![Microsoft](Microsoft logo)

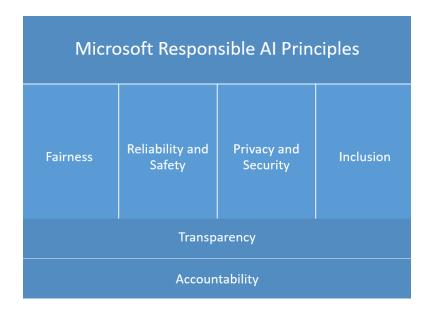# Open Education Analytics Data and AI Use Case

Template Published: February 2021

Introduction

This Use Case Template is intended for education systems and schools to use as they define, plan, and govern how their own data will be used for specific learning analytics and AI purposes. The template is informed by Microsoft's Responsible AI Principles and can help ensure that data and AI are used ethically and transparently.

It is recommended that a Use Case be documented for each specific learning analytics or AI purpose that will be developed using an education system's data estate.

Feedback on this template is welcome. Please send comments to openeduanaltyics@microsoft.com.

1) The Use Case Problem

The answer to these questions will inform decisions for the use case development through the rest of this document.

Defining the Problem: What problem does this use case seek to solve?

In a few sentences, describe:
- *The education or learning problem that needs a solution with this use case.*
- *The value the learning analytics or AI system is intended to deliver.*
- *The primary end user or users of the learning analytics or AI system.*
- *How the end user or users completes this task today.*
- *How this system can improve on today's solution.*

How does this use case align with the needs of learners in the education system?

Briefly describe how this use case will support learners or the learning process.

What is the overall context of this use case?

What is the background for this use case? How did it come to be defined? Who are the teams involved in developing it?

High risk uses to avoid:

Learning analytics and AI in education may be misused intentionally or unintentionally for purposes that introduce potential harm or unmitigated risk. It is important to consider what these high-risk uses might be so that they can be excluded or prevented. Provide a short list of the possible ways this use case could cause harm.

2) The Use Case Stakeholders

Considering potential harms and desired benefits of a system requires the consideration of important stakeholders. Stakeholders typically include the people who are responsible for, will use, or will be affected by the system. Stakeholders are defined by their role: their duties, contextual identity, or circumstances in relationship to the system. The table below may assist in identifying stakeholder groups in education.

- ***Direct stakeholders*** *interact with a data system directly. They include primary users, system owners, and even system maintenance staff.*
- ***Indirect stakeholders*** *are affected by the system but, unlike direct stakeholders, they do not have a role that requires them to use or maintain the system. Indirect stakeholders can include groups who may be affected by the downstream effects of the system.*
- ***Malicious actors*** *include hackers and others who may intentionally misuse the system. Considering malicious actors is important to supporting safe and reliable systems.*

Who are the stakeholder groups for this use case, and how are they involved in its development?

Stakeholder Group	Relationship to Use Case	Involvement in Use Case
Students	*For example, will students be providing data? Receiving recommendations?*	*For example, have students been asked if providing the data is acceptable? Or how they would like to receive recommendations?*
Parents or Guardians		
Educators (Faculty or Teachers)		
School or Institution Leaders		
System Leaders		
Researchers		

Please provide descriptions of what outputs each stakeholder group will receive from this use case, such as models, dashboards or other resources that will be created.

Stakeholder Group	Outputs
Students	*For example, will students see a personal dashboard?*
Parents or Guardians	
Educators (Faculty or Teachers)	
School or Institution Leaders	
System Leaders	
Researchers	

3) Datasets Description

What data will be used for this use case?

Please describe the datasets that are needed for this use case. What data exists that will solve the problems described above? What new data is needed? What is the process for new data generation or collection?

Dataset Name	Permission from Data Source Needed?	Storage Location	Classification
1.			
2.			
3.			
4.			

Data Governance.

See this resource for guidance on data governance. This course is also available.

Developing Classifications for Datasets.

See this resource for guidance on setting up data classifications. Data classification is often used to determine if anonymization or encryption is needed over the data.

What are the constraints of these datasets for this specific use case?

Please describe the limitations of these datasets. For example, are the datasets missing data for certain student populations? Is there bias in the data collection method?

Dataset Name	Constraints
1.	
2.	
3.	
4.	

4) Responsible AI Principles Applied

In these next sections, please answer the questions under each of the headings describing how responsible AI principles will be applied to this use case. Please write approximately a paragraph of text on each responsible AI principle and describe the planned steps to address each of these issues. Follow any other prompts to describe your plans in each of the sections.

Fairness Principle

AI systems should treat all people fairly

[Video](#) on Fairness Principle.

Who is most likely to be at risk of experiencing harms from this use case? For example, would any groups (immigrants, rural students) or subpopulations (gender, language group) face adverse consequences from the AI?

Planned Mitigations:

Are these groups and subpopulations clearly labelled in the dataset?

Reliability and Safety Principle

AI systems should perform reliably and safely

[Video](#) on Reliability and Safety Principle.

What are possible risks faced by learners or educators from the analytics of this use case?

1) Risk 1:

2) Risk 2:

3) Risk 3:

Planned Mitigations:

Transparency Principle

AI systems should be understandable

[Video](#) on Transparency Principle

What steps will the analytics or AI process include?

Describe analytics goals, definitions, design and method choices, and any assumptions made to provide documentation for traceability and communication.

Who will develop the analytics or models?

How will the limitations of the analytics or AI model be communicated to stakeholders and users?

What means will be built into the system for correction and model feedback by those who provide data and who use its outputs?

Privacy and Security

AI systems should be secure and respect privacy

[Video](#) on Privacy and Security principle

How will access to data be secured and protected in the data environment?

Does the dataset contain any personally identifiable information (PII) and how will that data be protected and governed?

Will data from young persons (e.g., "minors") be included in the dataset?

Accountability

People should be accountable for AI systems

[Video](#) on Accountability principle

Who is responsible for reviewing the Use Case documentation and ensuring that the implementation meets responsible AI principles?

How will stakeholders and end users be trained on the appropriate use of the system?

What is capacity of stakeholders to understand and make decisions over the data insights produced by the system? If capacities need to be developed to ensure the appropriate use of the data insights, how will that training be provided?

How will the analytics or AI system be monitored over time to ensure analytics and prediction performance reliability? Who will be responsible for this?

Inclusion

AI systems should empower everyone and engage people

[Video](#) on Inclusion principle

How does data collection ensure that data inputs are provided by all relevant populations, including diverse or traditionally marginalized groups?

How will the analytics or AI outputs from the system be provided to all relevant populations, including diverse or traditionally marginalized groups?