

AI Ethics and Governance

Day 3: AI Sustainability

23/11/2022

Professor David Leslie
Director of Ethics and Responsible Innovation



OVERVIEW

Introduction to Practical Ethics

01

AI Harms & Values

02

AI Sustainability

03

Fairness, Bias Mitigation & Accountability

04

Explainability, CARE & Act principles

05

OVERVIEW

Introduction to
Practical Ethics

AI Harms & Values

AI Sustainability

Fairness, Bias
Mitigation &
Accountability

Explainability,
CARE & Act
principles

01

02

03

04

05

CONTENTS

Day 2 recap

01 Stakeholder Engagement Process

1. Stakeholder analysis
2. Positionality reflection
3. Establishing and engagement objective and method

Q&A

Activity 1: Stakeholder analysis

Lunch break

Activity 2: Positionality reflection

02

Stakeholder impact assessment

Q&A

Activity 3: SIA

Recap from Day 2

AI Harms and Values



AI Harms

- ▶ Loss of autonomy
- ▶ Impairment of human connection
- ▶ Dangerous outcomes
- ▶ Bias & discrimination
- ▶ Widening global divides
- ▶ Security
- ▶ Biospheric harm



Bioethics and Human rights in context

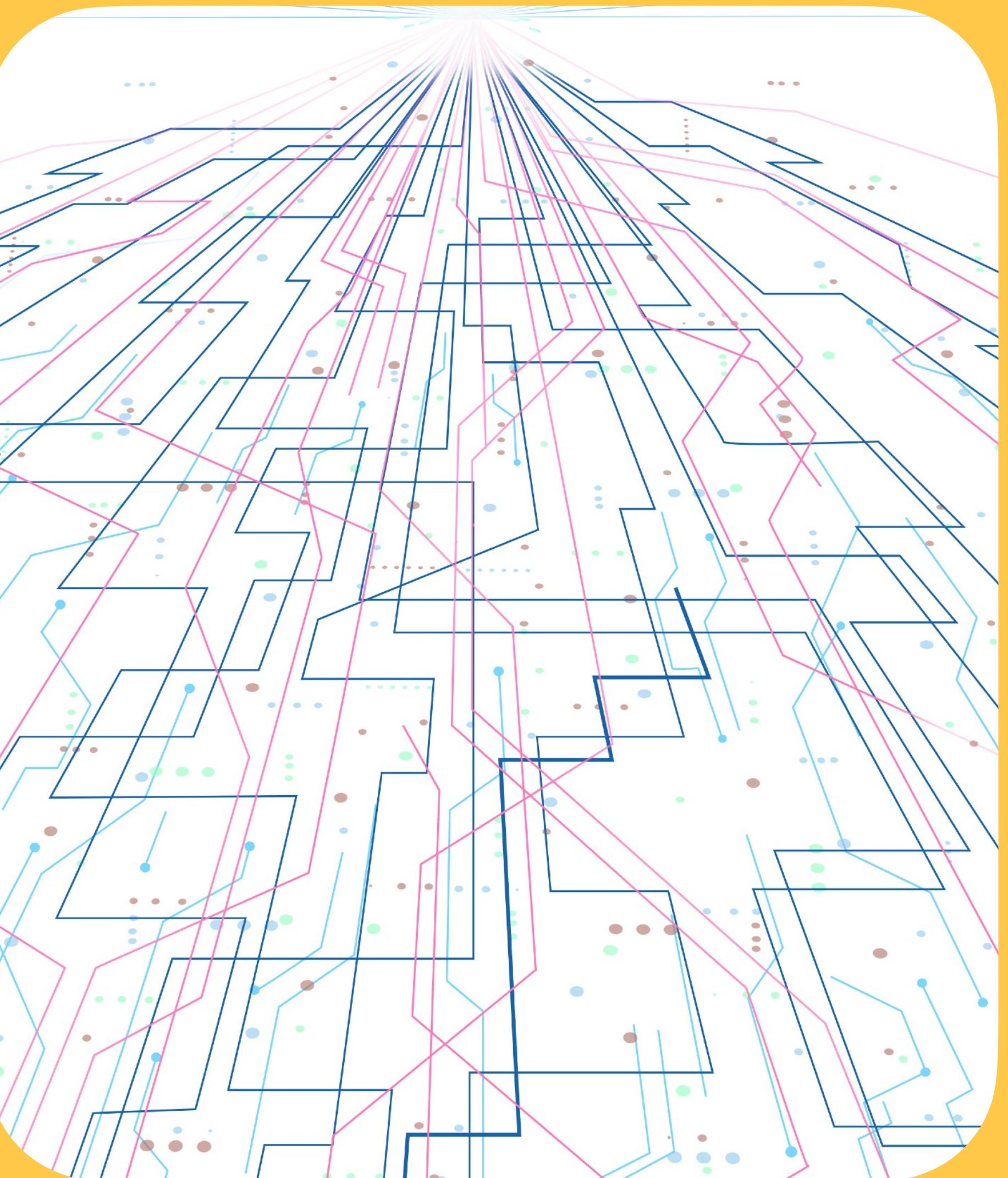


AI Values

- ▶ Respect
- ▶ Connect
- ▶ Care
- ▶ Protect
- ▶ Interconnectivity & Intergenerational Reciprocity
- ▶ Environmental Flourishing & Biospheric Rights

01

AI SUSTAINABILITY



Introduction

Moving from ‘principles to practice’

Introduction

Moving from ‘principles to practice’

- ➡ Approaches that focus on technical tools and documentation methods have fallen short

Introduction

Moving from ‘principles to practice’

- Approaches that focus on technical tools and documentation methods have fallen short
- Ex post methods of ‘auditability’ or ‘traceability’ show up a moment too late and remain too removed from design practices

Introduction

Moving from ‘principles to practice’

- Approaches that focus on technical tools and documentation methods have fallen short
- Ex post methods of ‘auditability’ or ‘traceability’ show up a moment too late and remain too removed from design practices
- AI sustainability requires end-to-end habits of critical reflection and deliberation across every stage of a research or innovation project’s lifecycle.

Introduction

So, what is AI sustainability?



Introduction

So, what is AI sustainability?

Designers and users of AI systems should remain aware that these technologies may have transformative and long-term effects on individuals and society. In order to ensure that the deployment of AI systems remains sustainable and supports the sustainability of the communities it will affect, AI project teams should proceed with a continuous sensitivity to the real-world impacts that their systems could have.

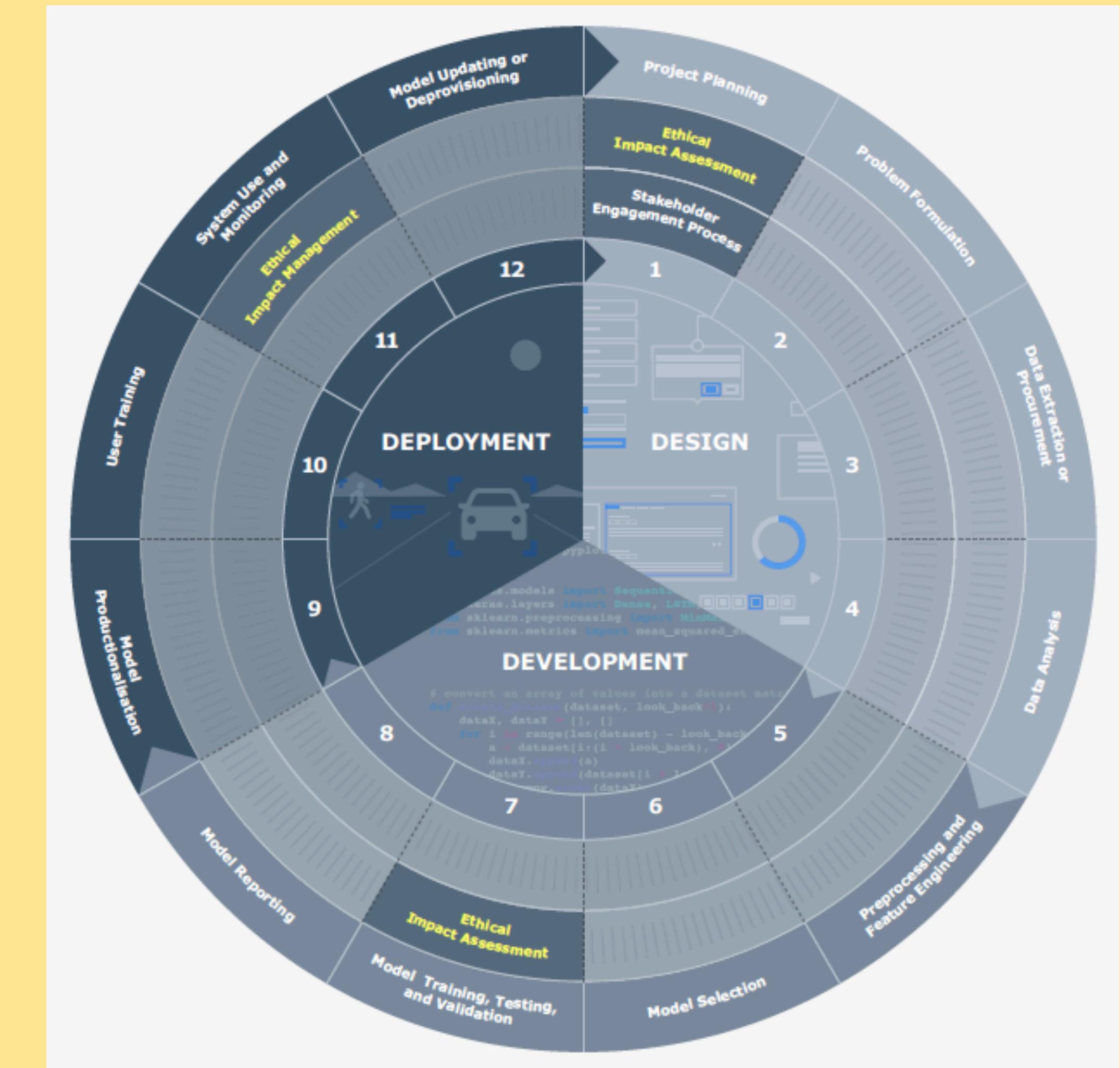
Introduction

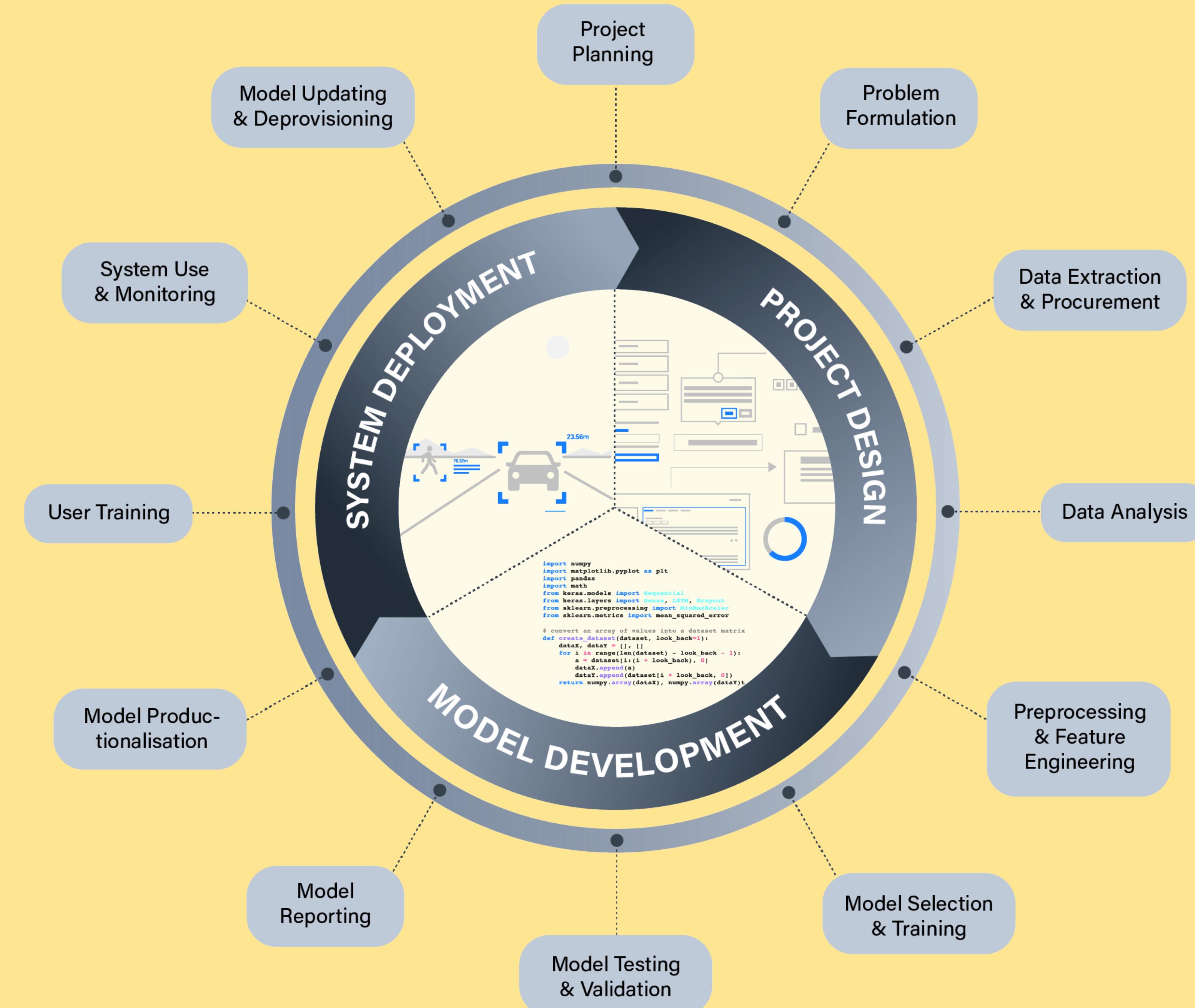
So, what is AI sustainability?

This involves engaging and involving impacted individuals and communities in the design and assessment of AI systems. AI sustainability also involves applying practices of critical reflection and deliberation and dedicating appropriate resources to ensure the system continues to operate safely and responsibly over time.

AI Project Lifecycle

- Socio-technical system
- (Project) Design
- (Model) Development
- (System) Deployment





Who are the stakeholders of a project?

Stakeholder are individuals or groups
that:

Stakeholder are individuals or groups that:

1//

Have **interests or rights** that may
be affected by the decisions and
activities of an organisation;

Stakeholder are individuals or groups that:

1/

Have **interests or rights** that may be affected by the decisions and activities of an organisation;

2/

May have the **power to influence the outcome** of such decisions and activities;

Stakeholder are individuals or groups that:

1/

Have **interests or rights** that may be affected by the decisions and activities of an organisation;

2/

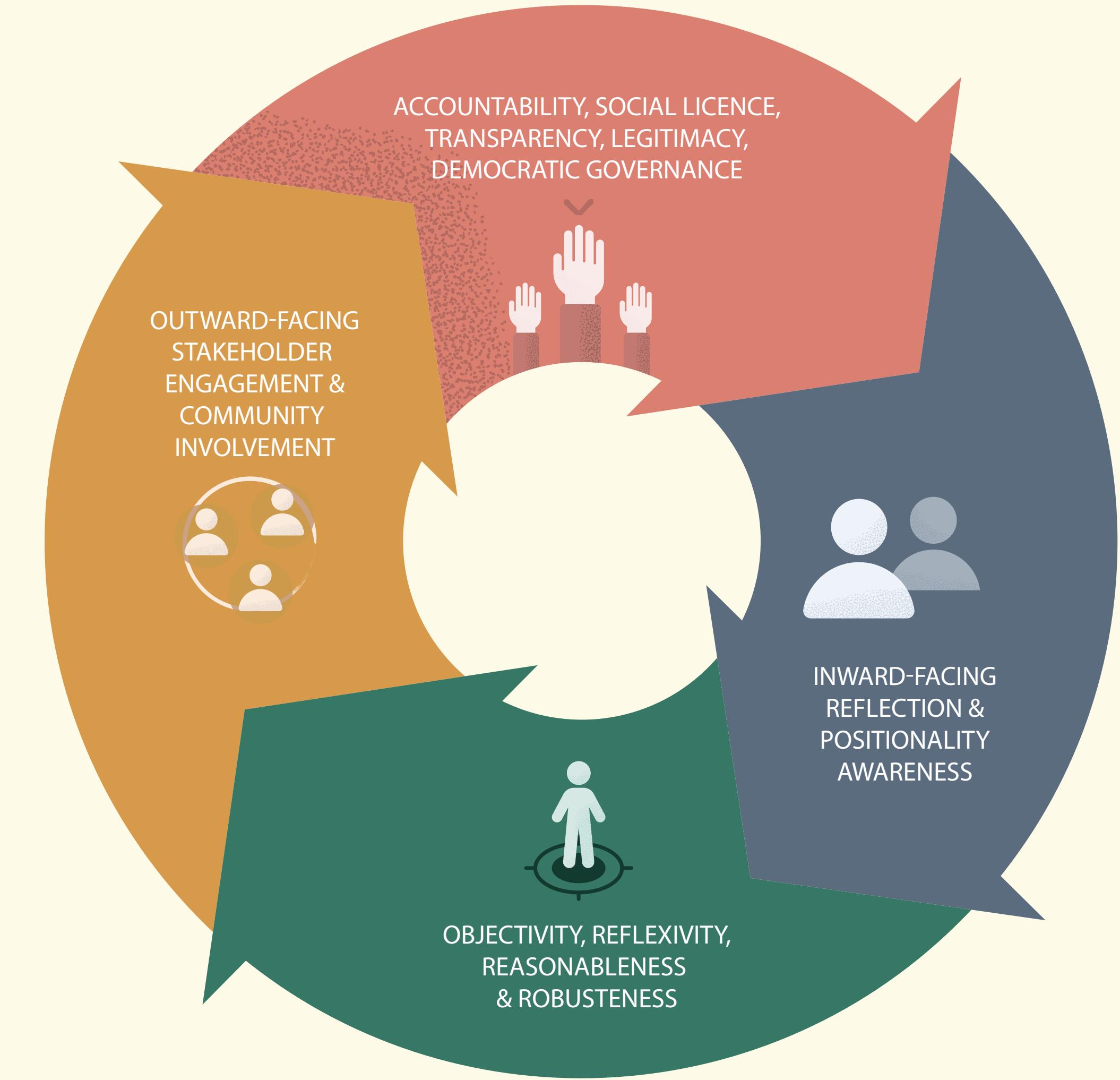
May have the **power to influence the outcome** of such decisions and activities;

3/

Possess **relevant characteristics** that put them in positions of **advantage or vulnerability** with regard to those decisions and activities.

Stakeholder engagement

Engagement and reflection cycle of sustainable AI innovation



What is the stakeholder engagement process?

Stakeholder engagement:

A three-step process

Stakeholder engagement:

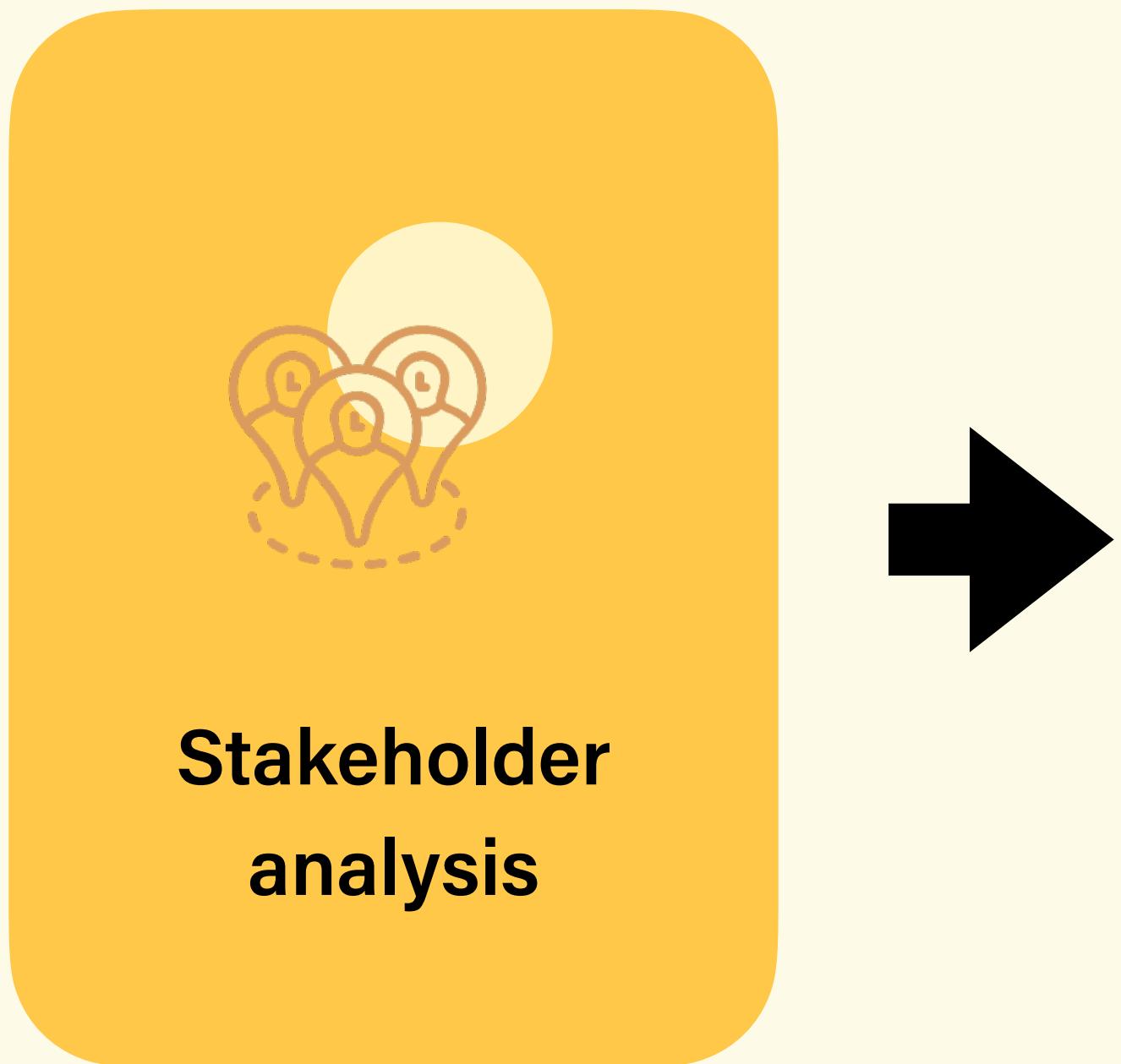
A three-step process



**Stakeholder
analysis**

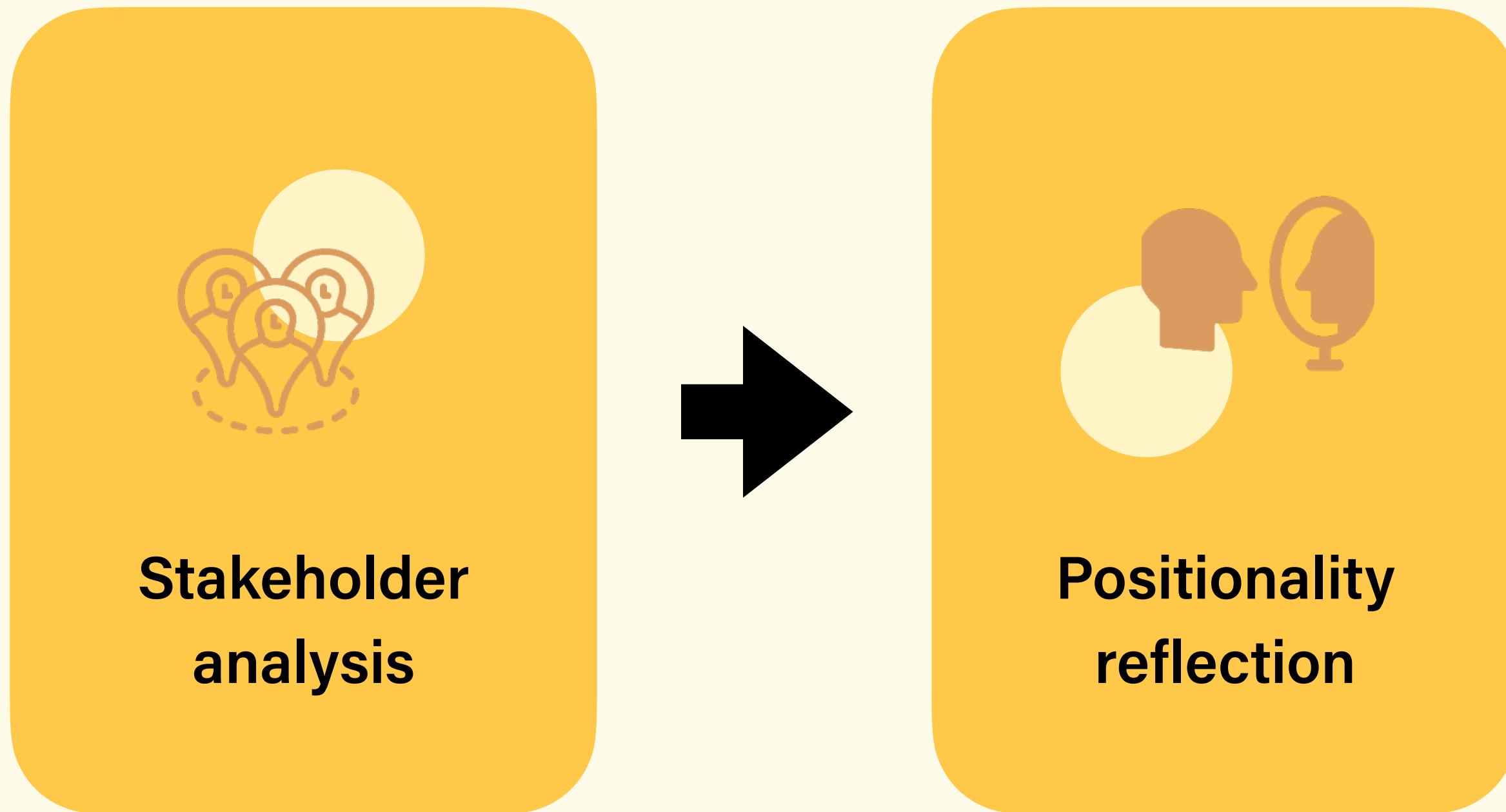
Stakeholder engagement:

A three-step process



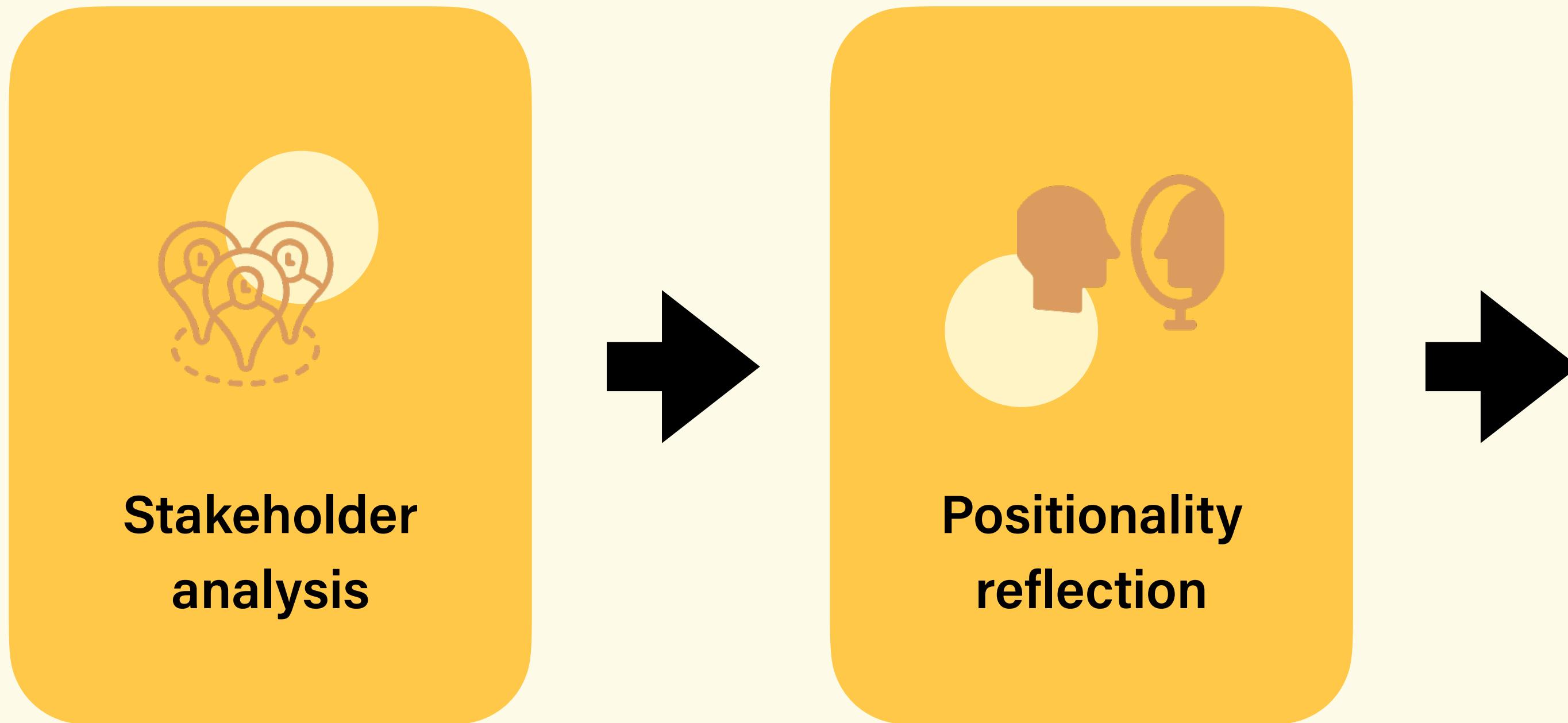
Stakeholder engagement:

A three-step process



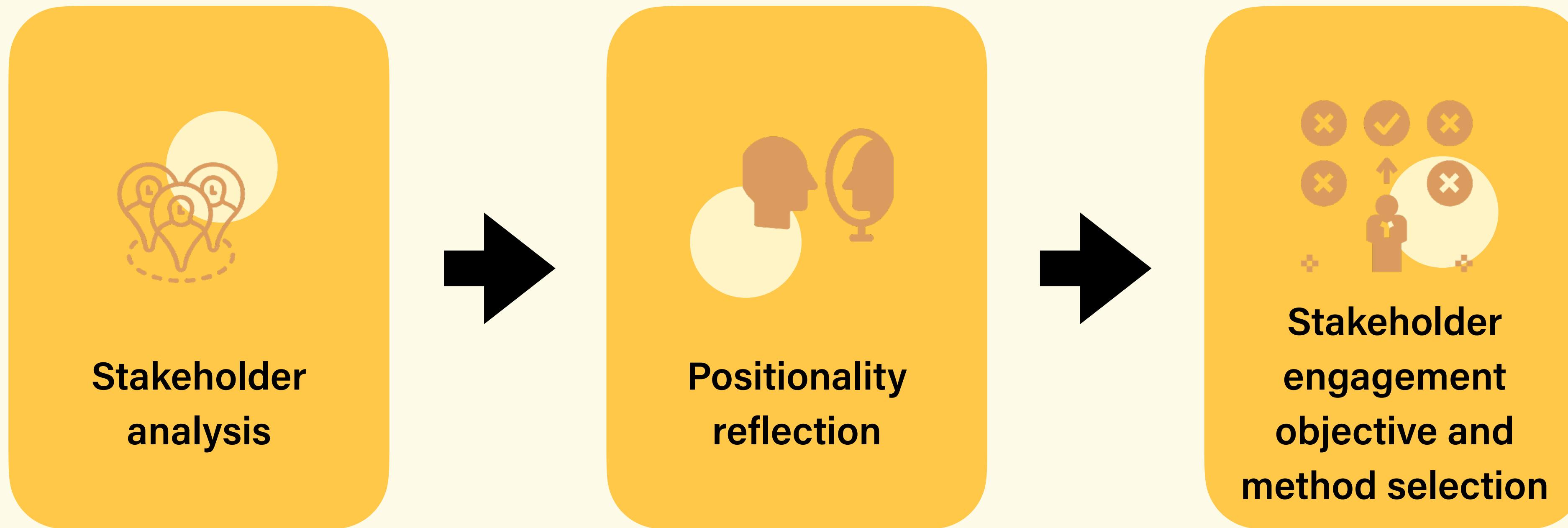
Stakeholder engagement:

A three-step process



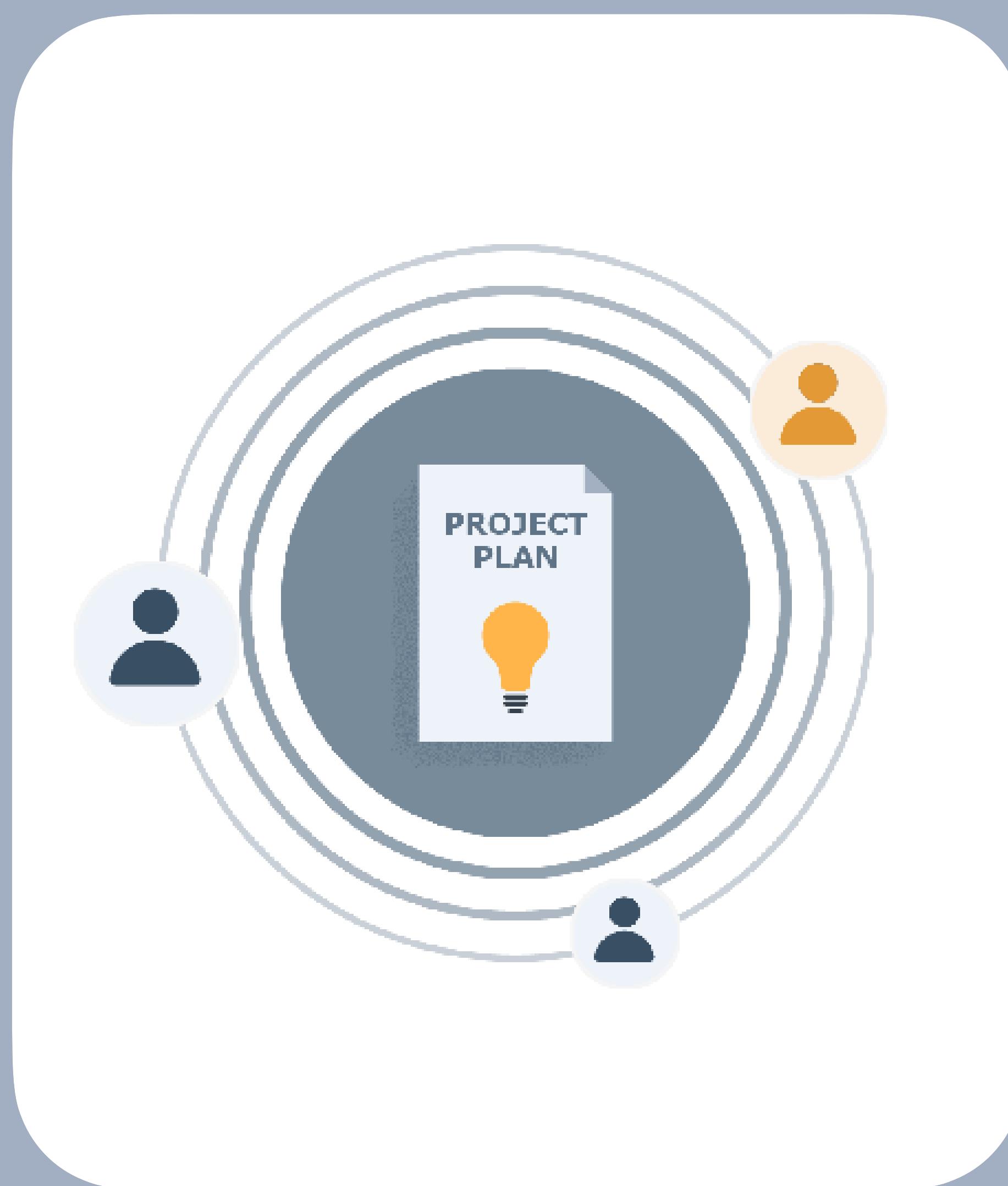
Stakeholder engagement:

A three-step process



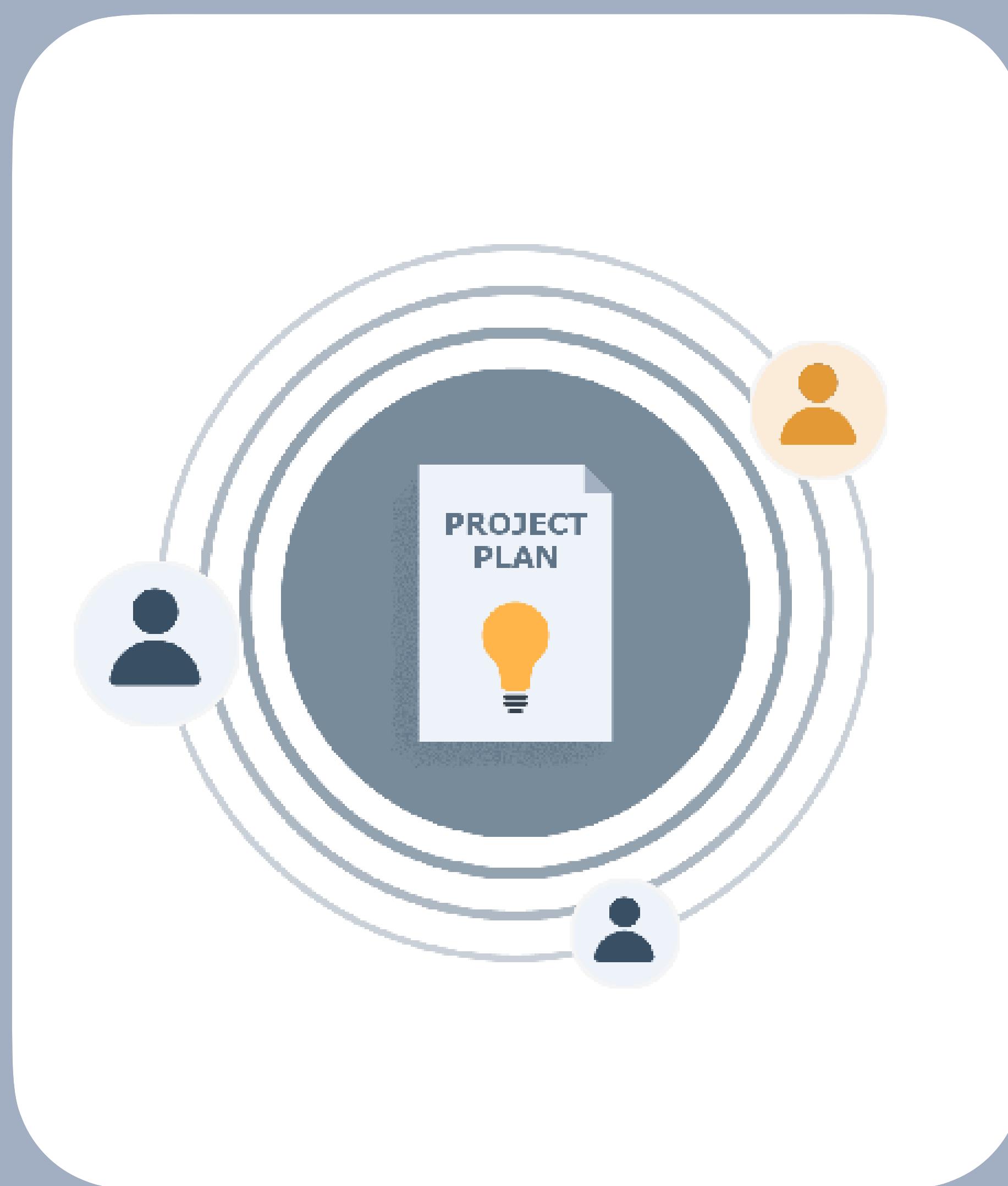


Stakeholder analysis and project scoping



Stakeholder analysis and project scoping

1. Outline the project: use context, domain, and data

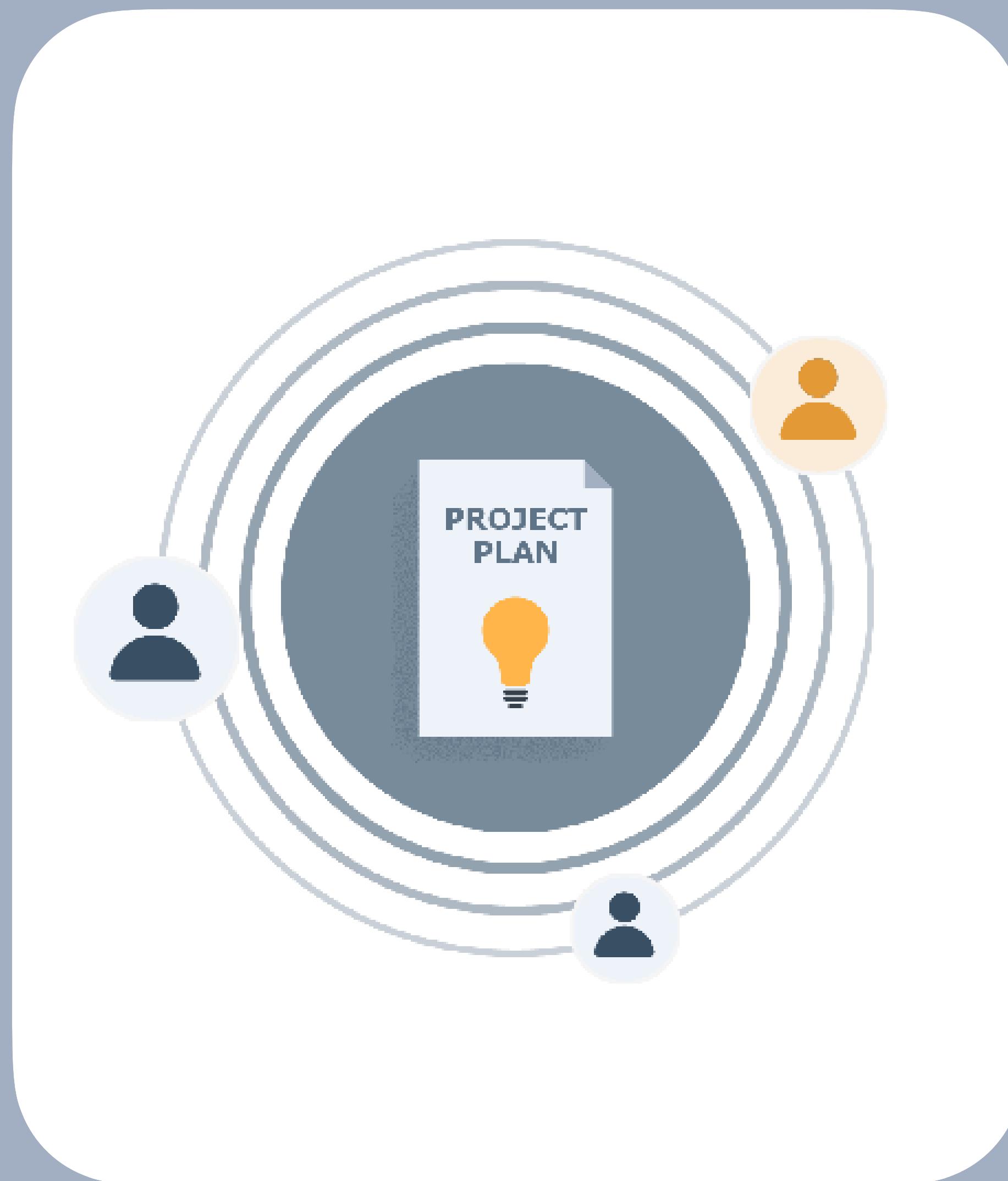


Stakeholder analysis and project scoping

1. Outline the project: use context, domain, and data
2. Identify stakeholders



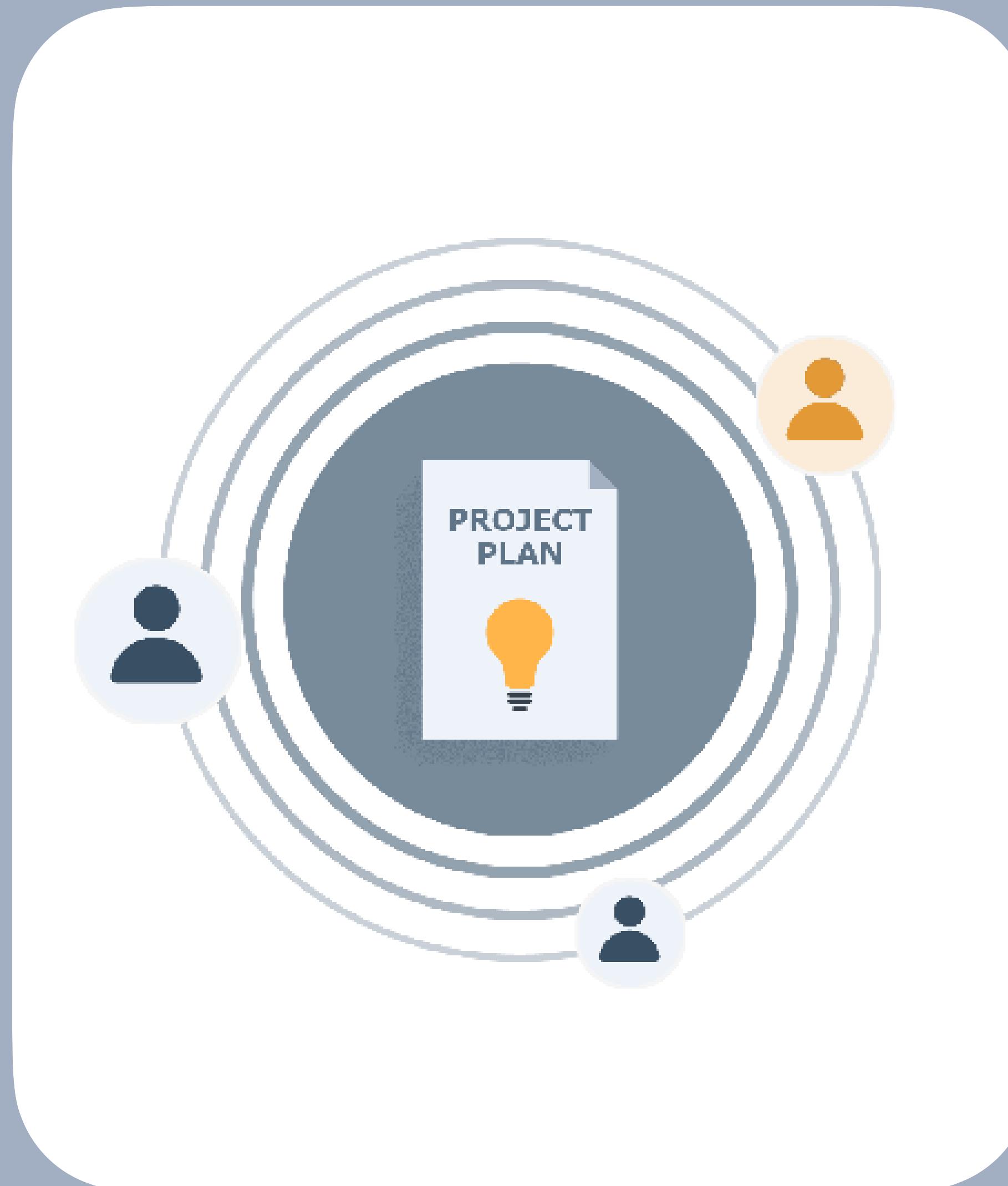
Stakeholder analysis and project scoping



1. Outline the project: use context, domain, and data
2. Identify stakeholders
3. Scope potential stakeholder impacts



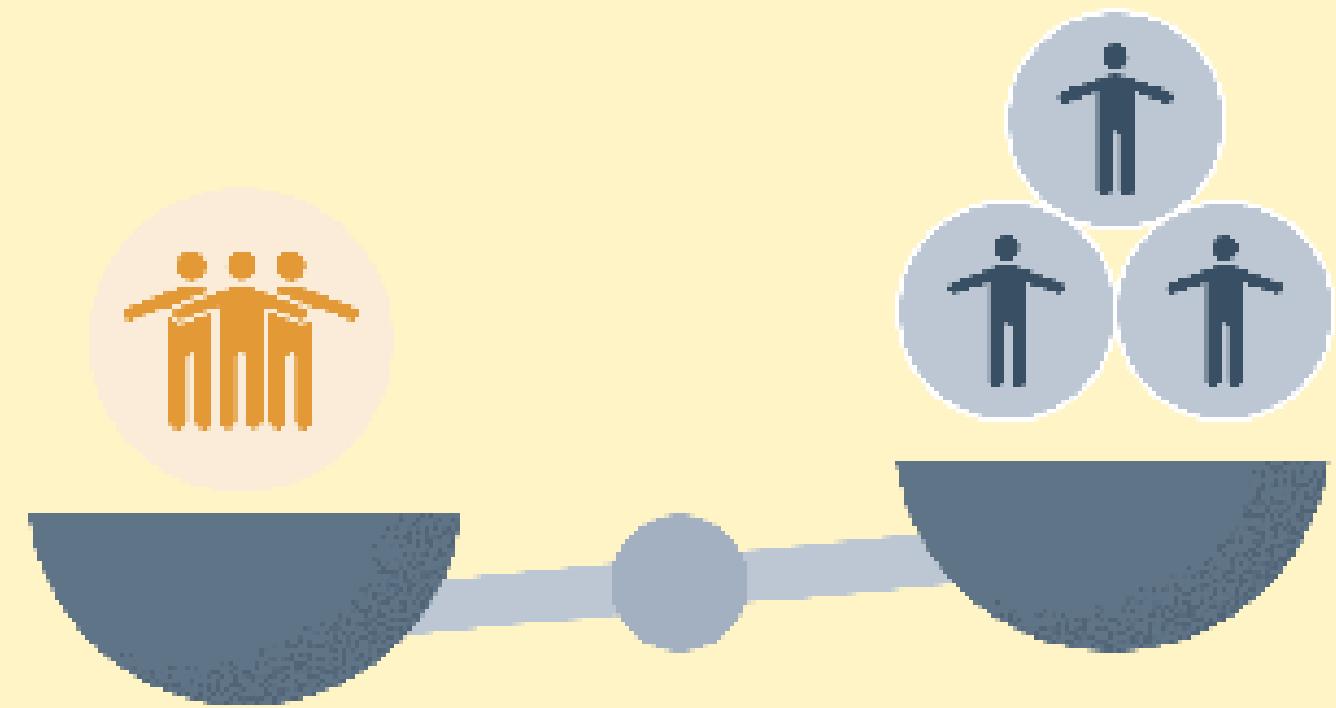
Stakeholder analysis and project scoping



1. Outline the project: use context, domain, and data
2. Identify stakeholders
3. Scope potential stakeholder impacts
4. Analyse stakeholder salience



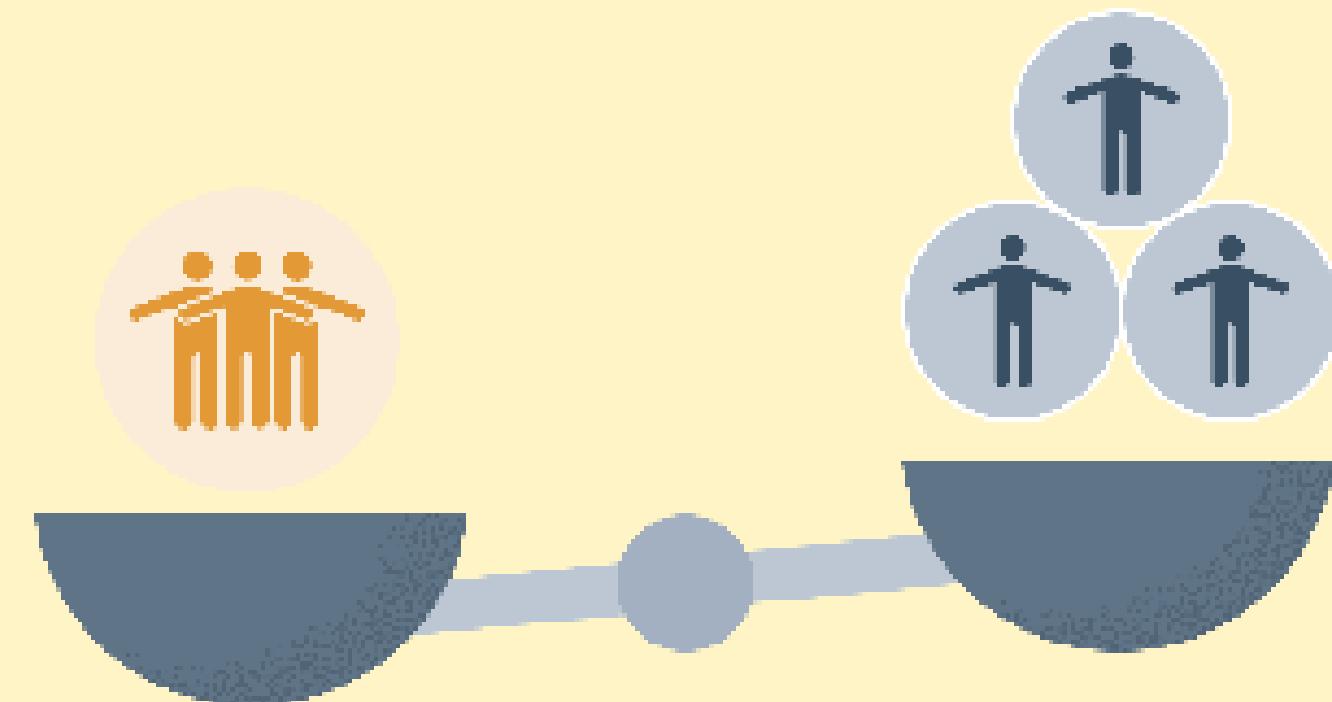
Positionality reflection





Positionality reflection

All individuals have **unique experiences** which shape their **thinking and perspectives**

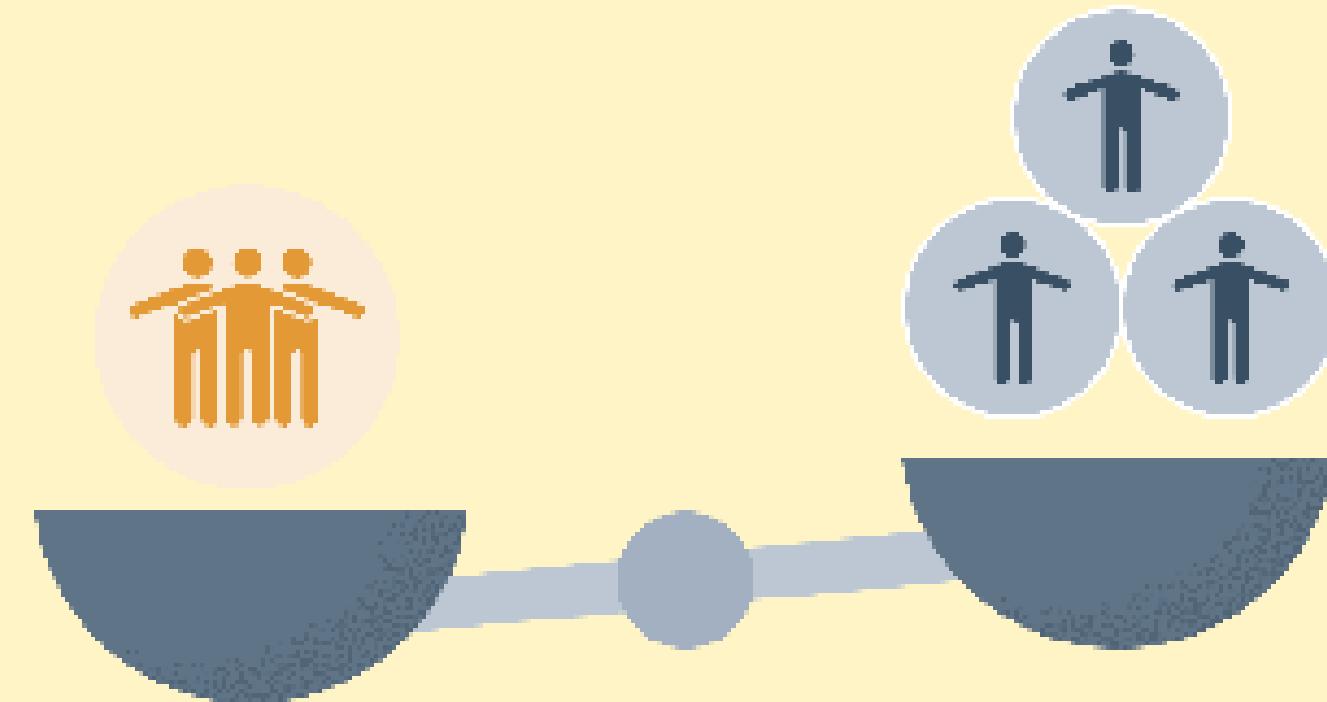




Positionality reflection

All individuals have **unique experiences** which shape their **thinking and perspectives**

This process of self-locating reflection is known as **positionality**



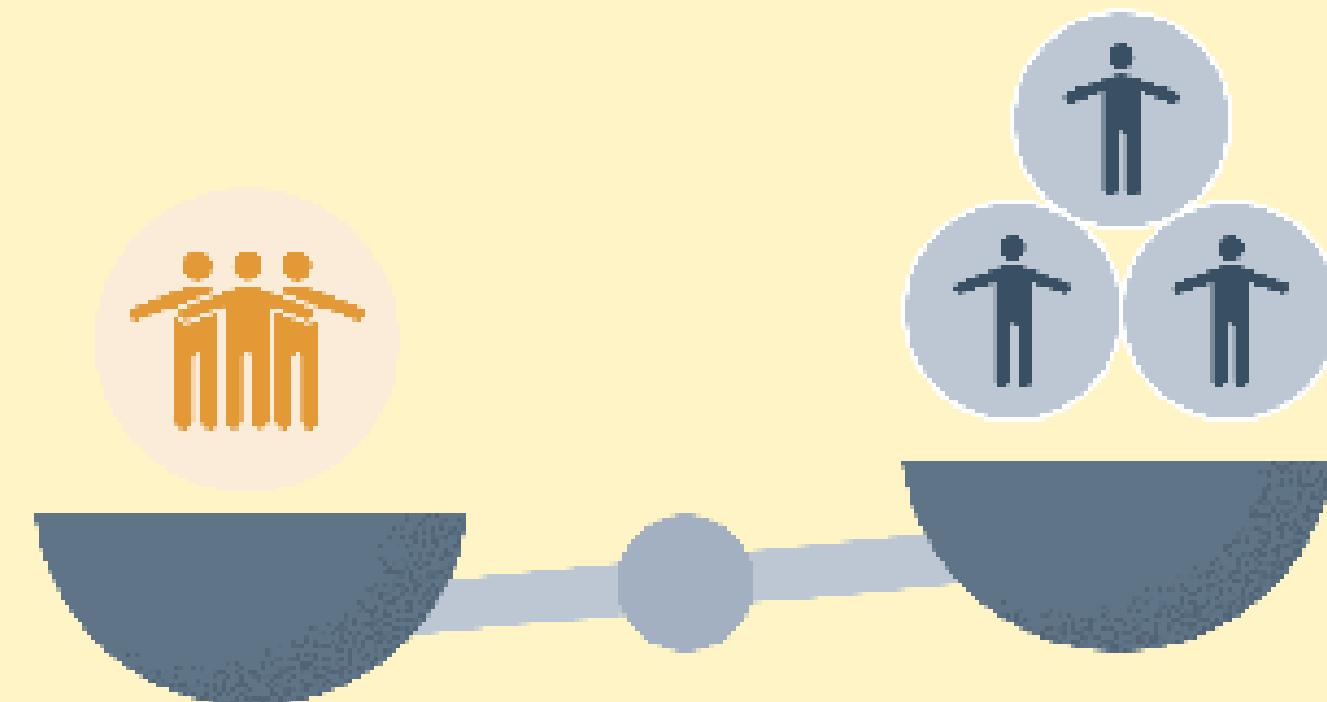


Positionality reflection

All individuals have **unique experiences** which shape their **thinking and perspectives**

This process of self-locating reflection is known as **positionality**

Understand how viewpoints differ from others with different **cultural and socioeconomic background**





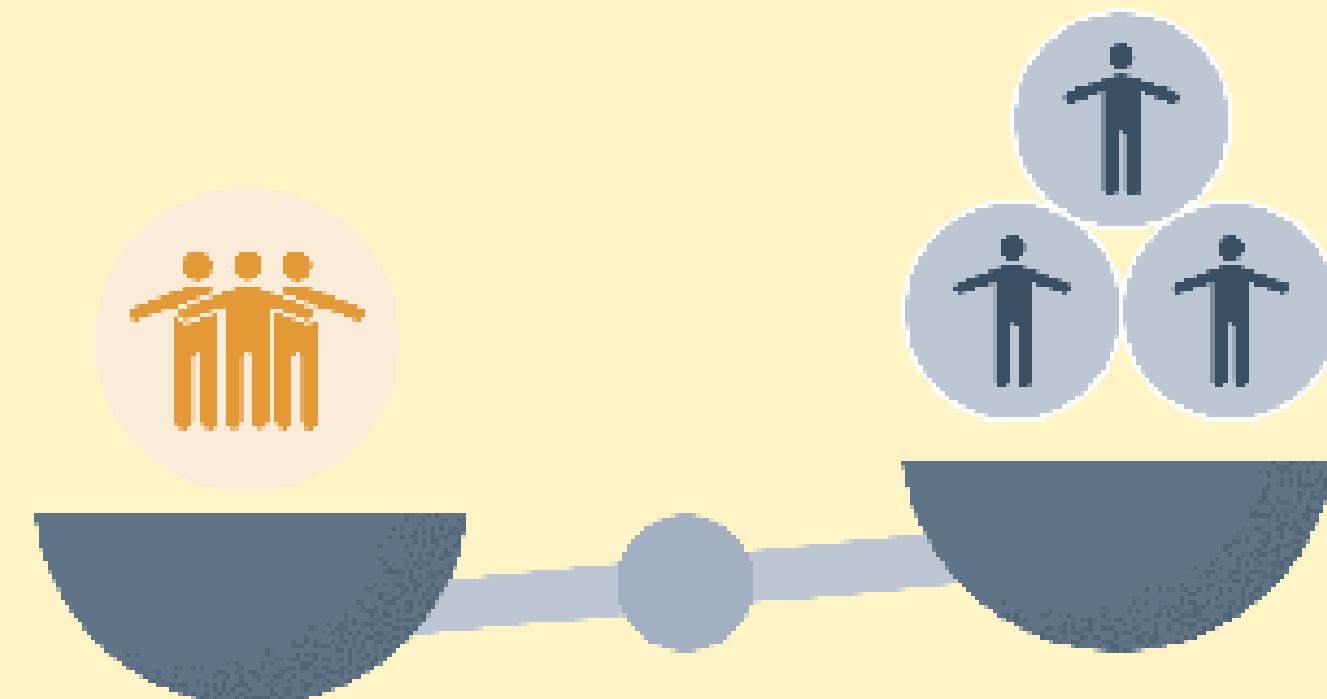
Positionality reflection

All individuals have **unique experiences** which shape their **thinking and perspectives**

This process of self-locating reflection is known as **positionality**

Understand how viewpoints differ from others with different **cultural and socioeconomic background**

Greater awareness of our own **biases** and **unconscious assumptions**





Positionality reflection

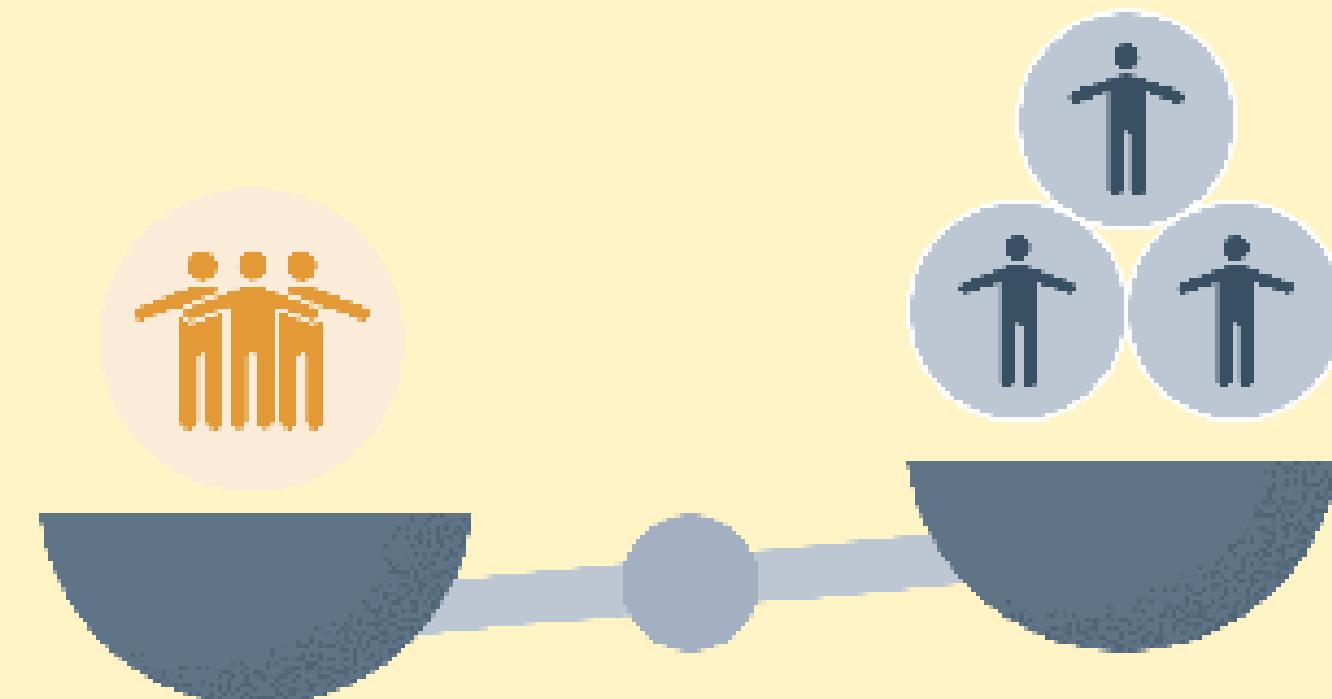
All individuals have **unique experiences** which shape their **thinking and perspectives**

This process of self-locating reflection is known as **positionality**

Understand how viewpoints differ from others with different **cultural and socioeconomic background**

Greater awareness of our own **biases** and **unconscious assumptions**

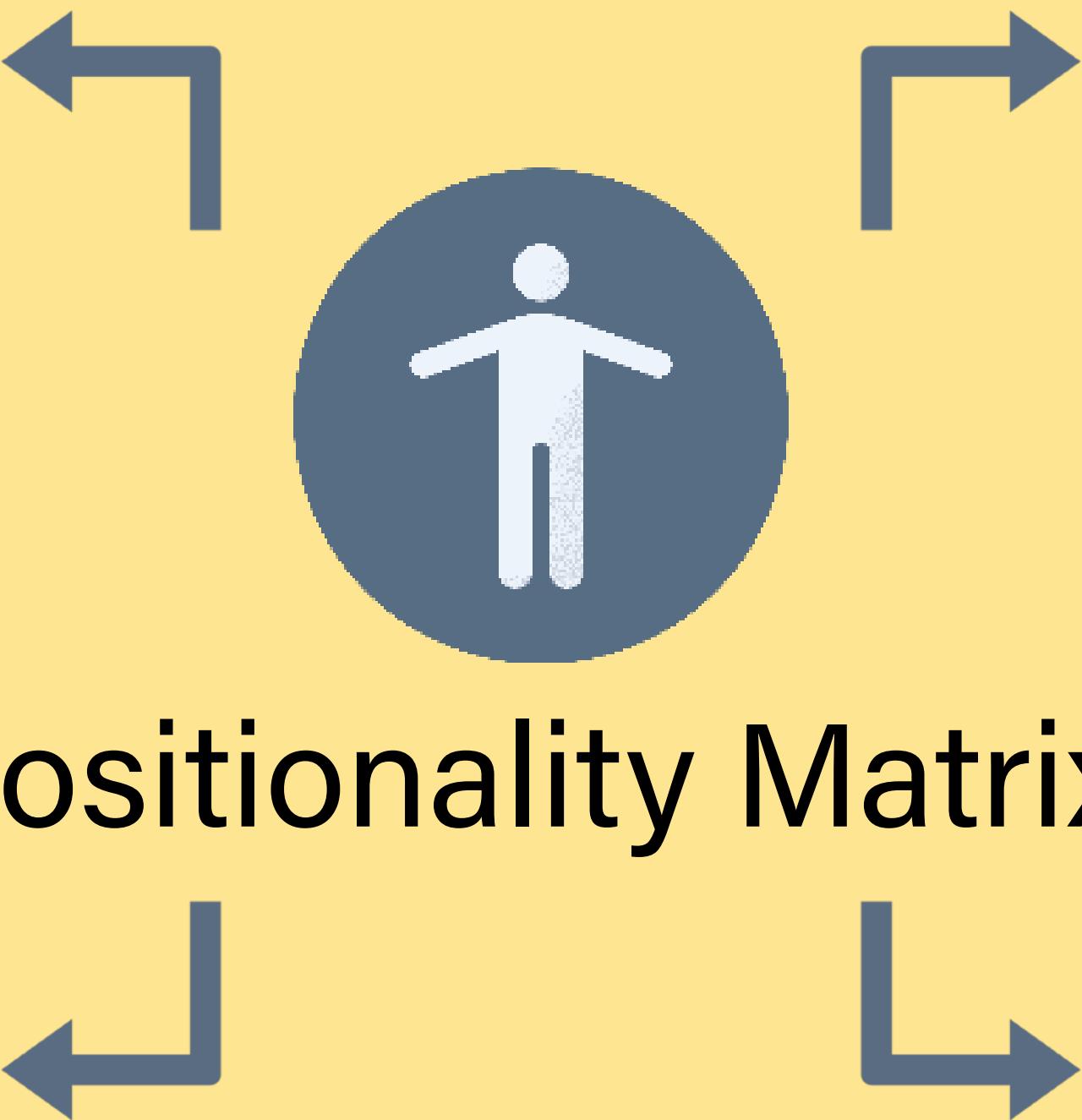
Influence of people's respective social and cultural positions creates **strengths and limitations**





Personal characteristics and group identifications

Education, training, and work background



Positionality Matrix



Institutional frame and team composition

Socioeconomic status



**what is my
socioeconomic
history?**
socioeconomic status
growing up, social
mobility over time,
present status,
socioeconomic
aspirations

Socioeconomic Status

how do I identify?

age, race & ethnicity, disability status, religion, gender,
sexuality, marital status, parental status, linguistic background

Personal Characteristics & Group Identifications

Positionality Matrix

To what extent do my personal characteristics, group identifications, socioeconomic status, educational, training, & work background, team composition, & institutional frame represent sources of power and advantage or sources of marginalisation and disadvantage? How does this positionality influence my (and my team's) ability to identify & understand affected stakeholders and the potential impacts of my project?

Education, Training, & Work background

**how have I been
educated and trained?**
schools attended,
level of education,
opportunities for
advancement and
professional
development,
employment history

Institutional Frame & Team Composition

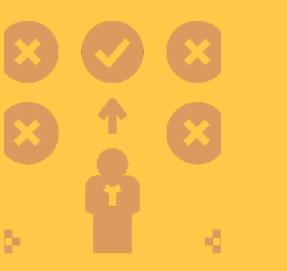
what does my institutional context and team composition look like?

authority structure within my project team, wider policy-ownership and power hierarchies in my organisation, levels of decision-making autonomy, opportunities to voice concerns & objections, team diversity, culture of inclusion or exclusion



Establishing a stakeholder engagement objective and engagement method





Establishing a stakeholder engagement objective and engagement method

Stakeholder engagement may be carried out in ways that involve **more-or-less stakeholder involvement**.



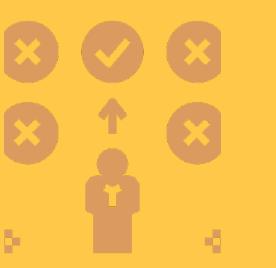


Establishing a stakeholder engagement objective and engagement method

Stakeholder engagement may be carried out in ways that involve **more-or-less stakeholder involvement**.

This should be **context-dependent** and vary from project to project





Establishing a stakeholder engagement objective and engagement method

Stakeholder engagement may be carried out in ways that involve **more-or-less stakeholder involvement**.



This should be **context-dependant** and vary from project to project

Take into account **a preliminary assessment of the potential risks and hazards** of the model or tool under consideration as well as the results of the positionality reflection.

INFORM



PARTNER



Establishing a stakeholder engagement objective and method

CONSULT



EMPOWER



Establishing a stakeholder engagement objective and method

SPECTRUM OF ENGAGEMENT OPTIONS

	DEGREE OF PARTICIPATION	MEANS OF PARTICIPATION	LEVEL OF AGENCY
EMPOWER	<i>Stakeholders are engaged with as decision-makers and are expected to gather pertinent information and be proactive in co-production.</i>	<i>Co-production exercises occur through citizens' juries, citizens' assemblies, and participatory co-design. Teams provide support for stakeholders' decision making.</i>	<i>Stakeholders exercise a high level of agency and control over agenda-setting and decision-making.</i>
PARTNER	<i>Stakeholders and teams share agency over the determination of areas of focus and decision making.</i>	<i>External input is sought out for collaboration and co-production. Stakeholders are collaborators in projects. They are engaged through focus groups.</i>	<i>Stakeholders exercise a moderate level of agency in helping to set agendas through collaborative decision-making</i>
CONSULT	<i>Stakeholders can voice their views on pre-determined areas of focus, which are considered in decision-making.</i>	<i>Engagement occurs through online surveys or short phone interviews, door-to-door or in public spaces. Broader listening events can support consultations.</i>	<i>Stakeholders are included as sources of information input under narrow, highly controlled conditions of participation.</i>
INFORM	<i>Stakeholders are made aware of decisions and developments.</i>	<i>External input is not sought out. Information flows in one direction. This can be done through newsletters, the post, app notifications or community forums.</i>	<i>Stakeholders are treated as information subjects rather than active agents.</i>

Questions?



Activity 1: Stakeholder analysis

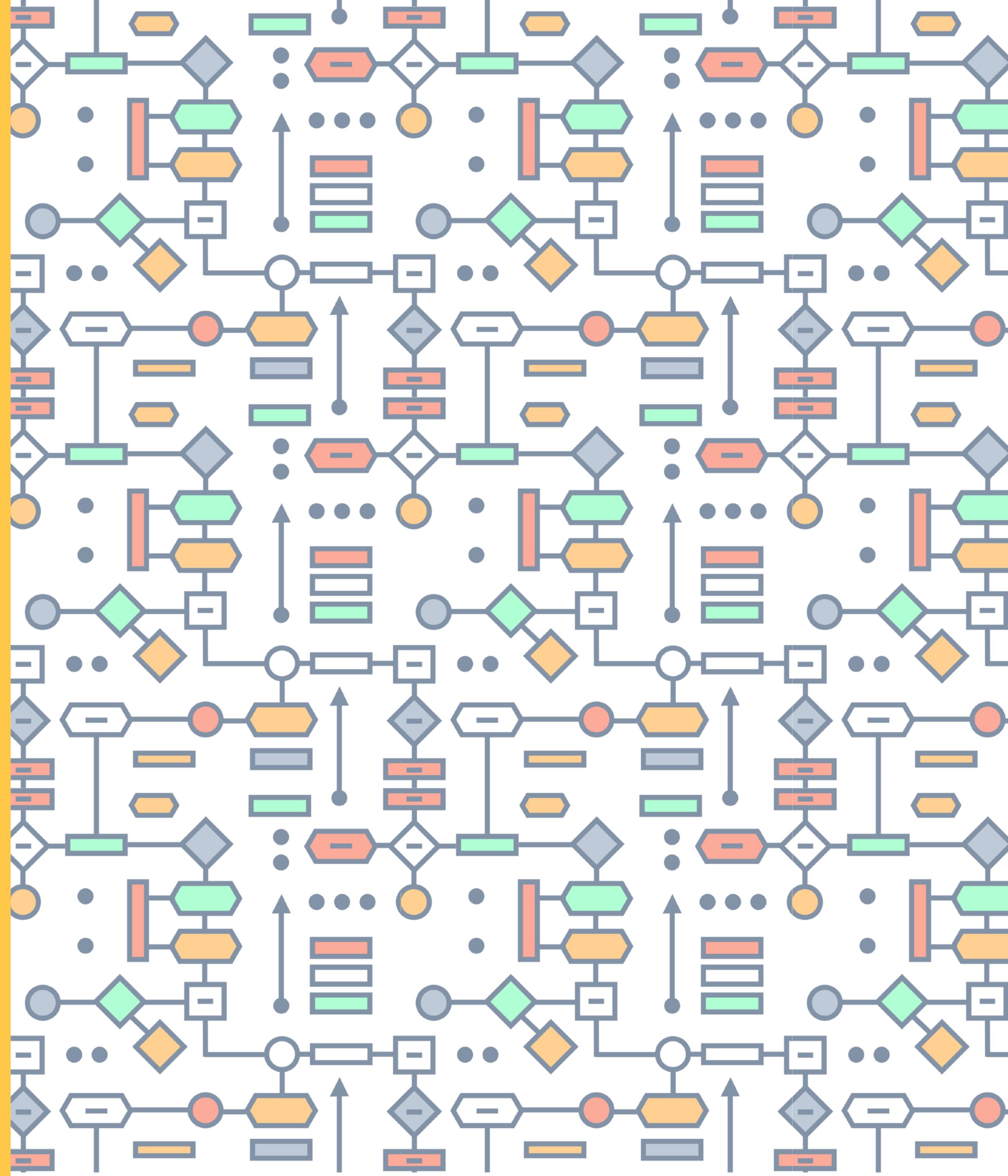
Go to this link [insert link for MIRO Board]

Activity 2: Positionality reflection

Go to link [insert link for MIRO Board]



2 STAKEHOLDER IMPACT ASSESSMENT



To ensure the deployment of an AI system remains sustainable, there should be a continuous monitoring of the real-world impacts that the system has.

The team should come together to evaluate the social impact and sustainability of an AI project through what is known as a Stakeholder Impact Assessment (SIA).

To ensure the deployment of an AI system remains sustainable, there should be a continuous monitoring of the real-world impacts that the system has.

The team should come together to evaluate the social impact and sustainability of an AI project through what is known as a Stakeholder Impact Assessment (SIA).

Objectives of a stakeholder impact assessment

Objectives of a stakeholder impact assessment

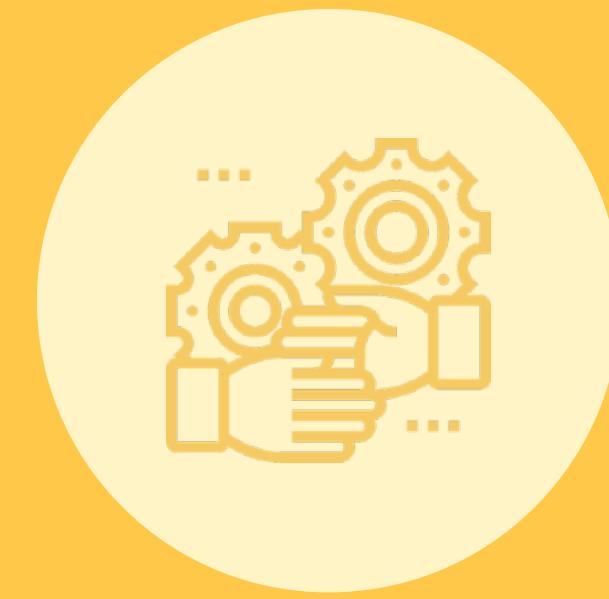


Build public trust

Objectives of a stakeholder impact assessment



Build public trust

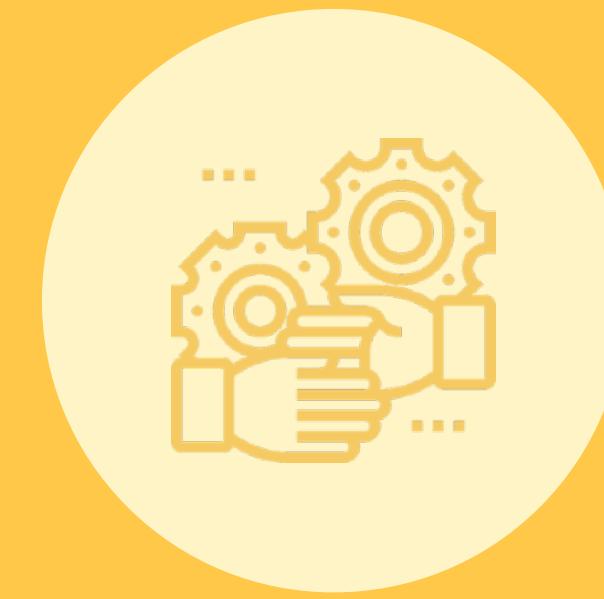


Facilitate and strengthen
accountability framework

Objectives of a stakeholder impact assessment



Build public trust



Facilitate and strengthen
accountability framework

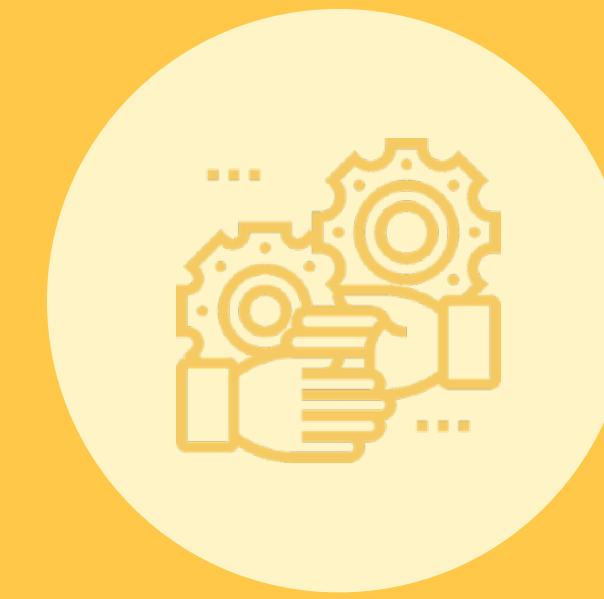


Highlight unseen
risks

Objectives of a stakeholder impact assessment



Build public trust



Facilitate and strengthen
accountability framework



Highlight unseen
risks

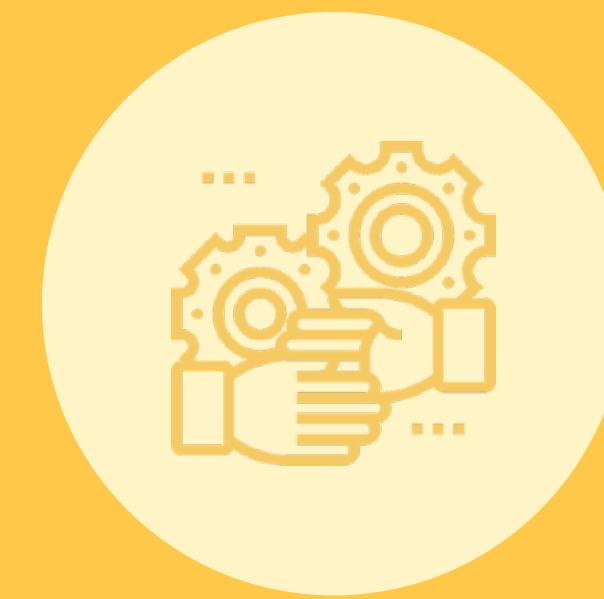


Support well-informed
decision-making and
transparent innovation

Objectives of a stakeholder impact assessment



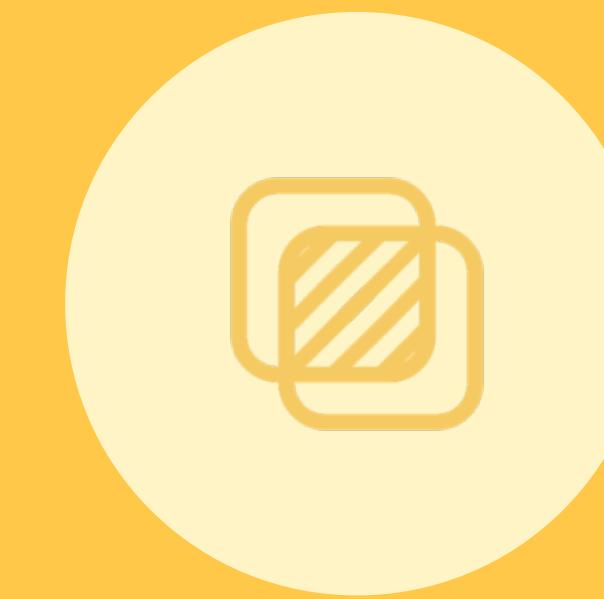
Build public trust



Facilitate and strengthen
accountability framework



Highlight unseen
risks



Support well-informed
decision-making and
transparent innovation



Demonstrate forethought
and due-diligence

Stakeholder Impact Assessment

- A SIA creates a procedure and documents for collaborative evaluation and reflective anticipation of the possible harms and benefits of AI innovation projects

Stakeholder Impact Assessment

- ▶ A SIA creates a procedure and documents for collaborative evaluation and reflective anticipation of the possible harms and benefits of AI innovation projects
- ▶ SIA's should be integrated into the wider impact assessment regime

Stakeholder Impact Assessment

- ▶ A SIA creates a procedure and documents for collaborative evaluation and reflective anticipation of the possible harms and benefits of AI innovation projects
- ▶ SIA's should be integrated into the wider impact assessment regime
- ▶ SIA should be carried out at the design, development, and deployment of the AI system

Skills for conducting SIA's



Weighing values and considering trade-offs

Consequences-based

Principles-based

Skills for conducting SIA's

Ensuring meaningful and inclusive deliberation

Impartiality

Consistency and coherence

Non-coercion

Mutual respect & egalitarian reciprocity

Sincerity

Inclusivity & publicity



Addressing and mitigating power dynamics that may obstruct meaningful discussion

Any questions?



Thank you!

See you tomorrow for Day 4: Fairness, bias,
accountability and governance