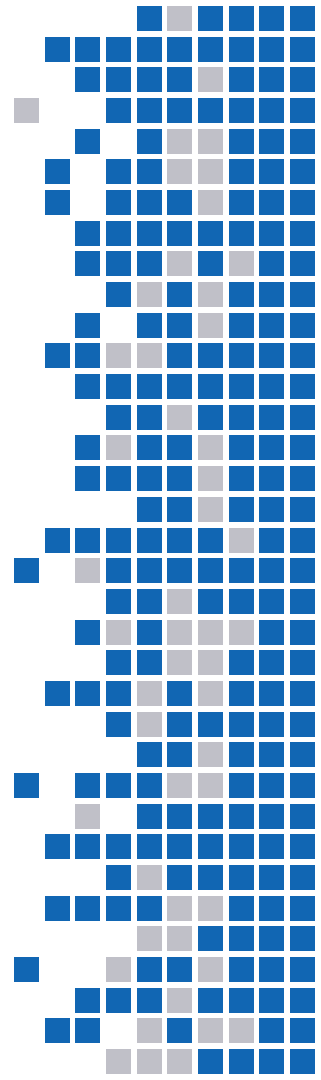


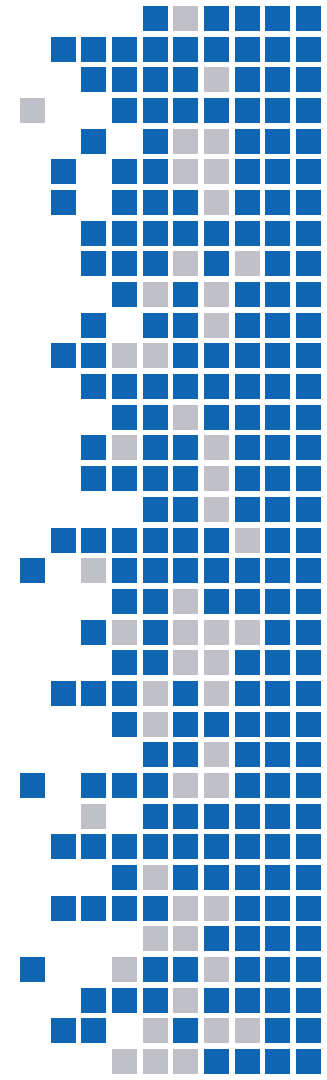
How the Autonomous Province of Trento intends to introduce AI into public administration

Maurizio Napolitano
Fondazione Bruno Kessler





Vision and Goal



AixPA Vision

*"This confirms the image of a country of inventors, innovators, experimenters, but fragile, if not incapable of bringing these results to the system, **powerless in the face of the need to transform individual innovations**, I would say individual, **into actions for the transformation of the entire Country.**"*

[Le opzioni tecnologiche per la digitalizzazione avanzata della Pubblica Amministrazione - TEH Ambrosetti e Salesforce, pag. 21]

BARRIER: **Fragmented and vertical adoption** of IA to support specific individual activities

- limited transformational impact
- technological and methodological fragmentation
- high operating costs

*'[Adopting AI in Public Administration] implies addressing the issue of the transformation of Public Administration itself, because **the main contribution of Artificial Intelligence is to transform the very entity that introduces it.**'*

[Ibid., p. 16]

In the accelerating transformation process that we are experiencing, we need a **coloured PA**: a PA that is able not only to **attract talent**, but also **to let them experiment**; a **courageous** PA that knows how to embrace divergent thinking **in order to create value for citizens and businesses**'.

[ForumPA 2024]

GOAL: **Structural adoption of IA** by the **public administration**, developing its **transformational value**...

- ... starting with the **Autonomous Province of Trento**, but creating the preconditions for nationwide reuse...
- ... with a specific focus on supporting **decision-making processes**

The “pillars” of the project

Systemic approach

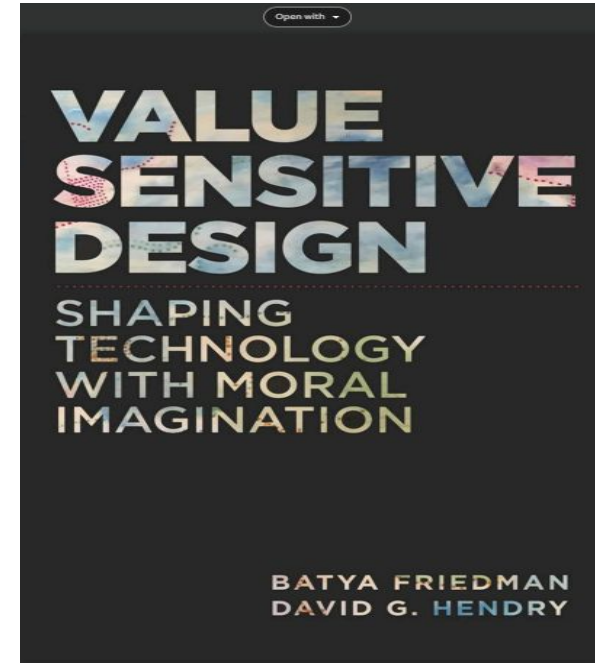
- Working on all aspects contributing to the solution:
 - technological (= platform)
 - methodological (= toolbox)
 - know-how, inside and outside PAT (= training, lab)
 - of operational capacity (= company involvement)
 - process (= process change)
 - governance, ethical and legal (= guidelines, supporting tools)

The “pillars” of the project

Methodology for engaging ‘PA users’

ADVANTAGES of Value Sensitive Design:

- It does not only consider ‘merely functional’ requirements but has a broader scope
-> **transformational value**
- Focuses on value aspects, offering strategies to include them in the design process
-> **greater impact, more responsible innovation**
- Encourages active involvement of users
-> **reduces the risk of disconnection** between developed solution and PAT needs



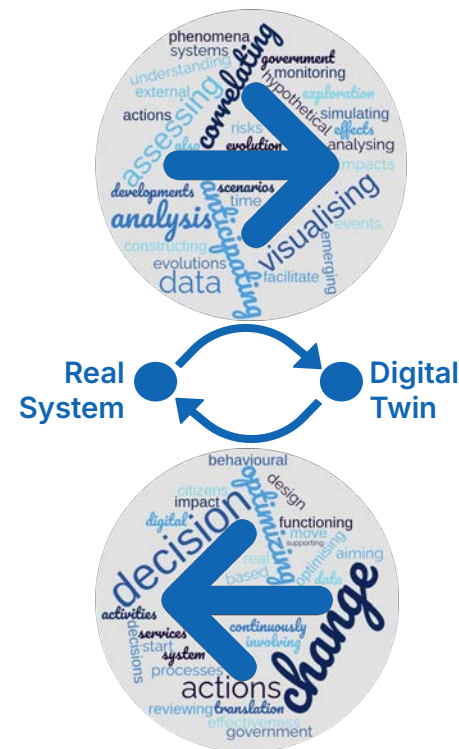
[Friedman & Hendry (2019). *Value sensitive design: Shaping technology with moral imagination*. MIT Press.]

The “pillars” of the project

Digital Twin for decision-making support

REFERENCE MODEL: Digital Twin

- **Complete and accurate digital model** of a system, fed by **data** collected (also in real time) by the system itself, **capable of supporting decision-making** processes through advanced analysis and forecasting functions, and of **adapting** to changes in its real counterpart
- **Maximalist approach** to system modelling
 - Complete system model
 - Large set of related Datasets
 - Support for various decision-making processes associated with the system
- Suitable for highly **complex systems**



AixPA in a nutshell



Financing

Complementary National Plan
'Digital Services and Digital Citizenship



Budget

5 M Euro



Duration

May 2023 - 31 October 2025

Autonomous Province of Trento
Digital Transformation
Project Promoter - Overall Coordination

Autonomous Province of Trento
Departments
Designing and testing
of use cases

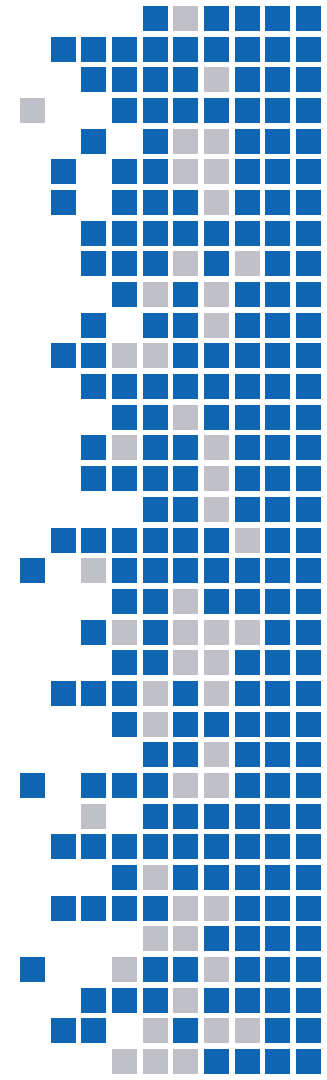


Bruno Kessler Foundation
Implementing entity - Scientific and
technical direction

Companies
Production and maintenance
of solutions

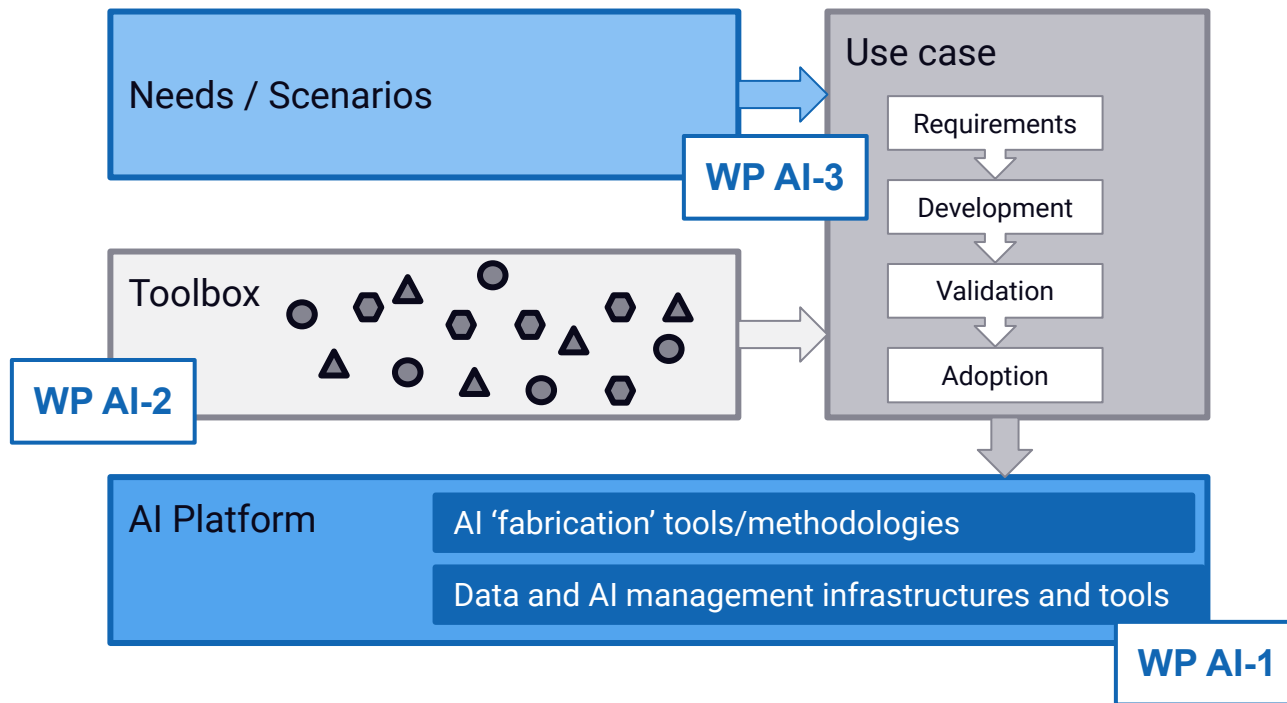


Project output and work plan



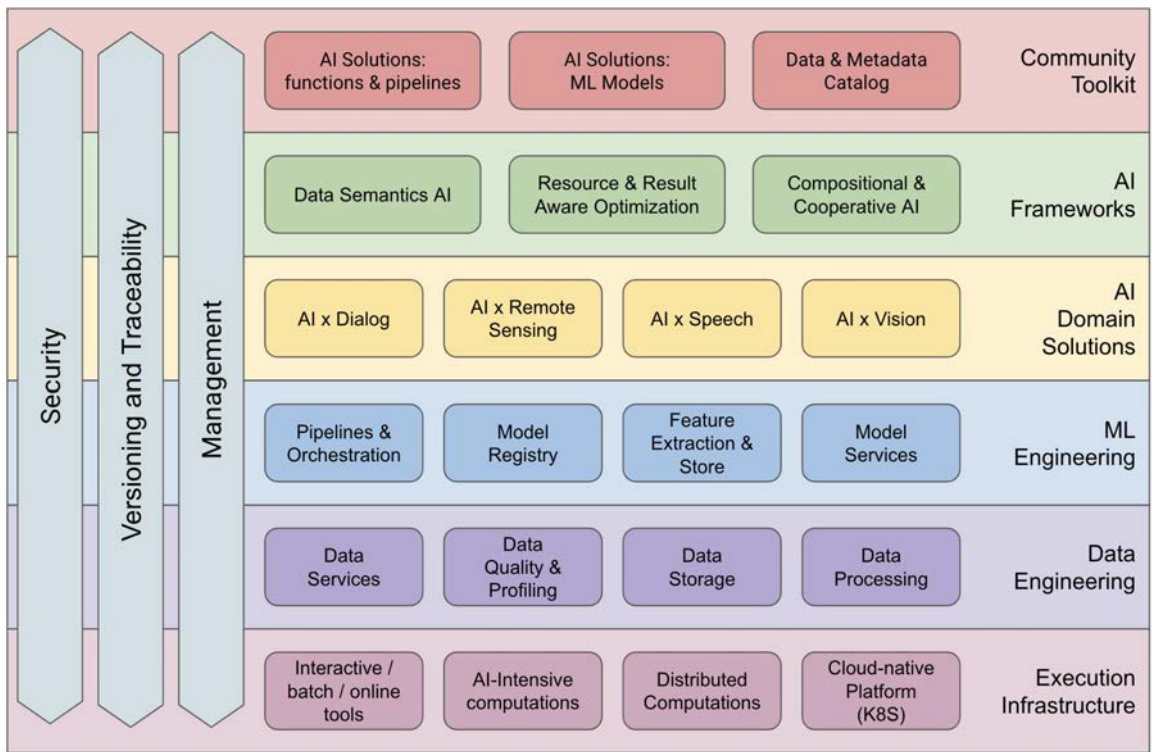
AIxPA

Project and Process Output



Output of the project

Technology platform



AI Factory

Support for creating, reusing, optimising, adapting, composing AI Products

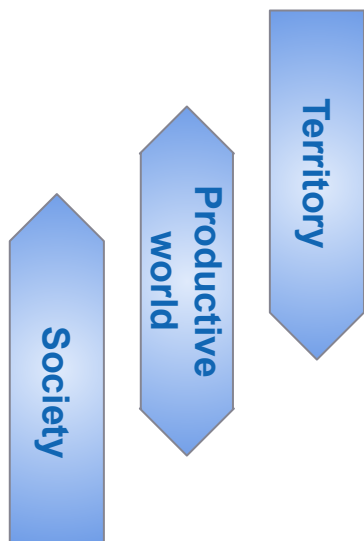
AI Management System

Support for the life cycle of AI products:

- Development
- Execution
- Monitoring
- Evolution

Output of the project

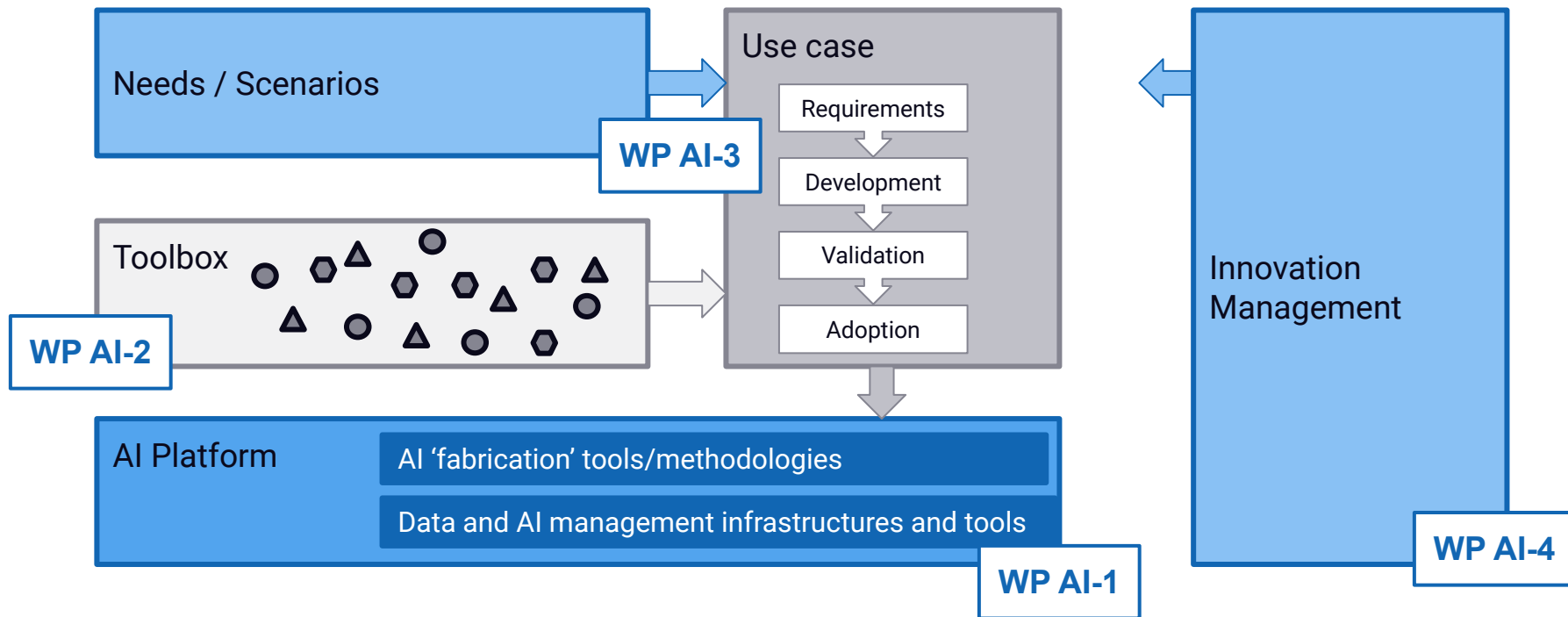
Use cases




- **Civil Protection**
 - Territorial risk management
- **Turismo**
 - Overtourism
- **Coesione sociale**
 - “Family plan”

AixPA

Project and Process Output




ver 1.0 - 04/04/2024




Use case

Canvas: AI for PAs


Data



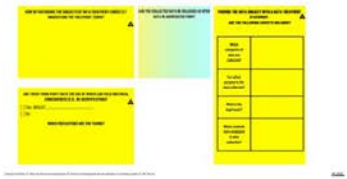
Algorithms




Methods of analysis



Socio-cultural elements



Technical requirements



Legend

- AI scientist
- AI engineer
- △ AI user

AixPA

AI Ethics Canvas - EU Regulations

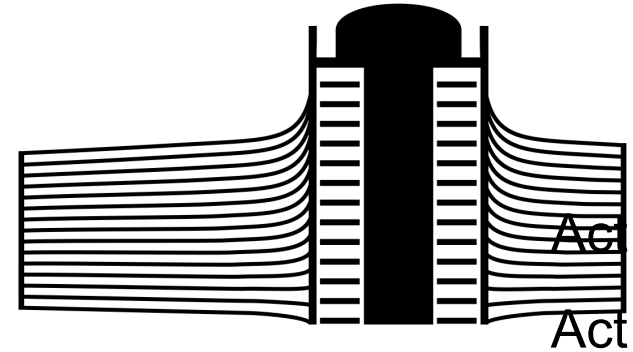
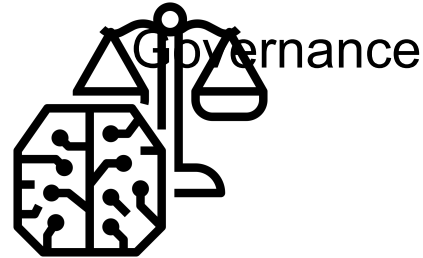
General Data Protection Regulation

Open Data Directive

Data

Data

AI Act



AixPA

AI Ethics Cavas - Ethical framework

EU Ethics guidelines for trustworthy AI

Italian deontological rules for data treatment for statistical or scientific research purposes

EU Data protection and privacy ethical guidelines

EU Ethics for Researchers

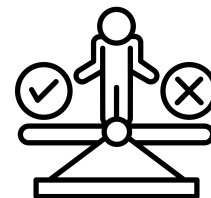
Open Data Institute: Data Ethics Canvas

Toronto Declaration: Protecting the right to equality and non-discrimination in machine learning systems

Ada Lovelace Institute's Algorithmic impact assessment: a case study in healthcare

Ethics as a service (**Digital Catapult**)

NIST Artificial Intelligence Risk Management Framework



AI Ethics Canvas - What does the canvas investigate?

Technical principles

Data

- GDPR, **governance**, transparency, quality, biases

AI Algorithms

- bias, **transparency**, responsibility, quality, review, norms

Methods of analysis

- transparency, **explainability**, evaluation, data collection, revision

Cultural and social principles

Transparency in AI use

Informing

training

Interdisciplinary working groups

Inclusion and social responsibility

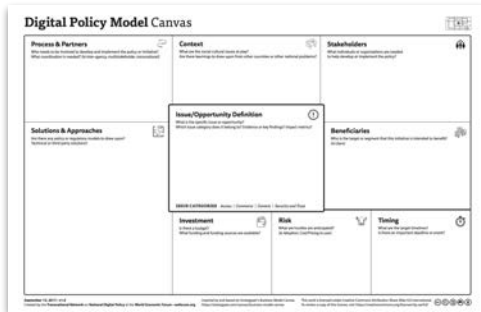
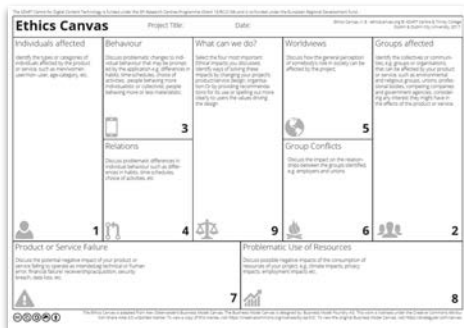
Inclusion, monitoring, collaboration, feedback, communication



towards
AI GOVERNANCE

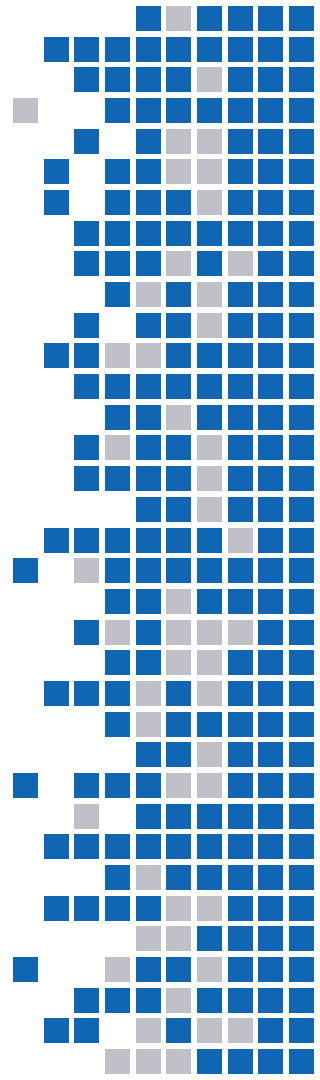
AixPA

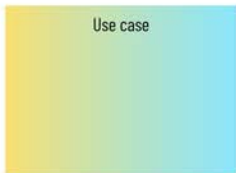
AI Ethics Canvas - What's canvas?





How use the canvas





Canvas: AI for PAs

A graphical tool to aid AI adoption in public administrations

AI scientist

Elaborates algorithms and develops AI instruments

AI engineer

Trains, implements, and maintains AI systems

AI user

Uses AI instruments in their daily activities

Use case

AI INSTRUMENT TO SUPPORT
TEACHERS IN STUDENT
EVALUATION

Canvas

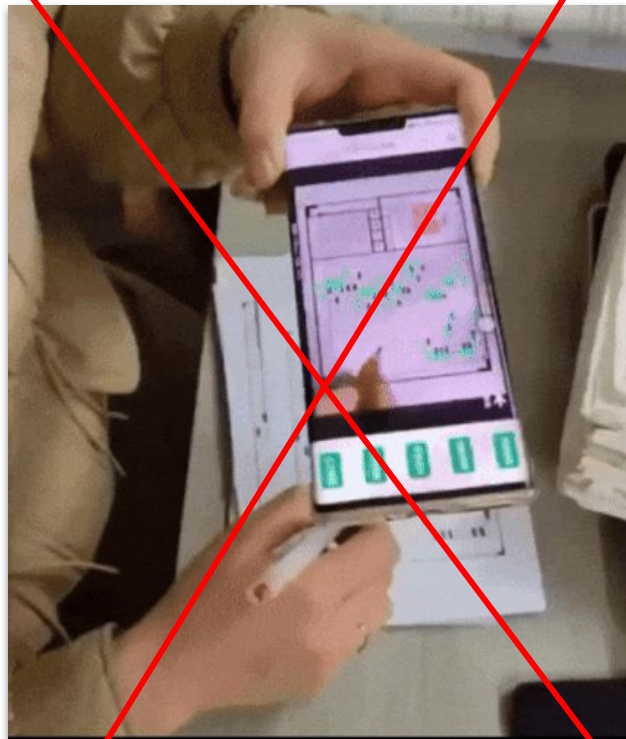
A graphical tool to a

Use case
AI INSTRUMENT TO SUPPORT
TEACHERS IN STUDENT
EVALUATION



Use case

AI INSTRUMENT TO SUPPORT TEACHERS IN STUDENT EVALUATION



Canvas: AI for PAs

1

Data

- AI scientist
- AI engineer
- AI user

ACTORS INVOLVED ▲

Which data is collected?	PUPILS' TESTS
Who are the subjects of the data collection?	PUPILS
Who collects the data?	TEACHERS

What is the purpose of the data collection? ▲

GRADING STUDENTS

Which groups of people are underrepresented in the data collected? ▲

Ethnicity [yes] [no]	FOREIGN STUDENTS
Gender [yes] [no]	SENDING PERFORMANCES?
Social class [yes] [no]	
Other: DSA [yes] [no]	STUDENTS WITH SPECIAL LEARNING NEEDS

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?

NO, TESTS AND GRADES ARE PERSONAL AND MUST BE IDENTIFIABLE

Can non-personal data lead to re-identifying data subjects? (e.g., if cross-checked with other databases) NOT APPLICABLE

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

STUDENTS' DIGITAL TRACES CAN DISCLOSE INFORMATION IF CROSSCHECKED WITH GRADES

Which precautions did you take?

ONLY TEACHERS, PARENTS AND SCHOOL AUTHORITIES HAVE ACCESS TO THE GRADES

Data protection

Anonymisation	
Obfuscation	
Clusterisation	
Other	

How is data access protected?

Cryptography [yes] [no]	
Password [yes] [no]	
Multifactor authentication [yes] [no]	

Data and process monitoring (data stewardship)

Who is the data steward? (training set and collected data)	PROVINCE SCHOOL OFFICE
Is the data collected accessible to all offices (except when legally forbidden)?	YES
Is data well documented? (metadata and shared vocabularies)	NA
When is the data updated?	AFTER EACH TEST

2

Algorithms

- AI scientist
- AI engineer
- AI user

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

Define the AI instrument type
 Chatbot Speech-to-text 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?
 PROVINCE SCHOOL OFFICE

Define the model training objectives
 SUPPORTING THE TEACHER IN STUDENT EVALUATION

WHICH BIASES CAN THE ALGORITHM HAVE?

Ethnicity [yes] [no]	NON-RECOGNITION OF LINGUISTIC BARRIERS
Gender [yes] [no]	GENDERED SCHOOL PERFORMANCE EVAL
Social class [yes] [no]	
Other: DSA [yes] [no]	SPECIAL LEARNING NEEDS

Does the training dataset contain biases?

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]	SPECIAL LEARNING NEEDS

Are the training dataset data subjects re-identifiable?

 No, DATA IS ANONYMISED PRIOR TO TRAINING

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

 NOT APPLICABLE

Is the algorithm used in the context for which it was originally conceived?

 YES, SCHOOL PERFORMANCE EVALUATION

ARE THE ALGORITHMIC OUTPUTS AND PROCESSES INTERPRETABLE?

 YES

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

 YES

3

Methods of analysis

- AI scientist
- AI engineer
- AI user

ARE DECISION-MAKERS CAPABLE TO INTERPRET THE ALGORITHMIC OUTPUT?

Understanding the technical functioning	[Xes][no] If the algorithm is unexplainable, do provide the user with explainability tools
Evaluating potential consequences	TEACHERS ARE TRAINED
Assessing feedback loops	[yes][no] Provide tools to prevent feedback loops

Which training activities should be implemented to facilitate bias detection?

TEACHERS RECEIVE AD HOC SECURITY TRAINING

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)

THE INSTRUMENT IS NEW FOR THE TEACHERS BUT IT AIDS THEIR EVERYDAY WORK.
THEY ARE GIVEN TRAINING ON HOW TO USE IT AND MITIGATE RISKS

Plan the following activities

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

WHO TAKES RESPONSIBILITY IN THE FOLLOWING CASES?

Non-discrimination	HEAD TEACHER
Privacy	TEACHER
Copyright protection	
Other: _____	

4

Socio-cultural elements

- AI scientist
- AI engineer
- AI user

HOW DO YOU ENSURE THE SUBJECTS OF DATA TREATMENT CORRECTLY UNDERSTAND THE TREATMENT TERMS? ▲

PARENTS ARE PROVIDED WITH PLAIN-LANGUAGE INFORMATION SHEETS, START OF SCHOOL YEAR MEETINGS ARE HELD, AND PARENTS SIGN DUE AUTHORISATIONS

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

PROVIDE THE DATA SUBJECT WITH A DATA TREATMENT STATEMENT.

ARE THE FOLLOWING ASPECTS INCLUDED? ▲

Which categories of data are collected?	STUDENT TESTS
For which purpose is the data collected?	STUDENT PERFORMANCE EVALUATION
What is the legal basis?	STUDENT EVALUATION FOR ADMISSION TO THE NEXT SCHOOL YEAR
Which methods were employed in data collection?	DATA IS COLLECTED MANUALLY BY THE TEACHERS THROUGH TESTS

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

Yes. Which? DIGITAL TRACES

No

WHICH PRECAUTIONS ARE YOU TAKING?

GRADES ARE NOT SHARED WITH THIRD PARTIES OUTSIDE SCHOOL AUTHORITIES. THEY ARE ENCRYPTED AND PW-PROTECTED

5

Canvas: AI for PAS

Functional requirements

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

- AI scientist
- AI engineer
- AI user

Implementing technical principles ▲

Data completeness notifications	YES
Copyright violation notifications	NOT APPLICABLE
Legal obligations notifications	YES
Other	NOTIFY PLAGIARISM

Implementing social and cultural principles ▲

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

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

Privacy	4
Equal distribution of (economic) benefits	1
Ethnic and gender equality	2
Other (e.g., environmental sustainability)	3 CORRECT EVALUATION OF STUDENTS WITH SPECIAL LEARNING NEEDS

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



README CC-BY-SA-4.0 license

canvas AI ethics

a canvas to plan the development of an AI solution

description

[A canvas to design and deploy ethical artificial intelligence for public administrations. The experience of the Autonomous Province of Trento, Italy](#)

download pdf

[Italian version](#)

sources

Italian version with license

A canvas to design and deploy ethical artificial intelligence for public administrations. The experience of the Autonomous Province of Trento, Italy

22 Pages · Posted: 7 May 2024 · Last revised: 14 May 2024

Pietro Giovanni Bizzaro
affiliation not provided to SSRN

Riccardo Nanni
Fondazione Bruno Kessler - Digital Commons Lab

Maurizio Napolitano
Bruno Kessler Foundation

Date Written: May 7, 2024

Abstract

Artificial Intelligence (AI) has reached a new hype in the last few years. However, its definition and development has been ongoing since the 1950s, while governments all over the world have started considering how to regulate it and deploy it in the everyday running of public administration (PA). Applying AI to PAs promises to accelerate bureaucratic processes and to provide functionalities with instruments to elaborate ad hoc solutions oriented to citizens' needs. However, slow decision-making mechanisms, lack of competence, and the novelty of AI-focused legislation constitute bottlenecks in the process. In the European Union (EU), supranational institutions are adopting a number of regulatory measures to regulate the (mis)use of AI and data, the latter being an essential raw material for the former. In Italy, where this study is conducted, the Agency for Digital Italy (AgID) has stressed the role of AI in the PA in its 2024-2026 Three-Year Plan for Informatics in the PA and is releasing guidelines at the time of writing. This article presents a canvas to guide PAs in the development, deployment, and monitoring of AI systems for their own everyday use. This canvas has been elaborated within the scope of a project involving the Autonomous Province of Trento (PAT), Italy, and is currently undergoing further development and validation. While being oriented towards the use by the Autonomous Province of Trento, the canvas is applicable to all PAs who wish to adopt AI in their operations. Its scope of application is mostly limited to the EU as it makes reference to requirements expressed in EU legislation.

Keywords: Artificial Intelligence, Public administration, European Union, Italy, Digitalisation

Suggested Citation:

Bizzaro, Pietro Giovanni and Nanni, Riccardo and Napolitano, Maurizio, A canvas to design and deploy ethical artificial intelligence for public administrations. The experience of the Autonomous Province of Trento, Italy (May 7, 2024). Available at SSRN: <https://ssrn.com/abstract=4819535> or <http://dx.doi.org/10.2139/ssrn.4819535>

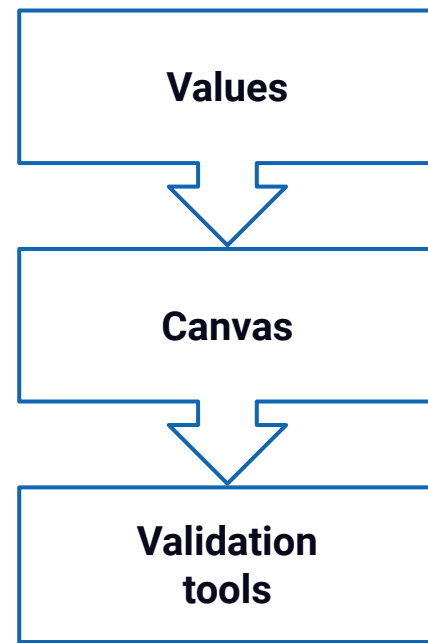
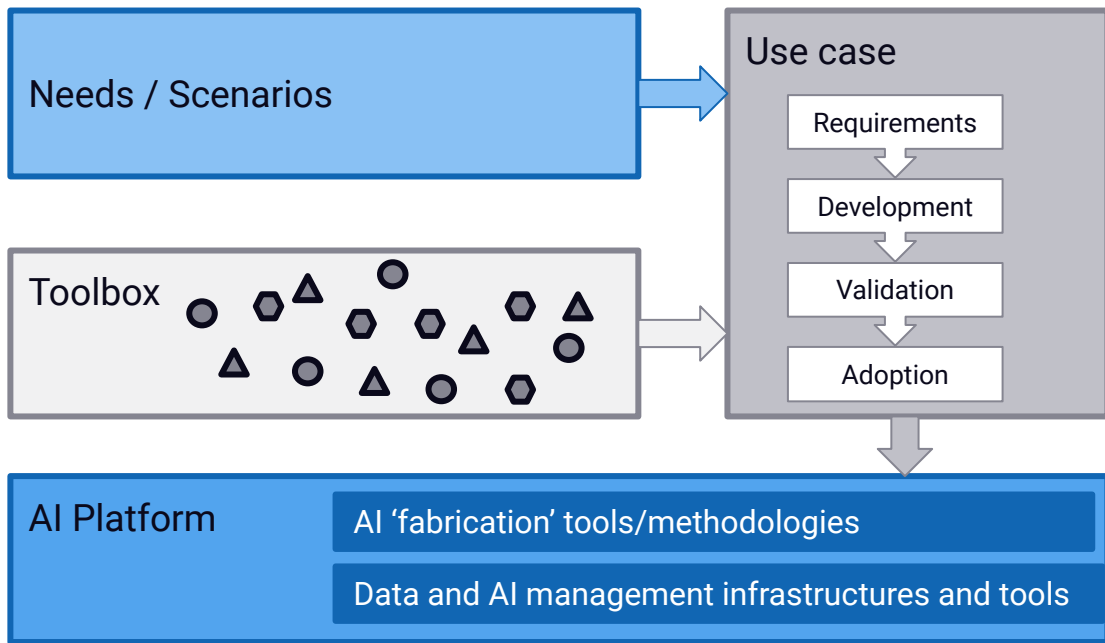
Show Contact Information >

Pietro Giovanni Bizzaro
<pbizzaro@fbk.eu>
Maurizio Napolitano
<napolitano@fbk.eu>
Riccardo Nanni
<rnanni@fbk.eu>
Munazza Usmani
<musmani@fbk.eu>

<https://github.com/dclfbk/canvasaiethics>
The article can be commented [here](#)

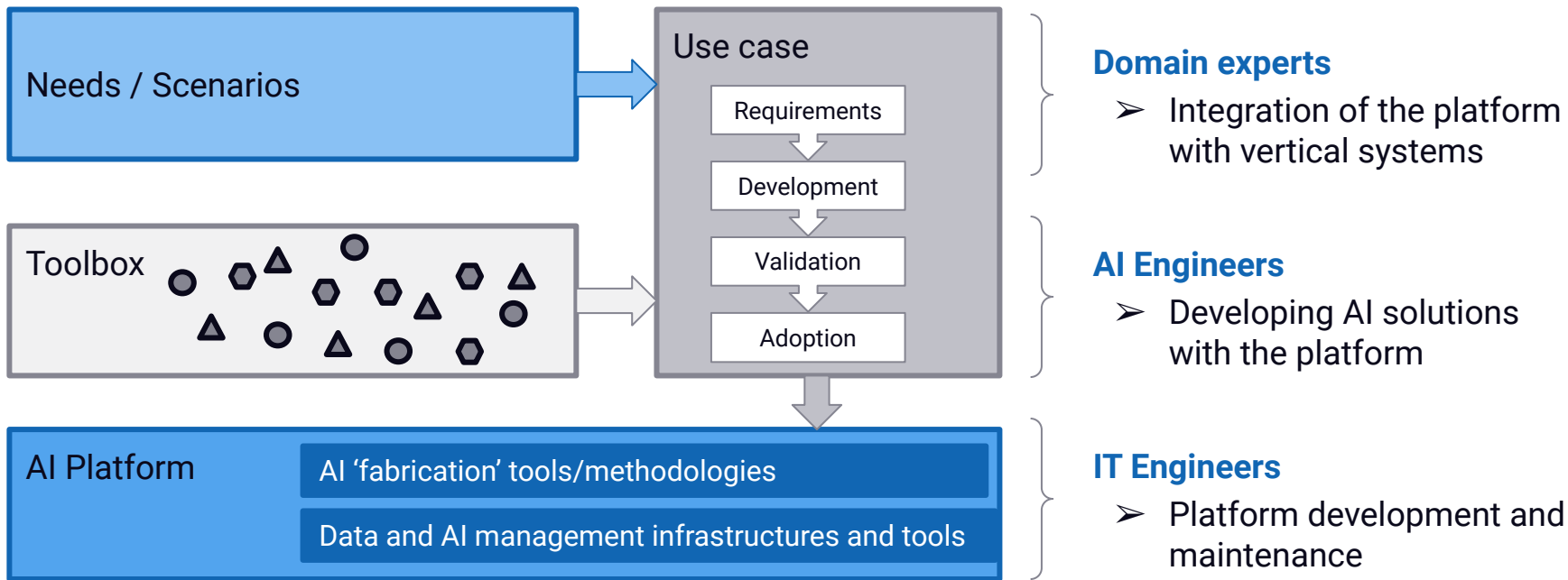
AIxPA

Project and Process Output



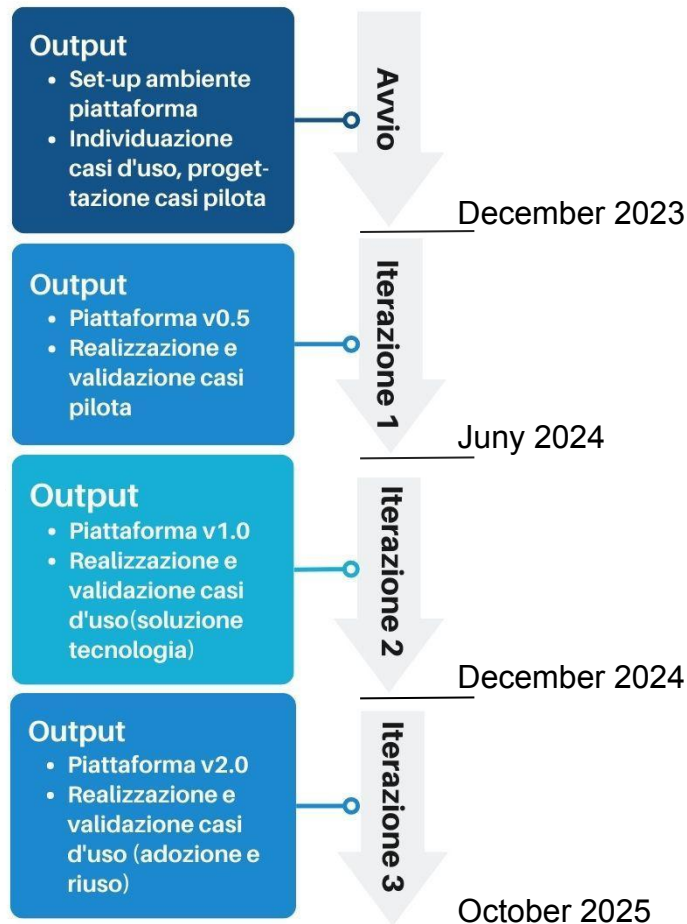
AixPA

Co-innovation ecosystem



AixPA Project Plan

- **Incremental approach** (iterative)
- **Involvement** of PAT end-users at all stages, progressive extension to other actors in the ecosystem
- **Parallel evolution** of technological and methodological components and **use cases**



Thanks

Presentation by
Maurizio Napolitano
<napolitano@fbk.eu>

FBK project coordinator
Marco Pistore - <pistore@fbk.eu>



In collaborazione con:

