4: IMPLEMENTATION

If the outcome of the assessment is positive, the learning analytics process is considered proportionate and no high residual risks have been identified, it is time to take the necessary legal, technical and organisational measures.

4.1 Conclude the necessary contracts

In the design phase (step 2), you outlined which suppliers and other third parties are involved in the learning analytics system. Immediately after this, you defined which parties will hold which roles (controller/processor) governed by privacy laws. Your institution must use this information to make arrangements and conclude agreements.

Draw up a Joint Responsibility Agreement

If your institution defines the purpose and means for the learning analytics process along with others it will result in a joint responsibility. Those bearing the joint responsibility must make mutual arrangements to transparently align their responsibilities with the GDPR. This particularly involves how individuals' rights are exercised and who will provide information on learning analytics. The essence of this arrangement must also be available to individuals.

Tip

- SURF has a template document for the Joint Responsibility Agreement
- See also SURF's Impact and Riskassessment portal

Draw up a Processor agreement

A Processor agreement describes the relationship between a service provider (the processor) who processes personal data on behalf of a different party (the controller). This agreement is mandatory. Large suppliers often have their own model agreements. A processor may only process personal data assigned to it by the controller under the Processor agreement. Check carefully whether the learning analytics processing and purposes are clearly described in the Processor agreement.

Tip

Institutions and SURF have jointly created a model Processor agreement as part of the SURF Legal Standards Framework for (Cloud) Services. This document creates standards for confidentiality, privacy, property, and availability for (cloud) suppliers.

Access to personal data outside the European Economic Area (EER): additional measures

Institutions are not obliged to store personal data with parties within Europe, but doing so is slightly easier from a legal perspective. This is because the GDPR grants individuals the same data protection standards in every country within the European Union, but different rules apply to passing personal data to parties outside the European Union, to bodies in so-called third countries. Third countries are all countries which are not EU member states plus Norway, Liechtenstein, and Iceland (the European Economic Area). It is likely that the United Kingdom will also become a third country, with separate rules, due to Brexit.

The main rule for third countries is that organisations may only pass on personal data if there is a suitable level of protection. The European Commission has created a Country List of third countries that it considers to have a suitable level of protection. In addition to the cases stated above, international data transfer is only permitted on the basis of legal provisions under the GDPR such as guaranteeing a level of protection with contractual provisions (model contracts or '*Standard Contractual Clauses*') or a manner designated by the European Commission (such as '*Privacy Shield*', a certificate for parties in the United States receiving European personal data). That is why you must check your agreements with the relevant parties to determine whether these require clarification or review.

Therefore you must check the solutions used by parties in third countries to facilitate the transfer of personal data. See more information from the Personal Data Authority regarding international circulation.

4.2 Draw up the necessary policy documents and procedures

List which policies may relate to learning analytics. Define the correct follow-up steps if your institution does not yet have relevant policies.

- Security policy, including user identity management (see also SURF publications on security).
- Procedures for handling data breaches in order to comply with reporting obligations in the event of a data breach. If a data breach occurs, your institution is required to report it in certain cases. Determine whether a data breach is one which you must report to the supervisory authority and, potentially, the parties involved.
- Procedures for handling requests from individuals to access or delete data.

4.3 Inform the individuals about the learning analytics processing

Anyone processing personal data must clearly inform persons affected of what happens with their data and why. Individuals must receive this information before or at the time the institution processes their data.

Institutions can use privacy statements to explain to individuals, such as students, what happens with their personal data. The ways institutions use data can be divided into categories. Assess the extent to which your institution can be transparent about learning analytics. The aim of transparency is to facilitate explaining the use and effect of learning analytics. In addition to the existence and effects of applying learning analytics, it is also important to clearly explain the consequences. In the privacy statement you explain to individuals exactly what learning analytics are, what data the institution uses towards this purpose and how and with what consequences learning analytics draws conclusions. For example: "We monitor how long it takes you to do the online exercises. If it takes you significantly longer than average, you will be given extra explanations and exercises to review before you can finish this module."

It is important that individuals can easily request the privacy statement and cookie declaration (where necessary) before their data is processed. Acknowledgement is not obligatory, and you are also not required to force individuals to read the privacy statement. It is not necessary to have individuals declare (e.g. by ticking a box) that they agree with the privacy statement.

If the institution invokes consent as a basis, then an individual may give consent by ticking a box on a consent declaration. This declaration must make clear what the student is consenting to; see also the 'Consent' basis, paragraph 3.1.

4.4 Security and risk mitigation measures

Apply the described security measures to guarantee the confidentiality, integrity, and availability of the personal data. Also implement the additional risk mitigation measures under paragraph 3.5. Then check whether the measures are having the desired effect.

4.5 Other measures

Implement a right to object to profiling
An individual may always object to being subjected to profiling or affected by a measure based on that individual's personality profile. The individual must be explicitly informed on the possibility of objecting . The person receiving the objection must then be able to reverse the measure. If collaboration with a software supplier is necessary to reverse an action taken by the learning analytics system, you must have made arrangements regarding this process with the supplier (in the processor agreement).

Communication
Define whether your institution must take other measures based on learning analytics and implement these. Examples include:
- establishing a communication procedure or help desk for questions from individuals and other stakeholders
- setting up an information page with an FAQ section
- training users of the learning analytics system.

4.6 Pilot

Consider carrying out a pilot to test whether all functionalities work properly. After the pilot, you can approve or reject the operation. If it is accepted, you can roll out learning analytics on a larger scale as the conclusion of the implementation phase.

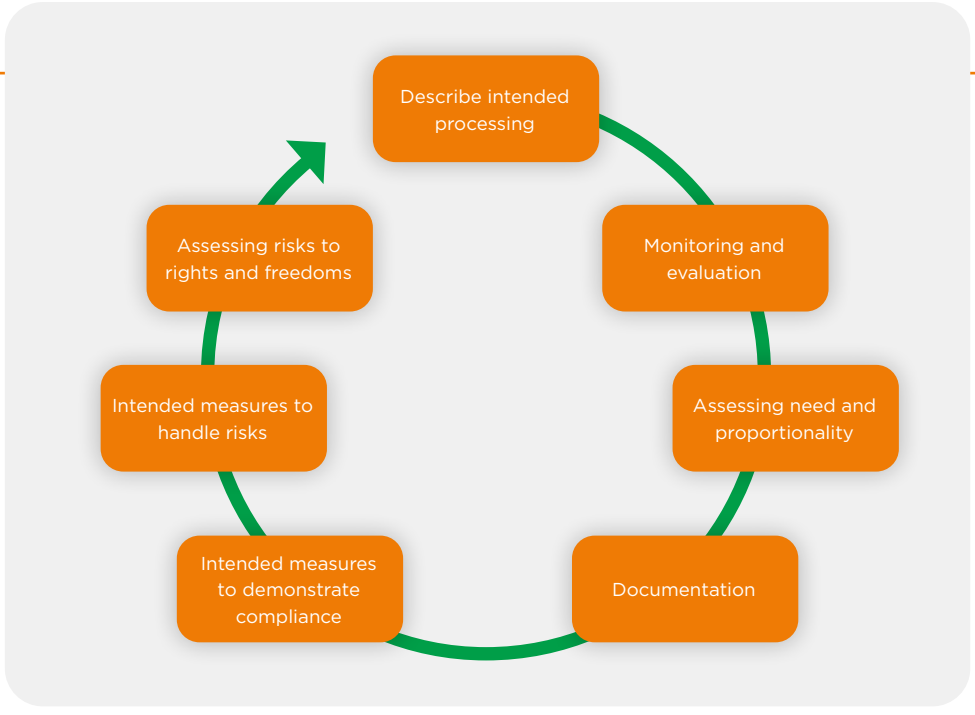
STEP 5: EVALUATION

5.1 Evaluate periodically

Assessing whether a learning analytics system is responsible from a privacy law perspective is not a one-time event. The organisation itself and external factors do change. This may impact the societal and legal frameworks of the system and therefore also its legitimacy. It is therefore important to periodically evaluate whether the system is still liable. By carrying out periodic evaluations you discover new risks on time and create a feedback loop that improves the application of learning analytics and makes it more effective.

Evaluations can be done at regular intervals (e.g. annually), but it is also a good idea to recognise situations that entail revaluation. Some examples are:
- The learning analytics system is used for a purpose other than that for which it was originally intended.
- New data sources or new technologies are used
- Existing data sources are adjusted or no longer used.

In such situations you should improve the process and where necessary, go through this step-by-step plan again. Carrying out this step-by-step plan is not a once-and-for-all assignment, it is an ongoing process.



Source: Visual from the DPIA opinion of the EPDB, p. 20.

5.2 Publish experiences

Consider saving and sharing your work methods. How can you ensure that teachers know where to find the description of the learning analytics process and which privacy considerations were made? Specialists, for example, often use a community platform where you can post your working methods, such as SURF's Privacy-wiki. You could also use your institution's internal portal.

ACKNOWLEDGEMENTS

Composition and editing
Niels Westerlaken, *Project Moore*
Jocelyn Manderveld, *SURF*
Floortje Jorna, *SURF*

With thanks to
Arnoud Engelfriet, *ICT Recht*
Evelijn Jeunink, *SURF*
Sebas Veeke, *SURF*

Project Management
Jocelyn Manderveld, *SURF*
Marieke de Wit, *SURF*

Layout
Vrije Stijl, Utrecht

Print
Drukkerij Libertas Pascal

Cover photograph
Annemiek van der Kuil, PhotoA.nl

Photography
p. 10, 21: Sicco van Grieken, Erasmus Universiteit Rotterdam
p. 18: Sicco van Grieken, Wageningen University & Research

May 2019

SURF
088 – 787 30 00
www.surf.nl/onderwijs
onderwijsinnovatie@surf.nl

Driving innovation together

Within SURF, universities (of applied sciences), vocational institutions, research institutions and university medical centres come together to work on ICT resources and innovation. Their aim: better and more flexible education and research. We do this by providing the best possible digital services, stimulating sharing and exchange of knowledge and most of all, constantly innovating! This is how we contribute to a strong and sustainable Dutch knowledge economy.

SURF