Algorithm Audit

Definition of an AI system under the AI Act

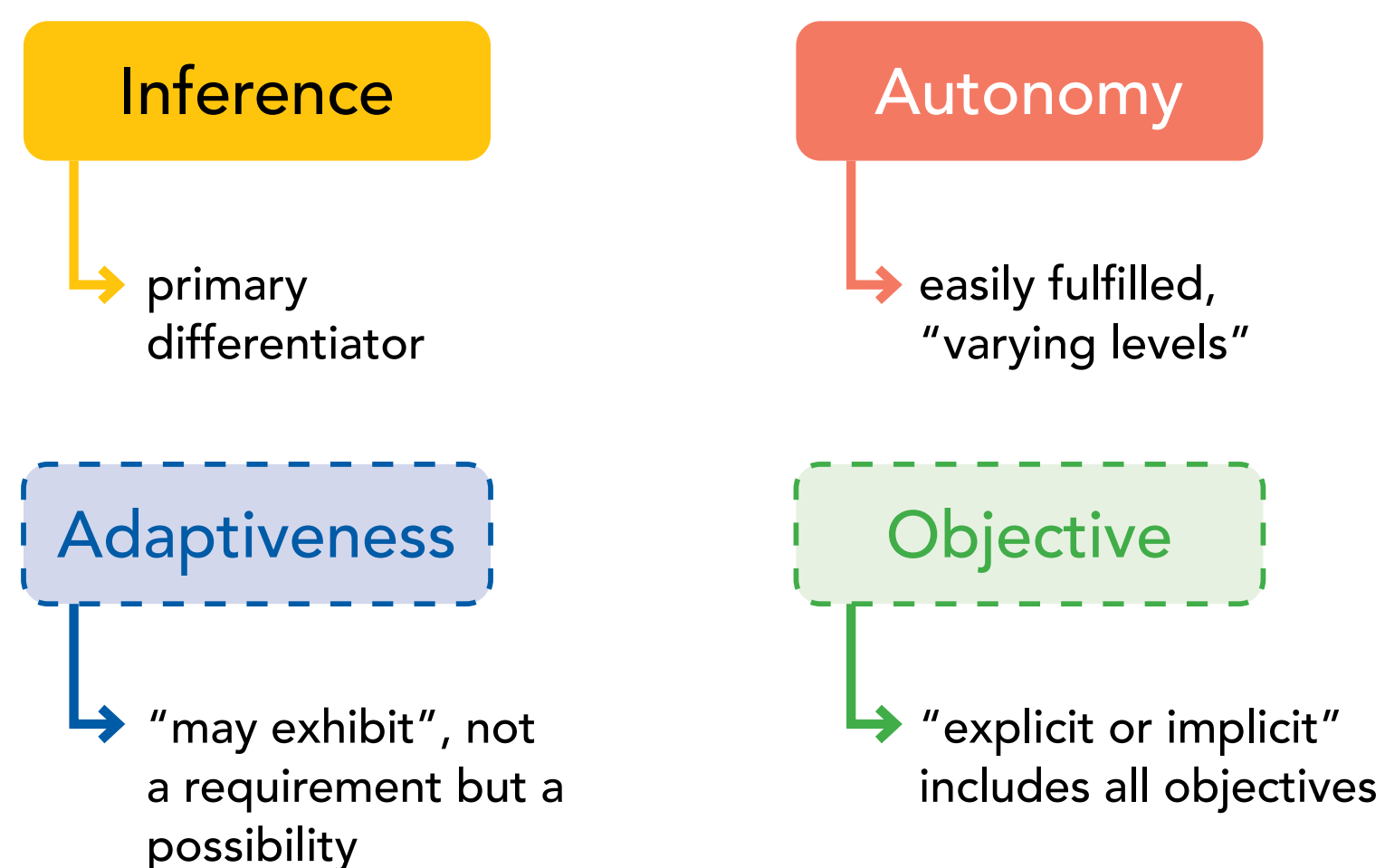*To be or not to be* an AI system

Discover the answer by answering three questions
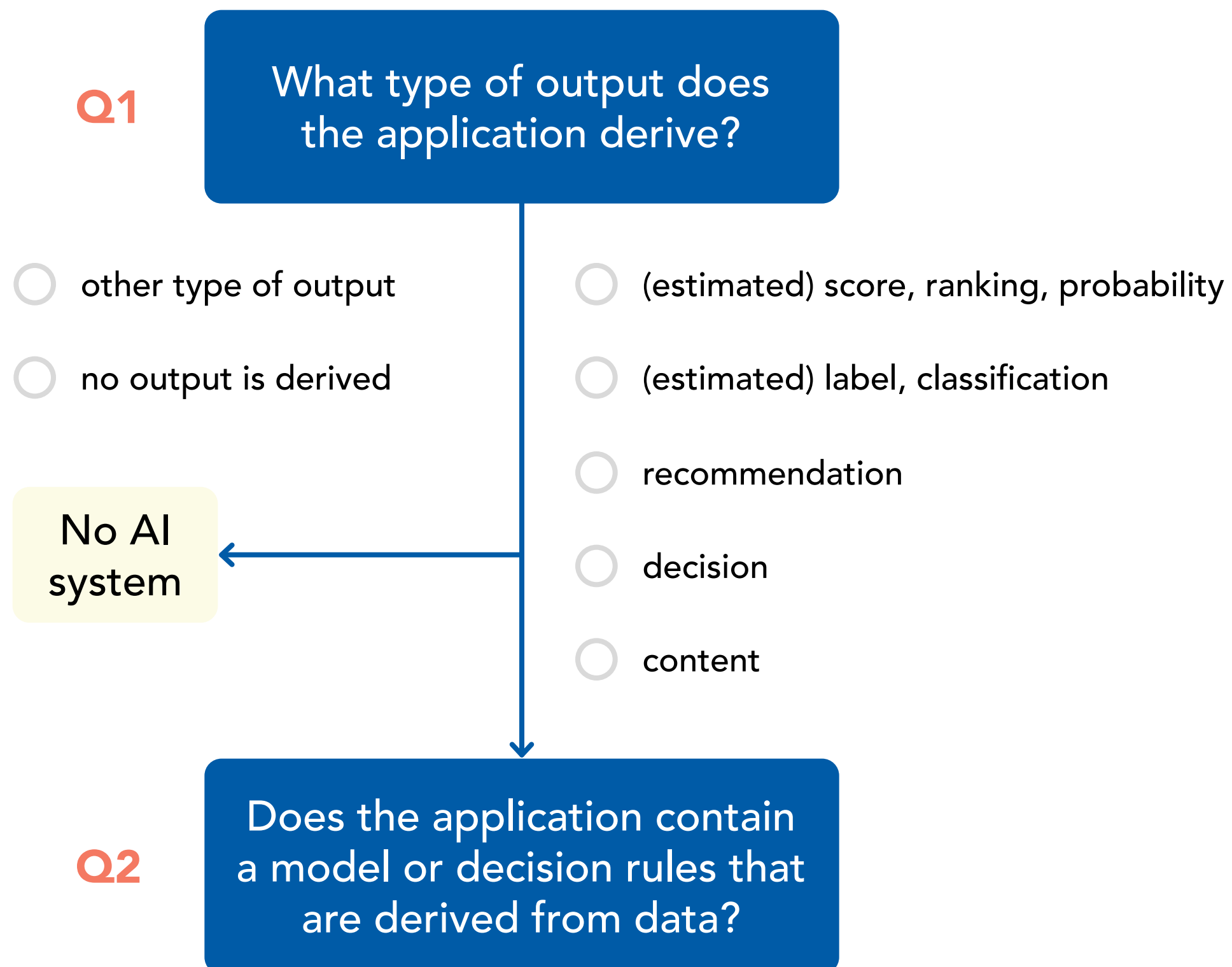
Just two aspects of the AI system definition matter

AI Act Article 3 – AI system definition

(1) 'AI system' means a machine-based system that is designed to operate with varying levels of **autonomy** and that may exhibit **adaptiveness** after deployment, and that, for explicit or implicit **objectives**, **infers**, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.
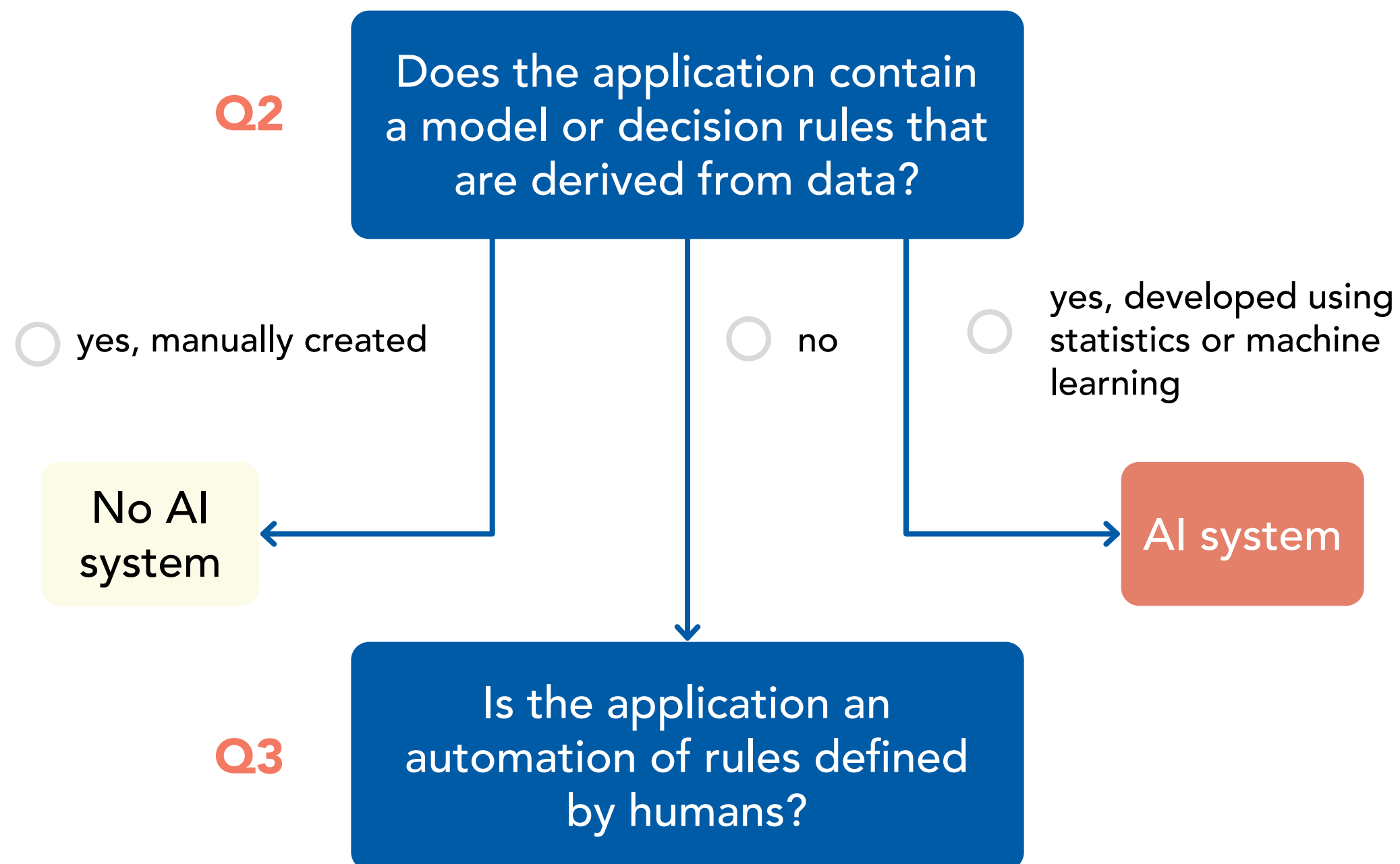
Inference
→ primary differentiator

Autonomy
→ easily fulfilled, "varying levels"

Adaptiveness
→ "may exhibit", not a requirement but a possibility

Objective
→ "explicit or implicit" includes all objectives

Q1 relates to the outcome of inference

Q1 — What type of output does the application derive?

- ○ other type of output
- ○ no output is derived
- ○ (estimated) score, ranking, probability
- ○ (estimated) label, classification
- ○ recommendation
- ○ decision
- ○ content

No AI system

Q2 — Does the application contain a model or decision rules that are derived from data?

Recital 12 states that inference relates to *"the process of obtaining the output such as predictions, content, recommendations, or decisions"*. Because 'prediction' is a broad term which is not understood by everybody in the same way, we include explanatory terms in the answer options.

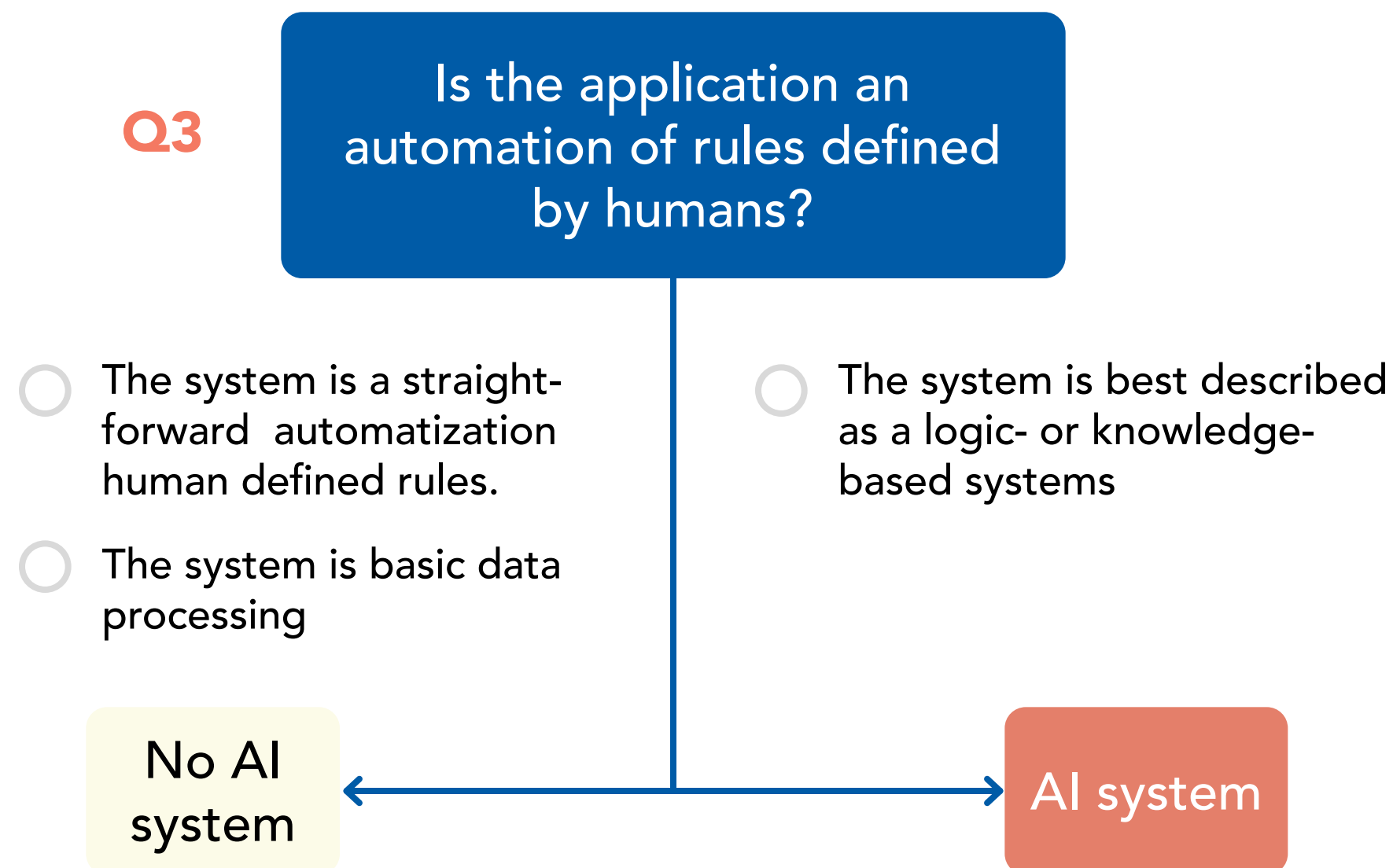# Q2 relates to inference of the model or algorithm from data

Q2
Does the application contain a model or decision rules that are derived from data?

○ yes, manually created

○ no

○ yes, developed using statistics or machine learning

No AI system

AI system

Q3
Is the application an automation of rules defined by humans?

Recital 12 states that inference refers to "*a capability of AI systems to derive models or algorithms, or both, from inputs or data*". This includes statistical approaches, because in statistical modelling a model is derived from (or: fitted to) data.

In some cases, rule-based algorithms are based on data-analyses, but the rules are selected and defined by humans. This happens for example in financial transaction monitoring. Based on recital 12, Algorithm Audit argues that these systems are not AI systems.

Q3 relates to explicit in- and exclusions in recital 12

Q3

Is the application an automation of rules defined by humans?

○ The system is a straight-forward automatization human defined rules.

○ The system is basic data processing

○ The system is best described as a logic- or knowledge-based systems

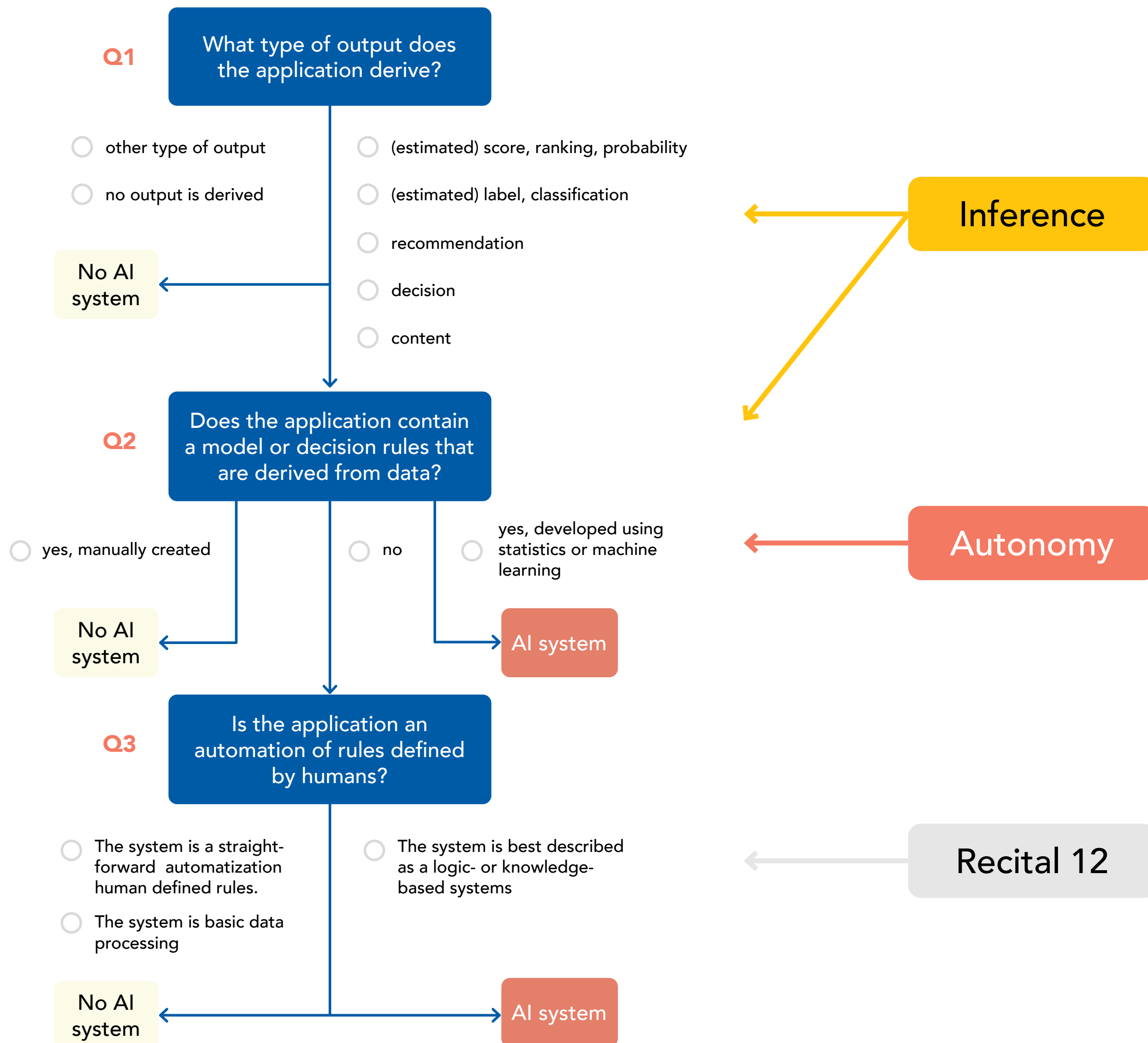**No AI system** ← → **AI system**

! Note that Q1 and Q2 are already answered.

🔍 Recital 12 explicitly includes logic- and knowledge-bases systems as AI systems.

In sum: answer 3 questions to identify AI systems

Q1 What type of output does the application derive?

- ○ other type of output
- ○ no output is derived

- ○ (estimated) score, ranking, probability
- ○ (estimated) label, classification
- ○ recommendation
- ○ decision
- ○ content

No AI system

Inference

Q2 Does the application contain a model or decision rules that are derived from data?

- ○ yes, manually created
- ○ no
- ○ yes, developed using statistics or machine learning

No AI system

AI system

Autonomy

Q3 Is the application an automation of rules defined by humans?

- ○ The system is a straight-forward automatization human defined rules.
- ○ The system is basic data processing

- ○ The system is best described as a logic- or knowledge-based systems

No AI system

AI system

Recital 12

Building *public knowledge*

for *ethical algorithms*

Join the discussion!

Algorithm Audit

🌐 www.algorithmaudit.eu

✉ info@algorithmaudit.eu

in www.linkedin.com/company/algorithm-audit

www.github.com/NGO-Algorithm-Audit