

AI-fairness: the FAIRBRIDGE approach to practically bridge the gap between socio-legal and technical perspectives

Giovanni Ciatto*, **Mattia Matteini***, Laura Sartori•, Maria Rebrean°, Catelijne Muller°, Andrea Borghesi*, and Roberta Calegari*

**Dept. of Computer Science and Engineering, University of Bologna, Italy*

•Dept. of Political and Social Sciences, University of Bologna, Italy

°ALLAI, Netherlands

*58th Hawaii International Conference on System Sciences
January 7-10, 2025*

Introduction

AI Fairness

Make AI systems **impartial, equitable, and just**

Fair-by-Design systems

Ensuring AI Fairness from the **outset**

Multidisciplinary challenge

Need to combine **social, legal, and technical** perspectives

Goals

1. Defining a meta-methodology (FairBridge)

Rules, methods, and best practices designers should follow to make fair AI systems

2. Reifying it into software system

Steering AI designers' decision-making, balancing automation and human-in-the-loop

Problems

Unclear guidelines

- Socio/legal guidelines are **vague** for technicians
- Many **notions** of fairness
- Many fairness **metrics**
- Too much room for **intepretation**

Different visions / jargons

- Integrating **multiple perspectives** is paramount...
- ... but harmonizing **multi-disciplinary expertises** is challenging

What is (un)fair?

- Legal and social prescriptions are **domain-specific**
- Potential **tensions** among **social, legal**, and **technical** dimensions

(Meta-)Methodology Desiderata

The methodology should:

- consider the **cultural context** and the **domain** AI system under design
- **assist experts** in translating the social-legal **requirements** into **practice**
- **adapt** to any **change** in the cultural context as it evolves
- account for **pre-existing** AI **systems, datasets, and algorithms**

The FairBridge Approach

Question–Answering

Questions & admissible answers steer decision-makers to **build fair AI systems**

What Q/A and Why?

Q/A **designed by experts** in law, sociology, statistics, and CS
to deepen decision-makers awareness.

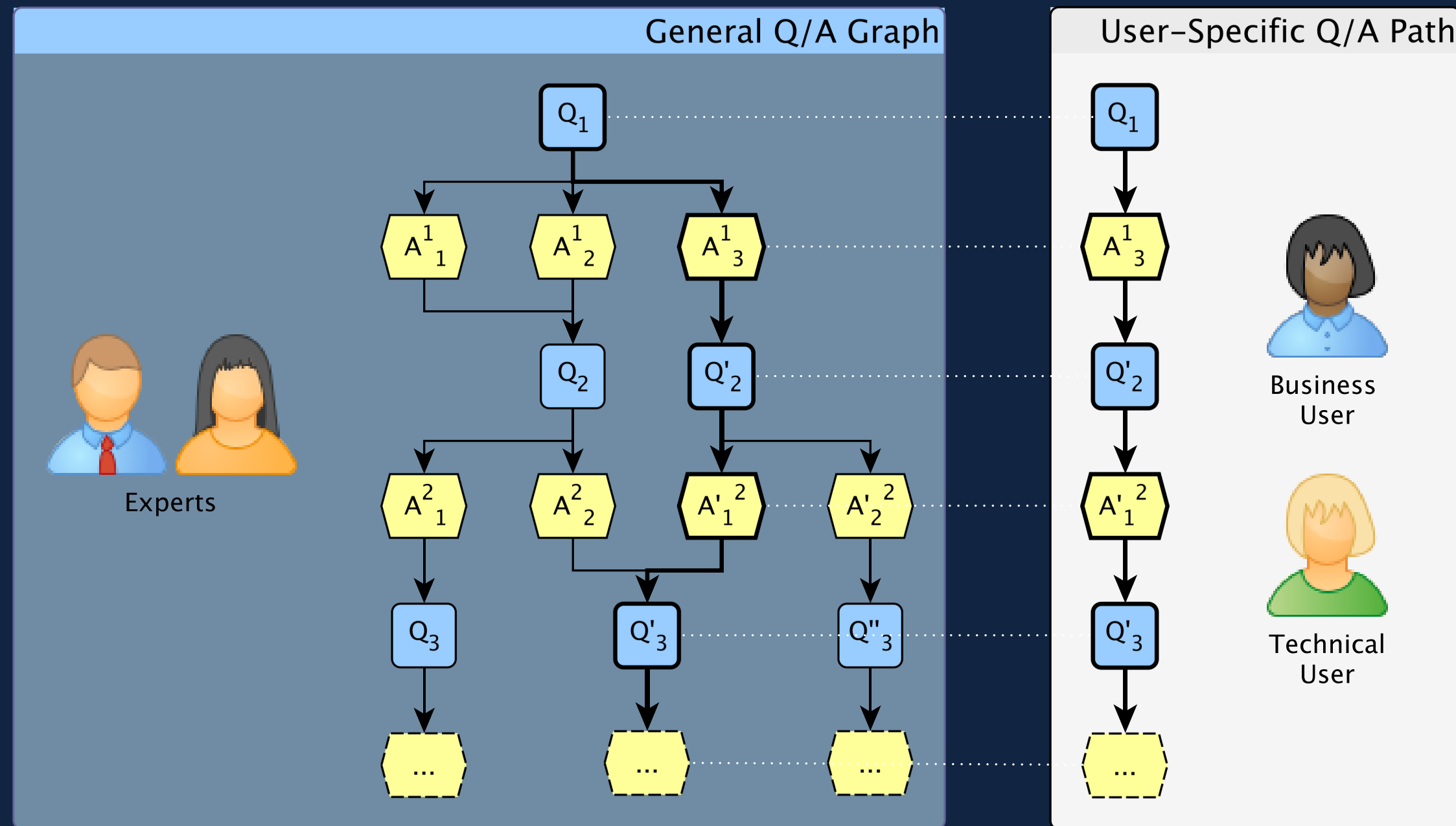
Order of questions

Hierarchical (most general first), answers impact the next questions

Decision Support

Advising the user, while keeping them responsible for choices

Question-Answering Flow



A practical example

Q1. “In which situation are you among the following ones?”

A. Data and algorithms available B. Only data available C. No data nor algorithms available

Q2. “Choose a dataset or load your own.”

...

...

Q3. “What are the sensitive and output features?”

**List of dataset features automatically
computed by the system**

Q4. “Specify the proxies for the sensitive features”

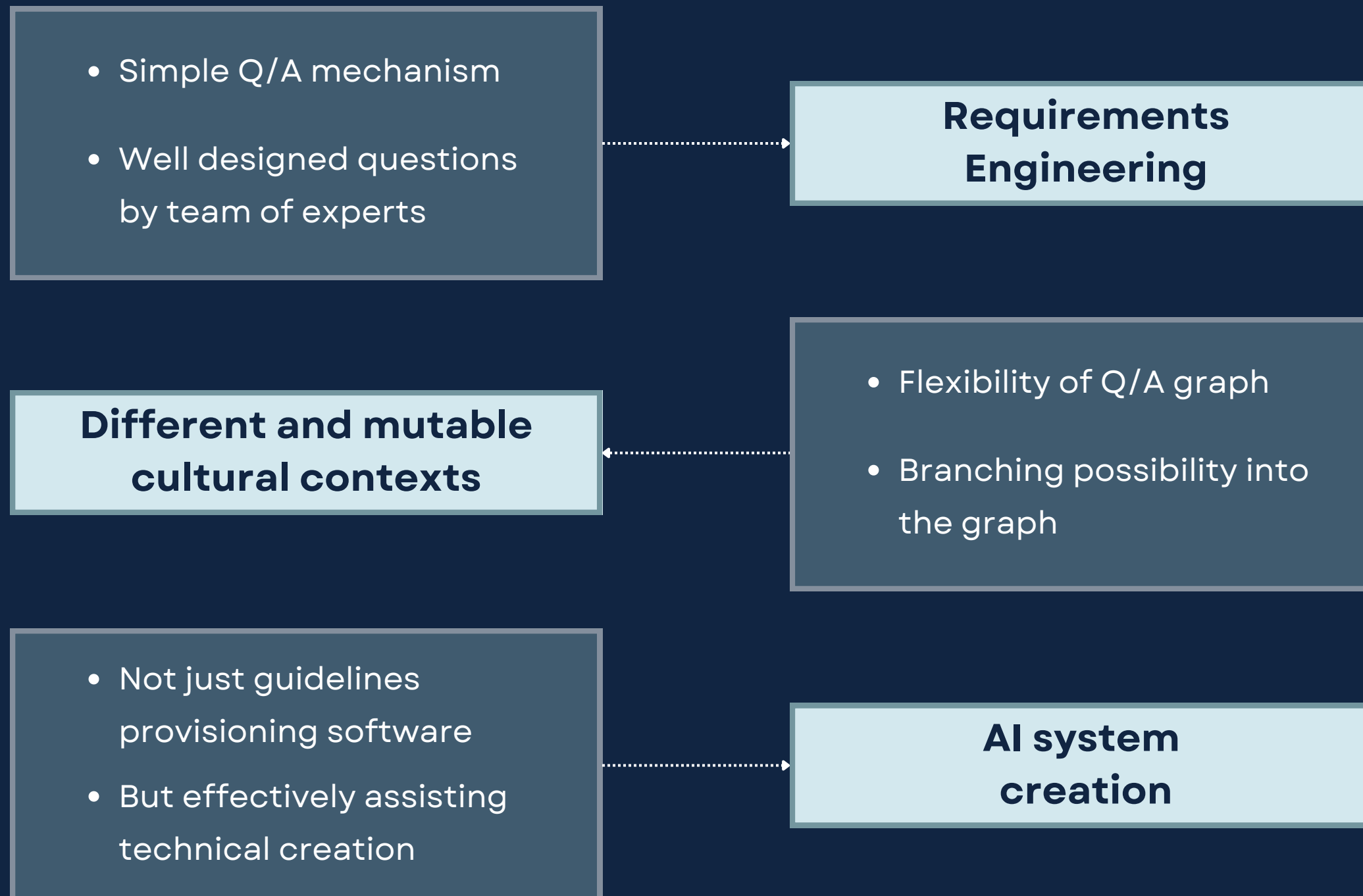
List of dataset features

...



Proxies suggested
by the system

Discussion



Conclusions

This contribution addresses the problem of building FbD methodologies merging social, legal, ethical, and technical perspectives.

The proposed approach does not directly provide a methodology, instead, provides a *meta-methodology* (using a Q/A based mechanism) to build FbD methodologies.

The latters, remain versatile and evolvable against law volatility and possible cultural changes.

Future works

- Concluding the development of the software system that reifies the meta-methodology proposed.
- Iterative refinement of Q/A graph design to improve decision-makers understanding, critical decision steps, etc...
- Automating even more computations that are consolidated and repetitive.

Thank you for listening

Also special thanks to:

