Data Innovation Toolkit

Digital innovationLab

European Commission

Publications Office of the European Union

ABOUT THIS PUBLICATION

This publication is an initiative of the Digital Innovation Lab (iLab), an open platform within the European Commission promoting the innovation of the public sector. The iLab collaborates with public administrations, academia, and industry to develop sustainable methodologies and tools for better policy-making and service delivery. Its mission is to incubate innovation and explore opportunities from business cases, new organisational models, and data technologies.

For more information about this publication, please contact:
European Commission, Directorate-General for Digital Services. Unit B.1 – Data, Artificial Intelligence & Web
Email: DIGIT-DIGITAL-INNOVATION@ec.europa.eu

Authors
Arianna Fini Storchi (https://orcid.org/0009-0002-6419-7888)
Maria-Claudia Bodino (https://orcid.org/0009-0001-2407-3760)
Nathan da Silva Carvalho (https://orcid.org/0000-0001-7824-8629)
Stefaan Verhulst (https://orcid.org/0000-0002-0107-2984)

LEGAL NOTICE

This document should not be considered as representative of the European Commission's official position. Neither the European Commission, nor any person acting on behalf of the European Commission is responsible for the use which might be made of the following information. The views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission. More information on the European Union is available on the Internet (http://europa.eu).
Luxembourg: Publications Office of the European Union, 2024

EPub	ISBN: 978-92-68-20779-6	doi: 10.2799/2716475	NO-01-24-000-EN-E
HTML	ISBN: 978-92-68-20780-2	doi: 10.2799/4555536	NO-01-24-000-EN-Q
PDF Web	ISBN: 978-92-68-20781-9	doi: 10.2799/8924292	NO-01-24-000-EN-N
PDF/X	ISBN: 978-92-68-20782-6	doi: 10.2799/6746549	NO-01-24-000-EN-C

© European Union, 2024
Reproduction is authorised provided the source is acknowledged.

The Commission's reuse policy is implemented under Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39, ELI: http://data.europa.eu/eli/dec/2011/833/oj). Unless otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (https://creativecommons.org/licenses/by/4.0/). This means that reuse is allowed, provided appropriate credit is given and any changes are indicated.

Digital innovationLab

European Commission

Publications Office of the European Union

GETTING IN TOUCH WITH THE EU

In person

Hundreds of Europe Direct centers are spread all over the European Union. You can find the center nearest to you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union. You can contact this service:
— by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
— at the following standard number: +32 22999696,
— via the following form: european-union.europa.eu/contact-eu/write-us_en.

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union is available in all EU official languages on the Europa website (europa.eu).

EU publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law in all official languages since 1951, go to EUR-Lex (eur-lex.europa.eu).

Open data from the EU

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wide range of datasets from EU Member States.

Digital innovationLab

European Commission

Publications Office of the European Union

The digital innovation Lab - iLab

The digital innovation lab, commonly referred to as iLab, is an open platform for digital innovation within the Commission and aims to create links and collaboration with external stakeholders including local and national public administrations, academia and the industry. The iLab aims at developing innovative, sustainable and responsive methodologies, models and tools for improved policy-making and service delivery. Its mission is to act as an incubator for innovation with a foresight capacity to assess opportunities stemming from concrete use/business cases, innovative organisational models, and a special focus on cutting-edge data technologies.

Digital innovationLab

Introduction

European Commission

What is the data innovation toolkit?

The Data Innovation Toolkit is designed to advance data-driven innovation in the public sector, with **four key stages to guide your work: Planning, Collect and Process, Share and Analyse, and Use and Evaluate.** Each stage focuses on practical steps to help you make meaningful use of data. Also, to support you along the way, the **Data Innovation Repository** offers an external collection of resources to provide you with extra support in each of these stages.

Who is it for?

It's designed to assist a diverse group of individuals involved in data innovation initiatives, **from novice analysts to experienced policymakers and tech-savvy innovators.** The toolkit addresses the different needs and challenges of its users, ensuring it delivers widespread benefits tailored to different audiences in the public sector. The toolkit was also created based upon many interviews and workshops with different data experts.

How to use it?

The toolkit is designed for use by teams or groups working through the tools together. These **resources can be utilised in a non-linear, interactive manner — moving back and forth between them**, combining them, or following the suggested structure. To facilitate browsing, the resources have been grouped into these four stages.

Data innovation toolkit personas

European Commission

Maria - Novice analyst

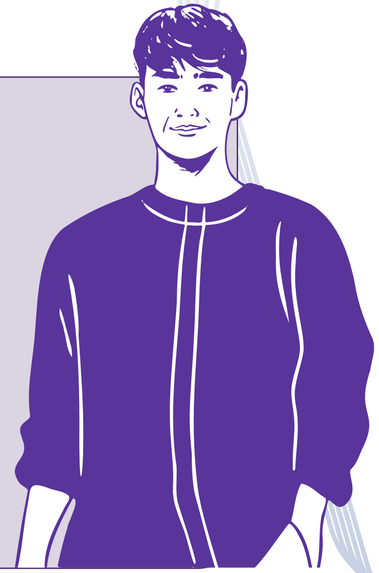
New to data analysis, Maria feels overwhelmed by technical aspects but is eager to learn data foundational skills for her project. Maria should start with **introductory resources,** focusing on the tools that break down the basics of data collection and analysis to build her confidence step-by-step.

Jean-Claude - Experienced policy maker

Jean-Claude is integrating data into environmental work but struggles to balance learning new skills with his heavy workload. **Concentrate on tools that streamline data insights into evaluation and policy recommendations**, such as the data innovation scalability plan - allowing for informed decisions without extensive technical involvement.

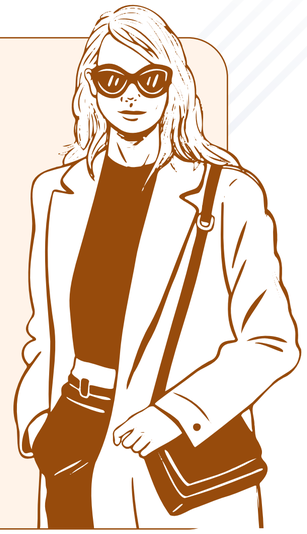
Alex - Tech-savvy innovator

Skilled in technology, Alex seeks to use advanced analytics for policy impact and faces challenges convincing non-technical colleagues of data's benefits. **Explore data processing and scalability tools,** focusing on resources that clearly demonstrate value and can serve as case studies to build buy-in from stakeholders.



Helena - Data sceptic

Cautious about adopting new data methods, Helena needs clear evidence of practical benefits to support her division's budget management. She may want to prioritise tools that to **help her in analysing and sharing the data that she already has** with the stakeholders to demonstrate that the project can be executed.



Data innovation toolkit personas

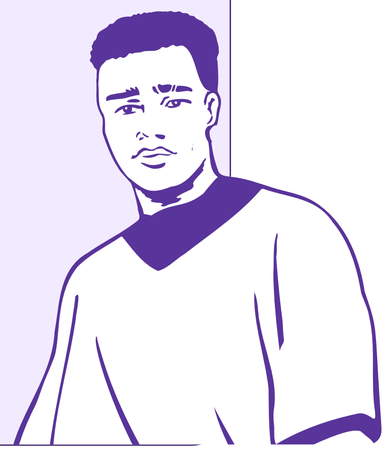
European Commission

Catherine - Collaborative networker

Focused on fostering EU partnerships, Catherine navigates diverse data landscapes across Member States. Catherine can utilise tools that enhance cross-border collaboration, **especially stakeholder mapping** and data-sharing frameworks that standardise and streamline cooperative efforts across varied datasets and actors.

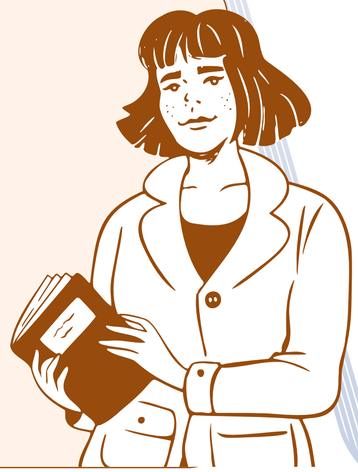
Thomas - Governance guardian

Thomas balances data accessibility with strict protection standards. He can emphasise tools that reinforce data quality and security, **such as the data gaps checklist** to ensure data use aligns with EU regulations and ethical guidelines for the project he is working on.



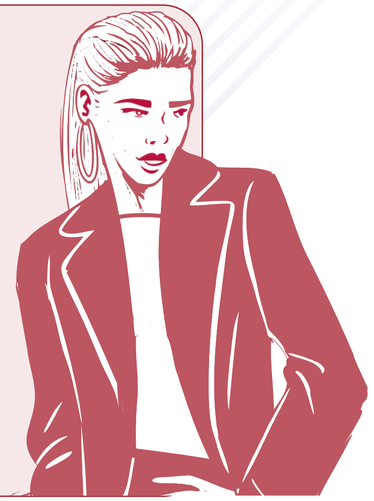
Sofia - Stewardship strategist

Sofia faces resistance to open data standards. She may want to focus on resources that promote interoperability and **data-sharing protocols, using case studies** and examples from research communities to advocate for open data practices and demonstrate collaborative benefits with the support of the available tools.



Olivia - Insight-driven decision maker

Olivia has experience with data but is primarily focused on high-level insights that directly support decision-making and strategic direction. If she is developing a project - she may want to prioritise tools that focus **on high-level analysis and impact evaluation resources.**



Disclaimer

About this publication

This toolkit is a work in progress and may be updated in the future with additional tools and insights related to data innovation projects. Developed through extensive research and numerous expert interviews, the toolkit is organized into sections, allowing you to follow the defined sequence or use tools as needed based on your specific requirements."

Please contact DIGIT-DIGITAL-INNOVATION@ec.europa.eu for more information or if you have any feedback on the material.

Data innovation toolkit structure

European Commission

4. Use and Evaluate

Are you ready to put your insights to work and see how well they're helping you and others?

This stage translates insights into action, communicates findings, manages data retention, measures success, and plans for scalability and future strategies.

3. Share and Analyse

Do you want to share what you've collected and start looking for patterns or insights?

Here, you ensure secure and compliant data transfers, addresses legal and ethical considerations, and involves exploratory and targeted data analysis to generate actionable insights aligned with your project's objectives.

1. Planning

Are you figuring out what you want to achieve and what you'll need to get there?

The planning stage involves defining the problem, identifying stakeholders, building a team, spotting data gaps, and understanding your users' perspective.

2. Collect and Process

Is it time to gather your data and make sure it's ready to use?

This stage involves the collection of data that will be used in your data innovation project, data quality, data storage, security, privacy concerns and data preparation to enable analysis.

Data Innovation Repository

Need additional support? The Data Innovation Repository is a compendium of resources designed to help you find materials to support your data innovation journey. Access it via: https://is.gd/datainnovationrepository

Data innovation toolkit: contents

European Commission

1 Planning 17

Challenge definition tool
How do I clearly define the challenge my data innovation project aims to address?

Stakeholder mapping tool
Who are the key stakeholders, and how do I map them effectively?

Data team creation
How do I build the right team for my data innovation project?

Role's definition tool (RACI matrix)
What roles and responsibilities should each team member have?

Persona definition tool
How do I create personas that accurately reflect the diverse behaviours and needs of my project's target users?

Data journey mapping - user perspective
How do I map the user journey to identify key actions and opportunities for improvement?

2 Collect and process 42

Data gaps checklist
Are there any potential gaps in the data I plan to use, and how can I address them?

Data collection template
How do I ensure that my data collection process is aligned with my challenge statement?

Data collection checklist
Can I access the correct data sources and the data itself to execute the project?

Data quality checklist exercise
Am I capturing high-quality data that meets my project's needs?

Data processing cards
Am I ensuring that all data processing activities are conducted securely, efficiently, and in compliance with regulations?

3 Share and analyse 68

Data sharing template
How can I ensure that my data-sharing process is secure, legally compliant, and transparent?

Data sharing checklist
Have I taken all necessary steps to share data securely, ethically, and in a way that supports project goals?

Data analysis template
Am I conducting thorough and accurate data analysis to generate meaningful insights for my project?

4 Use, Evaluate 83

Data using checklist
How can I ensure that the data and insights are used effectively and responsibly to achieve my project's objectives?

Data innovation KPI creation tool
What are the key performance indicators (KPIs) I need to measure success?

Measuring impact and evaluating success tool
How do I measure the impact and success of my data innovation project?

Data innovation scalability plan
What strategies should I use to scale my project effectively?

Data innovation toolkit keywords

European Commission

Term	Definition
Data analysis checklist	A tool to guide teams through the process of conducting thorough and accurate data analysis, ensuring the use of appropriate tools and techniques to generate meaningful insights.
Data collection checklist	A checklist designed to ensure that all critical factors in the data collection process are considered, including privacy, legality, and data quality.
Data cycle	A cyclical process in data projects that involves stages from planning to using data, where insights from one project inform the next, creating a loop of continuous improvement.
Data innovation repository	A consolidated resource that includes tools, templates, checklists, and models to guide project developers through each step of the data innovation process.
Data innovation risk assessment	A process used to identify and manage risks throughout a data innovation project, ensuring compliance with regulations and mitigating risks associated with data management and technology changes.
Data innovation scalability plan	A plan designed to guide the scaling of data projects, ensuring they grow effectively while maintaining quality, compliance, and impact.
Data innovation toolkit	A structured resource designed to simplify and manage data projects, breaking down the innovation process into clear steps like planning, data collection, analysis, and evaluation.
Data processing checklist	A checklist that helps teams ensure data is cleaned, transformed, securely stored, and efficiently prepared for analysis, maintaining compliance and security standards.
Data quality checklist	A tool used to assess and improve data quality, focusing on factors like accuracy, timeliness, and consistency, ensuring data is fit for its intended purpose.
Data responsibility journey	A framework for ensuring that all data innovation projects are conducted with ethical responsibility, ensuring data protection and security throughout the project lifecycle.
Data sharing checklist	A checklist used to ensure data is shared securely, legally, and transparently, facilitating collaboration while protecting sensitive information.
Data sharing template	A structured tool for organizing and managing the process of data sharing, ensuring compliance with privacy laws and security protocols while facilitating transparent data exchange.
iLab	The European Commission's Innovation Lab (iLab), focused on supporting data-driven innovation through methodologies, tools, and frameworks that enhance policymaking and service delivery.
Key performance indicators (KPIs)	Metrics used to measure the success and progress of a project, ensuring goals are met in a timely and quantifiable manner.
Measuring impact and evaluating success tool	A tool used to assess the overall impact and success of a data innovation project, focusing on both short-term outcomes and long-term effects.
Problem definition tool	A tool that helps teams clearly define and articulate the problem they aim to solve, identifying key factors and stakeholders involved.
Responsibility assignment matrix (RACI)	A responsibility assignment tool that defines team roles for tasks or decisions using four categories: Responsible, Accountable, Consulted, and Informed.
Scalability plan	A strategic framework to guide the expansion of a data innovation project, addressing challenges related to data quality, compliance, and resource management.
Stakeholder mapping tool	A tool used to identify, assess, and visualize stakeholders' impact and influence, categorizing them for effective engagement throughout the project.

EU data policy frameworks

European Commission

The table below introduces key EU policies that are shaping how data and digital tools are used across sectors. Understanding these policies can help you in having an overview of the EU landscape to drive data innovation, as they set the ground rules for how data can be accessed, shared, and protected. Each policy tackles a different part of the digital world—from data privacy to fair competition and AI safety—creating a setting that supports secure, fair, and innovative use of data in different sectors.

Policy name	Description
General data protection regulation (GDPR) (2016/679)	Sets a high standard for data protection and privacy, giving individuals greater control over their personal information. GDPR enforce stringent rules on data handling, aiming to protect citizens' privacy and standardize data protection across the EU.
Open data directive (2019/1024)	Supports transparency and innovation by facilitating access to high-value datasets held by public sector bodies. The Directive mandate open access to public data, encouraging its reuse to drive new digital products and services, stimulating economic growth across El sectors.
AI act (2021/0106)	Introduces regulations for artificial intelligence to mitigate risks associated with AI while promoting its development in Europe. The Ac seeks to establish safety standards, address ethical concerns, and ensure that AI development aligns with EU values.
Data act (2022/0047)	Establishes a comprehensive framework to enable fair access to and use of data across the EU. The Data Act introduces measures t ensure that businesses, especially SMEs, can benefit from the value of data generated in the EU, fostering innovation, competitiveness, an data-driven growth.
Data governance act (2022/868), 2022	Aims to create a trusted framework for secure data sharing and reuse within the EU. By encouraging transparency, data altruism, an secure data intermediation, it fosters a data-sharing culture while maintaining strict privacy and data protection standards.
Digital markets act (2022/1925)	Regulates the practices of large digital platforms, known as 'gatekeepers,' to promote fair competition in digital markets. By preventin anticompetitive practices and monopolistic control, it aims to create a more open and competitive digital economy in the EU.
Digital services act (2022/2065)	Strengthens accountability and transparency requirements for online platforms to create a safer digital space. By targeting illegal conter and improving content moderation practices, it aims to protect users' rights while supporting online freedoms and innovation.

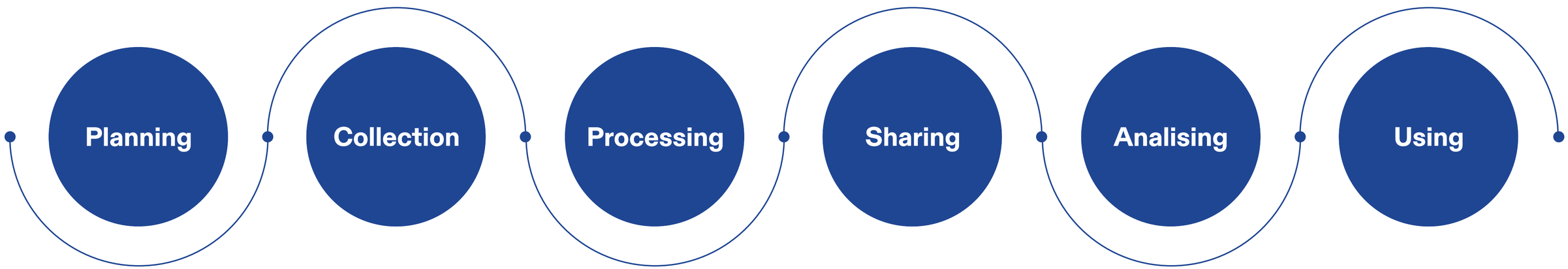
European Commission

The data lifecycle within data innovation projects

Which step are you in?

The data lifecycle shows how each project's insights shape the next, creating a cycle of continuous learning and improvement. It's a process where each project builds on the last, driving better data handling and analysis over time.

The data lifecycle



This lifecycle can be split into six stages

European Commission

Planning

Defines the expected value of the project, sets goals and metrics, identifies stakeholders, and develops a flexible plan focused on effective governance and data lifecycle management.

Collecting

Gathers and validates essential data, prioritising privacy protection and regulatory compliance, while excluding unnecessary information.

Processing

Involves storing, organising, and preparing data for analysis, ensuring security and efficiency, through strategies like data cleaning, transformation, and storage management.

Sharing

Coordinates operational activities, promotes collaboration, and ensures the secure and ethical transfer of data, benefiting stakeholders through data initiatives.

Analysing

Applies algorithms for exploratory and specific analyses, generating quantitative outcomes and uncovering insights using statistical modelling, machine learning, and artificial intelligence.

Using

Communicates results and makes decisions on data retention or disposal, focusing on resource optimisation and ensuring the responsible application of the insights gained.

What are the different types of data?

European Commission

Types of data and how to innovate with it

Qualitative data (e.g., user feedback, interview responses)

Quantitative data (e.g., statistics, metrics on service usage)

Nominal data: Categorical data without order (e.g., service types like healthcare or education). May be used to identify high-demand areas to tailor resources and improve service delivery.

Ordinal data: Ordered data (e.g., satisfaction ratings). You can use it for rankings to prioritise improvements in lower-rated areas.

Discrete data: Countable values (e.g., number of complaints). You may want to use it to focus on high-complaint areas for targeted enhancements.

Continuous data: Measurable values (e.g., response times). Track and optimise for faster, more efficient service delivery.

The data innovation toolkit guides you through each phase of the data lifecycle to support your project, following a four-step framework that spans from gathering initial insights to building on previous learnings. This toolkit fosters continuous improvement, enabling your team to make smarter, data-driven decisions at every stage.

Are you figuring out what you want to achieve and what you'll need to get there?

Planning

The planning stage involves defining the challenge, identifying stakeholders, building a team, spotting data gaps, and understanding the user perspective.

European Commission

What is happening at this stage?

Every data innovation project starts with laying the foundation: clarifying the value it will bring, setting clear goals and success metrics, identifying the people involved, and assigning roles. Then, you will build a flexible plan to keep it on track, ensuring good governance and smooth data management so the project can innovate and adapt as it grows.

Why?

This stage sets the direction and objectives, ensuring team alignment and clarity. Effective planning helps anticipate potential challenges, allocate resources efficiently, and integrates data governance from the outset, which is essential for compliance and sustainability.

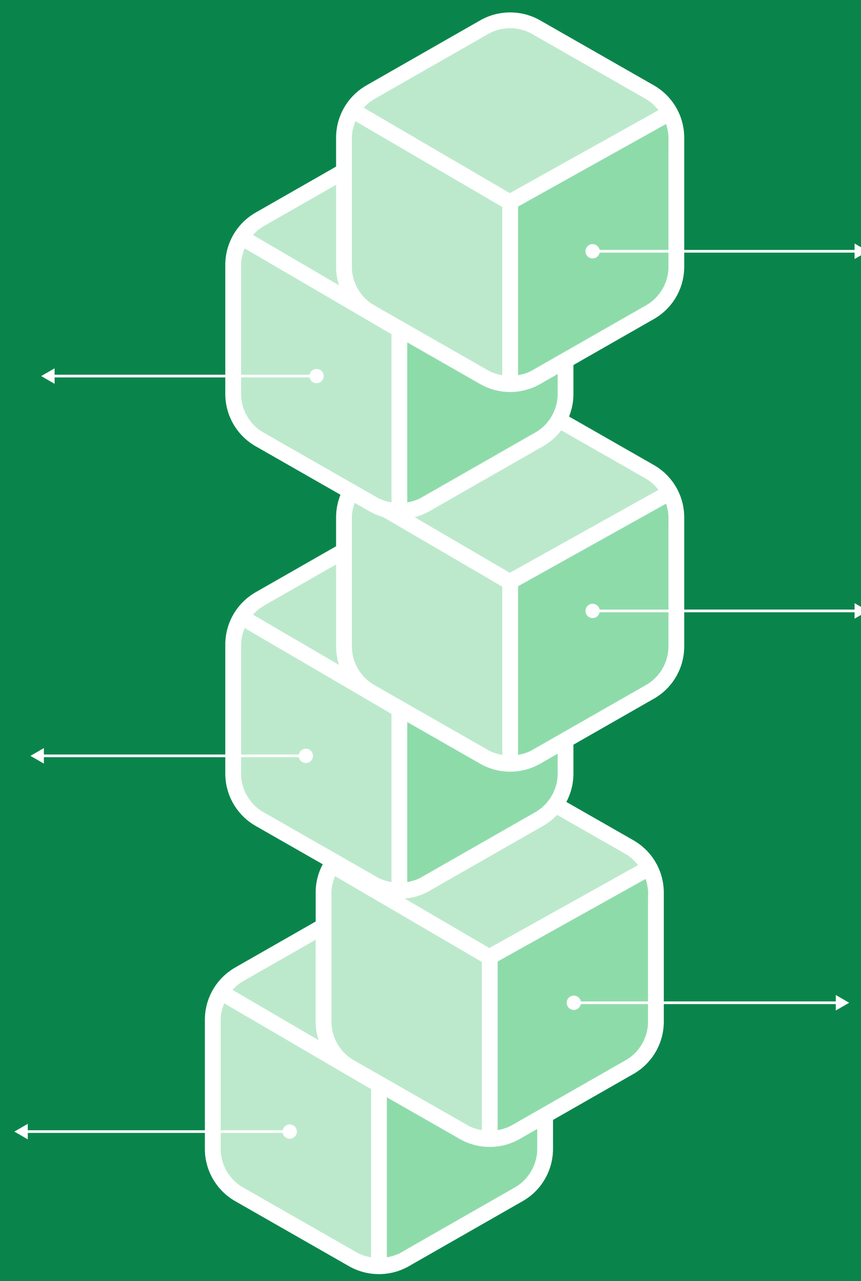
Considerations and risks

- Unclear objectives may lead to inefficiencies.
- Stakeholder misalignment can cause resistance or lack of support.
- Overlooking data gaps affects the quality of outcomes.
- Ambiguous roles hinder collaboration and accountability.

Success factors

- Clear, measurable goals.
- Effective stakeholder engagement.
- Comprehensive and flexible planning.
- Well-defined roles and responsibilities.
- Integration of data governance practices.

Tools in this chapter



Stakeholder mapping tool

Who are the key stakeholders, and how do I map them effectively?

Data Innovation project role's definition

What roles and responsibilities should each team member have?

Data journey mapping

How do I map the user journey to identify key actions and opportunities for improvement?

Challenge definition tool

How do I clearly define the challenge my data innovation project aims to address?

Data team creation

How do I build the right team for my data innovation project?

Persona creation tool

How do I create personas that reflect the diverse behaviours and needs of my project's users?

European Commission

How do I clearly define the challenge my data innovation project aims to address?

Pre-requisite tools:

None

Next tool:

Stakeholder mapping tool

European Commission

Challenge definition tool

How do I clearly define the challenge my data innovation project aims to address?

Resource type	Complexity	Time frame
Template	Easy	40 minutes

Who may support you in developing this resource?

Project manager, communication officer, data owner, data governance officer, legal adviser.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

- Toolkit - Problem definition toolkit - Nesta
- Toolkit – A Guide to Data Innovation for Development, from idea to proof-of-concept - UNDP
- How to Master the Practice of Defining Problems - GovLab

What is it for?

The Challenge Definition Tool is an instrument that helps teams precisely **identify and articulate the key issues** they aim to resolve, summarizing them in a single statement. It is characterised by a defined set of questions that delve into issues of relevance, contributing factors, affected stakeholders, existing initiatives, and potential data sources. The key objective is to capture, compare and discuss different viewpoints before focusing on a clear challenge on which you can base the rest of your data innovation project.

How to use it?

Step 1: Quickly review the challenge definition tool guiding questions

The canvas comes with different set of questions to help you formulate your challenge. Please read them and use post-its notes to write your answers.

Step 2: Have an open discussion with your team or even external stakeholders

It is recommended to validate and brainstorm your initial answers with more people with different knowledge/backgrounds. Having a very precise and clear understanding on the challenge you want to solve is imperative for innovation.

Step 3: Create your challenge statement

When you have filled all the fields in the canvas and discussed them with others, please create your challenge statement in the box provided.

What will you accomplish?

Identify the challenge to be addressed, achieve a shared understanding among team members, gain awareness of existing initiatives, and identify potential data sources.

European Commission

Challenge definition tool

European Commission

1. What is the key issue you are trying to solve and why is it relevant?

Clearly define the primary issue you aim to address and explain its significance to your organisation or stakeholders.

Example: Inefficient urban mobility causing congestion and longer commute times.

2. What factors contribute to this issue?

Identify and describe the contextual factors that contribute to the issue.

Example: Overreliance on private vehicles, limited public transport options, inadequate cycling infrastructure.

2. What factors contribute to this issue?

3. Who is already working on this issue? Are there any other initiatives addressing this issue?

Identify any existing efforts or initiatives, internal or external, that are addressing the same issue.

Example: City council transport departments, local cycling advocacy groups, public transport authorities.

4. What data, relevant to this issue, exist already?

Identify existing data sources that are relevant to the issue.

Example: Traffic flow statistics, public transport usage data, road network maps.

5. Based on the factors above, what potential new or additional data sources could provide insights into this issue?

Suggest new or additional data sources that could offer further insights into the issue.

Example: Real-time GPS data from vehicles, mobile app data on commuter preferences, environmental pollution levels.

6. What question are you trying to address with data?

Formulate the key question that your project seeks to answer using data. This question should be specific, focused, and directly linked to the challenge you identified.

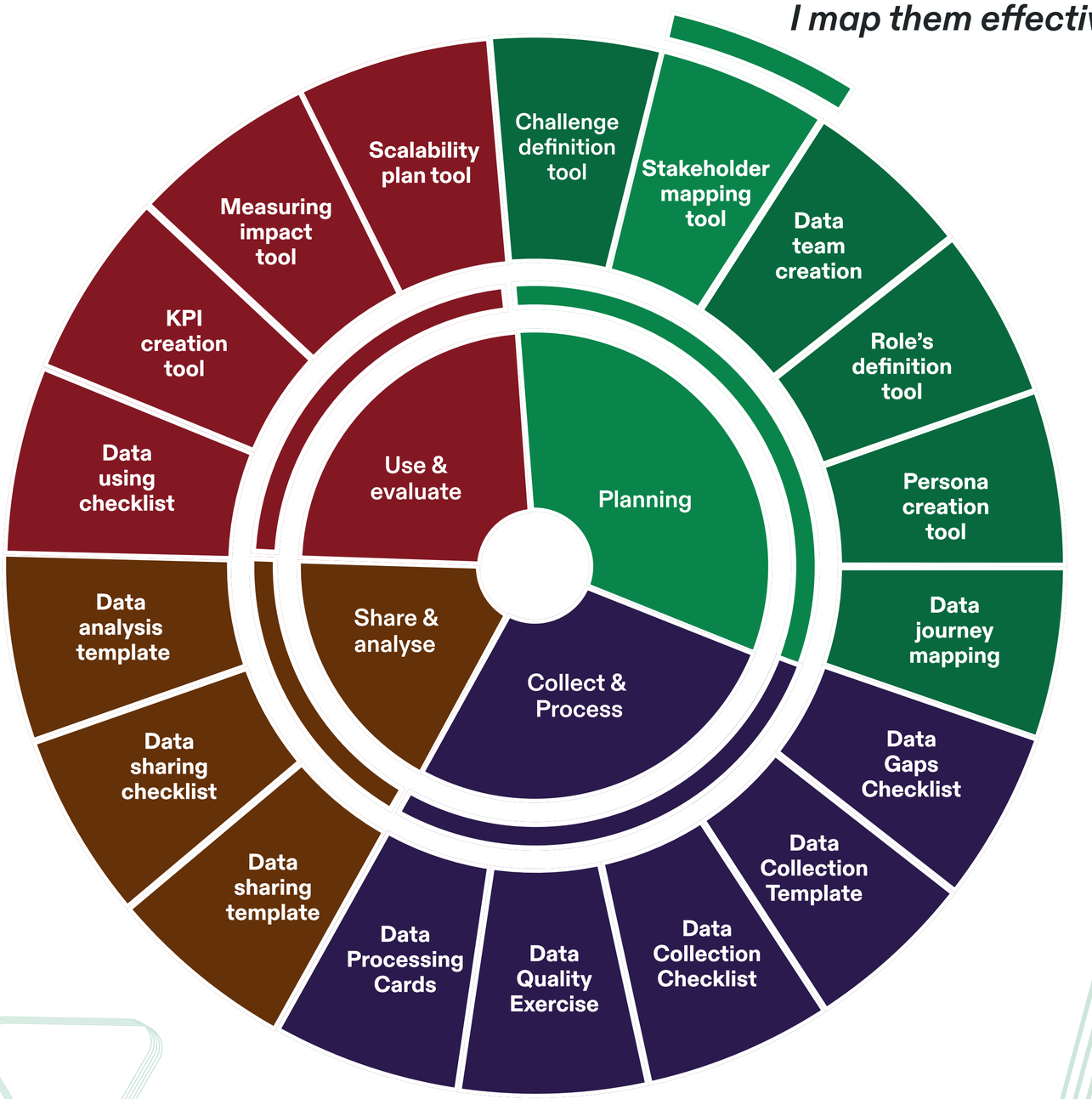
Example: How can we optimise urban transport networks to reduce congestion and improve commute times?

7. Reframe your initial issue as a challenge statement:

Finally, combine the insights from the previous answers to create a clear and concise challenge statement for the project.

Example: Develop a data-driven solution to enhance urban mobility by optimising transport networks and reducing congestion.

Who are the key stakeholders, and how do I map them effectively?



Pre-requisite tools:

Challenge definition tool

Next tool:

Data team creation

European Commission

Stakeholder mapping tool

Who are the key stakeholders, and how do I map them effectively?

Resource type	Complexity	Time frame
Canvas	Medium	40 minutes

Who may support you in developing this resource?

Communication officer, project manager, data owner, data custodian, data governance officer, data visualisation specialist.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

Guide - Ensuring Effective Stakeholder Engagement – UK Government Communication Service

What is it for?

Engaging the right people is crucial for data innovation projects. The Stakeholder Mapping Tool helps you by visualising the network of individuals and organisations involved, **identifying key stakeholders,** assessing their influence and interests, categorising engagement levels, and developing communication strategies.

How to use it?

Step 1: Read the examples of possible stakeholders

Check the potential list of stakeholders you can include in your data innovation project.

Step 2: Map on the identified stakeholders in the Stakeholder Radar

Place stakeholders on the matrix based on their impact and influence levels. Then focus on those in the top-right corner (high impact, high influence).

Step 3: Read the Stakeholder Mapping Table Guidelines

Review the guidelines to understand how you can fill in the Stakeholder Mapping Table.

Step 4: Use the Stakeholder Mapping Table

Organise stakeholder information clearly according with the cells provided in the table. Document their impact, interests, and how to engage them.

Note: Consider when mapping your stakeholders that enhancing data collection processes is essential, as member states often use predefined formats when sending data to national authorities, and similarly, national authorities may have standardised formats for transmitting data to EU institutions.

What will you accomplish?

Develop a comprehensive stakeholder map, **achieve a shared understanding of each stakeholder's impact and influence,** identify key stakeholders to prioritise, and plan effective engagement strategies.

European Commission

Examples of stakeholders for your data innovation project

European Commission

ROLE	DESCRIPTION
Data owner	Responsible for the overall management and oversight of data assets, ensuring data is used effectively and complies with regulations.
Data steward	Manages and ensures the quality, consistency, and accessibility of data within their domain, acting as a bridge between technical teams and business users.
Data custodian	Manages the technical environment, including data storage, security, and infrastructure, ensuring data integrity.
Data producers	Responsible for generating and collecting data, such as citizens and member states, providing raw data inputs that become valuable assets for organisations.
Project manager	Oversees the project, ensuring it aligns with strategic objectives, coordinating teams, and managing resources.
Data analyst	Examines and interprets data to extract meaningful insights that drive decision-making and innovation.
IT supporter	Provides the necessary technical infrastructure, including databases, servers, and data security measures.
Governance officer	Ensures adherence to data governance policies, standards, and regulations, promoting data quality and compliance.
Communications officer	Manages stakeholder communication, ensuring transparency and keeping all parties informed of progress and outcomes.
Legal adviser	Provides guidance on legal compliance, particularly related to data privacy, security, and usage regulations.
Finance manager	Oversees the financial aspects, ensuring the project stays within budget and evaluating its economic impact.
Upstream producers	Stakeholders responsible for gathering or acquiring data, ensuring its initial quality and relevance (e.g., national statistical offices, survey teams).
Downstream consumers	Individuals or groups who utilise the data for decision-making, analysis, or strategy development (e.g., healthcare providers, policymakers).
Data managers and transformers	Entities responsible for managing, transforming, and storing data to ensure it remains useful and accessible (e.g., data processing firms, cloud service providers).
Policymakers	Stakeholders who set data policies, standards, and regulations that guide how data is handled, protected, and used (e.g., EU Authorities).
Partner organisations	Organizations collaborating on the project, possibly providing expertise, resources, or additional data (e.g., academic institutions, NGOs).
Regulatory bodies	Organisations that set and enforce data standards and regulations (e.g., GDPR compliance authorities).

Data innovation stakeholder radar

Utilise the radar below to depict the to map stakeholders visually.

Data innovation stakeholder mapping table: Guidelines to help you fill in the information in the following table

1	**Who are the stakeholders?** *Populate with a stakeholder relevant to your project, e.g., "National Data Governance Committee - Oversees data governance policies across national health systems".*
2	**What is their impact over the project?** *Determine impact, e.g., "Medium - Influences the establishment of data governance standards but not directly involved in daily operations".*
3	**What is their level of influence over the project?** *Assess influence, e.g., "Medium - Has the authority to shape governance policies but relies on other bodies for implementation".*
4	**How might the different stakeholders contribute to the project?** *Describe contributions, e.g., "Can contribute by providing governance frameworks that ensure data quality and compliance".*
5	**What could they do to undermine the project?** *Consider risks, e.g., "Could impose restrictive governance policies that slow down project implementation.*
6	**How does data support these stakeholders' decision-making now?** *Explain current data use, e.g., "Uses data to track policy adherence and ensure compliance with national standards".*
7	**What could this person do with better data on the problem?** *Identify improvements, e.g., "With better data, could streamline governance processes and reduce compliance issues".*
8	**What is the best way to keep them engaged?** *Reflect on how you want to keep them engaged, e.g., "Create a feedback mechanism".*

Data innovation stakeholder mapping table:

Use the following table to organise your stakeholders based on your answers.

1. Who are the stakeholders?	2. What is their impact over the project?	3. What is their level of influence over the project?	4. How might the different stakeholders contribute to the project?	5. What could they do to undermine the project?	6. How does data support these stakeholders' decision-making now?	7. What could this person do with better data on the problem?	8. What is the best way to keep them engaged?

European Commission

How do I build the right team for my data innovation project?

Pre-requisite tools:

Stakeholder mapping tool

Next tool:

Role's definition tool

European Commission

Data team creation

How do I build the right team for my data innovation project?

Resource type
Canvas

Complexity
Easy

Time frame
25 minutes

Who may support you in developing this resource?

Project manager, data governance officer, legal adviser.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

- Template - Decision Provenance Mapping - RD4C
- Article - How to appropriately staff an organization to meet its data needs - Data.org

What is it for?

You have identified your stakeholders, but what about the team that will work on the data innovation project with you? With this tool, you will be able to **identify individuals within your organisation who can assist you.** There are some key types of roles you may want to engage in this process.

How to use it?

Step 1: Examine the three categories of roles and some examples for

- Data management and governance
- Data analysis and implementation
- Project coordination, communication, and administrative

Step 2: Identify essential roles for your data innovation project

- Determine which roles and expertise are necessary for your data innovation project. Check the examples but do not be limited by them - it might be you have different roles to add.

Step 3: Visualise your team in the Data Team Canvas

- Depict the selected roles in the diagram to visualise how your team will be structured.
- Illustrate the relationships and interactions between different roles.
- Use the diagram to ensure all critical functions are covered and to identify any gaps.

What will you accomplish?

Identify the **necessary roles and tasks for your data innovation project,** visualise how your cross-functional team will be structured, and ensure all essential functions are included for successful project execution.

European Commission

Data team creation - Examples of actors

Data management and governance

Data owner

Manages and oversees data assets, ensuring effective usage and compliance with EU regulations, including GDPR.

Data steward

Ensures data quality, consistency, and accessibility, acting as a bridge between technical teams and business users.

Data custodian

Supervises the technical environment, focusing on secure data storage and infrastructure in compliance with EU data protection laws.

Data governance officer

Make sure compliance with data governance policies, standards, and EU regulations, particularly GDPR.

Legal adviser

Provides legal guidance on data privacy, security, and usage, ensuring compliance with GDPR and other EU laws.

Data analysis and implementation

Data scientist

Analyses datasets to extract insights, develop models, and support data-driven decisions.

Data analyst

Interprets data to generate insights, creating reports and dashboards for stakeholders.

Machine learning engineer

Develops and deploys machine learning models, ensuring they meet EU standards for transparency and fairness.

Data architect

Designs and manages the data architecture, ensuring secure and efficient data flow across systems.

Data visualisation specialist

Creates visual representations of data, making complex information accessible and actionable for stakeholders.

Project Coordination, communication and administrative

Project manager

Coordinates the data innovation project, ensuring alignment with organizational goals and EU regulations.

IT infrastructure engineer

Ensures the technical infrastructure is robust, secure, and compliant with EU regulations.

Communications officer

Manages communication strategies, ensuring clear and consistent messaging across all stakeholders.

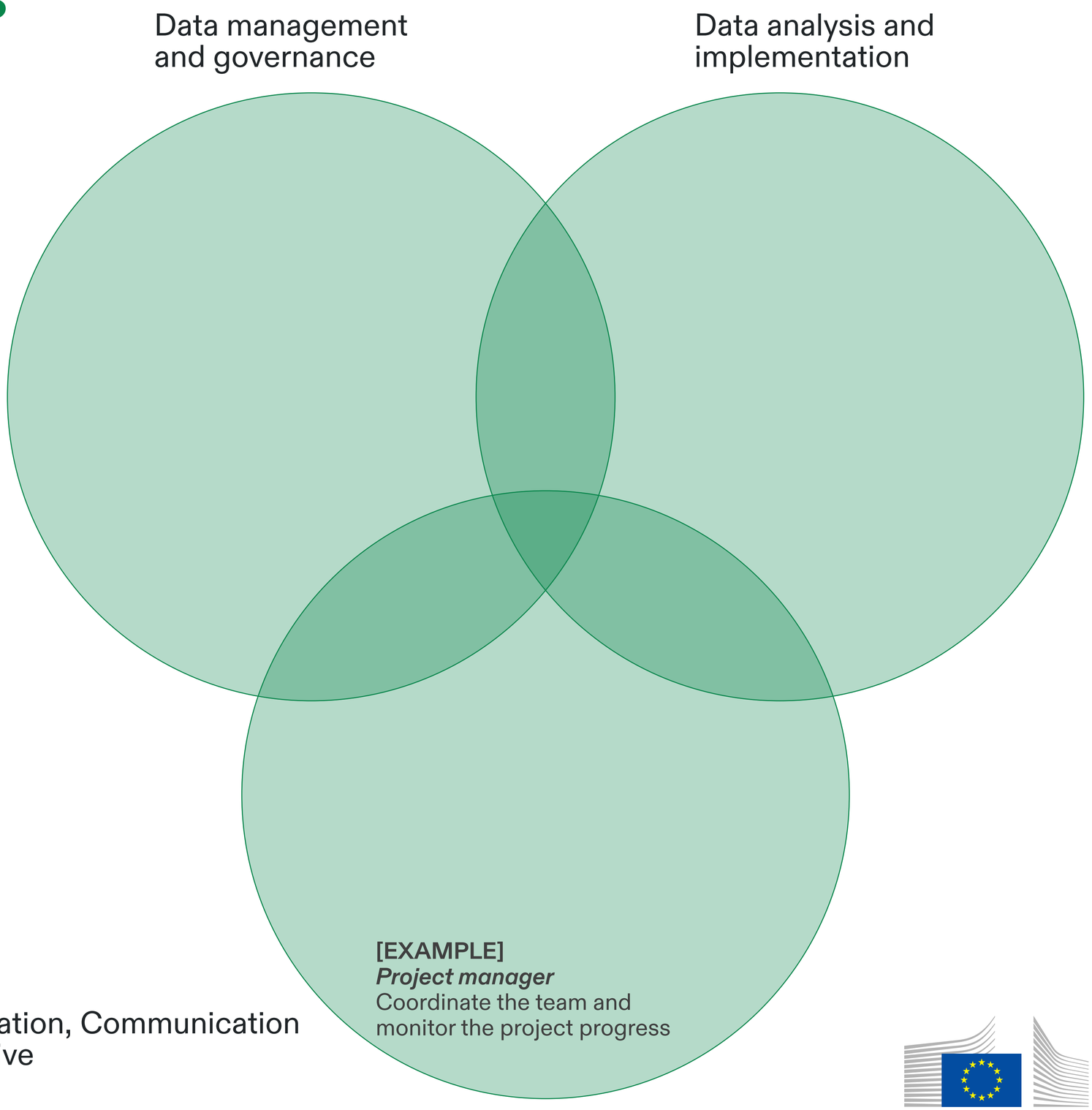
Financial manager

Manages the budgeting and financial aspects of the project, ensuring compliance with EU financial regulations.

European Commission

Data team canvas

Instructions:
Now that you have identified the roles and related tasks for your data innovation project, **indicate these roles in the diagram** to depict how your cross-functional team will look like in a visual way.



European Commission

Which roles and responsibilities should each team member have?

Pre-requisite tools:

Data team creation

Next tool:

Persona definition tool

European Commission

Role's definition tool (RACI matrix)

Which roles and responsibilities should each team member have?

Resource type	Complexity	Time frame
RACI Matrix	Medium	25 minutes

Who may support you in developing this resource?

Project manager, data owner, data governance officer, legal adviser.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

- Template - Decision Provenance Mapping - RD4C
- Article - How to appropriately staff an organization to meet its data needs - Data.org

What is it for?

Now that you have mapped your potential team, is time to define their shared responsibilities and their role with data in more detail, using a RACI matrix. The acronym RACI stands for:

R – Responsible: Individuals who complete the task or make the decision. For effective decision-making, ideally assign one person responsible per task.

A – Accountable: The person who ensures the task is finalised. They may not do the work themselves but are responsible for its completion. It's best to have one accountable person per task to avoid confusion.

C – Consulted: Those who provide input and feedback. They have two-way communication with the responsible individuals to assist in task completion.

I – Informed: Those who need updates on the task's progress. They don't engage in two-way communication but are affected by the project's outcome.

How to use it?

Step 1: Place the identified roles

Determine all the roles involved in your data innovation project. These roles should be listed on the columns of the matrix. Roles can be specific individuals or groups depending on your preference and project context.

Step 2: Assign RACI codes

For each task or activity, assign one of the RACI codes to each role.

What will you accomplish?

Clearly assign responsibilities and roles for each task, ensuring that **all team members understand their specific duties and foster accountability** — leading to streamlined communication.

European Commission

Data innovation RACI matrix

European Commission

TASK CATEGORIES	ROLE 1:	ROLE 2:	ROLE 3:	[EXAMPLE] Project Manager
Data collection and integration				C
Data quality and security				C
Data sharing and re-use				C
Data analysis and interpretation				A
Data usage and visualisation				C
Dissemination and stakeholder engagement				I
Project management, administration and financial				R

How do I create personas that accurately reflect the diverse behaviours and needs of my project's target users?

Pre-requisite tools:

None

Next tool:

Data journey mapping

European Commission

Persona definition tool

How do I create personas that accurately reflect the diverse behaviours and needs of my project's target users?

Resource type	Complexity	Time frame
Template	Easy	25 minutes

Who may support you in developing this resource?

Communication officer, project manager, data governance officer.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

Toolkit - Create Simple Personas - TTC Labs

What is it for?

The Persona Creation Tool is a reference model used to represent specific types of users, known as behavioural archetypes. These archetypes capture different behaviours without focusing on personality or demographics. As these archetypes take on more realistic characteristics—such as name, age, and household composition—they evolve into full personas. **A fully developed persona expresses the needs, desires, habits, and cultural backgrounds of a distinct group of users.**

How to use it?

Step 1: Answer the guiding questions

In the template you will find guiding questions to formulate your persona. Use them to create your profile and remember that this fictional person should reflect a personality type from the real world.

Step 2: Use post-its notes to write down your answers

Use post-it notes to mark down ideas for each step and place them in the appropriate sections. After completing the first draft, review your notes and write the final persona profile directly into the template.

Step 3: Present and discuss the profiles with your team

Are the personas created really reflecting your users? Are you considering the persona's needs and wants? Without a clear focus on your persona your project won't be user-centred, so consider discussing these profiles with your team.

Note: If possible, quantify the problem (challenge definition) and related needs of your persona.

What will you accomplish?

Create diverse persona profiles to utilise in your data innovation project.

European Commission

Persona definition tool

European Commission

Name:

Use a realistic name for the persona (e.g., John Smith, Sara Lee) to humanise the profile and make it relatable.

Age:

Specify the persona's age.

Role:

Define the persona's job title or primary role.

Short description:

Offer a concise summary highlighting the persona's key attributes and characteristics.

Context:

Describe the environment or circumstances in which the persona operates to grasp their situation and challenges.

Background:

Detail the persona's personal history, including education and experiences, to add depth and realism to their profile.

Motivations:

Identify what drives the persona, such as their needs, desires, and aspirations, to comprehend their actions and decisions.

Challenges:

Identify common pain points or obstacles they face (e.g., struggles with fragmented data systems).

Goals:

Outline the persona's objectives or what they aim to achieve to align strategies with their ambitions and expectations.

Data literacy:

Describe their attitudes toward data and their technical proficiency.

How do I map the user journey to identify key actions and opportunities for improvement?

Pre-requisite tools:

Persona creation tool

Next tool:

Data gaps checklist

European Commission

Data journey mapping - user perspective

How do I map the user journey to identify key actions and opportunities for improvement?

Resource type	Complexity	Time frame
Template	Medium	30 minutes

Who may support you in developing this resource?

Data owner, data steward, data custodian, data governance officer, data scientist, data analyst, data architect, data visualisation specialist, project manager, communication officer.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

Toolkit - Integrated Journey - AI4GOV

What is it for?

The Data journey Mapping tool is a synthetic representation that describes step-by-step how a user might interact with your data. The process is mapped from the **user perspective**, describing what happens at each stage of the interaction, what touchpoints are involved.

When developing the a data journey, consider data collection points, storage locations, processing steps, analysis methods and reporting mechanism that might affect your user.

How to use it?

Step 1: Identify the data interaction according with the data lifecycle

Break down the experience or service you are developing into distinct phases. You may use the data lifecycle phases, see the beginning of the Data Innovation Toolkit, as a reference.

Step 2: List user actions

For each phase, write down the specific actions your user or persona takes. How will your persona interact with you and your project in each of these phases?

Step 3: Identify data touchpoints

For each phase of the data lifecycle, pinpoint where data is generated, collected, or used. These touchpoints are critical moments where the user interacts with data.

Step 4: Identify data related opportunities and insights

Analyse each phase for potential opportunities or insights that could improve the user experience.

What will you accomplish?

Understand how your **user will interact and provide you with data** and also help you in having an overview of the user data journey.

European Commission

Data journey mapping - User perspective

European Commission

	Planning	Collecting	Processing	Sharing	Analysing	Using
User actions		*Example: The user enables location services and inputs their usual commute route.*				
Data touchpoints		*Example: Collection of GPS data, user preferences, and commute times.*				

Is it time to gather your data and make sure it's ready to use?

Collect and process

This stage involves the collection of data that will be used in the data innovation project, data quality, data storage, security, privacy concerns and data preparation to enable analysis.

European Commission

What is happening at this stage?

Here, you're gathering the key data needed for your data innovation project, making sure it's accurate, relevant, and focused on your goals. You'll clarify the main questions you want to answer, organise and clean the data for easy analysis, and keep an eye on any risks to ensure everything is secure, efficient, and compliant on the long term.

Why?

Quality data collection and processing are vital for valid results. Ensuring privacy and compliance protects your organisation and builds trust, while proper preparation enables meaningful analysis and insights that drive decision-making.

Considerations and risks

- Data privacy breaches.
- Non-compliance with regulations.
- Poor data quality compromising analysis.
- Security vulnerabilities in data handling.

Success factors

- Collecting only relevant data.
- Rigorous validation and quality checks.
- Adherence to legal requirements.
- Efficient data processing methods.
- Robust data security measures.

Tools in this chapter

Data gaps checklist

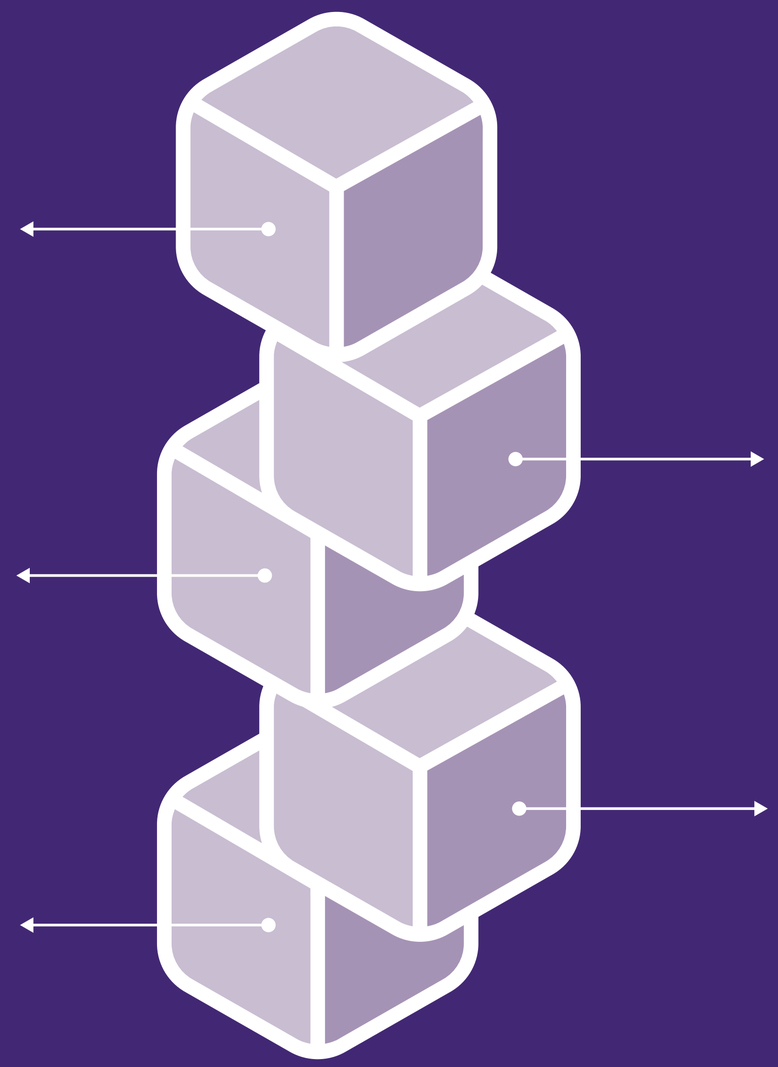
Are there any potential gaps in the data I plan to use, and how can I address them?

Data collection checklist

Can I access the correct data sources and the data itself to execute the project?

Data processing cards

Am I ensuring that all data processing activities are conducted effectively?



Data collection template

How do I ensure that my data collection process is aligned with my challenge?

Data quality exercise

Am I capturing high-quality data that meets my project's needs?

European Commission

Identify your data sources

Using the Challenge Definition Tool, you have already identified the key question you want to answer with data. Now, the next step is to determine where to collect the data you need. Consider which data sources will best support your analysis and answer your question effectively.

If you need support, you can also read more about data standards (rules for structuring information collected).

Understand your data source

When working with existing data, use these questions to better understand the context of the data source:

- *Why were the data collected? For what purpose?*
- *Who collected the data?*
- *When were the data collected?*
- *In what context were the data collected?*
- *How were the data collected?*
- *What methods were used? Were there any limitations?*
- *Are there any sources of error or bias in the data?*

Data that comes from inside your organisation.

Data that comes from outside your organisation (e.g. Statistics institutes, Academic researchers).

INTERNAL ⟷ **EXTERNAL**

Involves gathering original data directly from sources, such as surveys, experiments, or interviews, tailored to specific questions.

Refers to using data already collected by others, like reports, databases, or studies, to inform or support analysis.

PRIMARY (Collecting new data) ⟷ **SECONDARY (Using existing data)**

Are there any potential gaps in the data I plan to use, and how can I address them?

Pre-requisite tools:

Challenge definition tool

Next tool:

Data collection template

European Commission

Data gaps checklist

Are there any potential gaps in the data I plan to use, and how can I address them?

Resource type	Complexity	Time frame
Checklist	Medium	35 minutes

Who may support you in developing this resource?

Data owner, data custodian, data governance officer, data architect, data visualisation specialist.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

- Case study – An approach towards eliminating data silos within the Government - inGOV
- Toolkit – Data playbook – IFRC & Solferino Academy

What is it for?

Before starting a data innovation project, you will also need to **identify any potential gaps in your available data** that could hinder your progress. The data gaps identification tool is designed to guide you through a systematic process of evaluating your data landscape. This tool will also assist in exploring alternative data sources, improving data integration, and planning for effective data management throughout your project.

How to use it?

Step 1: Review the checklist

Go through each section of the tool's checklist to assess your current data situation.

Step 2: Identify gaps and consider actions to help with your data gaps

For each question, check the appropriate box and note any data gaps or issues. Consider the suggested action points to address these gaps - but do not be limited by them.

Step 3: Document and share findings

Record your findings and any decisions made. Discuss your responses with your team to address all potential data challenges.

What will you accomplish?

Identify missing or incomplete data points, assess the availability and quality of existing data, and develop a plan to address these gaps—ensuring your project is built on a solid foundation of reliable and relevant data.

European Commission

[STAGE 1] Break down the question into components

Start by dissecting your main question (defined in the Challenge Definition Tool), into smaller, more manageable sub-questions or components.

Identify the key aspects or variables that need to be addressed to answer the main question.

Guideline: At this stage, focus on discovering inconsistencies or omissions in how data is gathered across different agencies. For example, some agencies may collect data at varying frequencies (e.g., quarterly vs. annually) or use different definitions for similar metrics, such as "service utilisation" or "customer satisfaction." These inconsistencies create challenges for aggregating data at a national level or comparing across agencies, limiting its utility for strategic planning.

[STAGE 2] Existing data on the challenge

Identify existing data sources: Is there existing data available on your challenge / sub-questions?

◯ **Reliable**

Data is accurate, up-to-date, and validated (e.g., official statistics, peer-reviewed research).

◯ **Somewhat Reliable**

Data is somewhat accurate but may have minor issues (e.g., outdated, or inconsistent in parts).

◯ **Unreliable**

Data is outdated, incomplete, or from unverified sources.

Action: If the data is unreliable or somewhat reliable, consider seeking more reliable sources, validating the current data, or updating it to improve accuracy.

Example

Somewhat Reliable

Data is somewhat accurate but may have minor issues (e.g., outdated or inconsistent in parts). Some data, like Eurostat and OECD datasets, are reliable but might not fully capture the nuances of the adoption across all stakeholders, leading to possible reliability issues.

Assess data granularity: Is the data detailed enough for your needs?

◯ Sufficient Detail

Data is broken down to the required level of detail (e.g., by region, age group).

◯ Limited Detail

Data is available but lacks some granularity, which may limit its usefulness.

◯ Insufficient Detail

Data is too aggregated and lacks the necessary granularity.

Action: If detail is insufficient, consider combining datasets, using data processing techniques, or seeking more granular data sources.

Example

Limited Detail

While data exists on different EU sources, it may lack specific granularity related to the detailed implementation stages of my project, such as the level of interoperability one needs to achieve at local/ regional levels within the scope of my project.

Check data frequency: how often is the data collected and updated?

◯ Real-time

Data is continuously collected and updated.

◯ Frequent

Data is collected regularly (e.g., monthly, quarterly).

◯ Infrequent

Data is updated occasionally (e.g., annually).

◯ Outdated

Data is rarely updated or is out of date.

Action: If data is outdated or infrequently updated, explore other sources with more timely data or adjust the project scope to fit available data.

Example

Real-time

Data related to my project might be collected and updated in real-time as part of other ongoing initiatives, but the frequency of updates might vary across different stakeholders.

[STAGE 3] Accessibility of data

Assess data accessibility: do you have access to the necessary data?

◯ Open Access

Data is freely available without restrictions.

◯ Restricted Access

Data requires special permissions or licenses.

◯ No Access

Data is not accessible due to legal, proprietary, or other barriers.

Action: If access is restricted or unavailable, take steps to secure the necessary permissions or explore alternative data sources.

Example

Open Access

Much of the data related to my project implementation, especially from EU institutions like Eurostat, is openly accessible for public use.

Evaluate legal and ethical constraints: are there legal or ethical restrictions on using the data?

◯ No Restrictions

Data can be used without legal or ethical concerns.

◯ Some Restrictions

There are minor legal or ethical considerations (e.g., data anonymisation required).

◯ Significant Restrictions

Major legal or ethical constraints exist (e.g., privacy laws, data ownership issues).

Action: If significant restrictions exist, consult with legal experts, ensure compliance with regulations, or consider alternative approaches to data usage.

Example

Some restrictions

There are likely some legal constraints related to privacy and data protection that must be considered when using the data across different different stakeholders.

[STAGE 4] Other organisations data

Explore external data sources: could data from other organisations / departments help fill your gaps?

- ◯ **Yes, with full access**

External data sources are available and accessible.

- ◯ **Yes, but with limited access**

External data is available but requires permissions or agreements.

- ◯ **No relevant sources**

No external data sources are readily available.

Action: If relevant external data is available, initiate contact for data sharing; if not, consider other methods like partnerships or new data collection.

Example

Yes, with full access

Data from organisations like the World Health Organization (WHO) or private data sources could supplement the data for a more comprehensive view and adoption of my project.

Plan for data integration: can data from different sources be integrated in your project effectively?

- ◯ **Yes, seamlessly**

Data is compatible and can be easily merged.

- ◯ **With challenges**

Data can be integrated but may require significant transformation.

- ◯ **No**

Data sources are incompatible or require extensive work to integrate.

Action: If integration is challenging, plan for additional time and resources to address these issues and consider using specialised tools for data integration.

Example

With challenges

Integrating data from organisations with varying levels of digital infrastructure and standards can be challenging, requiring significant data transformation.

Validate and document data: have you validated the quality of all data sources?

- ◯ **Validated**

Data has been thoroughly checked for accuracy, consistency, and reliability.

- ◯ **Partially Validated**

Some data has been validated, but further checks are needed.

- ◯ **Unvalidated**

Data quality is uncertain or untested.

Action: Conduct comprehensive data quality checks, document the sources, and note any limitations or assumptions.

Example

Partially validated

While some data, particularly from well-known sources, may be validated, data from newer or less consistent sources might require additional validation to ensure reliability.

European Commission

Pre-requisite tools:

Data gaps checklist

Next tool:

Data collection checklist

How do I ensure that my data collection process is aligned with my challenge statement?

European Commission

Data collection template

How do I ensure that my data collection process is aligned with my challenge statement?

Resource type	Complexity	Time frame
Template	Medium	30 minutes

Who may support you in developing this resource?

Data owner, data custodian, data governance officer, data architect, data visualisation specialist, project manager.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

- Guide - Data Responsibility Journey - Collecting by GovLab
- Guide - Multidimensional Open Data Collection Tool - Open Data labs

What is it for?

The Data Collection Template is an instrument that guides your team in systematically **organising and planning the data collection process.** This template is designed to ensure clarity around the purpose of the data, the methods used to collect it, the people involved, and potential risks.

How to use it?

Step 1: Check the fields in the template and use post-it notes

Quickly review the tool's questions. Answer each question using post-it notes and place them in the respective sections.

Step 2: Review your answers with your team

After completing the first draft, review the answers and write the final responses. Follow the guidelines in each question to help with your decisions.

What will you accomplish?

Have a high-level view on **how to collect the data for your innovation, anticipate any risks** that may arise with this process, and ensure a smoother experience for you and your team.

European Commission

Data collection template

European Commission

What's the question that you want to answer with data? *Remember your challenge statement.*

What data are you collecting, why are you collecting it?

Specify the exact data you plan to collect and explain why this data is relevant to your project question. Be sure that every data point contributes to answering your question, and avoid unnecessary or irrelevant data to streamline your process.

Who will collect the information?

Assign responsibility for data collection. Determine which team member or group is responsible for collecting the data, ensuring accountability and clarity in the process. Utilise the **RACI Matrix** and **Data Team Creation** tools to support.

Who provides the data, have you obtained consent?

Identify the individuals or groups who will be providing the data. Ensure you have their consent, especially if the data is personal or sensitive, to comply with ethical and legal requirements. Consent should be specific to the purpose of your project.

How will you collect the data?

Describe the method you will use to collect the data (e.g., surveys, interviews, sensors, or databases). Make sure the method aligns with your project goals and is appropriate for the type of data being gathered.

What are the potential risks associated with collecting this data?

Identify any risks involved in data collection, such as privacy concerns, data security issues, or risks to the data providers. Plan how to mitigate these risks to ensure a secure and ethical data collection process.

How often will the data be collected?

Set a timeline for your data collection. Will the data be collected once, periodically, or continuously? Clearly define the frequency to ensure consistency and relevance to your project.

Pre-requisite tools:

Data collection template

Next tool:

Data quality checklist exercise

Can I access the correct data sources and the data itself to execute the project?

European Commission

Data collection checklist

Can I access the correct data sources and the data itself to execute the project?

Resource type	Complexity	Time frame
Checklist	Medium	1 hour

Who may support you in developing this resource?

Data owner, data custodian, data governance officer, data architect, data visualisation specialist, project manager.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

- Guide - Data Responsibility Journey - Collecting by GovLab
- Guide - Multidimensional Open Data Collection Tool - Open Data labs

What is it for?

This checklist helps you and your data teams cover all the essentials of **data collection,** including identifying the right data sources, validating and securing information, and ensuring privacy and compliance. If you've already used the **Data Gaps tool**, you've completed some of these steps, providing a strong foundation to keep data collection efficient, compliant, and aligned with best practices throughout your project.

How to use it?

Step 1: Review the checklist

Go through each section of the tool's checklist to assess your current data situation.

Step 2: Identify gaps and consider actions to help with your data gaps

For each question, check the appropriate box and note any data gaps or issues. Consider the suggested action points to address these gaps - but do not be limited by them.

Step 3: Document and share findings

Record your findings and any decisions made. Discuss your responses with your team to address all potential data challenges.

What will you accomplish?

Identify gaps in your data collection practices and receive actionable steps to improve compliance and efficiency.

European Commission

[STAGE 1] Break down the question into components

Start by dissecting your main question (defined in the Challenge Definition Tool), into smaller, more manageable sub-questions or components.

Have you ensured that the amount of data collected is the minimum necessary to address the problem or question? Write down sub-questions below to help you answer that.

[STAGE 2] Consent and compliance

Have you obtained meaningful consent to collect data?

◯ **Yes**

Consent has been obtained, with individuals fully informed about data use. Legal experts have also been consulted to ensure the consent was collected in the correct way.

◯ **Partially**

Consent obtained, but details on certain data uses are unclear or the responsible for collecting the consent of stakeholders are not experts on the matter and not aware of all possible legal issues.

◯ **No**

Consent was not obtained.

Action: If only partial or no consent was obtained, clarify data use and re-contact individuals internally or externally to ensure informed consent. If necessary, consult legal experts on data consent.

Example

Partially

Consent forms lacked specifics on third-party data sharing, so re-consent may be necessary to cover this issue.

Have you ensured data collection adheres to relevant data protection regimes (e.g., GDPR)?

- ◯ **Compliant**

Data handling complies with all regulatory requirements.

- ◯ **Partially Compliant**

Compliance in key areas, but some aspects need review.

- ◯ **Non-Compliant**

Practices do not meet data protection standards.

Action: Check GDPR or other relevant guidelines on data handling. Specifically verify data storage duration, user rights (e.g., right to access, right to erasure), and add any missing protocols. Implement or adjust as needed.

Example

Partially Compliant

Data storage periods exceed GDPR limits. Action: Adjust data retention policy to comply with GDPR, ensuring data is deleted after the legally specified timeframe.

[STAGE 3] Data privacy and security by design

Have you taken a "privacy-by-design" approach?

- ◯ **Yes**

Privacy built in from the very start of the data innovation project.

- ◯ **Somewhat**

Privacy considerations included but not fully integrated.

- ◯ **No**

Privacy added/ will be added later in response to issues.

Action: Integrate privacy measures like user-controlled data visibility or opt-out options. Review and address each potential point where user data is accessed to ensure it aligns with privacy requirements.

Example

Somewhat

Users can't currently request data deletion. So a possible action is to develop an easy-to-use "Delete My Data" option in your upcoming solution and integrate it into user settings, so users can directly request data removal - also state that during data collection procedures.

Have you ensured that any sensitive data is protected?

- ◯ **Fully protected**

Data is encrypted with access controls. The encryption of the data innovation project has been tested and validated with data legal experts.

- ◯ **Partially protected**

Some protections but with gaps. The gaps are known by your team.

- ◯ **Not protected**

Data lacks necessary protections or your team simply does not understand the full extent on how to secure sensitive data in your project.

Action: Encrypt sensitive data fields, set up multi-factor authentication (MFA) for system access, and implement regular security checks.

Example

Partially Protected

Sensitive data is stored securely, but access lacks MFA. An action could be to enable MFA on all systems handling this data and schedule quarterly security audits.

[STAGE 4] Data classification and governance

Have you classified your data based on sensitivity and value?

- ◯ **High Sensitivity**

Your data innovation project will deal with confidential information or with data that is very sensitive. Your team understands also the risks of working with the data that you need.

- ◯ **Moderate Sensitivity**

Your data innovation project involves somewhat sensitive data such as personal demographics. At the same time your team is not fully knowledgeable on all the risks.

- ◯ **Low Sensitivity**

Your data innovation project will only utilise public or non-sensitive data. Also, your team has no expertise in classifying sensitive/non-sensitive data.

Action: For high-sensitivity data, apply stricter access controls (e.g., limit to certain team members), encrypt data at rest and in transit, and regularly audit access. Use the **RACI Matrix and Data Team creation tools.**

Example

High Sensitivity

Health data classified as high sensitivity. Your team will need to store data in a secure server with encryption and restrict access to specific authorised personnel.

Have you established a data governance framework?

◯ **Yes**

The team working in this data innovation project has develop a framework with roles and policies established for the data.

◯ **In Progress**

There is a data governance framework in the making or at least the team understands the necessity of having one and how to develop it.

◯ **No**

No framework established and you/your team do not understand why you need one or how to develop it.

Action: Define key roles (e.g., Data Steward, Compliance Officer) and establish policies for data use, storage, and sharing. Hold regular team reviews to ensure compliance.

Example

No

Roles are unclear and the team is also not fully aware of your data policies. In this case, it may be necessary to assign data stewardship roles to team members, document responsibilities, and set monthly meetings to ensure data policies are followed.

[STAGE 5] Data quality and standardisation

Have you ensured that the data conforms to its syntax and definitions?

◯ **Fully Standardized**

Data conforms completely to the established syntax, definitions, and formats (e.g., all dates are in the same format, numerical data has consistent decimal places, and text data follows specified encoding).

◯ **Partially Standardised**

Data largely follows syntax and format standards but contains a few minor inconsistencies (e.g., occasional date or numerical format deviations, some fields with unexpected values).

◯ **Not Standardised**

Data is inconsistent and does not follow defined syntax or format rules, requiring extensive cleaning (e.g., text data lacks encoding standards, and numerical fields have varied decimal precision).

Action: Use a data validation tool or script to automatically flag and correct inconsistencies. Ensure that all fields meet defined standards (e.g., use systems for date formatting or set rounding rules for numerical data).

Example

Partially Standardised

Dates are mostly in “YYYY-MM-DD” format, but a few entries use “MM/DD/YYYY.” Action: Run a data cleaning script to convert all dates to “YYYY-MM-DD” format for consistency across the dataset.

Pre-requisite tools:

Data collection template

Next tool:

Data processing cards

Am I capturing high-quality data that meets my project's needs?

European Commission

Data quality checklist exercise

Am I capturing high-quality data that meets my project's needs?

Resource type	Complexity	Time frame
Canvas	Medium	20 minutes

Who may support you in developing this resource?

Data owner, data custodian, data governance officer, data architect, data visualisation specialist, project manager.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

- Report - Data.europa.eu data quality guidelines
- Toolkit - Data Governance Toolkit - Nesta
- Article - The Challenges of Data Quality and Data Quality Assessment in the Big Data Era

What is it for?

This activity is designed for teams, organisations, or projects that need to explore and establish **a clear understanding of data quality.** There is always the possibility of mistakes in the data lifecycle. When data contains too many inaccuracies, it can distort the reality of the situation. Therefore, collecting reliable and accurate data is essential for ensuring the overall success of your project.

How to use it?

Step 1: Explore the concept of data quality

Ask your team to share their own definition of "Data Quality" using post-its.

Step 2: Discuss and come to an agreement on data quality

Present a simple definition: "Data Quality is how well a dataset meets its intended purpose." Then ask: "How can we define Data Quality in our project?" Write down your consensus in the middle of the Canvas, in the white square box.

Step 3: Generate an insightful discussion and answer the questions

Guide the group to create questions for ensuring data quality based on the canvas (next page). Write in post-its notes the answer for each of the questions.

Step 4: Review your answers

Have participants pick the most relevant perspectives to include in a future Data Quality checklist. Try to comply with all of them.

What will you accomplish?

Create a **shared definition of data quality tailored to your project,** generate relevant questions for ensuring data quality, and identify key criteria for inclusion in a future Data Quality checklist.

European Commission

Data quality checklist exercise

WRITE HERE

Accuracy / Validity

Does the data accurately represent reality and measure what it is intended to measure?

Timeliness

Is the data updated frequently enough to maintain its relevance and usefulness?

WRITE HERE

Consistency

Are there processes in place to ensure uniform data entry and format across all datasets?

How do you define data quality in your project?
Why data quality is important for your data innovation project?

Uniqueness

Are there any duplicate records in the dataset?

WRITE HERE

Completeness

Of all provided records, what percentage of the available fields have a value?

WRITE HERE

WRITE HERE

Pre-requisite tools:

Data collection template

Next tool:

Data sharing template

Am I ensuring that all data processing activities are conducted securely, efficiently, and in compliance with regulations?

European Commission

Data processing cards

Am I ensuring that all data processing activities are conducted securely, efficiently, and in compliance with regulations?

Resource type	Complexity	Time frame
Cards	Easy	40 minutes

Who may support you in developing this resource?

Communication officer, data owner, data custodian, data architect, data visualisation specialist.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

- Guide - Data Responsibility Journey - Processing by GovLab
- Guide - Data Quality Guidelines - European Commission

What is it for?

The purpose of this card collection is to increase the **interactivity and manageability of the data processing stage.** With distinct actions and checkpoints, each card helps you concentrate on one activity at a time while representing a key stage in getting the data ready for your data innovation project. You can make sure that data is correct, well-organised, and prepared for insightful analysis in your data innovation project by following these steps and completing each card.

How to use it?

Step 1: Check the description and action in each card

Start with Card 1 and work through each card in order, completing the actions outlined in the description and action section.

Step 2: Track your progress

Use the “Completed?” checkbox to track your progress—check it off when you finish each task. Move to the next card only once you feel confident the task is complete and meets the requirements.

Step 3: Discuss and review your progress

Take notes on any challenges or questions that come up, so you can revisit or adjust as needed. Review your completed cards to ensure nothing was missed.

What will you accomplish?

Have **higher quality data** and ensure that your data is properly cleaned, securely stored, and ready for analysis; maintain data integrity, improve accessibility, and safeguard your data with redundancy measures to prevent loss or corruption.

European Commission

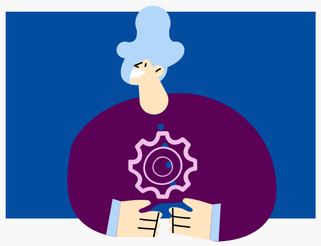
Data processing cards

European Commission

D ♠

DATA PROCESSING

ENSURE DATA COMPLETENESS



Description and Action: Verify that your dataset is complete, with no missing values. Explicitly mark null values to distinguish them from zeroes or empty strings. For numerical data, consider filling gaps with appropriate values.

☐ **Completed?** All my data is complete!

D ♠

D ♠

DATA PROCESSING

STANDARDISE THE DATA

Description and Action: Maintain consistency in data formats. Standardise date and time formats (e.g., ISO 8601), ensure uniformity in decimal numbers, and use standardised character encoding like UTF-8.

☐ **Completed?** Dates, categories, and formats are standardised for easy analysis.

D ♠

D ♠

DATA PROCESSING

REMOVE DUPLICATES

Description and Action: Identify and eliminate duplicate records and inconsistencies to improve data quality. This step ensures that each data point is unique and accurate.

☐ **Completed?** I have cleaned and standardised my data!

D ♠

D ♠

DATA PROCESSING

DATA ACCURACY

Description and Action: Confirm that the data is accurate, relevant, and aligns with your project's objectives. This involves validating data against trusted sources and ensuring it meets the required standards.

☐ **Completed?** My data is indeed valid and coming from trusted sources.

D ♠

D ♠

DATA PROCESSING

IDENTIFY OUTLIERS

Description and action: Find any extreme values that could skew your results. Use charts or basic calculations to spot outliers, and adjust or remove them as needed.

☐ **Completed?** Outliers are managed, ensuring data reflects true trends.

D ♠

D ♠

DATA PROCESSING

ENRICH THE DATA

Description and Action: Utilise controlled vocabularies to harmonise your data. Link and augment your data with external datasets to enhance its value and interoperability.

☐ **Completed?** Data is and enriched for better integration.

D ♠

D ♠

DATA PROCESSING

DOCUMENT PROCESSING

Description and action: Record each step taken to clean, organise, and prepare data. This card ensures that others in your team can follow and verify your process.

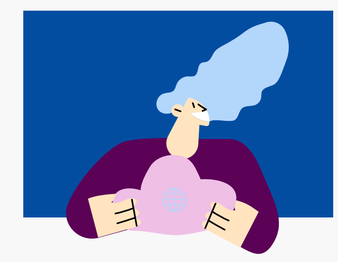
☐ **Completed?** All processing steps documented for transparency.

D ♠

D ♠

DATA PROCESSING

DATA ACCESSIBILITY



Description and Action: Improve the openness level of your data by using structured, non-proprietary formats and providing accessible download URLs. This facilitates data reuse and promotes transparency.

☐ **Completed?** Data is published in open formats and is easily accessible.

D ♠

Do you want to share what you've collected and start looking for patterns or insights?

Share and analyse

This stage ensures secure and compliant data transfers, addresses legal and ethical considerations, and involves exploratory data analysis to generate actionable insights aligned with the project's objectives. It also support you in preparing your data for once the project is finalised.

European Commission

What is happening at this stage?

In this step, data is shared securely and in compliance with ethical and legal standards to support collaboration for your data innovation project, once it is finalised. The data is then analysed thoroughly to uncover insights and answer key project questions, ensuring findings are accurate and aligned with project goals.

Why?

This stage transforms collected data into actionable insights. Secure sharing enhances collaboration, and sophisticated analysis reveals patterns that inform strategic decisions and drive innovation.

Considerations and risks

- Security risks in data sharing.
- Ethical concerns over data use.
- Analysis errors leading to flawed conclusions.
- Collaboration challenges among team members.
- Overreliance on technology without human insight.

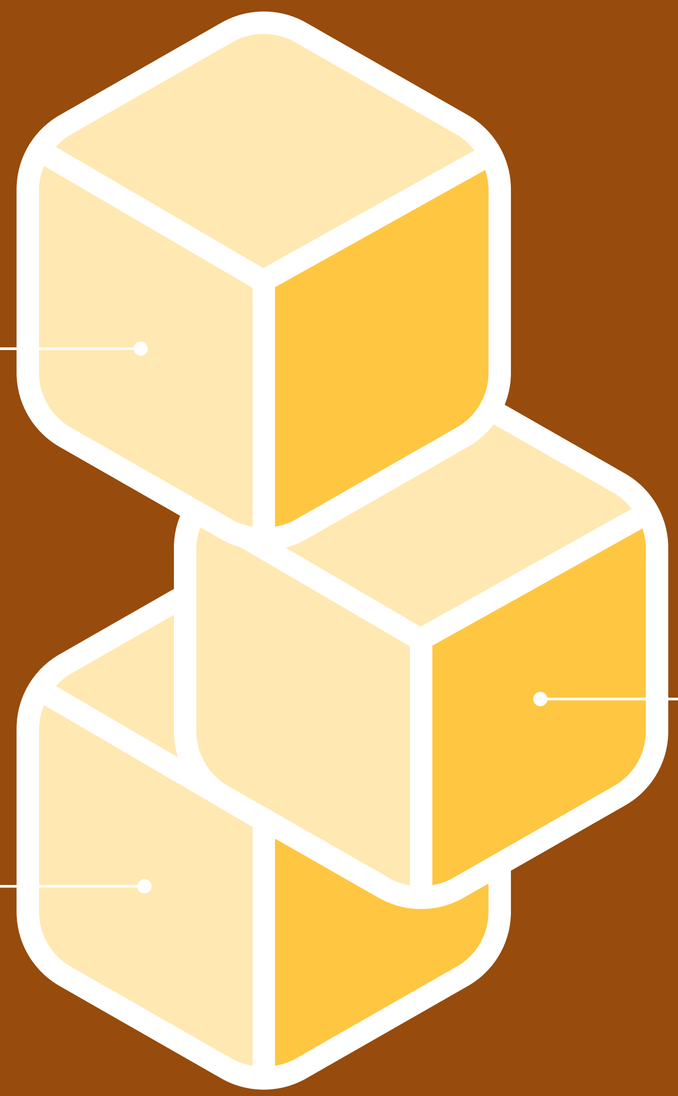
Success factors

- Secure data sharing protocols.
- Ethical data practices.
- Skilled application of analytical tools.
- Effective team collaboration.
- Balance between technology and human expertise.

Tools in this chapter

Data sharing template

How can I ensure that my data-sharing process is secure, legally compliant, and transparent?



Data sharing checklist

Have I taken all necessary steps to share data securely, ethically, and in a way that supports project goals?

Data analysis template

Am I conducting thorough and accurate data analysis to generate meaningful insights for my project?

European Commission

Pre-requisite tools:

Data processing cards

Next tool:

Data sharing checklist

How can I ensure that my data-sharing process is secure, legally compliant, and transparent?

European Commission

Data sharing template

How can I ensure that my data-sharing process is secure, legally compliant, and transparent?

Resource type	Complexity	Time frame
Template	Easy	30 minutes

Who may support you in developing this resource?

Data steward, data custodian, data governance officer, data scientist, data analyst, data visualisation specialist.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

- Article - Effective and Ethical Data Sharing at Scale - Global Partnership for Sustainable Development Data
- Book Chapter - Sharing Data - The Engine Room
- Guide - Data Responsibility Journey - Sharing by GovLab

What is it for?

The Data Sharing Template Tool is an essential instrument that helps teams clearly **define and organise the process of sharing data between** organisations or individuals. It provides a structured approach to ensure that data is shared securely, legally, and transparently. The template covers key considerations such as the nature of the data, the parties involved, timelines, security measures, and compliance with local laws.

How to use it?

Step 1: Check the fields in the template and use post-it notes

Quickly review the tool's questions. Answer each question using post-it notes and place them in the respective sections.

Step 2: Review your answers with your team

After completing the first draft, review the answers and write the final responses. Follow the guidelines in each question to help with your decisions.

What will you accomplish?

A detailed and **compliant data-sharing agreement is created,** ensuring data is transferred securely, legally, and transparently.

European Commission

Data sharing template

European Commission

What data are you going to share?

Be as specific as possible about the datasets you are planning to share. Ideally, list them out and separate 'personal and sensitive' data from 'non-personal' data to comply with local data protection and privacy laws.

Who's sending the data and who's receiving it?

List all the names and contact details of the organisations or individuals involved in sending and receiving the data. If some or all of the data belongs to another organisation, ensure you have permission to share it or include them in the agreement if they control the data.

When will you start sharing the data and when will you stop?

Specify the start and end dates for the data sharing. Clarify what will happen to the data at the end—will it be returned to the data provider, deleted, or archived? If you're unsure when the data sharing will end, include a clause to review the timeline at an appropriate point.

If there's personal data involved, how will you keep it safe during and after the transfer?

Review your data sharing plan with all data protection principles in mind: legal basis, minimisation, purpose limitation, data security, transparency, proportionality, and data subject rights.

Why are you sharing the data?

Clearly list out the reasons why the data is being shared. If personal or sensitive data is involved, ensure you document all the legitimate legal bases for sharing that data.

How are you sharing the data?

Specify how the data will be transferred—for instance, by email, granting remote access to a server, via the cloud, etc. Where possible, detail the standards and formats that apply to the data being shared.

Have I taken all necessary steps to share data securely, ethically, and in a way that supports project goals?

Pre-requisite tools:

Data sharing template

Next tool:

Data analysis template

European Commission

Data sharing checklist

Have I taken all necessary steps to share data securely, ethically, and in a way that supports project goals?

Resource type	Complexity	Time frame
Checklist	Medium	35 minutes

Who may support you in developing this resource?

Data owner, data steward, data custodian, data governance officer, legal adviser, data scientist, data analyst, machine learning engineer, data architect, data visualisation specialist.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

- Article - Effective and Ethical Data Sharing at Scale - Global Partnership for Sustainable Development Data
- Book Chapter - Sharing Data - The Engine Room
- Guide - Data Responsibility Journey - Sharing by GovLab

What is it for?

This checklist is designed to help you navigate the data-sharing stage of your project. It guides you through key steps like coordinating with partners, securing data transfers, and ensuring you meet ethical and legal standards. By following these steps, you'll be well-prepared to **share data safely and effectively, protecting sensitive information while supporting collaboration and transparency.** This material helps you in verifying that your data and insights are shared smoothly, amplifying the impact of your project.

How to use it?

Step 1: Review key questions

Go through each of the key questions provided, which are designed to help you assess critical considerations for effective data management during the data sharing stage.

Step 2: Select the best option

For each question, choose the option that best reflects your current situation.

Step 3: Take action

Based on your selected option, review the suggested actions in the corresponding column. Follow these steps to address any identified gaps and improve your data sharing practices.

What will you accomplish?

Facilitate secure and compliant data transfers, **ensuring your data is appropriately shared with others.**

European Commission

[STAGE 1] Purpose and lawfulness of data sharing

Have you clearly defined and documented the purpose for which you are sharing the data?

◯ **Yes**

The specific purpose for data sharing is clearly defined, documented, and aligns with the original reasons for data collection.

◯ **Partially**

The purpose is somewhat defined but lacks full clarity or documentation.

◯ **No**

There is no defined or documented purpose for sharing the data.

Action: If only partial or no purpose is defined, you should clearly articulate and document the specific purpose of the data sharing. Ensure it aligns with the original data collection purposes and complies with data protection laws.

Example

Partially

The organisation intends to share employee data with a benefits provider but hasn't fully documented how this aligns with their data protection policies.

Is the data sharing lawful, fair, and transparent under data protection regulations?

◯ **Yes**

A lawful basis for data sharing has been identified, and individuals have been informed accordingly.

◯ **Partially**

A lawful basis is considered, but its applicability is uncertain, or individuals are not fully informed.

◯ **No**

No lawful basis has been identified, and individuals are unaware of the data sharing.

Action: If only partial or no lawful basis is established, review data protection laws to identify an appropriate lawful basis. Ensure transparency by informing individuals about the data sharing in a clear and accessible manner.

Example

Somewhat

User data is shared with third parties for marketing without any lawful basis or customer awareness, breaching GDPR requirements.

European Commission

[STAGE 2] Data sharing agreements and risk management

Have you established and documented a data sharing agreement with the other party?

◯ **Yes**

A comprehensive data sharing agreement is in place, detailing responsibilities and data protection measures.

◯ **Partially**

An agreement exists but lacks detail on key aspects like security protocols or breach procedures.

◯ **No**

There is no formal data sharing agreement documented.

Action: If only partial or no agreement exists, develop a detailed data sharing agreement outlining the purpose, lawful basis, security measures, and roles of each party involved.

Example

Partially

An informal understanding is in place, but it doesn't specify how data will be securely transferred or who is responsible in case of a data breach.

Have you assessed the risks associated with the data sharing and taken steps to mitigate them?

◯ **Yes**

A thorough risk assessment or Data Protection Impact Assessment (DPIA) has been conducted, and mitigation measures are implemented.

◯ **Partially**

Some risks have been identified, but no formal assessment or full mitigation plan is in place.

◯ **No**

Risks have not been assessed, and no mitigation measures are considered.

Action: If only partial or no risk assessment is done, conduct a DPIA to identify potential risks. Implement appropriate measures to mitigate identified risks before proceeding with data sharing.

Example

Partially

Sensitive health data is shared without evaluating the risk of unauthorised access, leaving personal information vulnerable.

European Commission

[STAGE 3] Data minimisation and individual rights

Are you sharing only the minimum amount of data necessary for the intended purpose?

◯ **Yes**

Data sharing is limited to what is strictly necessary, adhering to the data minimisation principle.

◯ **Partially**

Some unnecessary data is included in the sharing process.

◯ **No**

Excessive data is shared beyond what is needed for the purpose.

Action: If only partial or no adherence to data minimisation, review the data being shared. Remove any non-essential information to comply with data protection principles and reduce the risk of data breaches.

Example

Partially

A government agency shares full citizen profiles with a municipal office when only contact details are needed, resulting in unnecessary data sharing.

Have you informed individuals about the data sharing and their rights in a clear and accessible way?

◯ **Yes**

Individuals are fully informed through privacy notices or direct communication, outlining how their data will be shared and their rights.

◯ **Partially**

Individuals have been informed, but the information is incomplete, unclear, or not easily accessible.

◯ **No**

Individuals have not been informed about the data sharing.

Action: If only partial or no information has been provided, update your privacy notices and communication strategies. Ensure that individuals are clearly informed about the data sharing activities and understand their rights under data protection laws.

Example

Partially

Employees are unaware that their personal data is being shared with external training providers, potentially breaching transparency obligations.

European Commission

Am I conducting thorough and accurate data analysis to generate meaningful insights for my project?

Pre-requisite tools:

Data sharing checklist

Next tool:

Data using checklist

European Commission

Data analysis template

Am I conducting thorough and accurate data analysis to generate meaningful insights for my project?

Resource type	Complexity	Time frame
Template	Easy	30 minutes

Who may support you in developing this resource?

Data analyst, IT infrastructure engineer, data visualisation specialist, data owner.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

Guide - Data Responsibility Journey - Analysing by GovLab

What is it for?

This template is here to support you through the analysis stage of your data innovation project. It breaks down the steps for exploring data and answering specific questions, helping you choose the right tools and methods to reach your project's goals. With this template, you can make sure your analysis is thorough and that the insights you gain are reliable and useful. Following these steps will help you **get the most value out of your data, making it easier to make informed decisions that support your organisation's goals.**

How to use it?

Step 1: Review the types of data analytics

Before answering the template, please check the types of data analytics to help you fill in the template that comes right after it.

Step 2: Start filling in the template

Go through each of the key questions provided, which are designed to help you assess critical considerations for effective data management during the data analysis stage.

Step 3: Review the answers with your team

After writing down your answers in each section of the template, take a moment to discuss them with your team and see if you need to take any necessary action to improve how you will develop the data analysis in your project.

What will you accomplish?

Your data will be analysed comprehensively and accurately - **helping you to produce meaningful insights with the data you have processed.**

European Commission

Types of data analytics

Data analysis involves different types of approaches — descriptive, diagnostic, predictive, and prescriptive — that each serve a unique purpose. These methods help you look at data in different ways: understanding past behaviours, identifying causes, predicting future trends, and recommending actions. By choosing the right approach and tools for each type, you can uncover patterns and insights that are specifically useful for your data innovation project.

Descriptive analytics

- **Description:** Descriptive analysis looks at past data to show what's happened. It reveals patterns or trends using simple charts, averages, or summaries, giving a quick overview of results.
- **How to do it:** Organise data by category (like dates or locations) and use tools to make charts or calculate averages. For example, create a bar chart of monthly sales to see trends over time.

Predictive analytics

- **Description:** Predictive analysis uses past patterns to make educated guesses about the future, helping answer, "What might happen next?"
- **How to do it:** Look at past data trends and extend them forward. For example, if sales have been increasing each month, you can expect them to keep growing at a similar rate. Use simple graphs to visualise such trends.

Diagnostic analytics

- **Description:** Diagnostic analysis goes deeper by asking, "Why did this happen?" It helps uncover reasons behind trends for a clearer understanding.
- **How to do it:** After spotting a trend, think about possible causes. If customer interest dropped in one month, check for events or changes (like product updates or promotions). Discuss with team members to explore ideas.

Prescriptive analytics

- **Description:** Prescriptive analysis suggests actions based on your findings, helping you decide the best steps forward by combining past insights and future expectations.
- **How to do it:** Based on the trends you've found, brainstorm actions (e.g., adjusting budgets or resources). Try "what if" questions to consider how different choices might affect outcomes, and discuss these options with your team.

Data analysis template

European Commission

Which type of analysis will you use: descriptive, diagnostic, predictive, or prescriptive?

Understand the focus of your analysis. Descriptive analysis summarises past data, diagnostic analysis explains why events happened, predictive analysis forecasts future trends, and prescriptive analysis recommends actions based on data insights. Identifying the correct type of analysis will guide your approach and methodology.

How are the data collected and what do they consist of?

Examine how the data were gathered, identify possible sources of errors, and ensure definitions and measurement units are consistent for accurate analysis.

How do you explore data distribution and trends?

Analyse the distribution of numerical and categorical variables, observe temporal trends, identify relationships, and discuss early findings with stakeholders for deeper insights.

How do you prepare data for analysis?

Ensure data are in the correct format, remove duplicates, check consistency across datasets, identify outliers, manage missing values, and create new variables to enhance the analysis.

How do you ensure the robustness of your findings?

Test how sensitive your results are to different assumptions, document your processes thoroughly, and seek peer reviews to validate the reliability and reproducibility of your analysis.

Use and evaluate

This stage translates insights into action, communicates findings, manages data retention, measures success, and plans for scalability and future strategies.

European Commission

What is happening at this stage?

Results are communicated to stakeholders, and decisions are made regarding the retention or disposal of data. The focus is on applying insights responsibly, optimising resources, and ensuring the benefits of your project are fully realised.

Why?

This stage realises your project's value. Effective communication enables informed decisions, thorough evaluation provides lessons for future initiatives, and planning for scalability amplifies successful outcomes.

Considerations and risks

- Ineffective communication of results.
- Inadequate evaluation of project success.
- Unnecessary data retention posing risks and costs.
- Challenges in scaling successful initiatives.
- Misapplication of insights without proper context.

Success factors

- Clear, actionable communication.
- Thorough evaluation against goals.
- Responsible data management decisions.
- Strategic scalability planning.
- Insightful application aligned with objectives.
- Commitment to continuous improvement.

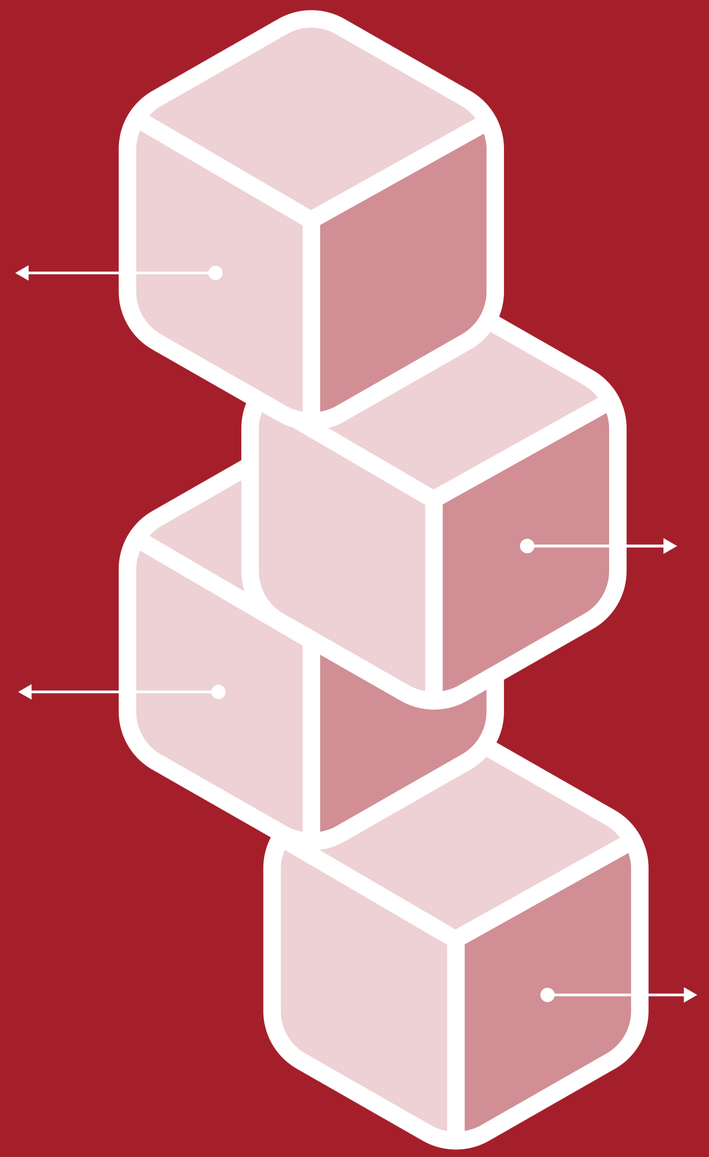
Tools in this chapter

Data using checklist

How can I ensure that the data and insights are used effectively and responsibly to achieve my project's objectives?

Measuring impact and evaluating success tool

How do I measure the impact and success of my data innovation project?



Data innovation KPI creation tool

What are the key performance indicators (KPIs) I need to measure success?

Data innovation scalability plan

What strategies should I use to scale my project effectively?

European Commission

How can I ensure that the data and insights are used effectively and responsibly to achieve my project's objectives?

Pre-requisite tools:

Data analysis template

Next tool:

Data innovation KPI creation tool

European Commission

Data using checklist

How can I ensure that the data and insights are used effectively and responsibly to achieve my project's objectives?

Resource type	Complexity	Time frame
Checklist	Easy	20 minutes

Who may support you in developing this resource?

Data governance officer, IT infrastructure engineer, data scientist, data architect, data visualisation specialist.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

Guide - Data Responsibility Journey - Using by GovLab

What is it for?

This checklist is set in visual roadmap, designed to guide you through the final "using" stage of your data innovation projects. The roadmap covers as a high-level overview on how to **responsibly manage data at the end of the project,** ensuring any outdated or redundant information is securely deleted. Try to complete each of these steps to use data responsibly and make sure your project's results are applied effectively.

How to use it?

Step 1: Review the steps in the roadmap

Go trough each of the steps provided in the roadmap. These are there to help you think about how to plan for your data usage .

Step 2: Mark if you have completed them

Simply mark down with an X the open "completed" field. Discuss with your team if different actions are needed in the roadmap before marking them. Feel free also to utilise post-its notes in the different steps to help you discuss each step with your team.

Step 3: If needed, evaluate how you can complete the step

You may utilise other resources in the Data Innovation Toolkit or the Data Innovation Repository to help you complete this high-level overview on your data usage process.

What will you accomplish?

Lead to well-informed decision-making, compliance with data management regulations, and **effective implementation of project outcomes.**

European Commission

Data using checklist

European Commission

1. Conduct a final review of the data to ensure it is accurate

☐ Completed?

2. Translate all of your data insights into actionable tasks

☐ Completed?

3. Document project successes and failures to inform iteration and future work

☐ Completed?

4. Establish a targeted plan for long-term data retention or eventual destruction of your data

☐ Completed?

5. Ensure your data retention complies with laws and policies on the right to be forgotten and privacy

☐ Completed?

6. Develop a communication strategy for sharing results

☐ Completed?

What are the key performance indicators (KPIs) I need to measure success?

Pre-requisite tools:

Data using checklist

Next tool:

Measuring impact and evaluating success tool

European Commission

Data innovation KPI creation tool

What are the key performance indicators (KPIs) I need to measure success?

Resource type	Complexity	Time frame
Worksheet	Medium	40 minutes

Who may support you in developing this resource?

Project manager, data scientist, data architect, data analyst.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

- Article - Taxonomising performance measurement systems' failures - Jelle Van Camp, Johan Braet
- Article - Measuring innovation using key performance indicators – Geanina Banu

What is it for?

Indicators define the type of data that you will track to **measure your solution's effectiveness over time**. They may be quantitative or qualitative depending on what you are trying to track in your data innovation project. This worksheet will help you structure indicators that are specific, measurable, realistic and time-bound, also known as SMART.

How to use it?

Step 1: Identify key focus areas for your data innovation project

Determine the critical aspects of your data innovation project that require monitoring, such as data collection, integration, quality, sharing, analysis, and usage.

Step 2: Set specific KPIs for each area:

For each of these areas, set KPIs that are specific, quantifiable, realistic, aligned with your broader goals, and time bound.

Step 3: Use the worksheet to structure your KPIs:

- Structure your KPIs using the provided worksheet to maintain clarity and coherence.
- Document each KPI, including its definition, measurement method, target values, and responsible team members.

What will you accomplish?

Develop a set of **meaningful and measurable indicators** that monitor both the implementation of your project and the progress towards achieving your short, medium, and long-term outcomes.

European Commission

KPI creation tool

European Commission

What question are you trying to address with data?

Establish KPIs linked to your challenge to measure if it's been answered. Use the question you formulated in the Challenge Definition Tool.

_____?

Process Indicators (Tracking the implementation and uptake of the project)

Think about tracking frequency, scale or quality.

# of _____ that have been _____ during _____.
resource/activity/product — delivered / distributed / accessed — period of time

# of _____ who have _____ a _____ during _____.
specific user group / population — been reached by / attended / purchased — activity / service / product — period of time

Examples

# of active users that have engaged with the new system by at least three logins per day during the next 12 months, aiming for a 20% increase.

# of distinct types of use cases that have been initiated using the project outputs during the first year, targeting at least five different sectors.

Outcome Indicators (Measuring the effectiveness and impact of the project)

Think about tracking and measuring the impact of the projects' outcomes.

% of _____ reached by _____ who _____.
specific user group / population — solution / intervention — shift in knowledge / attitude / behavior

Example

# of cost savings (in euros or % of budget) that have been achieved in data processing and analysis, aiming for a 15% reduction during the first 12 months.

How do I measure the impact and success of my data innovation project?

Pre-requisite tools:

Data innovation KPI creation tool

Next tool:

Data innovation scalability plan

European Commission

Measuring impact and evaluating success tool

How do I measure the impact and success of my data innovation project?

Resource type	Complexity	Time frame
Canvas	Medium	50 minutes

Who may support you in developing this resource?

Project manager, data scientist, data architect, data analyst.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

Article - Data Impact Assessments - The Centre for Humanitarian Data

What is it for?

The impact assessment tool is designed to help you evaluate the effectiveness and outcomes of your data innovation project. By systematically assessing each stage — from input to impact — this tool guides you to reach and deliver **meaningful and measurable benefits.** The goal is to provide a clear framework for understanding the broader effects of your project, both intended and unintended, and to facilitate continuous improvement in future initiatives.

How to use it?

Step 1: Review your data innovation project outputs, read the guidelines.

Examine the immediate results of the activities. The key questions will help you assess whether the outputs met the initial goals and expectations. The material comes with a couple of guidelines to support you in measuring the impact and success of your project.

Step 2: Analyse the current or planned outcomes

Assess the longer-term effects of the project's outputs. Use the questions to understand if the outcomes align with the intended objectives. You can use post-its notes and brainstorm each open field in the table with your team.

Step 3: Measure the long-term impacts of your project

Finally, evaluate the broader, long-term impact of the project. Consider the sustainability of the outcomes and their contribution to larger goals, such as societal benefits or policy changes. Discuss these, if possible, with your team too.

What will you accomplish?

Develop a comprehensive understanding of your **project's success across all stages,** from resources to long-term impact.

European Commission

Measuring impact and evaluating success, guidelines

European Commission

INPUT	ACTIVITY	OUTPUT	OUTCOME	IMPACT
What resources (time, budget, personnel) were invested in the project?	What activities were conducted to achieve the project's objectives?	What were the tangible outputs of the project?	What immediate outcomes did the project achieve?	What long-term impacts has the project achieved or is expected to achieve?
Were the necessary data and tools available and accessible?	Were the planned activities completed on time and within scope?	Did the outputs meet the quality standards and project objectives?	Were the outcomes aligned with the intended project goals?	How can the project's impact be sustained over time?
Identify the key resources allocated to the project, including budget, personnel, time, and data. Consider how these inputs align with the project's goals.	List the core activities, such as data collection, analysis, stakeholder engagement, and software development. Ensure these activities were aligned with the project plan.	Identify the direct products of the project, such as reports, datasets, or software tools. Ensure these outputs meet the quality standards and project objectives.	Evaluate the short-term results, such as changes in stakeholder behaviour or the adoption of new processes. Ensure these outcomes align with the intended objectives.	Consider the broader, long-term effects of the project. These may include changes in policy, improvements in public health, or sustained stakeholder engagement.
[Example] Allocated resources included a budget of 100.000 €, a team of 5 staff members, and 14 months to complete the proof of concept for my data innovation project.	**[Example]** Conducted workshops with EU Member State authorities, developed data integration protocols, and ran pilot tests for a dashboard that showcase data on public transportation.	**[Example]** The primary output was a dashboard to be used by EU Member States Authorities, along with comprehensive technical documentation and user guidelines.	**[Example]** The dashboard was adopted by EU Member States for monitoring issues related with public transportation, allowing them to have an overview on the main problems.	**[Example]** The dashboard is now contributing to better outcomes by facilitating cross-border access to records for public transports, leading to improved mobility in the EU.

Measuring impact and evaluating success tool

European Commission

INPUT	ACTIVITY	OUTPUT	OUTCOME	IMPACT
What resources (time, budget, personnel) were invested in the project?				
Were the necessary data and tools available and accessible?				
Identify the key resources allocated to the project, including budget, personnel, time, and data. Consider how these inputs align with the project's goals.				

What strategies should I use to scale my project effectively?

Pre-requisite tools:

Measuring impact and evaluating success tool

European Commission

Data innovation scalability plan

What strategies should I use to scale my project effectively?

Resource type	Complexity	Time frame
Canvas	Medium	1 - 4 hours

Who may support you in developing this resource?

Project manager, data scientist, data architect, data analyst, data governance officer, data owner.

Data innovation repository

The Data Innovation Repository is a compendium of resources designed to help you find **additional materials** to support your data innovation journey. The materials below can assist you in this specific stage.

Toolkit - Sustain and Improve Your Project - Federal Crowdsourcing and Citizen Science

What is it for?

As you prepare to launch your data innovation project, the next step is to expand its impact. This process involves applying it in new areas, adapting it for broader challenges, and ensuring data remains accurate and secure. Scaling requires careful planning to manage costs, maintain funding, and handle increased complexity—all while staying true to the project's goals. This material provides you with **an initial canvas for your long-term scalability.**

How to use it?

Step 1: Understand key guidelines

Review the guidelines to identify important perspectives to consider as you finalise your project, or as to prepare on how to make your outputs more scalable.

Step 2: Visualise your plan and check the accompanying checklist

Use post-it notes or write directly on the provided canvas to start sketching a scalability plan for your data innovation project. The checklist provided can support you to think on what to write in each of the sections of your plan.

Step 3: Refer to toolkit resources

If needed, revisit materials in the Data Innovation Toolkit and Data Innovation Repository to gather more insights for your scalability plan.

Step 4: Get team and stakeholder feedback

Review the plan with your team and, if helpful, seek input from external stakeholders.

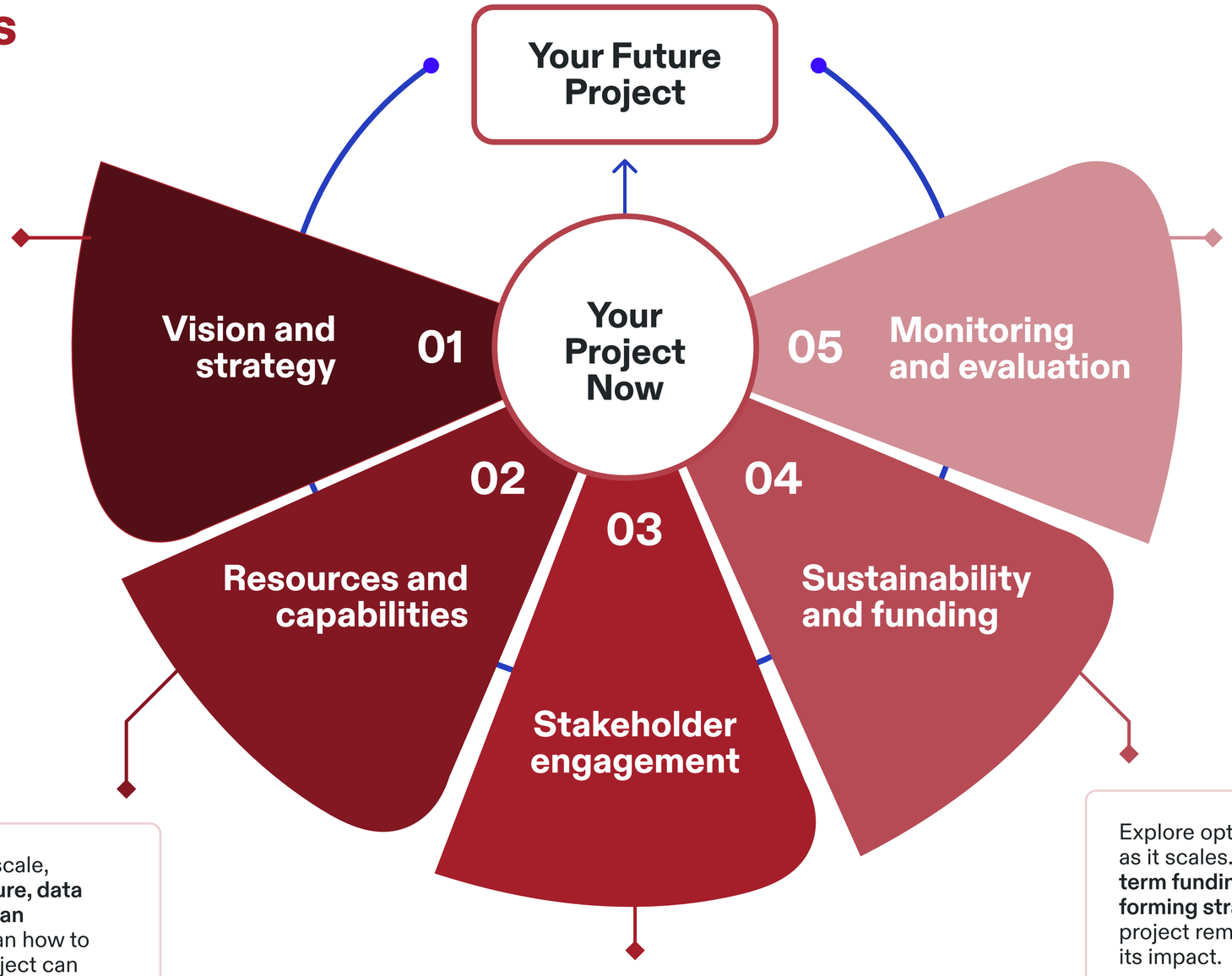
What will you accomplish?

Create a **clear strategy for expanding** your data innovation project.

European Commission

Data innovation scalability plan, guidelines

European Commission



Clearly define the long-term vision for scaling the project. Discuss **how expanding the project will continue to meet its objectives** and the broader goals of your organisation. Consider what changes might be needed in strategy as the project grows.

Plan how to **monitor and evaluate the scaled project to ensure it continues to meet its goals.** This includes setting up systems to track performance metrics, data quality, and compliance, and using this data to make informed decisions about further scaling or adjustments.

Evaluate the resources required to scale, including **technological infrastructure, data management capabilities, and human resources.** Identify any gaps and plan how to address them, ensuring that the project can scale without compromising on data quality or governance.

Assess the readiness and willingness of **current and potential new stakeholders to support the scaled project.** This includes partners, data providers, end-users, and regulatory bodies. Discuss how to maintain and expand stakeholder relationships as the project grows.

Explore options for sustaining the project financially as it scales. This process can involve **securing long-term funding, developing revenue models, or forming strategic partnerships.** Ensure that the project remains financially viable while expanding its impact.

Data Innovation scalability ilan

European Commission

WRITE HERE

WRITE HERE

WRITE HERE

WRITE HERE

WRITE HERE

Data innovation scalability plan: checklist for reflection

1. Vision and strategy

Key questions:

☐ Were you able to articulate a long-term vision for the project that aligns with the organisational goals?

☐ Is there a strategic roadmap that aligns with both short-term and long-term objectives?

☐ Have you identified potential challenges or scenarios that could affect scaling?

☐ Have you defined how the scaled project will continue to meet its objectives and address broader goals?

☐ Is there a strategy in place for adapting the project as it grows?

2. Resources and capabilities

Key questions:

☐ Have you assessed whether the current technology stack can handle increased data volume and complexity?

☐ Are there plans to expand or upgrade technological infrastructure as needed?

☐ Have you evaluated the skill sets required for scaling the project?

☐ Have you identified gaps in human resources and planned for training, hiring, or partnerships?

☐ Are data management practices scalable to ensure data integrity and security during growth?

☐ Is there a plan for maintaining or improving data quality as the project scales?

European Commission

3. Stakeholder engagement

Key questions:

☐ Have you assessed the readiness of current stakeholders to support the project as it scales?

☐ Have you identified new stakeholders that need to be engaged as the project grows?

☐ Is there a communication plan in place that includes regular updates and feedback mechanisms?

☐ Have you revisited the stakeholder map to reflect changes as the project scales?

☐ Have you developed strategies to maintain strong relationships with key stakeholders?

☐ Have you considered the impact of scaling on stakeholder roles and expectations?

4. Sustainability and funding

Key questions:

☐ Have you identified sustainable funding sources for the project as it scales?

☐ Have you developed a detailed financial plan that accounts for the costs of scaling?

☐ Have you explored potential partnerships to provide additional resources and expertise?

☐ Is there a plan for revenue generation or cost recovery as the project grows?

☐ Have you assessed the financial viability of the project over the long term?

European Commission

5. Monitoring and evaluation

Key questions:

☐ Have you established clear performance metrics for monitoring the scaled project?

☐ Is there a system in place to collect and analyse data for informed decision-making?

☐ Are you regularly reviewing and adjusting the project based on monitoring data?

☐ Have you set up compliance checks to ensure ongoing adherence to regulations?

☐ Have you planned for continuous improvement based on insights gained from monitoring?

European Commission

Publications Office
of the European Union