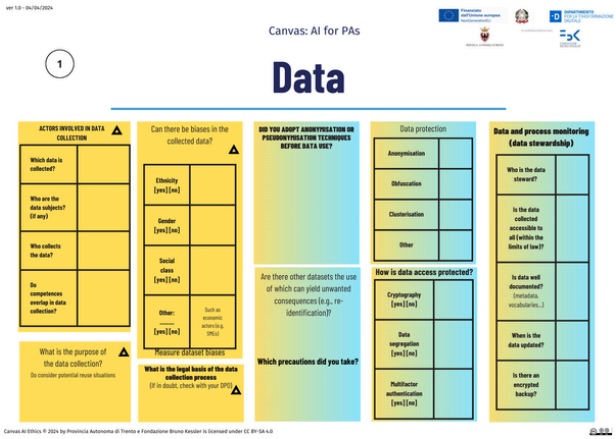
Finanziato dall'Unione europea NextGenerationEU
DIPARTIMENTO PER LA TRASFORMAZIONE DIGITALE
PROVINCIA AUTONOMA DI TRENTO
In collaborazione con: FONDAZIONE BRUNO KESSLER

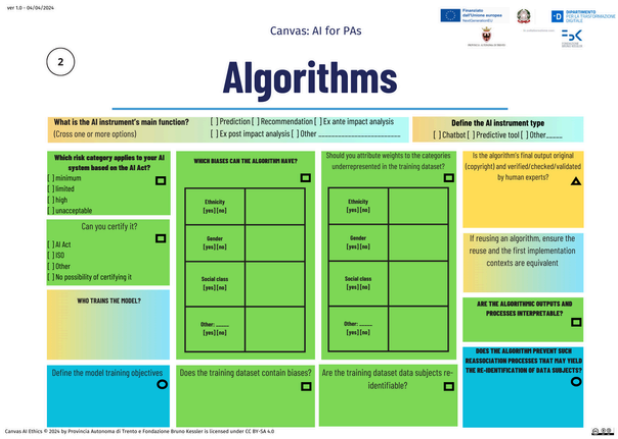
Use case

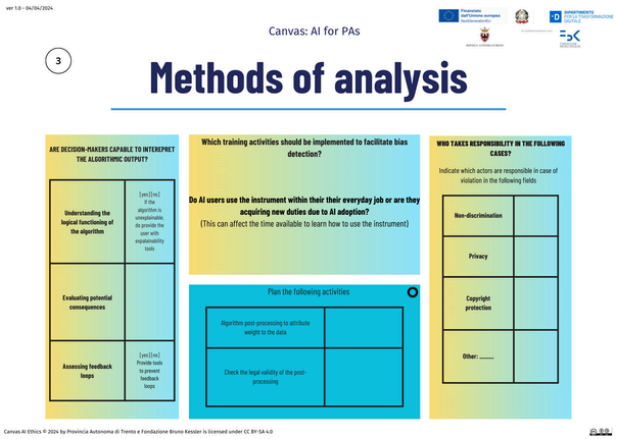
Do you need AI to achieve the objective?

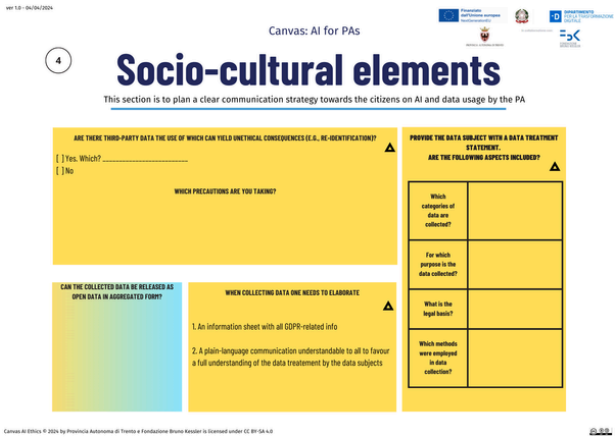
Canvas: AI for PAs

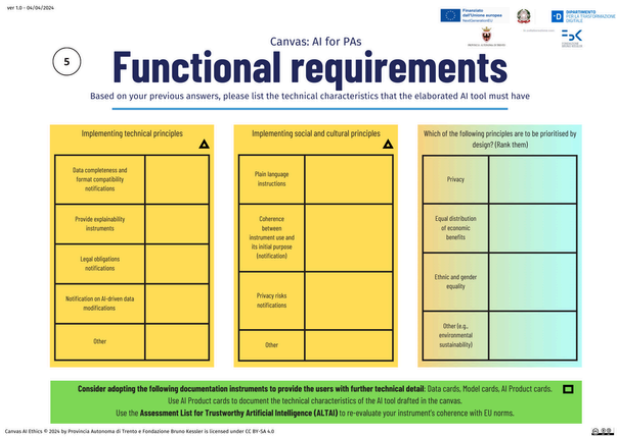
A tool to aid ethical AI adoption in support ot public administrations











Legend

- AI scientist
- AI engineer
- AI user

Finanziato dall'Unione europea NextGenerationEU
DIPARTIMENTO PER LA TRASFORMAZIONE DIGITALE
PROVINCIA AUTONOMA DI TRENTO
In collaborazione con: FONDAZIONE BRUNO KESSLER

How to fill in the canvas

Compile the canvas for each AI instrument at the start of its development phase

Legend

■ AI scientist: researches and develops AI systems (can be a third party)
● AI engineer: trains, deploys and maintains AI systems
▲ AI user: uses AI systems in their everyday job (e.g., a functionary).

How to read the canvas

1. Questions are in boxes of different colours. These match the colours in the legend to indicate which profile should address which questions.
2. Each colour is accompanied by a geometric shape. If the canvas user is colour-blind, they can refer to shapes rather than colours.
3. Mixed-colour boxes (with no geometric shape) indicate questions that should be addressed by all three profiles together.

Usage note

Appoint a moderator when you start your working on the canvas. When you cannot answer a question, do bring another professional profile into the discussion.

CANVAS SECTIONS

Data

To optimise data, reduce their bias, and protect privacy.

Algorithms

To define rules to optimise decision automation while reducing biases.

Methods of analysis

To critically interpret algorithmic outcomes and decide whether and how to apply them

Socio-cultural elements

To ensure that data treatment objectives and scope have been communicated correctly to the target audience.

Functional requirements

To define the AI system's characteristics.

ver 1.0 - 04/04/2024

Finanziato dall'Unione europea NextGenerationEU
DIPARTIMENTO PER LA TRASFORMAZIONE DIGITALE
PROVINCIA AUTONOMA DI TRENTO
In collaborazione con: FONDAZIONE BRUNO KESSLER

Canvas: AI for PAs

1

Data

ACTORS INVOLVED IN DATA COLLECTION

Which data is collected?	
Who are the data subjects? (if any)	
Who collects the data?	
Do competences overlap in data collection?	

What is the purpose of the data collection?

Do consider potential reuse situations

Can there be biases in the collected data?

Ethnicity [yes] [no]	
Gender [yes] [no]	
Social class [yes] [no]	
Other: _____ [yes] [no]	Such as economic actors (e.g, SMEs)

Measure dataset biases

What is the legal basis of the data collection process

(If in doubt, check with your DPO)

DID YOU ADOPT ANONYMISATION OR PSEUDONYMISATION TECHNIQUES BEFORE DATA USE?

Are there other datasets the use of which can yield unwanted consequences (e.g., re-identification)?

Which precautions did you take?

Data protection

Anonymisation	
Obfuscation	
Clusterisation	
Other	

How is data access protected?

Cryptography [yes] [no]	
Data segregation [yes] [no]	
Multifactor authentication [yes] [no]	

Data and process monitoring (data stewardship)

Who is the data steward?	
Is the data collected accessible to all (within the limits of law)?	
Is data well documented? (metadata, vocabularies...)	
When is the data updated?	
Is there an encrypted backup?	

Finanziato dall'Unione europea NextGenerationEU
DIPARTIMENTO PER LA TRASFORMAZIONE DIGITALE
PROVINCIA AUTONOMA DI TRENTO
In collaborazione con: FONDAZIONE BRUNO KESSLER

Canvas: AI for PAs

2

Algorithms

What is the AI instrument's main function?
(Cross one or more options)
[ ] Prediction [ ] Recommendation [ ] Ex ante impact analysis
[ ] Ex post impact analysis [ ] Other _____

Define the AI instrument type
[ ] Chatbot [ ] Predictive tool [ ] Other_____

Which risk category applies to your AI system based on the AI Act?
[ ] minimum
[ ] limited
[ ] high
[ ] unacceptable

Can you certify it?
[ ] AI Act
[ ] ISO
[ ] Other
[ ] No possibility of certifying it

WHO TRAINS THE MODEL?

WHICH BIASES CAN THE ALGORITHM HAVE?

Ethnicity [yes] [no]	
Gender [yes] [no]	
Social class [yes] [no]	
Other: _____ [yes] [no]	

Should you attribute weights to the categories underrepresented in the training dataset?

Ethnicity [yes] [no]	
Gender [yes] [no]	
Social class [yes] [no]	
Other: _____ [yes] [no]	

Is the algorithm's final output original (copyright) and verified/checked/validated by human experts?

If reusing an algorithm, ensure the reuse and the first implementation contexts are equivalent

ARE THE ALGORITHMIC OUTPUTS AND PROCESSES INTERPRETABLE?

DOES THE ALGORITHM PREVENT SUCH REASSOCIATION PROCESSES THAT MAY YIELD THE RE-IDENTIFICATION OF DATA SUBJECTS?

Define the model training objectives

Does the training dataset contain biases?

Are the training dataset data subjects re-identifiable?

Finanziato dall'Unione europea NextGenerationEU
DIPARTIMENTO PER LA TRASFORMAZIONE DIGITALE
PROVINCIA AUTONOMA DI TRENTO
In collaborazione con: FONDAZIONE BRUNO KESSLER

3

Methods of analysis

ARE DECISION-MAKERS CAPABLE TO INTEREPRET THE ALGORITHMIC OUTPUT?

Understanding the logical functioning of the algorithm	[yes] [no] If the algorithm is unexplainable, do provide the user with expalainability tools
Evaluating potential consequences	
Assessing feedback loops	[yes] [no] Provide tools to prevent feedback loops

Which training activities should be implemented to facilitate bias detection?

Do AI users use the instrument within their their everyday job or are they acquiring new duties due to AI adoption?

(This can affect the time available to learn how to use the instrument)

Plan the following activities

Algorithm post-processing to attribute weight to the data	
Check the legal validity of the post-processing	

WHO TAKES RESPONSIBILITY IN THE FOLLOWING CASES?

Indicate which actors are responsible in case of violation in the following fields

Non-discrimination	
Privacy	
Copyright protection	
Other: _____	

Finanziato dall'Unione europea NextGenerationEU
DIPARTIMENTO PER LA TRASFORMAZIONE DIGITALE
PROVINCIA AUTONOMA DI TRENTO
In collaborazione con: FONDAZIONE BRUNO KESSLER

Canvas: AI for PAs

4

Socio-cultural elements

This section is to plan a clear communication strategy towards the citizens on AI and data usage by the PA

ARE THERE THIRD-PARTY DATA THE USE OF WHICH CAN YIELD UNETHICAL CONSEQUENCES (E.G., RE-IDENTIFICATION)?

[ ] Yes. Which? _____
[ ] No

WHICH PRECAUTIONS ARE YOU TAKING?

CAN THE COLLECTED DATA BE RELEASED AS OPEN DATA IN AGGREGATED FORM?

WHEN COLLECTING DATA ONE NEEDS TO ELABORATE

1. An information sheet with all GDPR-related info

2. A plain-language communication understandable to all to favour a full understanding of the data treatement by the data subjects

PROVIDE THE DATA SUBJECT WITH A DATA TREATMENT STATEMENT.
ARE THE FOLLOWING ASPECTS INCLUDED?

Which categories of data are collected?	
For which purpose is the data collected?	
What is the legal basis?	
Which methods were employed in data collection?	

Finanziato dall'Unione europea NextGenerationEU
DIPARTIMENTO PER LA TRASFORMAZIONE DIGITALE
PROVINCIA AUTONOMA DI TRENTO
In collaborazione con: FONDAZIONE BRUNO KESSLER

Canvas: AI for PAs

5

Functional requirements

Based on your previous answers, please list the technical characteristics that the elaborated AI tool must have

Implementing technical principles

Data completeness and format compatibility notifications	
Provide explainability instruments	
Legal obligations notifications	
Notification on AI-driven data modifications	
Other	

Implementing social and cultural principles

Plain language instructions	
Coherence between instrument use and its initial purpose (notification)	
Privacy risks notifications	
Other	

Which of the following principles are to be prioritised by design? (Rank them)

Privacy	
Equal distribution of economic benefits	
Ethnic and gender equality	
Other (e.g., environmental sustainability)	

Consider adopting the following documentation instruments to provide the users with further technical detail: Data cards, Model cards, AI Product cards.

Use AI Product cards to document the technical characteristics of the AI tool drafted in the canvas.

Use the **Assessment List for Trustworthy Artificial Intelligence (ALTAI)** to re-evaluate your instrument's coherence with EU norms.