Annex 1. Project Design and Feasibility Sheet

1 Project name:

2 Organization name:

3 Problem definition:	
What is the problem to be solved?	
Describe the population(s) affected. Who or what is affected by this problem? (Certain types of people, organizations, neighborhoods, environment).	
How many people/ organizations/locations/ etc. are affected and to what extent? (For example, average waiting time for surgery, number of students dropping out of school, costs of tax evasion, etc.).	
Why is solving this problem a priority for your institution?	
Do you know of any similar AI use cases that have been implemented before? Which ones?	

4   Pre-feasibility analysis

Is it within our power to act on the problem? Will we have to partner with other public agencies? Do we have the necessary human and financial resources to carry out the project?	
Is the relevant data available (enough to be able to change the current way of responding to the problem)? Is it accessible?	
What are the risks of the project (ethical, social license, implementation, etc.)?	

5   Definition of objectives:

Objectives are usually expressed in terms of improving, maximizing, increasing or decreasing, mitigating, and/or reducing an outcome. The objective must be measurable, which requires establishing a metric or indicator that reflects progress. Achieving the objective should help solve the problem. The technical solution (e.g., a predictive model) is not the objective.

Typical limitations relate to budget, lack of human capital, legal constraints, political will and social license.

Keep in mind that when there are competing objective, you may have to sacrifice something to gain a benefit.

	Objective	Limitations
1		
2		
3		

6 Description of actions:

Actions are those activities that institutions carry out or can carry out in relation to a given problem, such as the specific programs they carry out according to their mission to society or their usual operating processes (hiring, user services, payment of salaries, etc.).

These actions can be improved when the institution has the information generated by the data science project. They should also have a connection with the results generated by the AI system and help achieve the objectives set (previous section).

	Complete		
	Action 1	Action 2	Action 3
Action E.g., responsible sexuality workshop for 13-year-old students/Contraceptive method delivery in the school infirmary. Note: Each action should be explained in a separate box.			
Who executes the action? E.g., psychology and psycho-pedagogy team of each establishment and/or school nursing team.			
On whom or what is the action being taken? E.g., students in schools currently in 7th grade and/or the general student body.			
How often is the decision made to perform this action? E.g., annually/monthly.			
What channels are being used or can be used to perform this action? E.g., face-to-face.			
Other useful information about the action			

7   Data mapping

If the institution is to achieve its objective, the data has to be connected to the actions it will support. Typical AI projects use administrative data as a primary source and enhance it with other data sources in the public domain (censuses, other open data, etc.). Partnering with the private sector or non-profit organizations can help obtain missing data internally.

What data is available internally?

	Complete		
	Data source 1	**Data source 2**	**Data source 3**
Name E.g., hospital discharge system.			
What does it contain? Describe attributes in as much detail as possible (e.g., hospital admission and discharge records nationwide, with patient sociodemographic data, diagnosis, days in hospital, type of health insurance, doctor's information).			
What level of granularity? E.g., transaction, person, organization, location.			
How often is information collected and/or updated once it is captured? E.g., real-time, daily, weekly, monthly, yearly, occasionally.			
Do you have unique and reliable identifiers that can be linked to other data sources? E.g., RUN, SSN, DNI, depending on the country.			
Who is in charge of the data? E.g., the hospital records department.			
How is it stored? E.g., in a database, PDF, Excel, SPSS.			
Additional comments			

What data can you obtain from external private or public sources?

	Complete		
	Data source 1	**Data source 2**	**Data source 3**
Name E.g., air quality record.			
What does it contain? Describe the attributes in as much detail as possible (e.g. concentration of pollutants—such as particulate matter of different sizes—in the air).			
What level of granularity? E.g., hourly geolocated monitoring station.			
How often is information collected and/or updated once it is captured? E.g., daily.			
Do you have unique and reliable identifiers that can be linked to other data sources? E.g., monitoring station code.			
Who is responsible for the data? Ministry of Environment			
Are legal agreements required for the exchange and/or access to the information?			
How is it stored? E.g., downloadable database via an API on an open data portal.			
Additional comments			

In an ideal world, is there additional data relevant to this problem that you would like to obtain (surveys, CCTV, phone records, DNA, range of frequency or granularity for currently available data, etc.)?

8   Analysis/Tool

Typical AI projects include a combination of several analyses, depending on the needs and particularities of each project. Analyses are tools, not the goal of the project.

Choose the right analyses for the right problem.

- The analyses or tools chosen should improve current actions in response to the problem.
- Analyses should be tested, and the validation process must match the objective.

	Complete		
	Analysis/Tool 1	**Analysis/Tool 2**	**Analysis/Tool 3**
Type of analysis/tool E.g., description, prediction, detection, behavioral change			
Purpose of analysis E.g., to understand the historical behavior of individuals; to estimate a patient's risk of disease; to identify actions that would reduce overfishing.			
For what type of actions will the information generated from this analysis be used? E.g., inspection of industrial and artisanal fishing vessels.			
How will you validate this analysis using existing data? E.g., using historical data, conducting a randomized controlled trial, etc.			

9   Ethical and legal considerations

Propor-tionality	**Do you think a data science/ AI system is the right way to solve the problem? Why? Have you evaluated other alternatives?**	
	What negative impacts might your project have? Review similar use cases identified in the section "Problem Definition".	
Social license	**Do you think the users or affected people will find the proposed use of data to solve the problem acceptable? Why?**	
	If the project's target population learns about it, will they approve? Why?	
	Have you identified the justification or legal basis for working with these data?	
	Have you identified the regulations that could affect the project?	
	Will it be necessary to have mechanisms to guarantee the quality of personal data, such as access, deletion or rectification mechanisms?	

Transpa-rency	**Which stakeholders should be aware of the project?** Stakeholders usually include policy makers, front-line workers, civil society organizations, government agencies, people who will be affected by the actions, etc. Please list specific organizations and/or types of people.	
	Have you considered any mechanisms for stakeholders to communicate with the institution about the project?	
	Will it be necessary to explain the decision-making mechanisms or the analysis to be carried out? Why?	
Discrimi-nation/ equity	**What structural inequalities are there in the process and/ or in the environment where the project is inserted?**	
	Are there specific (vulnerable) groups for which equity of outcomes or protection of rights is to be ensured? E.g., groups by gender, age, location, social class, educational level, origin (urban or rural), ethnicity?	
	What biases do you think the data might have?	

Accounta-bility	Who is responsible for providing information about the project if requested?	
	Who is responsible if the system is wrong?	
	Are monitoring, control and evaluation mechanisms in place? How will they be documented and how often?	
	Are training mechanisms in place to ensure that the team in charge understands the responsibilities, as well as the legal and ethical obligations of the project?	

10   Team composition

Generally, artificial intelligence projects require the involvement of various professionals from the same public entity, but also from other related organizations. This includes those responsible for the data, the IT infrastructure and the problem or process in question, as well as experts in data analytics, legal and communications. Add as many lines as required in the following table.

Organization/ Department	Description of desired participation	Counterparty name/role

This worksheet was originally developed by the Center for Data Science and Public Policy at the University of Chicago. For more information about our programs and work, please visit http://da-tasciencepublicpolicy.org or contact info@datascienceforsocialgood.org

Esta versión de la hoja de trabajo ha sido actualizada a través de una colaboración entre el Go-bLab UAI, Carnegie Mellon University y el Instituto Tecnológico de Monterrey.

GobLab UAI is the innovation laboratory of the School of Government of the Universidad Adolfo Ibáñez. Its mission is to contribute to innovation in public policy to benefit society. It works with public agencies, civil society organizations and researchers to achieve more effective, efficient and equitable public policies through data science. For more information, visit https://goblab.uai.cl or contact goblab@uai.cl.