Microsoft Responsible AI Impact Assessment Template

FOR EXTERNAL RELEASE

June 2022

The Responsible AI Impact Assessment Template is the product of a multi-year effort at Microsoft to define a process for assessing the impact an AI system may have on people, organizations, and society. We are releasing our Impact Assessment Template externally to share what we have learned, invite feedback from others, and contribute to the discussion about building better norms and practices around AI.

We invite your feedback on our approach:
https://aka.ms/ResponsibleAIQuestions

Microsoft

Responsible AI Impact Assessment for [System Name]

For questions about specific sections within the Impact Assessment, please refer to the Impact Assessment Guide.

Section 1: System information

System profile

1.1 *Complete the system information below.*

System name	
Team name	

Track revision history below.

Authors	
Last updated	

Identify the individuals who will review your Impact Assessment when it is completed.

Reviewers	

System lifecycle stage

1.2 *Indicate the dates of planned releases for the system.*

Date	Lifecycle stage
	Planning & analysis
	Design
	Development
	Testing
	Implementation & deployment
	Maintenance
	Retired

System description

1.3 *Briefly explain, in plain language, what you're building. This will give reviewers the necessary context to understand the system and the environment in which it operates.*

System description

If you have links to any supplementary information on the system such as demonstrations, functional specifications, slide decks, or system architecture diagrams, please include links below.

Description of supplementary information	Link

System purpose

1.4 *Briefly describe the purpose of the system and system features, focusing on how the system will address the needs of the people who use it. Explain how the AI technology contributes to achieving these objectives.*

System purpose

System features

1.5 *Focusing on the whole system, briefly describe the system features or high-level feature areas that already exist and those planned for the upcoming release.*

Existing system features	System features planned for the upcoming release

Briefly describe how this system relates to other systems or products. For example, describe if the system includes models from other systems.

Relation to other systems/products

Geographic areas and languages

1.6 *Describe the geographic areas where the system will or might be deployed to identify special considerations for language, laws, and culture.*

The system is currently deployed to:	
In the upcoming release, the system will be deployed to:	
In the future, the system might be deployed to:	

For natural language processing systems, describe supported languages:

The system currently supports:	
In the upcoming release, the system will support:	
In the future, the system might support:	

Deployment mode

1.7 *Document each way that this system might be deployed.*

How is the system currently deployed?	
Will the deployment mode change in the upcoming release? If so, how?	

Intended uses

1.8 *Intended uses are the uses of the system your team is designing and testing for. An intended use is a description of who will use the system, for what task or purpose, and where they are when using the system. They are not the same as system features, as any number of features could be part of an intended use. Fill in the table with a description of the system's intended use(s).*

Name of intended use(s)	Description of intended use(s)
1.	
2.	
3.	

Section 2: Intended uses

Intended use #1: [Name of intended use] – repeat for each intended use

Copy and paste the Intended Use #1 section and repeat questions 2.1 – 2.8 for each intended use you identified above.

Assessment of fitness for purpose

2.1 *Assess how the system's use will solve the problem posed by each intended use, recognizing that there may be multiple valid ways in which to solve the problem.*

Assessment of fitness for purpose

Stakeholders, potential benefits, and potential harms

2.2 *Identify the system's stakeholders for this intended use. Then, for each stakeholder, document the potential benefits and potential harms. For more information, including prompts, see the Impact Assessment Guide.*

Stakeholders	Potential system benefits	Potential system harms
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		

Stakeholders for Goal-driven requirements from the Responsible AI Standard

2.3 *Certain Goals in the Responsible AI Standard require you to identify specific types of stakeholders. You may have included them in the stakeholder table above. For the Goals below that apply to the system, identify the specific stakeholder(s) for this intended use. If a Goal does not apply to the system, enter "N/A" in the table.*

Goal A5: Human oversight and control
This Goal applies to all AI systems. Complete the table below.

Who is responsible for troubleshooting, managing, operating, overseeing, and controlling the system during and after deployment?	For these stakeholders, identify their oversight and control responsibilities.

Goal T1: System intelligibility for decision making
This Goal applies to AI systems when the intended use of the generated outputs is to inform decision making by or about people. If this Goal applies to the system, complete the table below.

Who will use the outputs of the system to make decisions?	Who will decisions be made about?

Goal T2: Communication to stakeholders
This Goal applies to all AI systems. Complete the table below.

Who will make decisions about whether to employ the system for particular tasks?	Who develops or deploys systems that integrate with this system?

Goal T3: Disclosure of AI interaction
This Goal applies to AI systems that impersonate interactions with humans, unless it is obvious from the circumstances or context of use that an AI system is in use, and AI systems that generate or manipulate image, audio, or video content that could falsely appear to be authentic. If this Goal applies to the system, complete the table below.

Who will use or be exposed to the system?

Fairness considerations

2.4 *For each Fairness Goal that applies to the system, 1) identify the relevant stakeholder(s) (e.g., system user, person impacted by the system); 2) identify any demographic groups, including marginalized groups, that may require fairness considerations; and 3) prioritize these groups for fairness consideration and explain how the fairness consideration applies. If the Fairness Goal does not apply to the system, enter "N/A" in the first column.*

Goal F1: Quality of service

This Goal applies to AI systems when system users or people impacted by the system with different demographic characteristics might experience differences in quality of service that can be remedied by building the system differently. If this Goal applies to the system, complete the table below describing the appropriate stakeholders for this intended use.

Which stakeholder(s)	For affected stakeholder(s) which	Explain how each demographic group

Goal F2: Allocation of resources and opportunities

This Goal applies to AI systems that generate outputs that directly affect the allocation of resources or opportunities relating to finance, education, employment, healthcare, housing, insurance, or social welfare. If this Goal applies to the system, complete the table below describing the appropriate stakeholders for this intended use.

Which stakeholder(s)	For affected stakeholder(s) which	Explain how each demographic group

Goal F3: Minimization of stereotyping, demeaning, and erasing outputs

This Goal applies to AI systems when system outputs include descriptions, depictions, or other representations of people, cultures, or society. If this Goal applies to the system, complete the table below describing the appropriate stakeholders for this intended use.

Which stakeholder(s)	For affected stakeholder(s) which	Explain how each demographic group

Technology readiness assessment

2.5 *Indicate with an "X" the description that best represents the system regarding this intended use.*

Select one	Technology Readiness
	The system includes AI supported by basic research and has not yet been deployed to production systems at scale for similar uses.
	The system includes AI supported by evidence demonstrating feasibility for uses similar to this intended use in production systems.
	This is the first time that one or more system component(s) are to be validated in relevant environment(s) for the intended use. Operational conditions that can be supported have not yet been completely defined and evaluated.
	This is the first time the whole system will be validated in relevant environment(s) for the intended use. Operational conditions that can be supported will also be validated. Alternatively, nearly similar systems or nearly similar methods have been applied by other organizations with defined success.
	The whole system has been deployed for all intended uses, and operational conditions have been qualified through testing and uses in production.

Task complexity

2.6 *Indicate with an "X" the description that best represents the system regarding this intended use.*

Select One	Task Complexity
	Simple tasks, such as classification based on few features into a few categories with clear boundaries. For such decisions, humans could easily agree on the correct answer, and identify mistakes made by the system. For example, a natural language processing system that checks spelling in documents.
	Moderately complex tasks, such as classification into a few categories that are subjective. Typically, ground truth is defined by most evaluators arriving at the same answer. For example, a natural language processing system that autocompletes a word or phrase as the user is typing.
	Complex tasks, such as models based on many features, not easily interpretable by humans, resulting in highly variable predictions without clear boundaries between decision criteria. For such decisions, humans would have a difficult time agreeing on the best answer, and there may be no clearly incorrect answer. For example, a natural language processing system that generates prose based on user input prompts.

Role of humans

2.7 *Indicate with an "X" the description that best represents the system regarding this intended use.*

Select One	Role of humans
	People will be responsible for troubleshooting triggered by system alerts but will not otherwise oversee system operation. For example, an AI system that generates keywords from unstructured text alerts the operator of errors, such as improper format of submission files.
	The system will support effective hand-off to people but will be designed to automate most use. For example, an AI system that generates keywords from unstructured text that can be configured by system admins to alert the operator when keyword generation falls below a certain confidence threshold.
	The system will require effective hand-off to people but will be designed to automate most use. For example, an AI system that generates keywords from unstructured text alerts the operator when keyword generation falls below a certain confidence threshold (regardless of system admin configuration).
	People will evaluate system outputs and can intervene before any action is taken: the system will proceed unless the reviewer intervenes. For example, an AI system that generates keywords from unstructured text will deliver the generated keywords for operator review but will finalize the results unless the operator intervenes.
	People will make decisions based on output provided by the system: the system will not proceed unless a person approves. For example, an AI system that generates keywords from unstructured text but does not finalize the results without review and approval from the operator.

Deployment environment complexity

2.8 *Indicate with an "X" the description that best represents the system regarding this intended use.*

Select One	Deployment environment complexity
	Simple environment, such as when the deployment environment is static, possible input options are limited, and there are few unexpected situations that the system must deal with gracefully. For example, a natural language processing system used in a controlled research environment.
	Moderately complex environment, such as when the deployment environment varies, unexpected situations the system must deal with gracefully may occur, but when they do, there is little risk to people, and it is clear how to effectively mitigate issues. For example, a natural language processing system used in a corporate workplace where language is professional and communication norms change slowly.
	Complex environment, such as when the deployment environment is dynamic, the system will be deployed in an open and unpredictable environment or may be subject to drifts in input distributions over time. There are many possible types of inputs, and inputs may significantly vary in quality. Time and attention may be at a premium in making decisions and it can be difficult to mitigate issues. For example, a natural language processing system used on a social media platform where language and communication norms change rapidly.

Section 3: Adverse impact

Restricted Uses

3.1 *If any uses of the system are subject to a legal or internal policy restriction, list them here, and follow the requirements for those uses.*

Restricted Uses

Unsupported uses

3.2 *Uses for which the system was not designed or evaluated or that should be avoided.*

Unsupported uses

Known limitations

3.3 *Describe the known limitations of the system. This could include scenarios where the system will not perform well, environmental factors to consider, or other operating factors to be aware of.*

Known limitations

Potential impact of failure on stakeholders

3.4 *Define predictable failures, including false positive and false negative results for the system as a whole and how they would impact stakeholders for each intended use.*

Potential impact of failure on stakeholders

Potential impact of misuse on stakeholders

3.5 *Define system misuse, whether intentional or unintentional, and how misuse could negatively impact each stakeholder. Identify and document whether the consequences of misuse differ for marginalized groups. When serious impacts of misuse are identified, note them in the summary of impact as a potential harm.*

Potential impact of misuse on stakeholders

Sensitive Uses

3.6 *Consider whether the use or misuse of the system could meet any of the Microsoft Sensitive Use triggers below.*

Yes or No	Sensitive Use triggers
	Consequential impact on legal position or life opportunities The use or misuse of the AI system could affect an individual's: legal status, legal rights, access to credit, education, employment, healthcare, housing, insurance, and social welfare benefits, services, or opportunities, or the terms on which they are provided.
	Risk of physical or psychological injury The use or misuse of the AI system could result in significant physical or psychological injury to an individual.
	Threat to human rights The use or misuse of the AI system could restrict, infringe upon, or undermine the ability to realize an individual's human rights. Because human rights are interdependent and interrelated, AI can affect nearly every internationally recognized human right.

Section 4: Data requirements

Data requirements

4.1 *Define and document data requirements with respect to the system's intended uses, stakeholders, and the geographic areas where the system will be deployed.*

Data requirements

Existing data sets

4.2 *If you plan to use existing data sets to train the system, assess the quantity and suitability of available data sets that will be needed by the system in relation to the data requirements defined above. If you do not plan to use pre-defined data sets, enter "N/A" in the response area.*

Existing data sets

Section 5: Summary of impact

Potential harms and preliminary mitigations

5.1 *Gather the potential harms you identified earlier in the Impact Assessment in this table (check the stakeholder table, fairness considerations, adverse impact section, and any other place where you may have described potential harms). Use the mitigations prompts in the Impact Assessment Guide to understand if the Responsible AI Standard can mitigate some of the harms you identified. Discuss the harms that remain unmitigated with your team and potential reviewers.*

Describe the potential harm	Corresponding Goal from the Responsible AI Standard (if applicable)	Describe your initial ideas for mitigations or explain how you might implement the corresponding Goal in the design of the system

Goal Applicability

5.2 *To assess which Goals apply to this system, use the tables below. When a Goal applies to only specific types of AI systems, indicate if the Goal applies to the system being evaluated in this Impact Assessment by indicating "Yes" or "No." If you indicate that a Goal does not apply to the system, explain why in the response area. If a Goal applies to the system, you must complete the requirements associated with that Goal while developing the system.*

Accountability Goals

Goals	Does this Goal apply to the system? (Yes or No)
A1: Impact assessment *Applies to:* All AI systems.	
A2: Oversight of significant adverse impacts *Applies to:* All AI systems.	
A3: Fit for purpose *Applies to:* All AI systems.	
A4: Data governance and management *Applies to:* All AI systems.	
A5: Human oversight and control *Applies to:* All AI systems.	

Transparency Goals

Goals	Does this Goal apply to the system? **(Yes or No)**
T1: System intelligibility for decision making *Applies to:* AI systems when the intended use of the generated outputs is to inform decision making by or about people.	
T2: Communication to stakeholders *Applies to:* All AI systems.	
T3: Disclosure of AI interaction *Applies to:* AI systems that impersonate interactions with humans, unless it is obvious from the circumstances or context of use that an AI system is in use, and AI systems that generate or manipulate image, audio, or video content that could falsely appear to be authentic.	

If you selected "No" for any of the Transparency Goals, explain why the Goal does not apply to the system.

Fairness Goals

Goals	Does this Goal apply to the system? **(Yes or No)**
F1: Quality of service *Applies to:* AI systems when system users or people impacted by the system with different demographic characteristics might experience differences in quality of service that can be remedied by building the system differently.	
F2: Allocation of resources and opportunities *Applies to:* AI systems that generate outputs that directly affect the allocation of resources or opportunities relating to finance, education, employment, healthcare, housing, insurance, or social welfare.	
F3: Minimization of stereotyping, demeaning, and erasing outputs *Applies to:* AI systems when system outputs include descriptions, depictions, or other representations of people, cultures, or society.	

If you selected "No" for any of the Fairness Goals, explain why the Goal does not apply to the system below.

15

Reliability & Safety Goals

Goals	Does this Goal apply to the system? **(Yes or No)**
RS1: Reliability and safety guidance *Applies to:* All AI systems.	
RS2: Failures and remediations *Applies to:* All AI systems.	
RS3: Ongoing monitoring, feedback, and evaluation *Applies to:* All AI systems.	

Privacy & Security Goals

Goals	Does this Goal apply to the system? **(Yes or No)**
PS1: Privacy Standard compliance Applies when the Microsoft Privacy Standard applies.	
PS2: Security Policy compliance Applies when the Microsoft Security Policy applies.	

Inclusiveness Goal

Goals	Does this Goal apply to the system? **(Yes or No)**
I1: Accessibility Standards compliance Applies when the Microsoft Accessibility Standards apply.	

Signing off on the Impact Assessment

5.3 *Before you continue with next steps, complete the appropriate reviews and sign off on the Impact Assessment. At minimum, the PM should verify that the Impact Assessment is complete. In this case, ensure you complete the appropriate reviews and secure all approvals as required by your organization before beginning development.*

Reviewer role and name	I can confirm that the document benefitted from collaborative work and different expertise within the team (e.g., engineers, designers, data scientists, etc.)	Date reviewed	Comments

Update and review the Impact Assessment at least annually, when new intended uses are added, and before advancing to a new release stage. The Impact Assessment will remain a key reference document as you work toward compliance with the remaining Goals of the Responsible AI Standard.

Scan this code to access responsible AI resources from Microsoft: