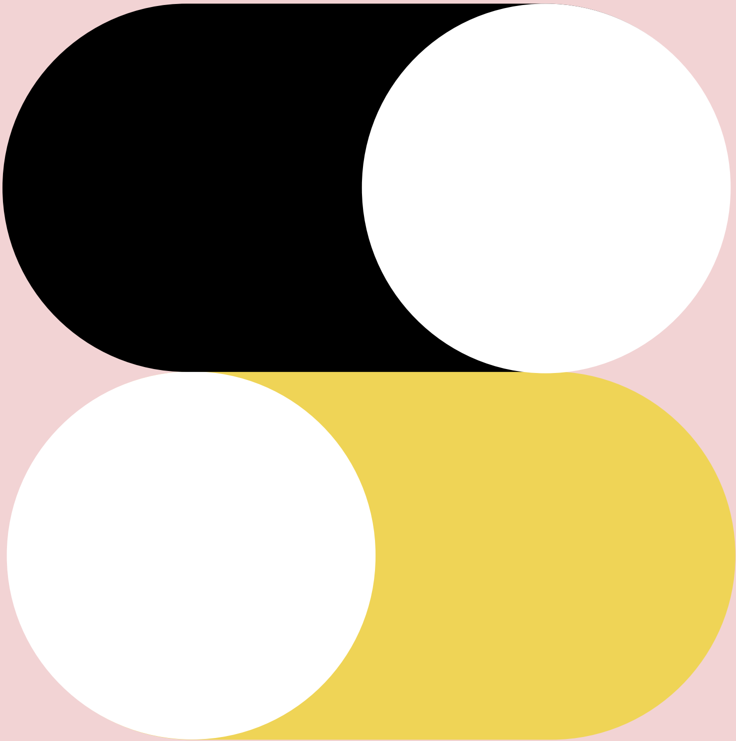
Npuls

npuls.nl



Reference Framework 2.0 - Responsible use of education data and AI.

Guide to weighing options and making choices

Guide

Reference Framework 2.0 - Responsible use of education data and AI

Guide to weighing options and making choices

April 2025

Contents

Summary

1. Introduction

There is much scope for the use of education data and AI in educational institutions. This offers many opportunities and potential benefits. However, there are also risks inherent to using education data and AI. How can these opportunities and benefits be harnessed while mitigating or eliminating risks? The key question in this Reference Framework is therefore: how can an institution use education data and AI responsibly?

To answer this question, clarity is needed on **values**, **legal frameworks** and **responsibilities**. This Reference Framework serves as a guide and a tool to help institutions determine where and how to start using education data and AI responsibly.

This Reference Framework is aimed at professionals in tertiary education in the Netherlands (research universities, universities of applied sciences and vocational education and training schools) who wish to work with education data and AI. This Reference Framework applies to cases where education data is used both with and without algorithms or AI.

2. Values

This Reference Framework views values as desirable general, abstract ideas or ideals that guide the use of education data and AI. The foundation for the values in the Reference Framework lies in public values for digitalisation in education. These values have been supplemented and enriched through close contact and dialogue with education experts, resulting in the values described below.

Fairness. The use of education data and AI is inclusive and promotes equality. Its use does not lead to unintended inequity or bias towards certain individuals and groups. Additionally, transparency, accountability and system reliability and security are essential for responsible use of education data and AI.

Human in the loop. The use of education data and AI takes due consideration of meaningful contact between teaching staff and students and among students themselves as well as the possibility of students' self-development in a safe and respectful environment. Human oversight is crucial, so automated use of education data and AI should always have a human in the loop.

Autonomy. The use of education data and AI respects students' self-determination to make choices for themselves, act independently and have agency over their lives and education. Autonomy is also essential for teaching staff and other education professionals; they can influence substantive and pedagogical choices within education. Furthermore, there should be a good balance between the use of education data and AI on the one hand and protection of privacy of data subjects on the other.

Institutions make their own considerations when interpreting these values and their own institution-specific values. In doing so, institutions can consider the purpose and impact of using education data and AI on students, teaching staff and other stakeholders as well as the context in which education data and AI are used and the risks associated with their use.

3. Legal frameworks

Besides values, there are also Dutch, European and international laws and regulations on the use of education data and AI, such as intellectual property rights, personal data protection (GDPR) and recent legislation specifically focused on AI (the AI Act). This Reference Framework focuses on the GDPR and the AI Act.

GDPR. Education data is almost always personal data. This means that the GDPR applies to the use of education data and AI. For each use of education data and AI, all educational institutions apply this triad: purpose & purpose limitation – legal basis – due care.

Purpose/purpose limitation	Legal basis	Due care
The purpose of using education data and AI is *well-defined, explicitly described* and *justified*.	The use of education data and AI requires the existence of a valid legal basis, namely:	To notify data subjects about the use of education data and AI.
In principle, an institution may ***not*** use education data for an incompatible purpose.	Consent of the data subject, where consent is informed, specific, unambiguous and freely given	Only education data necessary for achieving the intended purpose may be used.
An institution ***may*** use education data for archiving in the public interest, for scientific or historical research and for statistical purposes.	Necessary for the performance of a contract to which the data subject is party	When using education data, attention is paid to *privacy by design* and *privacy by default*.
Institutions must implement appropriate safeguards for protecting the rights and freedoms of data subjects.	Necessary for compliance with a statutory obligation	The education data is accurate and complete.
	Necessary for the institution to comply with a statutory obligation	Conducting a Data Protection Impact Assessment (DPIA) may be required.
	Necessary to protect the vital interests of the data subject	The education data may not be kept longer than necessary for the purpose for which this data is used.
	Necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the institution	The institution implements appropriate technical and organisational measures to protect the education data.
	Necessary for the legitimate interests of the institution or a third party	The rights of data subjects are respected, including the right not to be subject to solely automated processing of education data.
		The institution also applies the GDPR when sharing education data with other institutions and parties and concludes the relevant GDPR agreement (such as the processing agreement).

AI Act. The AI Act contains rules for developing and using AI systems. The AI Act also applies to institutions The nature and scope of these rules varies depending on whether the institution qualifies as a provider or a deployer of the AI system. It is crucial that institutions take the following actions:

- **Identify**. Institutions should identify and list which AI they use and for what purposes.
- **Classify**. Institutions should classify these AI applications based on the AI Act; what is the risk classification of the AI system (unacceptable, high, transparent or minimal)? What is the institution's role in the respective AI system?
- **Implement**. Based on the classification, institutions should implement prescribed measures to mitigate risks.

Onderstaande tabel geeft in één opslag de systematiek weer van de AI Act.

Risk	Consequences of the AI Act	Examples of AI Systems in education
Unacceptable risk	Prohibited	Emotion recognition in the workplace and in education
High risk	Permitted but with requirements for providers and – to a lesser extent – deployers. Requirements include: – Risk and quality management system – Data and data governance – Technical documentation and user instructions – Logging and monitoring – Human in the loop – Conformity assessment – Informing data subjects – Performing a Fundamental Rights Impact Assessment – Accuracy, robustness and cyber security	The AI Act defines four high-risk applications in education: 1. Access, admission and assignment to educational and vocational training institutions. 2. Evaluating learning outcomes, including for steering the learning process. 3. Assessing the appropriate level of education. 4. Monitoring and detecting prohibited behaviour during assessments/exams. Note: there are a number of exceptions in the AI Act
Transparency risk	Permitted, but with the obligation to inform data subjects	Chatbots and deepfakes
Minimal risk	Permitted without additional obligations under the AI Act	

4. Responsibilities

Responsible use (and continued responsible use) of education data and AI requires clarity on who makes decisions and what the roles and tasks are regarding all aspects of using education data and AI. During the entire life cycle of that use, various stakeholders (including students and teaching staff) and experts must be involved in the considerations to be made on responsible use of education data and AI. Organising dissenting voices provides new perspectives and other insights, allowing for values and legal frameworks to be discussed in a cohesive way.

5. Where and how to start?

Institutions can take the following steps for responsible use of education data and AI.

1. Develop a business case to gain insight into the purpose, advantages and disadvantages of using education data and AI but also the values at stake. Perform a check on the GDPR and AI Act to determine what must be arranged.
2. Involve stakeholders in a multidisciplinary approach. This way, the institution gains new insights when fleshing out values as well as identifying the opportunities and risks of using education data and AI.
3. Establish policies, guidelines and procedures for clarity on the various aspects of responsible use of education data and AI.
4. Ensure that education professionals have the necessary data and AI literacy.
5. Play the Dilemma Game to gain insight into opportunities and risks, interpretation of values, requirements under legal frameworks and measures to be taken to prevent or mitigate risks.

The Reference Framework as a dynamic tool

It is difficult to provide ready-made answers to many of the questions institutions have about education data and AI. Many aspects, especially those concerning the AI Act, are still in full development and require further interpretation Moreover, institutions are expected to gain increasing experience with the use of education data and AI in the coming years. In light of this, this Reference Framework is intended as a dynamic tool rather than an endpoint. We encourage institutions to learn from each other's positive and negative experiences – openly and in mutual trust.

1. Introduction

1.1 Background

How could institutions use education data? The desire to explore this has existed for some time. It is important that students, teaching staff and other stakeholders have confidence in responsible use of education data. To support institutions in the Netherlands in the use of education data, the *Reference Framework for privacy and ethics in education data (version 1.0)* was published in November 2021.

Since the release of Reference Framework 1.0, there has been a significant increase in the potential applications of education data, mainly in relation to developments in the possibilities of AI (especially generative AI).

> *Turbulent developments around Artificial Intelligence (AI) related mainly to the possibilities of big data. See the 'Exploration of data-driven educational research in the Netherlands' carried out in 2017 by the University of Twente (Verkenning_datagedreven_onderzoek.pdf, in Dutch only)*

Although AI is not a new phenomenon, ChatGPT's launch in November 2022 marked a new phase of AI development – a phase characterised by substantial attention for both its possibilities and its risks. And the tumultuous development of ever better and new applications and possibilities in many areas is far from over. Hardly a day goes by without a company or organisation introducing a new AI application. In the same vein, hardly a day goes by without someone publishing a report, study, guideline or framework that highlights the risks and dangers of AI.

> *The website theresanaiforthat.com publishes a list of new AI applications brought onto the market daily. A handy overview of all possible risks (more than 750) associated with the use of AI to date is the MIT Risk Taxonomy Report, see: The AI Risk Repository*

1.2 Reference Framework 2.0: a deeper and broader interpretation

Developments in AI naturally also affect tertiary education in the Netherlands. The impact of using education data and AI is significant for students, teaching staff, supporters, policymakers and boards of institutions. Many institutions are also just starting to navigate how to interact with AI – both in light of public values in education and the recently enacted AI Act. Given all these developments, there is sufficient reason to expand and deepen the Reference Framework on responsible use of education data by adding the highly relevant topic of AI.

There is also another reason for revising the Reference Framework. Reference Framework 1.0 focused on universities of applied sciences and research universities in the Netherlands. However, many of the topics and challenges addressed are also relevant to vocational education and training schools (*mbo*). However, themes specific to vocational education and training – such as issues surrounding minors, relationships with parents or guardians, duty of care towards students and collaboration with different parties (e.g., internship companies and support organisations) – are not included. This new Reference Framework has therefore been expanded to include vocational education and training in the Netherlands.

Reference Framework 2.0 draws on an evaluation of experiences gained since 2021 as well as numerous discussions with subject-matter experts from all tertiary education institutions (vocational education and training schools, universities of applied sciences and research universities – *mbo*, *hbo* and *wo* respectively in Dutch).

1.3 Purpose, target group and structure

The purpose of this Reference Framework is to help tertiary education institutions in the Netherlands use education data and AI responsibly. This Reference Framework is intended for professionals in institutions that work with education data and AI in practice.

Tertiary education institutions face a number of questions when it comes to the responsible use of education data and AI: Where do you begin if you want to start using education data and AI and what must you pay attention to? What values are relevant to the institution? What can and – in light of the GDPR and the AI Act – what must you do and arrange? Who do you involve in your considerations and choices? As an institution, how do you ensure continued responsible use of education data and AI?

Providing answers to these questions and the decisions an institution makes require careful consideration. If an institution is too cautious, students and teaching staff may miss out on opportunities. But if an institution does not take due care, students and teaching staff run unnecessary risks of exclusion and unreasonable treatment or inequity.

This Reference Framework aims to provide guidance in answering these questions and to provide inspiration for further development of institution-specific policy frameworks, working methods and processes regarding the responsible use of education data and AI. This Reference Framework can provide that guidance by presenting values and legal frameworks but also creating clarity around responsibilities.

The title of this Reference Framework is Responsible use of education data and AI, which means that the framework applies to three scenarios:

1. Use of education data without algorithms or AI.
2. Use of education data with algorithms (which are not AI).
3. Use of education data with AI (including generative AI).

In Chapter 2, we explain what the Reference Framework means by education data, algorithm and (generative) AI.

1.4 The Reference Framework as a dynamic framework

It is difficult to provide a ready-made answer to a large number of questions that educational institutions may have. Many aspects, especially those relating to the AI Act, require further clarification and interpretation. In the coming period, we expect more clarity from the European Commission, the European AI Office and national supervisors such as the Dutch Data Protection Authority, among others.

For that reason alone, this Reference Framework is a dynamic framework; further clarification can and will have consequences for the content of the Reference Framework.

Furthermore, the number of successful (and unsuccessful) applications and best practices involving the use of education data and AI is expected to increase in the coming period. It is important that educational institutions learn from each other’s positive and negative experiences and can do this openly and in mutual trust. We therefore encourage all stakeholders to actively engage with the Reference Framework and continue sharing experiences to keep it relevant and up-to-date.

> A good example of sharing experiences is the publication of MBO Digitaal on various hits and misses in the area of digitalisation.

Finally, we encourage you to discuss the use of education data and AI in your own institution. Consider questions such as: What do we want achieve? What are the advantages and disadvantages? How does the use of education data and AI contribute to public values in education? How do we mitigate the risks? These conversations with students, teaching staff, support staff, policymakers and administrators are crucial for a better understanding of the possibilities and impossibilities of education data and AI, making a balanced choice about the use of education data and AI and for continued responsible use of education data and AI. This Reference Framework is a tool, not an endpoint; actively engage all stakeholders in your institution and closely follow other initiatives related to education data and AI.

An overview of ongoing initiatives and activities by Npuls, SURF and MBO Digitaal:
- *Exploration of an algorithm register for education*
- *ChatGPT guidelines*
- *Joint procurement negotiations for AI tools*
- *Exploration of the AI Ethics maturity model*
- *Policy and AI course*

1.5 Reading guide

Chapter 2 defines and describes key concepts used in this Reference Framework. It also includes applications and potential applications in tertiary education institutions as well as the opportunities and risks of using education data and AI.

Chapter 3 provides an overview of values for the responsible use of education data based on public values in education, as included in the Value Compass for Digitalisation in Education provided by Kennisnet and SURF (*WaardenWijzer*). Institutions have greater flexibility to interpret these values as compared to the legal frameworks discussed in Chapters 4 and 5.

Chapters 4 and 5 outline the legal frameworks an educational institution in the Netherlands must consider, namely the General Data Protection Regulation (GDPR) and the AI Act. These two pieces of legislation determine what each institution *must* in any case arrange when using education data and AI, or certain forms of AI. In view of the more binding nature of this legislation, these chapters are longer than the other chapters.

Chapter 6 describes the importance of properly arranging responsibilities and a multidisciplinary approach to using education data and AI.

Chapter 7 concludes with the steps an institution can take to start using education data and AI responsibly.

Almost all chapters include examples, practical tips and focus areas.

2. Education data and AI

2.1 Definitions

Education data

The Education Data Zone[1] defines education data as a collective term for a wide range of structured and unstructured data that can be used in institutions to improve the quality, effectiveness and efficiency of education.

Examples of education data are students' educational programmes (and prior education), grades obtained, work submitted, figures on student intake, progression and exit, recommendations on continuing the programme, socio-economic background of students, lectures/working groups attended, educational resources, timetables, lecturer evaluations, and so on.

Education data is created in various ways, such as when organising and providing education and through collaborations between institutions and third parties (such as those Studielink, DUO and CBS have with internship companies and civil society organisations).

Most education data in the Netherlands is held in information systems for teaching, administration of enrolments and academic results, quality assurance and administering of assessments and exams.

Artificial Intelligence (AI)

It is striking that in descriptions and discussions about AI, it is often used as a broad umbrella term with descriptions such as generative AI, machine learning, neural networks, algorithms, AI systems, AI models and so on. In brief, what exactly are we talking about when we use the term AI in this Reference Framework?

Where this Reference Framework mentions AI, it is referring to AI systems as defined in the European regulation on artificial intelligence (2024/1689), hereinafter referred to as the AI

[1] See: doe-meer-met-studiedata.nl/datagedreven-werken/

Act[2]. The Reference Framework is thus consistent with this legal framework. In using this definition, the AI Act aims to be technology-neutral so that the AI Act remains relevant as new technology and learning techniques emerge. This also applies to the Reference Framework.

AI-systems are defined as follows:
"An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment."

An AI system is a system specifically designed to analyse data and recognise patterns and use that knowledge to make informed decisions or predictions. AI systems can learn from data and adapt over time. Learning takes place through machine-learning techniques such as neural networks and deep learning.

No AI system

Under the AI Act and this Reference Framework, there is no AI system if the system is based on rules that have been established solely by humans to perform actions automatically. See below under algorithms.

AI systems can be applied both specifically and generically. An example of a specific application is an AI system that makes a prediction about students' academic success. A generic application involves the generation of text, audio, video or a combination thereof. This is also called generative AI. Examples of generative AI include ChatGPT, MidJourney and DALL-E.

Where this Reference Framework refers to AI, it refers to both specific and generic/generative AI systems.

It is important to understand that the application of the specific and generic AI system determines the risk classification of the AI Act. This can be a difficult distinction when it comes to generative AI. For example, a chatbot in itself is not a high-risk application, but if that chatbot is used as part of a high-risk application, it is a different matter. Chapter 5 provides further details on this topic.

[2] The AI Act definitions are based on the OECD definitions (see: Explanatory Memorandum on the updated OECD definition of an AI system, OECD Artificial Intelligence Papers, March 2024, No. 8. (Explanatory memorandum on the updated OECD definition of an AI system | OECD

The development and commissioning of AI systems is characterised by a number of phases (life cycle). Roughly speaking, these are: :

1. Defining the problem and determining the purpose of the AI system.
2. Obtaining and processing data.
3. Developing and training the AI model.
4. Evaluating and adjusting the AI model.
5. Putting the AI system into use.
6. Monitoring the use of the AI system.

Algorithms

The AI Act does not define the concept of an algorithm. Yet this is crucial for a proper understanding of AI systems. Algorithms are, in essence, a set of instructions or rules to complete a task or solve a problem. Examples of algorithms include mathematical formulas, recipes for preparing a meal or rule-based systems established to perform tasks automatically (if-then systems). Descriptive reports on, for example, trends in numbers of students can also be algorithms.

These examples of algorithms are not AI or AI systems according to the AI Act; nevertheless, the impact on students of the use of education data with algorithms (which are not AI) can be significant. That is why the Reference Framework also applies to the use of education data with algorithms that are not AI.

AI model

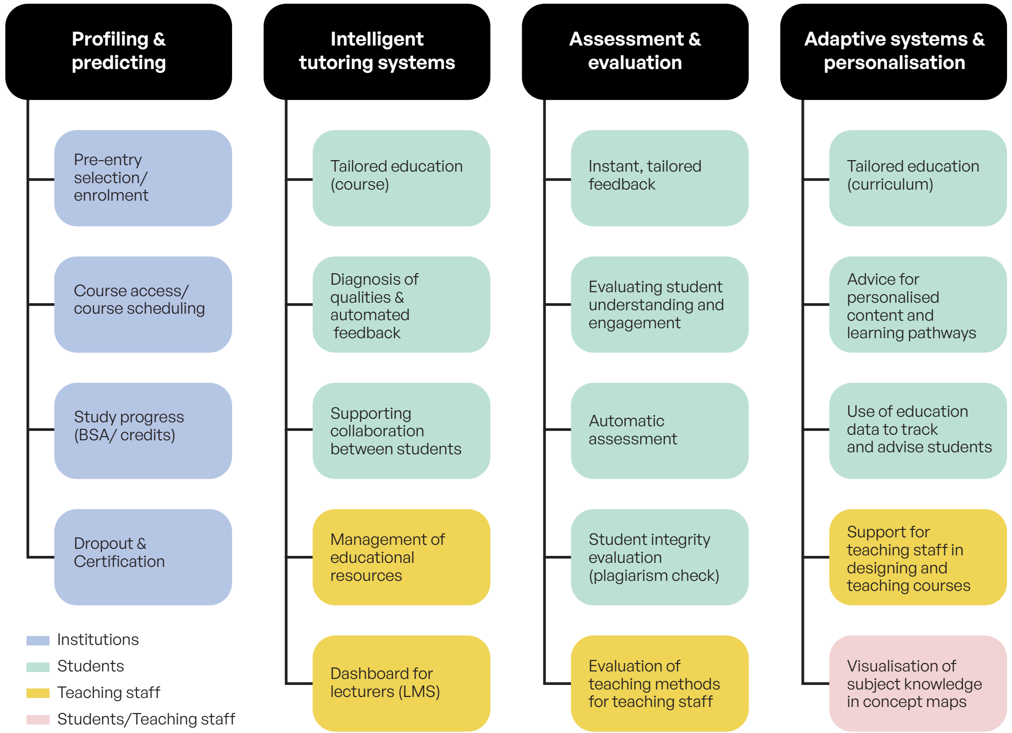
The AI Act also includes the term "AI models for general purposes". These models are explained in further detail in Chapter 5. For now, it is important to emphasise that these models are not an AI system in themselves and that the risk classification for AI systems under the AI Act does not apply to these AI models.

Data subjects

This Reference Framework regularly refers to "data subjects". These are human beings who can be identified directly or indirectly (data subjects as referred to in the GDPR) as well as those who are affected by the use of education data and AI. These are mostly students and teaching staff.

2.2 Applications and opportunities of education data and AI

Bducation data and AI can be widely used in tertiary education (i.e. vocational, higher professional and research-oriented education). This is shown in the overview below.


Profiling & predicting
Pre-entry selection/ enrolment
Course access/ course scheduling
Study progress (BSA/ credits)
Dropout & Certification
Intelligent tutoring systems
Tailored education (course)
Diagnosis of qualities & automated feedback
Supporting collaboration between students
Management of educational resources
Dashboard for lecturers (LMS)
Assessment & evaluation
Instant, tailored feedback
Evaluating student understanding and engagement
Automatic assessment
Student integrity evaluation (plagiarism check)
Evaluation of teaching methods for teaching staff
Adaptive systems & personalisation
Tailored education (curriculum)
Advice for personalised content and learning pathways
Use of education data to track and advise students
Support for teaching staff in designing and teaching courses
Visualisation of subject knowledge in concept maps
🟦 Institutions
🟩 Students
🟨 Teaching staff
🟥 Students/Teaching staff

AIEd typology based on Bond (2024). The colours indicate the main target group of the AI application. Bond et at., International Journal of Educational Technology Higher Education (2024), 21:4 (doi.org/10.1186/s41239-023-00436-z).

The Reference Framework 1.0 already included examples of the use of education data (with or without AI): well-being monitoring, distance learning and standard reports for lecturers from the Learning Management System (LMS). Education data (with or without AI) can also be used to gain insights into the academic results of individual students, find out how students are progressing, predict student dropout and prepare assessments and exams.

Examples of existing applications (including pilot applications) of education data with AI in education (both in the Netherlands and abroad) include:

- Helping students manage stress (see ixperium.nl/ai-helpt-studenten-omgaan-met-stress, in Dutch).
- Supporting students living with special needs (see the BeMyEyes app).
- Assisting students in selecting courses tailored to their individual needs (Magazine-AI-in-onderwijs-NL-AIC.pdf (nlaic.com)).
- Predicting academic success (see kennisnet.nl/onderzoek/kan-ai-studie-uitval-voorspellen-in-het-mbo).
- Providing automatic feedback on students' written (feedbackfruits.nl).

See, among others:

- *Praktijkgids AI in het onderwijs* by Zuyd University of Applied Sciences (1734605532357)
- SURF, Promises of AI in Education, discussing the impact of AI systems in educational practices, June 2022 (Promises of AI in Education | SURF.nl)

The possibilities for using education data and AI are vast and will continue to grow in step with increasing data volumes and ever new technologies.

2.3 Risks of using education data and AI

The examples mentioned make it clear that the use of education data and AI has great potential. At the same time, the use of education data and AI may lead to unintended and undesirable consequences, such as:

- loss or limitation of the human dimension (and human contact);
- loss or limitation of autonomy and the right to self-determination;
- loss or limitation of agency, the freedom to make one's own choices and the possibility to fail; in other words – changing a learning environment into a performance environment;
- excessive or unnecessary monitoring of students (insight into behavioural, living and learning patterns);
- risk of educational institutions becoming less inclusive and of increasing inequity by excluding certain groups on the basis of available data;

- exclusion or discrimination[3].
- misunderstandings due to incorrect/incomplete data or misuse/misinterpretation of data.

There are also general concerns about the privacy and protection of the personal data of data subjects.[4] And also about the desire to capture everything with data in models, which ignores other relevant dynamics.

In short, responsible use of education data and AI allows for positive interventions in education while minimising negative consequences.

[3] See the complaint a VU student filed with the Netherlands Human Rights Board about discrimination by AI software of online proctoring. The student claimed that the software did not recognize her face because of her dark skin colour. This complaint was eventually dismissed, but the Board held that it could not be ruled out that the use of such software could lead to discrimination in other situations. See the news item of 17 October 2023 on the website of the Netherlands Institute for Human Rights.

[4] For example, the Dutch Data Protection Authority has warned that the use of AI chatbots can lead to data breaches (because employees enter personal (or special personal) data into AI chatbots, thereby giving unauthorised access to that personal data. See the news report of 6 August 2024 on the AP website.

3. Values

3.1 Introduction

What values guide the responsible use of education data and AI in institutions? This chapter focuses on values.

> The Reference Framework views values as aspirational general, abstract ideas or ideals that guide action.

Besides values, there are laws and regulations on the use of education data and AI, in particular legislation on intellectual property rights, privacy and protection of personal data (the GDPR) and recent specific legislation on AI (the AI Act). This legislation – as well as the Dutch Constitution, the European Convention on Human Rights (ECHR) and the Charter of Fundamental Rights of the European Union – contains rules and regulations that are also based on values (that apply in Europe). A strict separation between values and laws and regulations is therefore not always easy to make.

> The protection of personal data is elaborated in the GDPR and the values of transparency and human in the loop are elaborated in the GDPR and the AI Act.

Separate focus on values

Why a separate focus on values? Specifically citing values as a starting point for the responsible use of education data and AI is essential for two reasons. First, not everything that is legally permissible is consistent with values. In other words, what is legally permissible is not always ethically responsible. Moreover, laws and regulations sometimes allow for an individual interpretation that can be given substance with values. Values can therefore contribute to an institution's ambition regarding the use of education data and AI, for example.

Values in this Reference Framework

What values does this Reference Framework apply? It applies the public values for digitalisation in education, namely **fairness**, **human in the loop** and **autonomy**. These core values represent a number of other values. Based on interviews with, answers to the questionnaire and reviews from a large number of experts in tertiary education institutions in the Netherlands, these values have been elaborated in more detail.

Selecting and weighing values

Not all values are applicable to the same extent in every case. Moreover, there may be tension between these values and their application in achieving a specific goal. It is up to the institutions to select and weigh these values carefully. This choice is determined by a number of factors, such as the size and organisation of the institution, the impact of the use of education data and AI, the context in which education data and AI are used, the risks associated with the use of education data and AI (in which a high-risk AI system is a significant factor according to the AI Act) and the values of an institution. Institutions can (and must) therefore interpret these values themselves. Chapter 7 offers tips on how educational institutions can address these challenges.

The structure of this chapter

This chapter on values is structured as follows:

- a description of the values of fairness, human in the loop and autonomy according to the Value Compass for Digitalisation in Education;
- a description of what those values imply for the use of education data and AI;
- examples of questions/dilemmas that can help educational institutions interpret and elaborate those values.

Contributing ideas for further elaboration of values

The questions/dilemmas outlined here are a selection of examples. Stakeholders are invited to contribute ideas on other questions and dilemmas and on further elaboration of the values. The Dilemma Game accompanying this Reference Framework can be of help in this respect. In the Dilemma Game, values and dilemmas are elaborated further based on the specific applications of education data and AI in education, namely: (i) profiling and prediction, (ii) intelligent tutoring systems, (iii) assessment and evaluation and (iv) adaptive systems and personalisation (see Section 2.2).

3.2 FAIRNESS

3.2.1 What does fairness entail?

The Value Compass for Digitalisation in Education provides a description of what fairness entails.

> Fairness includes concepts such as equality, inclusiveness and integrity. In education, everyone should have equal opportunities without specific groups being disadvantaged or excluded. Social or cultural background or gender should not affect the treatment that pupils and students receive in education. This means they should be treated without prejudice by teaching staff but also by algorithms. Equality is paramount in education. Inclusiveness in education means accessibility for all pupils and students. Everyone can participate, is included and is prepared for a good life in society. Integrity means, among other things, that what happens in education is reliable, transparent and verifiable – just like the data and the systems that are used. Integrity also means that public funds are spent efficiently and with a view to sustainability to ensure fairness for future generations.

3.2.2 Fairness and the use of education data and AI

But what does fairness entail in terms of using education data and AI?

Equity

Institutions ensure that the use of education data and AI does not lead to (unintended) inequity and discrimination against individuals and groups. Especially when using AI systems, there is a risk of (unintentional) bias in the various phases of development of the AI system. This may be because the training data is not representative or complete, the algorithm itself is biased, or the developers of the AI system are unintentionally biased. Bias could potentially lead to unfair decisions that discriminate against certain groups.

> **Bias**
>
> The Rhite agency has published a practical guide together with Radboud University. This guide not only describes the different forms of bias during the life cycle of an AI system but also includes practical recommendations on how to address bias. See: From-Inception-to-Retirement-Addressing-Bias-Throughout-the-Lifecycle-of-AI-Systems.pdf
>
> See also the *Handbook on non-discriminating algorithms* of Tilburg University: Handbook on non-discriminating algorithms | Tilburg University
> Specifically for AI and equity, see: *AI and equity, Vision document on the impact of AI on equity in tertiary education*, Npuls, November 2024 (AI and equity - Npuls)

Preventing and managing bias in the use of education data and AI requires students and teaching staff (or their representatives) to be involved in identifying risks and the measures to be taken to mitigate these risks.

Procedure for organising equitable use

Weighing up and assessing risks for equitable use requires a clear procedure for developing, using and monitoring the use of education data and AI. In this procedure, all stakeholders play a role in a multidisciplinary approach. Establishing a new procedure may not always be necessary; existing risk procedures or the DPIA procedure can be enhanced or supplemented to address these aspects. Institutions can also enhance or supplement existing risk procedures or their DPIA processes to address these aspects effectively.

Procedure for handling complaints

Ensuring fair use also requires a clear and easily accessible complaints procedure so that data subjects' complaints are dealt with quickly and effectively.

Inclusiveness

Students with a functional disability, a migration background or a socio-economic disadvantage should not be disadvantaged by the use of education data and AI. Institutions should carefully assess whether the use of education data and AI might have adverse effects on specific groups or individuals, for example through scenario analyses.

Integrity

Reliability and security of systems. This means that institutions should ensure that sufficient relevant data is, becomes and remains available for the predetermined purpose. Institutions should also ensure that data is reliable, is securely stored and used and meets high methodological quality standards.

Focus areas to bear in mind when using education data and AI:

- Inaccuracies in the education data are understood and minimised.
- The implications of incomplete data sets are clear.
- An appropriate set of data sources is used.
- Techniques such as anonymisation and pseudonymisation are understood and correctly applied.
- False correlations are avoided.
- Results of previous studies are taken into account.
- Results are tested for bias.
- Processing, analysing and using education data is always considered in a broader context and, where necessary, combined with other knowledge and approaches.

Reliability also requires robust information and cyber security. These values are embedded in the GDPR and the AI Act. At the same time, specific security risks associated with AI require tailored security measures. These risks include both an increase in existing threats, such as data breaches, and new risks that are unique to AI systems. Examples of these new risks are the manipulation of models and jailbreaking (bypassing safety mechanisms). This requires, among other things, new control measures, adjusted governance frameworks and special tooling. At the same time, AI also offers opportunities for improved information and cyber security.

Cyber Security Council (Cyber Security Raad, CSR)

Besides an overview of the specific opportunities and risks of AI in the field of cyber security, the CSR has also provided an overview with the conditions of responsible use of AI in this context. See: Informerende+brief+aan+de+staatssecretaris+van+BZK+over+(generatieve)+AI+en+cybersecurity+DEF (2).pdf

Transparency. Transparency means that institutions ensure that it is clear to everyone involved in education data and AI for what purpose education data and AI are used, what data has been relied on, how the data was obtained, what results they wish to achieve and by what means and how a specific result was obtained, whether automated or not. These choices and considerations have been documented by the institution to make them traceable and explainable. In addition, data subjects have the specific right to be informed that they are dealing with an AI application.

However, providing full transparency on the use of AI systems – which are often complex and opaque – poses challenges. The challenges lie in keeping information comprehensible while providing transparency about which data contributes to specific outcomes – a task often complicated by the fact that many AI systems are self-learning. Opening this "black box" requires explainable AI.

Transparantie

Institutions can create transparency in various ways, such as:

- Ensuring that employees in relevant positions are sufficiently data and AI literate to handle questions, requests and complaints from data subjects quickly and satisfactorily.
- Involving suppliers in providing clear information about how their AI "works" (which input is used to develop the output, etc.), for example through visualisations that show which elements weighed most heavily in developing the output.
- Identifying and recording (e.g. in an algorithm register) the applications of AI in an educational institution.

Accountability. Accountability means taking responsibility. Especially in cases of conflicting interests, it is important for an institution to weigh the options carefully and document the considerations and choices made (in the Register of Processing Operations or in an algorithm register, for example). Institutions provide insight into who is responsible or accountable if there are doubts about specific uses of education data and AI.

Accountability also means being receptive to the question of usefulness and necessity of using education data and AI and how the institution takes into account the legitimate interests of data subjects and other stakeholders as well as societal and social aspects. In the context of accountability, institutions must continue to assess whether the intended purpose of using education data and AI has been achieved and/or whether adjustments are necessary and desirable.

The educational institution must be able to explain why and on what basis certain choices are made – regardless of whether these are policy-related choices or choices that affect an individual directly – and to do this on a regular basis. Moreover, the use of new but also existing techniques should never be a goal in itself but a means to achieve a higher purpose.

Plan-Do-Check-Act (PDCA)

A proven method is to maintain a PDCA cycle. This phased approach involves the following steps:

- **Plan**: Consider the activities involved and draw up a plan for performing them.
- **Do**: Perform the activities set out in the plan.
- **Check**: Analyse the progress of the activities performed and identify the related risks/bottlenecks.
- **Act**: Adjust the activities/plan based on the outcome

Accountability for what can and may be done with education data and AI is not a one-way street. Students, teaching staff, support staff and other stakeholders can expect the institution to actively involve them in the choices surrounding the use of education data and AI and – where relevant – actively inform them about achievement of the intended purposes, any associated risks and the measures needed to mitigate those risks. In this way, institutions maintain an ongoing dialogue with all data subjects and thus contribute to promoting a culture of responsible use of education data and AI.

In certain situations, involving bodies such as a student council and/or a participation council in the use of education data and AI may be a statutory obligation.

3.2.3 Questions and dilemmas about fairness as a value

- Are there applications of education data and AI in education for the benefit of the student that must be considered unacceptable, irrespective of the potential benefits of using education data and AI for certain groups of students? If so, what are the factors that determine whether an application is unacceptable?
- Is the purpose for which the institution wants to use education data and AI sufficiently clear and measurable? Do we have sufficient and representative education data to achieve this purpose?
- How are students and teaching staff who are affected by the use of education data and AI informed about this? Do they receive meaningful explanations when important decisions are made? How is that guaranteed?
- Which persons or departments can students and teaching staff address their questions, concerns and complaints to? How does the institution handle them?
- How will the education data and the use of AI remain safe and reliable? How does the institution keep track of the use of education data and AI and how does it monitor its operations? And who in the institution does that?

3.3 HUMAN IN THE LOOP

3.3.1 What does ‘human in the loop’ entail?

The Value Compass for Digitalisation in Education provides a description of what ‘human in the loop’ entails.

> Human in the loop means having an eye for people in education. It is about social cohesion, meaningful contact, respect, safety, health, well-being and self-development. Educational institutions provide opportunities for social connection, encounters and meaningful contact. In doing so, the educational institution respects the unique nature of each pupil and student, who is seen and heard as a human being and is not treated as a number or cog in a system. They provide a safe environment – online and on-campus, physically and mentally – where the health and well-being of pupils and students is safeguarded. In this safe environment, it is possible to make mistakes without this having an impact outside the educational context. The human factor must remain paramount in education: no decisions or judgements are made about pupils or students based purely on the analysis of data. Education contributes to self-development – expressing the individual nature of pupils and students in relation to the world.

3.3.2 Human in the loop and the use of education data and AI

What does 'human in the loop' mean in the context of using education data and AI?

Meaningful contact

Meaningful contact is about interactions between teaching staff and students and between students in areas such as their studies and development. The institution ensures that the use of education data and AI does not stand in the way of meaningful contact between teaching staff and students and between students themselves. These interactions contribute to the learning and development process.

Self-development

Self-development refers to the individuals' ability to develop themselves in a safe and respectful environment. This applies to both students and teaching staff. The institution ensures that the use of education data and AI is focused on self-development and contributes to helping students determine their own learning pathways while broadening their skills and interests in different areas. The institution also ensures that teaching staff have sufficient scope and freedom in developing educational resources and teaching methods when using education data and AI.

Human in the loop

Maintaining human oversight is essential, especially when using education data and AI in automated processes. Institutions ensure that there is always a human being involved in the automated use of education data, the 'human in the loop'. This is the case for automated processes with potential consequences for individual students or small groups of identifiable students and lecturers. However, it also applies to monitoring input, functionality and output of the education data and AI used. The institution ensures that a data subject can object to a decision in an easy way and that a competent decision-maker responds to this objection quickly and with reasons.

Procedures for processing, analysing and using education data and AI and for interventions are carefully designed and regularly reviewed, for instance through the PDCA cycle. Institutions also recognise that automated analyses of educational data and AI are unlikely to provide a complete picture of a person's learning process and that personal circumstances cannot always be taken into account.

3.3.3 Questions and dilemmas about 'human in the loop' as a value

- What is the role of using education data and AI in educational decisions? Should that role only be supportive or can it extend beyond that? And when and to what extent should there be a human in the loop?
- How can we strike a balance between the use of education data and AI for individual optimisation of study and learning processes on the one hand and maintaining meaningful contact and self-development in education on the other?
- How do we prevent students' perspectives being shaped too much by education data and AI, thus removing the human behind the data?
- Do teaching staff members and other professionals have the necessary training and information to use education data and AI effectively and ensure that it is safe and does not harm or infringe students' rights?
- Is conscious thought given to how the institution approaches students, teaching staff and other staff as human beings? And to what extent does the use of education data and AI detract from this or does it actually contribute to it?

3.4 AUTONOMY

3.4.1 What does autonomy entail?

The Value Compass for Digitalisation in Education provides a description of what autonomy entails.

Autonomy literally means: prescribing the law for yourself. Autonomy includes values such as self-determination, protection of privacy, independence and freedom of education. Pupils and students have self-determination: they are given freedom of choice to follow learning pathways that are suited to their needs and have autonomy over their development and choices. Protection of privacy is an important part of autonomy: teaching staff, lecturers, pupils and students must be able to trust that their privacy is safeguarded when they work with the digital resources in their institution and decide for themselves what happens to their data.

Independence of education means that institutions can design their education and curricula free from external influence. Professionals in education have scope to make their own assessments and choices when supervising students based on their professional autonomy. Freedom of education means that educational institutions can design their education on the basis of their own identity and convictions within the limits of the law and appropriate in a free, pluralistic and democratic society.

3.4.2 Autonomy and the use of education data and AI

But what does autonomy mean in the context of the use of education data and AI?

Self-determination

Self-determination refers to a student's ability and freedom to make their own choices, act independently and have agency over their own life and education. It encourages students to actively participate in learning processes and also prepares them for life after their studies. Institutions ensure that the use of education data and AI does not lead to disproportionate restrictions in self-determination when it comes to choices in the study process, learning pathways and development avenues. The use of education data and AI should be a tool that supports students, not a system that makes decisions for them without their input or choice.

Independence

The independence of education and teaching staff in relation to the use of education data and AI raises important questions about how technology can be integrated without undermining the core values of education and the professional autonomy of teaching staff. Institutions should use education data and AI as a tool that supports teaching staff, not as a replacement.

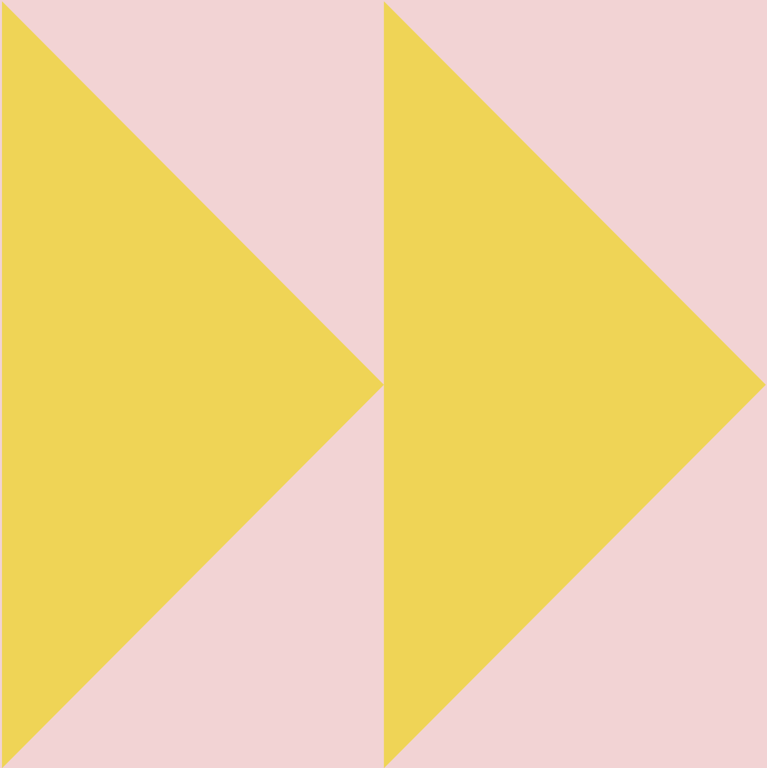
A frequently cited purpose of using education data and AI is to relieve teaching staff of routine tasks so that they can spend more time and energy interacting with students. Professional autonomy is important for a meaningful role as teacher or lecturer. This means that education professionals must always be able to influence substantive and pedagogical choices within the educational process.

Protection of privacy

The use of education data and AI means collecting and using a lot of data and personal data, especially from students and teaching staff. This raises many questions about the protection of privacy and the protection of personal data of students, teaching staff and other data subjects. Much education data can be considered sensitive because it is information about students' intellectual capacities, learning disabilities and sometimes even emotional and psychological states. Institutions only collect and use education data that is suitable for achieving the intended, pre-defined purpose. It is also essential that institutions are transparent about the use of education data and that students are well informed about that use and the rights they can exercise.

3.4.3 Questions and dilemmas about autonomy as a value

- How do we prevent the use of education data and AI from being overly relied on or users becoming too dependent on that use?
- To what extent is a student's self-development at odds with the interests of the institution and monitoring the quality of education?
- Should students have the right not to participate in the use of education data and AI, even if this has a negative impact on their studies and development?
- How much scope do students still have to choose their own learning pathways and make decisions without interference from, for example, predictive or recommendation algorithms or AI?
- How far can an institution go in monitoring student behaviour for educational purposes? Should there be a limit to what education data the institutions can collect, even if more data can lead to better educational outcomes?
- How much scope do teaching staff still have to determine the content and direction of education?



4. Legal framework: GDPR

4.1 Relevant definitions

Education data is **personal data** in information that directly or indirectly leads to the identification of students, prospective students, interested parties, former students, teaching staff, supervisors and all other persons whose personal data is processed (data subjects under the GDPR).

The GDPR distinguishes between the following personal data:

- "ordinary" personal data, such as name, email address, study programme, student number, academic results achieved, recommendations on continuing the programme, learning materials used, study progress and socio-economic information;

> Although the GDPR does not make a formal distinction, certain personal data (e.g. academic results and recommendations on continuing the programme) can be regarded as sensitive personal data. Protection of *sensitive* personal data requires extra attention.

- "special" personal data[5], for example data about disabilities or data that reveals political opinions, religious or philosophical beliefs or trade union membership;
- personal data concerning criminal convictions, offences and Citizen Service Number (BSN).

> **Pseudonymisation and anonymisation**
> The terms pseudonymisation and anonymisation are frequently used when talking about privacy and the protection of personal data The main difference between these terms is that pseudonymous data is still personal data to which the GDPR applies. Data which is anonymised, for example through aggregation, cannot be reasonably traced back to an individual. The GDPR does not apply to the use of anonymous data.

[5] Special personal data includes personal data that reveals racial or ethnic origin, political opinions, religious or philosophical beliefs or trade union membership. It also refers to genetic data, biometric data for the purpose of unique identification of a person, data concerning health or data concerning a natural person's sex life or sexual orientation.

> Removing a person's name or replacing it with a number does not by definition mean that the data is anonymous, especially if a document still exists somewhere by means of which the number can be traced back to a name.
>
> Even if such a document does not exist, there is a good chance that the remaining data, whether or not in combination with other information or documents, can be traced back directly or indirectly to a person. In practice, therefore, it is more likely to be pseudonymous data than anonymous data.
>
> For an overview of the concept of pseudonymisation and different techniques of pseudonymisation, see the ENISA recommendations on shaping technology according to-TP0618398ENN.pdf

The GDPR is formulated in a technology-neutral way and does not refer to specific technological applications, including AI or machine learning. The GDPR focuses on personal data and provides different rules, depending on the nature of the personal data.

For example, "ordinary" personal data may only be used for specific, explicitly described and legitimate purposes if there is a legal basis and the personal data is processed carefully.

> **Special personal data, criminal personal data and BSN (burger service nummer = Dutch citizens service number) may not be used as education data**
> The GDPR prohibits the processing of special personal data with a limited number of clearly defined exceptions to that prohibition. This Reference Framework is based on the principle that special personal data and criminal personal data may not, in principle, be used as education data. If an institution is considering using this type of data nonetheless, it must contact the institution's data protection officer, privacy officer or legal expert.
>
> The BSN may only be used if the law explicitly provides for this. Institutions may process the BSN for enrolment and communication with government authorities. Under the Reference Framework, the BSN may not be used for other purposes, including as a characteristic to link databases or files.

If education data concerns personal data, then the use by an institution of that education data is considered **processing** of personal data under the GDPR. The term 'processing' is a broad concept and includes storing, viewing, modifying, linking with other files, forwarding, publishing and actually using personal data.

In short, the GDPR applies to the use of education data (personal data) and AI. Where this Reference Framework refers to education data, this is also considered to be personal data, unless otherwise indicated.

Under this Reference Framework, educational institutions (individually or collectively) determine the purpose and means of using personal data. These institutions are therefore called **controllers or joint controllers**.

> **Who is the controller: the institution or the employee?**
> Under the GDPR, the institution is the controller. Employees of an institution who use personal data/education data in the context of their work are not themselves controllers under the GDPR.

4.2 Purpose - Legal basis - Due care

The impact of the GDPR on the use of education data can be briefly summarised as follows:

PURPOSE
For which purpose do you want to process personal data?

LEGAL BASIS
What is the legal basis for processing?

DUE CARE
How do you ensure due care when processing the data?

4.2.1 Purpose and purpose limitation

Education data must be used for a specific, explicit and legitimate purpose and may not be subsequently processed for other purposes. First of all, this means that it must be determined what the intended purpose (or purposes) of using education data is and what education data is necessary for this.

The purpose (or purposes) should be formulated as precisely and specifically as possible. Merely referring to general purposes such as: improving education, optimising educational processes or carrying out individual interventions is not specific enough.

Education data can therefore only be used for a specific purpose (purpose limitation). In reality, the use of education data almost always involves *further* processing of data. Education data will originally not have been collected for a specific purpose such as providing insight into student intake, progression and exit or providing insight into factors for academic success.

Incidentally, further processing of education data for archiving purposes in the public interest, scientific or historical research or statistical purposes is permitted, *provided* that a number of appropriate measures to safeguard the rights and freedoms of data subjects have been taken. This includes technical and organisational measures to ensure data minimisation or pseudonymisation.

Whether further processing of education data for a new purpose (other than that stated above in the box) is possible requires consideration by the institution. In making these considerations, the following factors come into play: how close the new purpose is to the initial purpose, the expectations of the data subjects, the context and nature of further use, the nature of the education data and the potential impact on the data subjects. An institution must document these considerations in writing.

If the new purpose is not compatible, it is not permitted to continue using the education data for the new purpose, unless there is an independent legal basis for processing the education data for the new purpose. Of course, this basis must also meet the requirements set by the GDPR (see below).

4.2.2 Legal basis

The use of ordinary education data requires the existence of a valid legal basis as set out in the GDPR. These legal bases are:

- **Consent** of the data subject. This consent must be informed, specific, unambiguous and freely given. Consent can be withdrawn at any time. Given these requirements, it is important that an institution not only knows with whom and where such consent is stored, but also has a procedure for where requests for revocation can be submitted and how they are processed.

Consent for children
The protection of children's personal data deserves special attention and is subject to a number of additional requirements. For example, children under the age of 16 cannot give valid consent themselves. Only their parents or guardians can do that. Furthermore, the educational institution cannot simply assume the consent of the child's parent or guardian. The institution has a duty to make reasonable efforts to verify that the parent or guardian has actually given that consent, for example by carrying out checks to verify the age of the child.

- Necessary for the performance of a **contract** to which the data subject is party.
- Necessary for the institution to fulfil a **statutory obligation**. The statutory obligation, including subordinate laws and regulations, must be clear and precise and the application sufficiently predictable.

Example of a statutory obligation: progress report

An example of a statutory obligation is Section 7.1.5 of the Dutch Adult and Vocational Education Act (*Wet educatie en beroepsonderwijs*, WEB). This section of the law obliges vocational education and training schools (*mbo*) to report on the progress of students to their parents, guardians or carers or to the students themselves if they are of age (and legally competent).

- Necessary for the protection of the data subject's **vital interests**.
- Necessary for a task carried out in the **public interest** or in the exercise of official authority vested in the institution.

Example of public interest

Institutions have a duty of care to their students, for instance in promoting well-being. In principle, the basis of statutory obligation qualifies if the statutory obligation in question is clear, precise and sufficiently predictable (with regard to the use of education data) for the students. Usually this is not the case and the basis of public interest (or consent) is more obvious.

Institutions could use this basis for processing educational data for the purpose of providing (or improving) education. After all, the provision of education by institutions qualifies as performing a task in the public interest.

- Necessary for the **legitimate interests** of the institution or a third party. It is essential, however, that a legitimate interest is well-founded, that a balance is struck between the interests of the institution or a third party on the one hand and the interests of data subjects on the other and that data subjects are informed of this.

Example of a legitimate public interest

Processing of education data that relates only to improvements in the institution's business processes can be based on this basis.

For all legal bases, with the exception of consent, the requirement applies that the processing of personal data is necessary. But what does 'necessary' mean? In brief, it means that proportionality and subsidiarity apply.

Proportionality: the purpose of the processing must be proportionate to adverse privacy consequences for data subjects. Key perspectives are the extent to which the processing contributes to the purpose, the nature and scope of the personal data to be processed and the possible consequences for data subjects.

Subsidiarity: if the purpose of processing can be achieved with fewer or no personal data, then this other method should be chosen. Being able to resort to alternatives is crucial in this context. Are there any alternatives? Has the institution looked into this?
How were these alternatives assessed and weighed?

The basis on which an institution can process education data depends on the specific application of the use of that data. The institution determines which basis is most appropriate for each application.

Notes on the basis of consent

When processing on the basis of consent, it is important to consider the following:

- Has consent truly been "freely" given? In other words, without coercion or pressure. Given the relationship between the institution on the one hand and teaching staff/ students on the other, this remains a point of attention. If consent has not been freely given, that consent cannot serve as a basis.
- Informed consent must be given. This means that the data subject should have all the information needed to make an informed choice about whether or not to consent to the intended processing of personal data.
- Giving consent means that a data subject can always withdraw that consent, after which the relevant education data may no longer be used. Institutions must arrange this adequately.
- If some of the students give permission for the use of education data and some do not, there is a risk that inequality will arise between different groups of students. That touches on the aspect of equity.

An institution's choice for the basis of using education data and AI requires careful deliberation. It is important that an institution consults with the institution's privacy expert on the matter; this may be the privacy officer, the privacy lawyer or the data protection officer (DPO).

4.2.3 Due care

The broad rubric of due care refers to all obligations (incumbent on institutions) to ensure that education data is handled in a responsible manner.

Of these obligations, the **right to information** of data subjects is crucial. Institutions must inform data subjects in a clear manner about, among other things, which of their education data is used, for what purposes this is done, what the basis is for that use and what security measures the institution takes for the safe and careful use of education data. This information is best given at such time when the personal data will actually be used.

Informing children

When informing children (under the age of 16), the fact that they are children must be taken into account. This also means informing children in a way that they can understand. In other words, difficult words and long sentences should be avoided.

In addition, the following obligations apply:

Data minimalisation; only education data that is necessary to achieve the intended purpose may be processed. If it is not or no longer necessary to use directly identifying data, the data should be anonymised or pseudonymised as soon as possible.

Pseudonymisation is an important measure for the protection of personal data, and the GDPR explicitly mentions this measure. In general, it is advisable to pseudonymise personal data as much and as quickly as possible.

Privacy by default en privacy by design; the required technical and organisational security measures to protect personal data must be considered as early as possible when devising and developing new applications or technologies. This is called privacy by design. Furthermore, institutions should ensure that the standard settings of applications and systems are as privacy-friendly as possible. This is called privacy by default.

SURF has information available on privacy by design and privacy by default; see (only available in Dutch): surf.nl/privacy-by-design-en-privacy-by-default.

Accuracy; the education data that is processed must be correct and complete.

Data Protection Impact Assessment; a *Data Protection Impact Assessment* (DPIA) is an assessment of the risks to data protection in the event of intended use of education data. If a processing operation is likely to present a high risk to data subjects, a DPIA is obligatory. This requires the institution to analyse the nature, scope, context and purposes of the processing. A DPIA must be performed in the following cases:

- A systematic and extensive assessment of personal aspects of natural persons which is based on automated processing – including profiling – based on which decisions are made that have legal consequences for or otherwise significantly affect the person.
- If special personal data is processed on a large scale.
- If public areas are monitored systematically or on a large scale.

> This Reference Framework is based on the principle that in all cases in which education data is used with the help of (or in combination with) AI, this use will be qualified as high-risk according to the AI Act (see below in Chapter 5). In such cases, conducting a DPIA is always necessary.

The Dutch Data Protection Authority (*Autoriteit Persoonsgegevens*, AP) and the European privacy authorities (European Data Protection Board, EDPB) have also drawn up a list of criteria to assess whether there is a high risk.[6]
Although a DPIA can be carried out in a free format, it must include a number of obligatory sections.

Retention periods; education data may not be kept longer than necessary for the purpose for which this data is used. Insofar as necessary for historical or academic research purposes, education data may be stored for longer provided that appropriate technical and organisational safeguards are in place.

Information security; the educational institution must take appropriate technical and organisational measures (such as pseudonymisation and encryption of education data, guaranteeing the confidentiality, integrity, availability and resilience of the systems and regularly testing and evaluating the measures taken). These measures are based on a risk assessment of the processing so that there is appropriate security of education data.

> SURF has more information about this on:
> surf.nl/en/themes/cybersecurity

Rights of data subjects; data subjects have a number of essential rights regarding their education data. These are: the right of access (and, where applicable, also the right to copies of that education data), the right to rectification, the right to erasure, the right to restrict processing, the right to object to processing and the right to data portability.

[6] See the website of the AP and the EDPB Guidelines on Data Protection Impact Assessments (WP 248).

In addition, the data subject has the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him/her. There are a number of exceptions to this rule, for example if the data subject gives explicit consent.

In this Reference Framework, the main focus is the aspect of *solely* automated processing, for example in decisions about access and admission to education. Case law shows that solely automated processing applies if there is no meaningful human intervention by persons with sufficient competencies and powers to influence and even reverse decisions.

> ### Solely automated decision-making and children
> There is some ambiguity as to whether children (under 16) should not be subject to solely automated decision-making (including profiling) at all. Although the text of the relevant article in the GDPR does not seem to support such an absolutist position, European supervisors united in Working Group 29 (now the EDPB) apply the principle that children should not be subject to solely automated decision-making and therefore no appeal can be made to legal exceptions under the GDPR.
>
> To complicate matters further, the predecessor of the European Data Protection Board (EDPB), the so-called Working Group 29, does provide for the possibility of solely automated decision-making regarding children if, for example, this is to protect their well-being. In these cases, appropriate protection measures must be taken for children.
>
> Educational institutions are advised to ensure that there is no solely automated decision-making in the case of children. If an institution does wish to make use of this, it is advisable to contact the DPO, a privacy officer or a legal expert.

Sharing education data*: sharing education data with other institutions, internship companies, municipal and youth care institutions, municipalities, school attendance officers, RMC trajectory counsellors, benefits agencies and the Safety Domain.*

Many institutions work together with other institutions and internship companies. Internship companies play a role in the education of pupils and students to a greater or lesser extent. And, apart from the fact that they can generate education data themselves, they can also receive education data from institutions. Such sharing of data is considered processing under the GDPR and is permitted provided that the triad of: purpose – legal basis – due care is met. In the context of due care, the institutions and internship companies make agreements about the use, security and method of sharing education data.

In certain situations, the GDPR also prescribes a specific agreement. Consider, for instance, a processing agreement or an arrangement that regulates a number of matters when two or more parties are joint controllers.

In addition, institutions may have to deal with various government agencies, such as municipalities, youth care institutions and benefit agencies. Sharing education data is also possible in these cases, provided that the triad of: purpose (and purpose limitation) – legal basis – due care is met.

For vocational education and training schools (*mbo*), see also the following service document: *Sharing of personal data between education and care institutions with place of work, MBO Raad (in Dutch only)* Privacy | MBO Raad.

4.3 AI, generative AI and the GDPR

DThe GDPR applies to the use of education data. This is also the case if the education data is used with the help of AI, including generative AI. The principles of the GDPR, as described in this chapter, apply in full to the processing of education data using AI.

But where and how exactly do the GDPR principles apply to AI? Especially with generative AI, where personal data is used in different phases of the life cycle, there may be particularities and difficulties in each phase, including with regard to purpose and purpose limitation, legal bases, provision of information, rights of data subjects and data minimisation.

SURF discussion board

SURF has a discussion board that illustrates the role of the GDPR, including an explanation (in Dutch); see: AI and Privacy discussion board - Privacy Expertise Centrum.

Supervisors

The European Data Protection Board (EDPB) has issued an opinion that takes a closer look at various aspects of data protection in the context of AI models. See: edpb_opinion_202428_ai-models_en.pdf.

Furthermore, the French and UK supervisors – the CNIL and the ICO – have published recommendations/guidelines for the development phase of AI systems and how to mitigate data protection risks in AI projects See: AI: CNIL publishes its first recommendations on the development of artificial intelligence systems | CNIL en Guidance on AI and data protection | ICO.

The Belgian Data Protection Authority has an information brochure on systems and the GDPR (and the interaction between the GDPR and the AI Act in the context of the development of AI systems). See: informatiebrochure-over-artificiele-intelligentiesystemen-en-de-avg.pdf (in Dutch).

Although there is considerable debate and uncertainty about how the requirements of the GDPR are to be implemented, particularly in relation to generative AI, prohibiting the use of generative AI in institutions is not always an option This is not only due to the widespread, easy, and often free availability of generative AI, but also because it is recognised that it is beneficial for students and teaching staff to be able to experiment with and learn from generative AI.

Most institutions therefore primarily use guidelines, recommendations, and other rules to foster the responsible use of generative AI in education. The core of these guidelines is that they set conditions for the use of generative AI. Consider, for example, the following:
(i) For what purposes may generative AI be used (and for what purposes may it not)?
(ii) What are the conditions if generative AI is permitted (for instance, no personal data, students remain responsible for their own work, etc.)?

Generative AI in education

A good example of such a guideline is that of Amsterdam University of Applied Sciences, *Generatieve AI in onderwijs: regels en adviezen*, Hogeschool Amsterdam, April 2024 (in Dutch only), see: generatieve-ai-in-onderwijs-regels-en-adviezen-extern.pdf

5. Legal framework: AI Act

5.1 Introduction

On 2 August 2024, the AI Act came into force. The AI Act is part of a broader European regulatory framework on product regulation which, at its core, focuses on safe and reliable products.

The principal aim of the AI Act is to ensure that AI systems are used responsibly and safely in both the private and public sectors. The AI Act applies if AI systems are commercialised or used in the EU or if the output of the AI system is used in the EU.

> **When does the AI Act not apply?**
> The AI Act does *not* apply if it only concerns personal and non-professional use of a generic AI system.
>
> This means that the AI Act does not apply to students who generate texts for essays, create summaries, or seek information and clarification using ChatGPT or similar tools for their own use.
>
> Similarly, the AI Act also does not apply to teaching staff who experiment with AI (including generative AI) for their own professional development. But *it is different* if a lecturer or teacher (and therefore their institution) uses generative AI, for example to select educational resources, formulate learning objectives, make assignments or design a lesson plan. In such cases, it is important to consider how that specific application is classified under the AI Act.

The AI Act looks at the *application of an AI system*; what is the AI system used for? The AI Act assumes a risk-based approach; in other words, what are the risks for data subjects in terms of health, safety and fundamental rights (such as the right to privacy and not to be discriminated against) when applying an AI system? Based on these risks, the AI Act prohibits certain applications and classifies others as high-risk. In addition, the AI Act covers AI systems that have transparency risks.

> Note: AI systems that are not high-risk according to the AI Act and are not otherwise covered by the AI Act can still pose risks. Examples include privacy violations, inadequate information security or unjustified favouritism. In these cases, too, it is important to properly identify the risks and take measures to mitigate these risks if necessary.

The AI Act sets out rules and obligations for AI systems. What these rules and obligations are depends on the risk classification of the AI system and the institution's role with regard to those AI systems. An institution can be a provider or a deployer[7].
The premise is that most obligations apply to providers of high-risk AI systems.

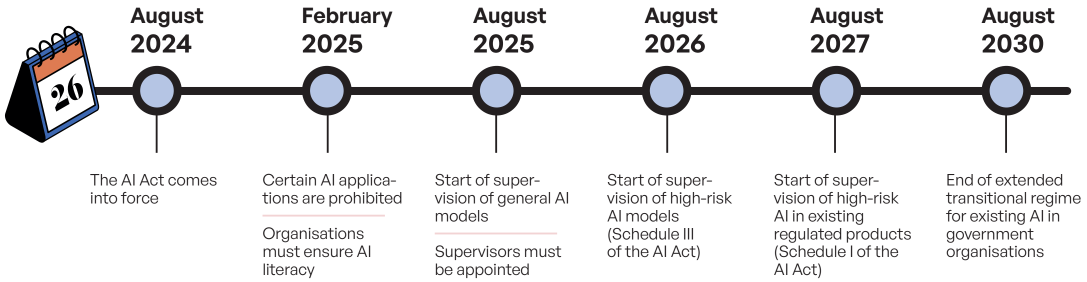
A **provider** of an AI system is a natural or legal person who develops the system or has it developed, commercialises it or puts it into use and does so under their own name or trademark.

A **deployer** of an AI system is a legal or natural person who uses the system under their own responsibility. Note: the deployer is not the same entity as the person affected by the AI system.

> ### The role of the institution under the GDPR and the AI Act
> Both the GDPR and the AI Act consider the role that a party, such as an institution, plays in determining the nature and scope of its obligations. These roles are determined independently by the GDPR and the AI Act.
>
> An institution will usually qualify as a data controller under the GDPR. The role of an institution in the AI Act (provider or deployer) is less clear. It is advisable to carefully determine which role the institution fulfils in relation to specific uses of education data and AI. The privacy lawyer, privacy officer, AI Compliance Officer and/or the DPO can provide assistance with this.

AI Act timelines
The obligations of the AI Act will be implemented in phases. The Dutch Data Protection Authority has provided a useful overview of the relevant timelines and associated obligations.

[7] The AI Act also defines the roles of importer, distributor and authorised representative. This Reference Framework assumes that educational institutions do not fulfil these roles. These roles demonstrate that the AI Act is also a regulation that is consistent with the broader legislative landscape of product regulation.


26
August 2024
The AI Act comes into force
February 2025
Certain AI applications are prohibited
Organisations must ensure AI literacy
August 2025
Start of supervision of general AI models
Supervisors must be appointed
August 2026
Start of supervision of high-risk AI models (Schedule III of the AI Act)
August 2027
Start of supervision of high-risk AI in existing regulated products (Schedule I of the AI Act)
August 2030
End of extended transitional regime for existing AI in government organisations

The AI Act and personal data
For the purposes of the AI Act, it is not relevant whether the education data contains personal data. The AI Act may apply to the use of education data, even if this does not include personal data.

> ### Relationship between GDPR and AI Act
> When using education data and AI, both the GDPR and the AI Act may apply concurrently. This is the case if education data qualifies as personal data. Note that one of these pieces of legislation does not replace the other.
>
> This means that compliance with the GDPR does not automatically mean compliance with the AI Act, and vice versa.

5.2 What does the AI Act regulate for education?

This section explores the implications of the AI Act for education and outlines when specific provisions come into effect.

From 2 February 2025: Prohibited AI in education

The AI Act prohibits certain applications of AI systems because the application entails unacceptable risks for data subjects. These include systems that exploit vulnerabilities of people on the basis of age, disability or social circumstances. AI systems intended for emotion recognition in the workplace and in education are prohibited, unless this is done for medical or safety reasons.

Section 5 of the AI Act lists the prohibited AI applications. The European Commission is entitled to amend and supplement the list of prohibited AI applications; this does not require an extensive legislative process.

Prohibited AI applications are not expected to be used in education, or at least only very rarely. It is important that institutions identify whether they use prohibited AI, regardless of whether the institution is a provider or a deployer. If they find that this is the case, they should stop this use.

From 2 February 2025: AI literacy

The AI Act requires providers and deployers to take measures to ensure a sufficient level of AI literacy of their employees who are involved in the operation and use of AI systems. In itself, the AI literacy obligation does not automatically apply to *all* employees. However, it is advisable to inform and keep all employees generally informed about the advantages and disadvantages of using AI. This involves knowledge, understanding and skills regarding the responsible use of AI systems and awareness of the opportunities, risks and possible harmful consequences of using AI systems.

How extensive that AI literacy must be depends, among other things, on the context in which the AI system is used. The risk classification of an AI system determines the degree of AI literacy required.

Steps in promoting AI literacy

- Find out which employees are or will be involved in the use of AI.
- Make AI literacy (but also data literacy) part of existing training and onboarding programmes.
- Organise AI training (e-learning) for the Executive Board, senior management and participation councils (of students and employees).
- Organise regular workshops.

From 2 August 2025: AI models

The assumption of the European legislator was that an AI system is developed for a specific application. However, the launch of ChatGPT in November 2022 proved that AI can be used for many different purposes. This was inconsistent with the idea of one AI system for one particular application. The European legislator therefore wanted to provide for this by introducing specific rules for providers of AI *models* for general purposes. The nature and scope of these rules depend on whether an AI model entails systemic risks or not.

According to the AI Act, general-purpose AI models are – in short – models that can automatically create content (text, audio, video or a combination of these) based on user requests (prompts). Examples include GPT4 by OpenAI, Google Gemini, Meta LLama or Claude Sonnet. These models are trained with large amounts of data based on machine learning, such as deep neural networks. Big Tech companies, in particular, are highly active in developing these models and they are increasingly integrating them into their own products and services (such as Microsoft Copilot).

For a proper understanding of these developments, it is important to realise that these AI *models* are not AI applications in themselves; a model is merely a part of a system. In the AI Act, these AI models have their own risk classification system, depending on whether or not the model has systemic risks.

Providers of AI models for general purposes must meet the following requirements:

- Technische documentatie leveren
- Gebruiksinstructies leveren
- Voldoen aan de auteursrechtrichtlijn
- Samenvatting publiceren over de content die voor training is gebruikt

Providers who make their AI model publicly available under a free and open licence only need to comply with the Copyright Directive and publish a summary about the content used for training, unless it concerns an AI model with systemic risks.

An AI model with systemic risk is one that meets the following criteria:

- It has high impact capabilities, evaluated on the basis of appropriate technical tools and methods – including indicators and benchmarks.
- In the case of a GPAI model, it has capabilities or effects equivalent to those in the previous point pursuant to an official decision of the European Commission or a qualified warning by the scientific panel.

All providers of AI models with systemic risks must:

- Conduct evaluations of these models, including conducting and documenting contradiction tests to identify and mitigate systemic risks.
- Assess and mitigate potential systemic risks, including the sources of these risks.
- Track, document and report serious incidents and possible corrective measures to the AI Office and relevant national competent authorities without undue delay.
- Ensure an adequate level of cyber security.

From 2 August 2026: High-risk AI in education

The most important provisions of the AI Act for institutions arguably pertain to those AI systems that qualify as high-risk under the AI Act.

Part 3 of Annex III of the AI Act lists high-risk applications that specifically apply to education. There are also other high-risk applications that potentially apply to education. The AI systems listed in Section 1 relate to:

- remote biometric identification, unless the system is used for verification only;
- systems for biometric categorisation based on sensitive properties or characteristics;
- emotion recognition systems.

Since these applications are irrelevant for the use of education data, they will not be taken into consideration here.

What applications qualify as high-risk in education? For now, these are only systems that have the following purposes:

1. Granting access, admission and assignment to educational and vocational training institutions.
2. Evaluating learning outcomes, including for guiding the learning process.
3. Assessing the appropriate educational level.
4. Monitoring and detecting unauthorised behaviour during assessments and exams.

The AI Act does not provide much more by way of clarification; for example, no use cases are included. It is expected that this clarification will eventually be provided by the European Commission and/or national supervisors. Incidentally, the European Commission can change or add new applications to the list at a later date. In other words, it is a dynamic list.

In light of these four categories, the following applications may be high-risk:

- screening applications
- ranking of candidates
- assessment (evaluation?) of exams, essays, and so on
- school recommendations and recommendations on continuing the programme
- online proctoring
- fraud detection

It is notable that only these AI applications qualify as high-risk. This means that other applications that – for whatever reason – could also pose major risks to data subjects are not high-risk according to the AI Act.

Application of the Reference Framework

Institutions are not legally obliged to make a risk assessment for AI applications that are not high-risk. The Reference Framework is based on the principle that institutions apply the Reference Framework even if the AI application does not qualify as a high-risk AI under the AI Act or is otherwise exempted from the AI Act!

Although the AI Act does not provide any further clarification on high-risk applications, the AI Act does list some exceptions. The reason for this is that the AI system is not high-risk if it does not substantially influence a decision due to the fact that the system is intended to:

- perform limited procedural tasks;
- improve the outcome of a previously completed human activity;
- identify decision-making patterns or deviations from them;
- carry out preparatory tasks for assessments relevant to high-risk areas of application.

However, these exceptions cannot be relied on if the AI system is profiling.

What is profiling?

The GDPR defines profiling as follows: *"any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular for the purpose of analysing or predicting performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements."*

From 2 August 2026: AI applications in education with transparency risks

In addition to the categories of prohibited and high-risk, the AI Act also includes the category of AI systems with transparency risks. These are, for example, AI systems such as chatbots and virtual assistants that interact directly with data subjects (natural persons including students and teaching staff). But also systems that generate audio, video, image or text content, deep fakes and applications that recognise emotions or biometric characteristics.

What are the obligations for these AI systems?

Institutions that use these applications have notification obligations to ensure that the data subjects are informed and aware, when first using an AI system, that it is an AI system that is generating the output and not a natural person. The institution must always identify deep fakes as such.

Finally, there are AI systems for which the AI Act does not provide further rules, for example AI applications like spam filters and spelling checkers.

5.3 What must institutions regulate for the use of high-risk AI?

Most of the obligations under the AI Act relate to providers and deployers of high-risk AI. It is mainly the providers of these AI systems that have to meet many of the legal requirements.

Providers of high-risk AI systems must meet certain requirements to ensure that their AI systems are reliable, transparent and auditable. To this end, they must:

- Set up a risk management system during the entire life cycle of the AI system.
- Implement data governance and ensure that the training, validation and testing datasets are relevant, sufficiently representative and – as far as possible – complete and error-free in line with their intended purpose.
- Prepare technical documentation to demonstrate compliance and provide authorities with the information needed to assess that compliance.
- Design their AI system to automatically record events relevant to identifying risks at the national level as well as significant changes during the system's life cycle.
- Give instructions for use to the deployer to ensure that they meet the requirements.

> The Knowledge Centre Data & Society (KCDS) has developed a working template for user instructions; T.Gils and W. Ooms (Knowledge Centre Data & Society), "Instructions for use (IFU) for high-risk AI systems under the EU AI Act – working template", October 2024. See: KCDS-Template-instructions-for-use_PDF.pdf

- Design their AI system so that deployers can implement human in the loop.
- Design their AI system in such a way that it achieves the right levels of accuracy, robustness and cyber security.
- Establish a quality management system to ensure compliance.
- Ensure that a valid EU declaration conformity is present.

Deployers are also subject to specific obligations regarding the use of high-risk AI. They must:

- Take appropriate technical and organisational measures so that the use of the relevant AI is in accordance with the provider's instructions.
- Monitor the operation of the AI system based on the instructions for use and, in the event of serious incidents, stop the system and notify the providers. Log files of use of the relevant AI must be kept for at least 6 months.
- Ensure sufficient relevant input data (insofar as they have control over this).
- Implement human in the loop.
- Inform data subjects that they have been subjected to an AI system with a high-risk application. Data subjects also have the right to a clear and meaningful explanation of the role of the high-risk AI system in the decision-making process when high-risk decisions have been made.
- Comply with certain reporting and registration obligations towards supervisors.
- Government organisations – including educational institutions – must perform an assessment of the potential impact on fundamental rights when using the high-risk AI system. Such an assessment is called a Fundamental Rights Impact Assessment (FRIA).

> ### The FRIA
> - These impact assessments are about identifying specific risks in the area of human rights and possible control measures.
> - A FRIA performed by a provider may be used (although a provider is not obliged to do perform a FRIA); in that case, the user remains responsible.
> - A FRIA can be part of a DPIA and can be carried out simultaneously with its execution.
> - The findings of the FRIA must be reported to the supervisory authority (AP).
> - The government has developed a template for performing a FRIA. See: Impact Assessment Fundamental Rights and Algorithms. The AI Office of the European Commission intends to develop a special template/tool for this.

5.4 From deployer to provider; when is this the case and what are the consequences?

The AI Act makes a clear distinction between the roles of provider and deployer. The AI Act also contains a provision that can be of great importance in practice, namely that under certain circumstances an institution that is merely a deployer becomes a provider. Because this shift in role has major implications for the institution, it is addressed in this Reference Framework.

Shift in role
Wanneer is sprake van een dergelijke transformatie (rolverandering)? Dat is het geval als de instelling als gebruiksverantwoordelijke op enig moment:

1. Puts their own name or trademark on an existing high-risk AI system.
2. Makes a substantial change to an existing high-risk AI system in such a manner that it remains a high-risk AI system.
3. Adapts the intended purpose (as envisaged by the provider) of a non-high-risk AI system so that it becomes a high-risk AI system.

> A number of aspects require clarification:
> 1. How do you know when a change is a substantial change?
> 2. What is the intended purpose? That depends greatly on how the provider of the AI system has described this, for example in the guide.
> 3. Can a change of intended purpose carried out by just one employee lead to a shift in role?

Consequences of a shift in role
What are the consequences of a shift in role? The most important consequence is that the institution is then considered a provider who must comply with all the obligations incumbent on a provider of a high-risk AI system.

This is a considerable burden for the educational institution (see above). Moreover, the question arises whether the institution can meet all these obligations. For example, the institution may not have the technical documentation relating to that new intended purpose. While it is true that the original provider has a duty to make information available and to provide the institution with technical access and information, it is unlikely that it has that information for that new purpose. After all, the institution will have consciously (or unconsciously) determined the new purpose.

Please note

A shift in role has major consequences for an institution and the Reference Framework therefore recommends that institutions:

1. Establish a clear process for the use and changes of the use of AI systems so that a well-considered and informed assessment can be made. That way, an institution can avoid inadvertently and unintentionally shifting its role from deployer to provider.
2. In the agreement with the provider, clearly agree on the intended use of the AI system. Which use is covered? Which is not?
3. In the agreement with the provider, also agree on what support and assistance they must provide if the deployer changes roles.

5.5 Procurement and AI

Educational institutions can use AI systems developed by external parties. These can be stand-alone AI systems as well as systems that are part of existing services that institutions purchase. Before starting a procurement process, the institution must consider whether using AI when making use of education data in a specific case is in line with the values of the institution, regardless of which supplier is chosen.

If this consideration results in the decision to procure an AI system, it is advisable (in view of the requirements of the GDPR and the AI Act) for institutions to include the institution's requirements and wishes regarding responsible use of education data and AI in the procurement/tendering procedures at an early stage, irrespective of whether the AI system is high-risk or not.

In the case of high-risk AI systems, an institution can suffice by stipulating in the contract that the high-risk AI system (to be provided by the supplier) meets the requirements of the AI Act (and the GDPR) and that it is designed in such a way that the institution can use the AI system for its intended purposes. Moreover, the provider of a high-risk AI system is already required to provide technical documentation and user manuals.

Nevertheless, the Reference Framework considers it desirable to engage with suppliers so as to acquire a responsible product or service, irrespective of whether it is a high-risk AI application under the AI Act. This has consequences for the contracts that institutions conclude with suppliers.

Check / questionnaire for suppliers

The University of Amsterdam (UvA) has drawn up a useful check/questionnaire for the procurement of AI. The questionnaire consists of a 'pre-flight check' (questions that must be answered by the institution itself), 'required checks' and 'optional checks'. See: ai-checklist-vu-uva-taskforce.pdf

Contractual clauses

The European Commission has drawn up standard contractual clauses for contracts between deployers (such as institutions) and providers of AI systems. There are standard clauses for high-risk AI systems and for non-high-risk AI systems.
See: public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/eu-model-contractual-ai-clauses-pilot-procurements-ai

5.6 Preparing for the AI Act

Although many provisions will be introduced in phases, it is important that institutions prepare for this legislation now. This will not be straightforward, as many obligations in the AI Act will in the coming period be elaborated with various standards and further clarification by the European Commission and/or the AI Office.

In the Reference Framework, preparing for the AI Act is deemed to consist of the following three steps:

1. **Inventariseer** binnen de instelling welke AI-systemen worden gebruikt en leg dit goed vast, bijvoorbeeld in een AI- of algoritmeregister. Dat kunnen zelfstandige AI-systemen zijn, maar ook AI-systemen die onderdeel zijn van al bestaande diensten. Kijk ook of deze AI-systemen gebruikmaken van persoonsgegevens.

2. **Classify** these AI systems. Are they high-risk AI systems as defined in the AI Act? What are the potential risks for data subjects associated with the use of these AI systems? Has a risk analysis or a DPIA been carried out?

3. **Implement** the relevant provisions of the AI Act by:
   - Establishing a process/procedure to continuously monitor and control the risks of AI systems. The institution can align this with existing risk processes, such as the DPIA procedure, and complement them.
   - Setting up an AI and data literacy programme.
   - Institutions may need to establish human in the loop in the use of AI systems.
   - Adapting information provision for data subjects.
   - Supplementing procurement and tendering procedures with requirements/wishes regarding AI.
   - Adjusting model contracts, SLAs and general terms and conditions.
   - Supplementing the procedure for questions from and complaints by data subjects.

5.7 Other legal doctrines in the use of education data and AI

In addition to the GDPR and the AI Act, other legal doctrines may be relevant in the use of AI, such as provisions on cyber security, liabilities, copyright (especially in the case of generative AI) and the protection of consumers. Reference Framework 2.0 does not discuss this in detail.

6. Clarity on responsibilities

The basis for ongoing responsible use of education data and AI requires clarity on responsibilities in all choices regarding whether or not to use education data and AI (including changing the type of use or stopping this use).

There should be clarity about **who** does **what** and **when** in the educational institution. This includes topics such as accountability, stakeholder involvement, monitoring and tracking outcomes, AI literacy, implementing FRIA, monitoring correct assignment of roles in the use of AI, explainable AI (XAI), requests from data subjects, and so on.

It may be necessary to expand or modify existing roles and positions, or even create new ones.

Pooling knowledge and expertise across multiple institutions

Many institutions have limited staffing and cannot organise the necessary expertise in their own institution. Institutions are free to share specific expertise with each other. Organisations such as the MBO Council could facilitate this.

Incidentally, the AI Act does not mandate the appointment of AI-specific roles (unlike the GDPR, which requires the appointment of a data protection officer under certain conditions).

AI governance officer

Utrecht University and Tilburg University have jointly appointed an AI governance officer; this is a new position/role. The AI governance officer supports, among other things, the development of a holistic approach to AI.

AI compliance officer

Partly as a result of the AI Act, the position of AI compliance officer is on the rise. Specialised consultancy firms organise programmes/courses for this position.

A multidisciplinary approach remains crucial throughout the life cycle of education data and AI use. This approach involves various stakeholders and experts, including educators, educational support staff, information managers, IT professionals, data stewards/engineers, architects, information security specialists, privacy/legal advisers, procurement officers, contract managers and data protection officers (DPOs). Involve staff and students as well as their representatives in this multidisciplinary approach.

RACI-model

An institution can lay down the tasks and responsibilities in a so-called RACI model. This outlines the roles and responsibilities of the different officers:
R – responsible: the person responsible for executing the task and who reports to the person who is accountable.
A – accountable: the person who has ultimate responsibility for the result and approves it.
C – consulted: the roles/persons who have to be consulted beforehand in performing a task and therefore co-determine the result.
I – informed: the roles/persons who receive information about the decisions, progress and results afterwards.

Through this approach, the values and legal frameworks can be discussed in a coherent way so that the institution can make informed choices about the use of education data and AI. Actively organise dissenting opinions in this approach. Involving sceptics and critics provides new perspectives and different insights, which ultimately enhances the decision-making process around AI.

Ethics committees

Does your institution have an ethics committee for reviewing research proposals? If so, this committee can play a role in organising the aforementioned multidisciplinary approach and providing a countervoice.

Education institutions that do not have an ethics committee can set up a broader committee, ad hoc if necessary.

7. Where and how to start?

After taking a closer look at the values, legal frameworks and responsibilities, the question remains: where and how to start? This Reference Framework proposes the following steps.

1. Develop a business case

Developing a business case helps the educational institution to fully understand the purpose, advantages and disadvantages of using education data and AI as well as the values at stake. It provides insight into the measures that the institution must take for the responsible use of education data and AI. An example of a business case is included as an annex to the Reference Framework.

2. Perform a GDPR and AI Act check

A GDPR and AI Act check makes it clear what the institution is *obliged* to arrange.

3. Involve stakeholders in a multidisciplinary approach

It is essential to involve stakeholders in the discussion on the use of education data and AI at an early stage. This way, the institution can gain new insights when elaborating the values, the opportunities and the risks of using education data and AI. Various tools are available to foster engagement and dialogue about values:

1. The Dilemma Game. This game was developed in the context of Framework 1.0 and has recently been updated. The game outlines the various dilemmas surrounding the use of education data and AI. Players are invited to come up with solutions to these dilemmas in a collaborative dialogue.

2. SURF has an overview of tools and resources that institutions can use in this dialogue, see: Aan de slag met Publieke Waarden: een overzicht van praktische hulpmiddelen | SURF Communities (in Dutch only).

4. Establish (or supplement) policies, guidelines and procedures

Establish (or supplement) policies, guidelines and procedures for various aspects of the responsible use of education data and AI. This concerns transparency, stakeholder involvement, assessments, data quality and data governance, logging and monitoring of use and adjustments of use based on monitoring and lessons learned, incident management, human intervention, supervision, handling of questions, concerns and complaints, among others. Use the PDCA cycle to periodically adjust policies, guidelines and procedures.

5. Promote data literacy and AI literacy

Ensure that education professionals understand why education data and AI are used, what the intended purposes are, how the algorithms and AI work, what the limitations are, what the potential risks of using education data and AI are when it comes to fundamental human rights as well as possible ways to mitigate or eliminate these risks.

> A useful tool for this is The AI Maturity in Education Scan (AIMES) by VU Amsterdam. Educators can use this tool to assess their current level of AI literacy and enhance it accordingly.

ANNEX – Business case

The questionnaire below has been compiled for the purposes of this business case. This list of questions is not exhaustive.

General: purpose and use of education data and AI

- What do you want to do with the education data? What do you want to achieve? For whom do you want to achieve this? Are these goals clear and measurable?
- How can AI help achieve the goals? Are other alternative options present and included in the assessment?
- What are the expectations in terms of performance and accuracy of the AI?
- Is the AI already available or does the institution have to develop it itself or purchase it from a third party?
- What education data do you need to achieve the goal? What is the source of this education data?
- Is there a legal basis for the use of this education data? Have the data subjects been informed about this?
- Do you have sufficient representative education data to achieve the goals?
- How is the education data for the specific use secured?
- How is human oversight (human in the loop) arranged?
- What other values are at stake? And how do you interpret this?

Risks of using education data and AI

- What foreseen and unforeseen disadvantages are associated with the use of education data and AI? Are there possible specific risks for certain groups of data subjects? If so, specify these risks.
- How can the use of education data and AI be prevented from leading to (unintentional) discrimination or exclusion?
- How can the use of education data and AI be prevented from leading to inequity?
- What are the possible causes of discrimination, exclusion and/or inequity?
- What are the potential consequences if a risk occurs? Can any disadvantages be remedied quickly and easily?
- How will the education data and the operation of the AI remain safe and reliable?
- How is the use of education data and AI tracked and its operation monitored? What exactly is being looked at? And how often and by whom is monitoring carried out?
- How will stakeholders be involved in the evaluation of the use of education data and AI?

- Which persons or departments can data subjects address their questions, comments and complaints to? How are they handled?
- Have data subjects, students (and prospective students), alumni, current and former employees or other stakeholders been informed about the use of their education data with AI? How were they informed?
- Do data subjects receive meaningful explanations when important decisions are made based on the use of AI? How is that guaranteed?

Credits

Project team

- Bram Enning (Npuls)
- Dominique Campman (Npuls)
- Menno Jehee (Npuls)
- George Wurpel (MSG Strategies)
- Mariken Betsema (MSG Strategies)
- Niek Reijmers (MSG Strategies)

Support group

- Irene Eegdeman (Windesheim University of Applied Sciences)
- Duuk Baten (SURF)
- Marlon Domingus (Erasmus University Rotterdam)
- Nigel Steenis (ROC Mondriaan)
- Frank van Tatenhove (University of Amsterdam)
- Floortje Jorna (SURF)

Focus group

- Frank van Tatenhove (University of Amsterdam)
- Annemiek Mandemaker (Aeres)
- Theo Nelissen (Avans University of Applied Sciences)
- Jan van den Berg (Amsterdam University of Applied Sciences)
- Pascal Tielkens (ROC Nijmegen)
- Annie Slotboom-Memelink (Graafschap College)
- Abdullah Yerebakan (Utrecht University)
- Lex Freund (Rotterdam University of Applied Sciences)
- Marlon Domingus (Erasmus University Rotterdam)
- Duuk Baten (SURF)
- Marlies van Hal (Koning Willem 1 College)
- Tri Hartanto (ROC Mondriaan)
- Claartje Uitterhoeve (Graafschap College)
- Shane Bozelie (Deltion College)
- Iglika Vassileva-van der Heiden (Radboud University Nijmegen)
- Quinta Dijk (SURF)
- Ineke Stoop (Tilburg University)
- Rowin Smits (ROC Mondriaan)
- Tom van Munster (Intercity Student Consultation)

- Harry van de Vis (Albeda College)
- Yorick van der Heijden (Intercity Student Consultation)
- Barbara Gerretsen (University of Amsterdam)
- Daphne Peters (JobMBO)
- Charlotte Vonk Noordegraaf (Deltion College)
- Jan Tjeerd Groenewoud (University of Groningen)
- Irene Eegdeman (Windesheim University of Applied Sciences)
- Nigel Steenis (ROC Mondriaan)
- Mark Trimpe (VISTA College)
- Dominique Campman (Npuls)
- JaapJan Vroom (SURF)
- Simon Jong (Deltion College)
- Wladimir Mufty (SURF)

Review DPO

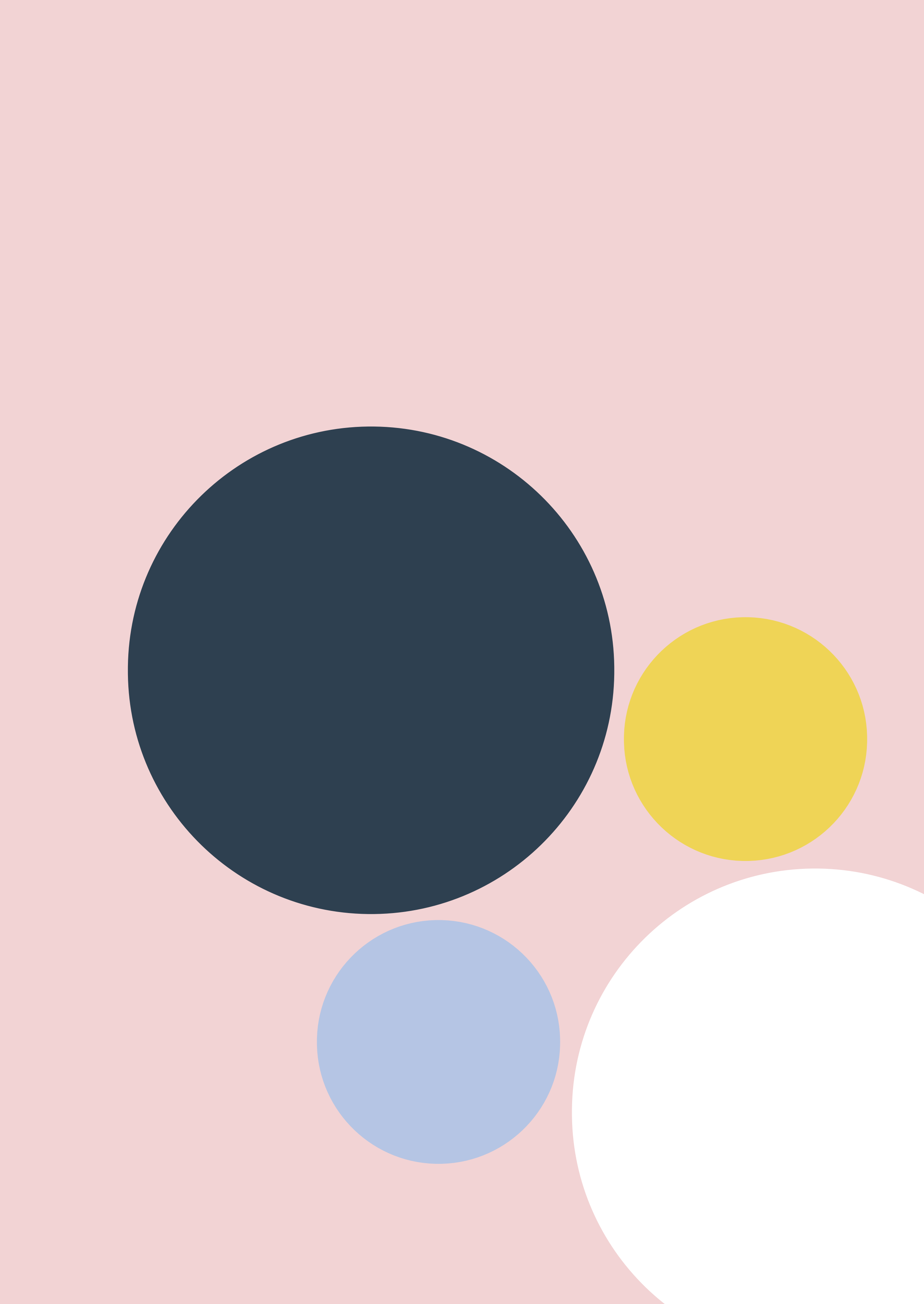
- Peter Vermeijs (MBO Raad)
- Moswa Herregodts (Tilburg University)
- Raoul Winkens (Maastricht University)
- Esther van der Ent (HAN University of Applied Sciences)
- Ingeborg ten Oever (Breda University of Applied Sciences)
- R. Kronieger (Iselinge University of Applied Sciences)
- Silvia van Dijk (The Hague University of Applied Sciences)
- Kees-Jan van Klaveren (Rotterdam University of Applied Sciences)
- Machiel van Velzen (ROC Midden)

Interviewees

- Andres Steijaert (SURF)
- Annie Slotboom (Graafschap College)
- Duuk Baten (SURF)
- Frank Benneker (University of Amsterdam)
- Frank van Tatenhove (University of Amsterdam)
- Han Werts (Radboud University Nijmegen)
- Irene Eegdeman (Windesheim University of Applied Sciences)
- Jaap Jan Vroom (SURF)
- Justian Knobhout (Utrecht University of Applied Sciences)
- Karianne Vermaas (SURF)
- Mark Trimpe (VISTA College)
- Marlon Domingus (Erasmus University Rotterdam)
- Niek Kersten (Graafschap College)
- Yvette Roman (Utrecht University of Applied Sciences)

Survey

- Annie Slotboom (Graafschap College)
- Barbara Gerretsen (University of Amsterdam)
- Elisa Jochims (Graafschap College)
- Frank Benneker (University of Amsterdam)
- Ineke Stoop (Tilburg University)
- Irene Eegdeman (Windesheim University of Applied Sciences)
- Jan Tjeerd Groenewoud (University of Groningen)
- Jan van den Berg (Amsterdam University of Applied Sciences)
- Jorrit Arntzen (Deltion College)
- Kim Schildkamp (University of Twente)
- Lex Freund (Rotterdam University of Applied Sciences)
- Marco van Leeuwen (Breda University of Applied Sciences)
- Marlon Domingus (Erasmus University Rotterdam)
- Memon Boukiour (The Hague University of Applied Sciences)
- Raoul Winkens (Maastricht University)
- Ronald Sarelse (Radboud University Nijmegen)
- Theo Bakker (The Hague University of Applied Sciences)
- Wim Siemann (Albeda College)
- Alan Berg (University of Amsterdam)



Moving
education.