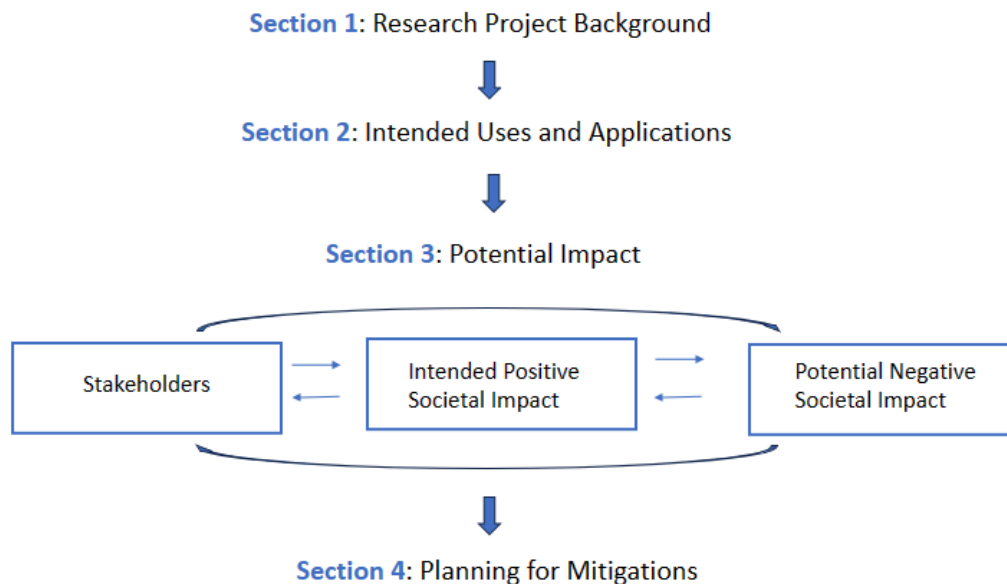


Work in Progress. Please do not distribute.

Overall Structure of Societal Impact Assessment Template



In total, this template contains **16 questions**, attempting to guide you through different aspects of the potential societal impact of your research. If you find some of these questions difficult or are struggling to come up with answers, **rather than skipping the questions or providing an unsatisfying answer, use the text box to note what it is that you're finding difficult or challenging**. If you come back to the template in the future, perhaps once you're further along in your research, you may find that you're then able to overcome these difficulties or challenges.

Why would you want to use this template

Based on interviews with industry researchers across organizations and research areas, we have found that completing this template could potentially:

- Help you surface and concretize the **positive societal impacts** of your work
- Help you consider the potential **negative societal impacts** of your work
- Help you strengthen your research **design and methods**
- Help you surface potential **future research directions**
- Help you prepare an impact statement to include in your **publications**, as is increasingly suggested or required by publication venues like conferences, journals, etc.
- Help you prepare an impact statement to share with **product** teams.

Work in Progress. Please do not distribute.

Societal Impact Assessment Template

Section 1. Research Project Background

Describe your project in a paragraph or two. If you intend to share a completed version of this template with other researchers, a review panel, a compliance team, a product team, and/or general audiences, please describe your project in sufficient detail and in sufficiently accessible language that the intended audience can understand your project.

1A: Project Description

<Your Answer Here>

What are some of the assets you are planning to release? Think *separately* about the assets you plan to release *internally* and *externally*, as this might affect how you assess and address the potential impact of your work.

Some potential assets you might consider include: research papers, blog posts, demos, source code, model(s), datasets, patents, project web pages, white papers, policy briefs, etc.

1B: What are some assets you are planning to release *publicly* as intended outcomes of this research?

<Your Answer Here>

1C: What are some assets you are planning to release *internally* as intended outcomes of this research?

<Your Answer Here>

Work in Progress. Please do not distribute.

Section 2. Intended Uses and Applications

This section is intended to help you think through the **intended uses and imagined applications** of your research. These may include, for example, direct use in the context of products and services or use by others in the research community.

2A: What are the intended uses and imagined applications of this research?

<Your Answer Here>

2B: Does this research project relate to any existing or potential new products or services in your organization? Please elaborate.

<Your Answer Here>

For example, does your research aim to improve certain aspects of existing products or services within your organization? Would your research contribute to the design and development of a new product or service within your organization?

Section 3. Potential Impact

Research is a complex process whose ultimate impact may depend on various contingencies, making it sometimes difficult to contemplate its potential impacts. This section nevertheless serves as a **starting point** for introspection, prompting you to think through various scenarios that might result from the application of your research findings.

Stakeholders

Before diving into the concrete societal impacts, let's think about **who are the possible stakeholders that would make use of or be most impacted by** both your research and any future applications of it? **To help you brainstorm potential stakeholder groups, you could consider:**

Work in Progress. Please do not distribute.

Individuals:

- Think about end users or others who will be **directly** impacted by the research outcomes, but also those who might be **indirectly** impacted.

Community:

- Think about the **marginalized subgroups** within society that may be especially helped by or especially vulnerable to certain uses of the research (e.g., LGBTQ+ individuals, people from low-income backgrounds, people with disabilities, children, and the elderly).

Companies:

- Which companies might be influenced by your research findings? These could include companies developing products and services that may be (positively or negatively) impacted by the advancements or discoveries from your research.

Industries:

- Consider not only the industries to which your company belongs or the field to which your team belongs, but also **other industries** or fields that may be **interconnected or relevant**.

Governments:

- Consider how your research might influence policies and regulations in governments **beyond just the country** you are located in.

Civil society:

- Consider how your research might influence the participation and empowerment of **non-profit organizations** within civil society, influencing social movements, and guiding **community** actions towards more informed and effective practices.

3A: Who are the possible stakeholders that would **make use of your research findings?**

<Your Answer Here>

3B: Who are the possible stakeholders that would **be most impacted by both your research and future applications of it?**

<Your Answer Here>

Note: It is hard to enumerate all stakeholders. You can always **come back to these two questions** and add more stakeholders throughout the process of completing this template and throughout your research.

Work in Progress. Please do not distribute.

Intended Positive Societal Impact

You could think through this by sharing the **best-case scenario for all the impacted stakeholders you envisioned in 3B**. Below are some potential types of impact you could think about.

Economic Impact: This reflects how research contributes to economic growth, such as through the creation of new industries, jobs, or products, or by making existing industries more efficient or sustainable.

Policy Impact: This reflects how research influences policy making at various levels, from local to international, leading to more evidence-based and effective policies.

Health and Wellbeing Impact: This considers how research contributes to improving human health and quality of life, such as through the development of new medical treatments or public health interventions.

Environmental Impact: This assesses how research contributes to protecting the environment, like advancing renewable energy technologies or understanding and mitigating the impacts of climate change.

Cultural Impact: This reflects how research influences cultural understanding, social structures, or behaviors, which could range from improving education practices to challenging societal norms.

Knowledge Impact: This reflects how your research influences researchers, practitioners, and the general public by deepening their understanding or transforming their perspective on a particular subject.

You could copy pastes the impacted stakeholders you surfaced in 3A–B and use the scaffold above to think concretely about the positive societal impact of your research.

3C. In the best-case scenario, what would be the positive societal impact of your research?

<Stakeholders from 3A–B>

<Your Answers Here about the positive societal impact>

Remember that the process of envisioning stakeholders and considering potential impact is iterative. If you think of any additional stakeholders, feel free to add them here and in 3A–B!

Potential Negative Societal Impact

Note: Potential negative societal impacts are **different from** risks to human subjects that may arise during the research process itself. If your research involves human subjects, you should (and often must) consult an institutional review board (IRB).

Work in Progress. Please do not distribute.

Known Limitations

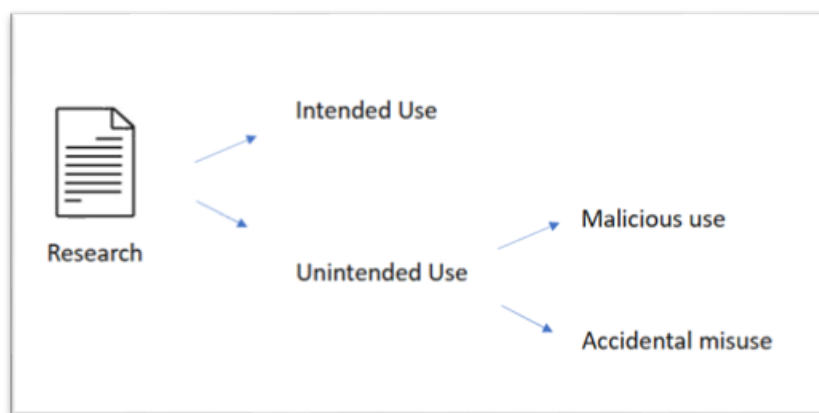
In this section, please think about and document the limitations of your research. Based on the research discipline, these could include but are not limited to:

- **Generalizability/Transferability:** To what extent do the study results or findings hold in different contexts?
- **Reproducibility:** To what extent could other researchers reproduce the study?
- **Robustness:** How sensitive are the results to minor violations of assumptions (e.g., small tweaks to a mathematical model, metrics, or hyperparameters)?
- **Subjective viewpoints** of researchers: How might researchers' backgrounds and identities affect the way they interpret the study results?

Please note that these limitations can potentially result in unintentional misuse and may also be abused by malicious parties, as we will delve into in the following subsections.

3D: What are the current limitations of the results or findings of your research?

<Your Answers Here>



Research can be used according to researchers' intentions, but it can also have unintended uses, either accidentally or maliciously. In the following sections, we will help you think through these different scenarios.

Misuse (Accidental unintended use)

Here are some potential misuses to help you brainstorm:

Work in Progress. Please do not distribute.

- **Over-trust/Over-reliance:** For example, doctors could over-trust the information shared by a medical chatbot and accidentally share incorrect medical advice or diagnosis with patients.
- **Usage under incorrect operating conditions:** For example, an AR headset designed to be used indoors could be used outdoors, resulting in physical injuries to users wearing the headset while crossing the road.
- **Incorrect generalization:** For example, policymakers might use the study results from economic research about small business owners in urban areas to inform policy making. However, without proper understanding of regional differences, these well-intentioned policies could end up being counterproductive towards small business owners in rural areas.
- **Misinterpretation:** For example, research results showing that practitioners misuse machine learning interpretability tools could be misinterpreted as implying that all data scientists over-trust and misuse interpretability tools, discrediting data scientists with different training, backgrounds, or experience levels than those in the study sample.

3E: What are some scenarios in which your research could be misused unintentionally?

<Which group(s) of stakeholders you identified in question 3A might misuse your research outcomes?>

<Which group(s) of stakeholders you identified in question 3B might be harmed. In what ways?>

Remember that the process of envisioning stakeholders and considering potential impact is iterative.

If you think of any additional stakeholders, feel free to add them to 3A–B!

Abuse (intentional malicious use)

Here are some potential abuses to help you brainstorm:

- **Discrimination:** For example, fitness trackers designed to help users monitor their health and fitness levels could be sold to private healthcare firms to set insurance rates without users' consent.
- **Political manipulation:** For example, research findings aimed at understanding public discourse on COVID vaccination could be selectively used or misinterpreted by politicians. This intentional misuse might serve to manipulate voter perceptions and opinions.
- **Security threats:** For example, a paper disclosing the vulnerability of a cryptographic algorithm could be exploited by a hacker to engage in criminal activities.
- **Misinformation:** For example, deepfakes could be used by malicious parties to spread misinformation on social media.

3F: How might a malicious entity use the results and artifacts from your research to cause harm?

What are the different scenarios that you could imagine?

Work in Progress. Please do not distribute.

<Which group(s) of stakeholders you identified in question 3A might abuse your research outcomes?>

<Which group(s) of stakeholders you identified in question 3B might be harmed. In what ways?>

Remember that the process of envisioning stakeholders and considering potential impact is iterative.
If you think of any additional stakeholders, feel free to add them in 3A–B!

Harms from intended use

Here are some potential harms to help you brainstorm:

- **Addiction:** For example, the design and development of a recommendation system might successfully increase the time people spend on a platform, but cause harm to people’s mental wellbeing.
- **Harm to trust:** For example, research aimed at highlighting the uncertainty in AI-generated outputs may unintentionally be perceived as a demonstration of the system's lack of credibility or reliability. This could lead to a decline in the usage or trust of these AI systems, despite their exceptional performance under certain conditions.
- **Social stigma:** For example, research results focused on understanding social media usage patterns among the LGBTQ community, revealing how these individuals seek HIV medical information on social media, may further exacerbate the social stigma towards LGBTQ community members.
- **Financial harm/job displacement:** For example, text-to-image AI system generating high quality artwork may cause financial harm to artists.

3G: Refer back to Section 2 where you described how you intend for your research to be used in practice. Do any of these intended uses present risk of harm?

<Your Answers Here. While thinking about the scenarios, consider which group(s) of stakeholders you identified in questions 3A–B might be harmed. In what ways?>

Remember that the process of envisioning stakeholders and considering potential impact is iterative.
If you think of any additional stakeholders, feel free to add them in questions 3A–B!

Work in Progress. Please do not distribute.

Section 4. Planning for Mitigations

Now that you've identified the potential impacts of your research and the relevant stakeholders, you may be wondering how to **determine the seriousness of the risks posed by potential negative societal impacts and how to prioritize them**. To do so, we encourage you to consider the following questions.

- **Likelihood:** What is the likelihood that your research will be used in this way?
 - **Ease:** How easy would it be for others to use it in this way?
 - **Resources:** Would they need considerable resources (e.g., money, computational infrastructure, etc.)?
 - **Expertise:** Would they need specialized expertise?
 - **Capabilities:** How many actors would be capable of doing this?
 - **Incentives:** How many actors would have incentives to do this?

- **Magnitude:** What is the likely magnitude of the impact?
 - **Size:** For everyone that might be impacted, how large would the impact be?
 - **Scope:** How many individuals might be impacted?

- **Concentration:** How concentrated is the impact?
 - **Distribution of benefits:** Which populations are likely to experience the beneficial impacts?
 - **Distribution of costs:** Which populations are likely to experience the negative impacts?
 - **Relative impacts:** What is the *relative* beneficial or negative impact on different populations given differences in their starting positions?

- **Time horizon:** Are the impacts likely to be immediate or long-term?

4A: In light of these considerations, which potential negative societal impacts do you think are high priority to mitigate and which do you think are low priority?

<Your Answer Here>

Here is a list of methods to help you plan **possible mitigations throughout different stages of the research life cycle**: from early-stage research design, to executing and publishing the research, to monitoring research outcomes. Please note that this is not an exhaustive list of potential mitigations that could be put in place. We strongly recommend discussing the appropriateness and effectiveness of these methods with the stakeholders you identified in questions 3A–B.

Work in Progress. Please do not distribute.

Altering your research design and process

- If you are in an early stage of your project, reconsidering the research design to circumvent the potential negative societal impact
- **Reaching out and discussing** your research with researchers in other relevant areas or with subject matter experts within your organizations
- **Reaching out and discussing** your research with relevant stakeholders you identified in questions 3A–B

Releasing clear documentation

- Clearly **documenting limitations and intended use cases** in the public-facing artifacts you listed in Section 1
- Adding ethics and/or societal impact statements to your research paper
- Releasing documentation like model cards and datasheets-
- Clearly indicating **future research directions** that could overcome current limitations or help prevent potential harms
- Releasing relevant **education and training materials** to encourage the appropriate use of the research outcomes

Limiting open release

- Choosing **not** to publicly release code/models/data
- Choosing to release code/models/data to limited groups of researchers or practitioners
- Leveraging a process like the ACM artifact evaluation process, i.e., providing code/models/data only to reviewers without public release
- Providing a demonstration web site in lieu of releasing high risk models/code/data publicly

Monitoring use for abuse

- Releasing in such a way that usage can be tracked.
- Actively setting up alerts and escalations when abusive patterns are discovered

4B: How can the identified high-priority risks be mitigated?

<Your Answer Here>

4C: What actions can you take to ensure these mitigations are put in place?

<Your Answer Here>

Work in Progress. Please do not distribute.

4D: If you can't put these mitigations in place yourself, what steps can you take to increase the likelihood that others will?

<Your Answer Here>