Table1 Overview: practitioner-informed recommendations and the challenges they mitigate in AI-enabled perception system (AIePS) development.

(*Each entry includes the challenge theme, relevant annotation phase, and empirically derived metadata such as priority, severity, frequency of mention across interviews, and expected benefit. This mapping supports traceable and actionable insights for improving data annotation requirements.*)

Recommendation	Challenge Mitigated	Challenge Theme	Annotation Phase	Priority	(ID) (Explicitly Mentioned)	Challenge Coverage Summary	Challenge Severity	Frequency (Interview Support)	Expected Benefit
R1	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Grounds annotation requirements in regulatory frameworks, reducing ambiguity and aligning updates.	High	High (15/19)	Improves annotation consistency and reduces bias
R1	C3	Evolving Requirements	Guideline Evolution	Medium	ID2, ID3	Supports proactive adaptation by integrating legal and ethical standards.	High	Medium (11/19)	Enables stable and traceable requirement evolution
R1	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Reduces inconsistent interpretation through harmonized implementation.	High	High (16/19)	Reduces rework and improves model training quality
R1.1	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Clarifies privacy-sensitive annotations via legal compliance.	High	High (15/19)	Improves annotation consistency and reduces bias
R1.1	C3	Evolving Requirements	Guideline Evolution	Medium	ID2, ID3	Stabilizes evolving requirements through early legal involvement.	High	Medium (11/19)	Enables stable and traceable requirement evolution
R1.1	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Improves consistency via documentation and data governance.	High	High (16/19)	Reduces rework and improves model training quality
R1.1	C5	Resource Limitations	All Phases	Medium	ID2, ID3	Ensures resource-efficient compliance from the start.	Major	High (17/19)	Improves efficiency across the annotation pipeline
R1.1	C1.2	Defining Requirements for Edge Cases	Requirement Planning	High	ID6, ID9	Provides legal clarity for edge case definitions.	Critical	High (13/19)	Reduces ambiguity in critical requirement definition
R1.2	C1	Edge Case Coverage Gaps	Requirement Planning	High	ID6, ID9	Identifies and prioritizes edge cases through safety assessment.	Critical	High (14/19)	Improves rare scenario detection and model generalisation
R1.2	C1.2	Defining Requirements for Edge Cases	Requirement Planning	High	ID6, ID9	Defines critical scenario requirements early.	Critical	High (13/19)	Reduces ambiguity in critical requirement definition
R1.2	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Contextualizes ambiguous scenarios via safety framing.	High	High (15/19)	Improves annotation consistency and reduces bias

R1.2	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Aligns annotations with consistent safety-centric expectations.	High	High (16/19)	Reduces rework and improves model training quality
R1.3	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Promotes fairness and reduces ambiguity in edge-case handling.	High	High (15/19)	Improves annotation consistency and reduces bias
R1.3	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Introduces bias-awareness to improve consistency.	High	High (16/19)	Reduces rework and improves model training quality
R1.3	C5.1	Strict Budgets	Planning and Execution	Medium	ID2, ID3	Justifies investment in fairness despite budget limits.	Major	High (12/19)	Minimises quality loss due to budget constraints
R2	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Simplifies complex instructions to reduce ambiguity.	High	High (15/19)	Improves annotation consistency and reduces bias
R2	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Supports standardisation through clearer guidelines.	High	High (16/19)	Reduces rework and improves model training quality
R2	C5.2	Limited Workforce and Scalability	Training and Execution	Medium	ID2, ID3	Improves scalability by reducing annotator training burden.	High	High (14/19)	Increases annotation accuracy and throughput
R2.1	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Reduces ambiguity via iterative validation cycles.	High	High (15/19)	Improves annotation consistency and reduces bias
R2.1	C3	Evolving Requirements	Guideline Evolution	Medium	ID2, ID3	Updates guidelines adaptively as requirements evolve.	High	Medium (11/19)	Enables stable and traceable requirement evolution
R2.1	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Improves consistency via continuous improvement.	High	High (16/19)	Reduces rework and improves model training quality
R2.2	C1.2	Defining Requirements for Edge Cases	Requirement Planning	High	ID6, ID9	Clarifies edge case definitions using expert feedback.	Critical	High (13/19)	Reduces ambiguity in critical requirement definition
R2.2	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Reduces ambiguity in rare scenario labeling.	High	High (15/19)	Improves annotation consistency and reduces bias
R2.2	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Improves consistency with feedback and review.	High	High (16/19)	Reduces rework and improves model training quality

R2.3	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Improves clarity through atomic labels and visuals.	High	High (15/19)	Improves annotation consistency and reduces bias
R2.3	C3	Evolving Requirements	Guideline Evolution	Medium	ID2, ID3	Refines evolving requirements using A/B testing.	High	Medium (11/19)	Enables stable and traceable requirement evolution
R2.3	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Reduces inconsistency by validating guideline usability.	High	High (16/19)	Reduces rework and improves model training quality
R2.4	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Provides context to reduce subjective interpretation.	High	High (15/19)	Improves annotation consistency and reduces bias
R2.4	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Improves consistency using domain-specific tools.	High	High (16/19)	Reduces rework and improves model training quality
R2.4	C5.2	Limited Workforce and Scalability	Training and Execution	Medium	ID2, ID3	Compensates workforce expertise gaps.	High	High (14/19)	Increases annotation accuracy and throughput
R2.5	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Reduces inconsistencies in repetitive annotations.	High	High (16/19)	Reduces rework and improves model training quality
R2.5	C5	Resource Limitations	All Phases	Medium	ID2, ID3	Improves scalability under constrained resources.	Major	High (17/19)	Improves efficiency across the annotation pipeline
R2.5	C1.2	Defining Requirements for Edge Cases	Requirement Planning	High	ID6, ID9	Refines edge case annotations with human-AI review.	Critical	High (13/19)	Reduces ambiguity in critical requirement definition
R2.6	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Resolves ambiguity through interdisciplinary input.	High	High (15/19)	Improves annotation consistency and reduces bias
R2.6	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Improves requirement clarity and cross-role alignment.	High	High (16/19)	Reduces rework and improves model training quality
R2.6	C5	Resource Limitations	All Phases	Medium	ID2, ID3	Optimises resource use via collaborative review.	Major	High (17/19)	Improves efficiency across the annotation pipeline
R2.6	C1.2	Defining Requirements for Edge Cases	Requirement Planning	High	ID6, ID9	Enables expert-informed edge case definitions.	Critical	High (13/19)	Reduces ambiguity in critical requirement definition

R2.7	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Reduces ambiguity via co-creation.	High	High (15/19)	Improves annotation consistency and reduces bias
R2.7	C3	Evolving Requirements	Guideline Evolution	Medium	ID2, ID3	Enables proactive handling of evolving requirements.	High	Medium (11/19)	Enables stable and traceable requirement evolution
R2.7	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Ensures consistent requirement definition.	High	High (16/19)	Reduces rework and improves model training quality
R2.8	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Reduces ambiguity using repeatable structures.	High	High (15/19)	Improves annotation consistency and reduces bias
R2.8	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Improves consistency with modular templates.	High	High (16/19)	Reduces rework and improves model training quality
R2.8	C5	Resource Limitations	All Phases	Medium	ID2, ID3	Saves effort and ensures cross-team clarity.	Major	High (17/19)	Improves efficiency across the annotation pipeline
R2.9	C3	Evolving Requirements	Guideline Evolution	Medium	ID2, ID3	Stabilises requirement evolution through planned allocation.	High	Medium (11/19)	Enables stable and traceable requirement evolution
R2.9	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Improves requirement consistency with early structure.	High	High (16/19)	Reduces rework and improves model training quality
R2.9	C5.3	Time Constraints	Planning and Execution	Medium	ID2, ID3	Addresses planning-related time constraints.	Moderate	Medium (9/19)	Reduces rushed errors and debugging delays
R2.10	C3	Evolving Requirements	Guideline Evolution	Medium	ID2, ID3	Enables consistent handling of evolving changes.	High	Medium (11/19)	Enables stable and traceable requirement evolution
R2.10	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Reduces inconsistency via version control.	High	High (16/19)	Reduces rework and improves model training quality
R2.10	C5.4	Tool/Process Limitations	Tooling and QA	Medium	ID2, ID3	Improves traceability and process control.	Moderate	High (15/19)	Enables scalable, auditable annotation workflows
R3	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Eliminates ambiguity using formal quality review structures.	High	High (15/19)	Improves annotation consistency and reduces bias
R3	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Ensures consistency across annotations through validation.	High	High (16/19)	Reduces rework and improves model training quality

R3	C5.4	Tool/Process Limitations	Tooling and QA	Medium	ID2, ID3	Improves reliability through process QA mechanisms.	Moderate	High (15/19)	Enables scalable, auditable annotation workflows
R3.1	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Improves disambiguation via training and benchmarks.	High	High (15/19)	Improves annotation consistency and reduces bias
R3.1	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Improves consistency via expert review.	High	High (16/19)	Reduces rework and improves model training quality
R3.1	C5.2	Limited Workforce and Scalability	Training and Execution	Medium	ID2, ID3	Addresses workforce skill limitations.	High	High (14/19)	Increases annotation accuracy and throughput
R3.2	C2	Ambiguity in Annotation Requirements	Guideline Design	High	ID2, ID3	Reduces ambiguity using consensus-driven approaches.	High	High (15/19)	Improves annotation consistency and reduces bias
R3.2	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Improves consistency across annotators.	High	High (16/19)	Reduces rework and improves model training quality
R3.2	C5.2	Limited Workforce and Scalability	Training and Execution	Medium	ID2, ID3	Addresses expertise variation.	High	High (14/19)	Increases annotation accuracy and throughput
R3.3	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Ensures consistent performance at scale.	High	High (16/19)	Reduces rework and improves model training quality
R3.3	C5.1	Strict Budgets	Planning and Execution	Medium	ID2, ID3	Reduces cost of quality failures.	Major	High (12/19)	Minimises quality loss due to budget constraints
R3.3	C5.4	Tool/Process Limitations	Tooling and QA	Medium	ID2, ID3	Supports efficient QA under process constraints.	Moderate	High (15/19)	Enables scalable, auditable annotation workflows
R3.4	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Aligns annotator behavior with accuracy goals.	High	High (16/19)	Reduces rework and improves model training quality
R3.4	C5.1	Strict Budgets	Planning and Execution	Medium	ID2, ID3	Reduces errors caused by cost-saving rush.	Major	High (12/19)	Minimises quality loss due to budget constraints
R3.4	C5.2	Limited Workforce and Scalability	Training and Execution	Medium	ID2, ID3	Improves attention to edge cases.	High	High (14/19)	Increases annotation accuracy and throughput

R3.5	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Improves consistency via traceable workflows.	High	High (16/19)	Reduces rework and improves model training quality
R3.5	C5.3	Time Constraints	Planning and Execution	Medium	ID2, ID3	Links quality issues to specific batches.	Moderate	Medium (9/19)	Reduces rushed errors and debugging delays
R3.5	C5.4	Tool/Process Limitations	Tooling and QA	Medium	ID2, ID3	Supports debugging and process transparency.	Moderate	High (15/19)	Enables scalable, auditable annotation workflows
R3.6	C3	Evolving Requirements	Guideline Evolution	Medium	ID2, ID3	Aligns tools with changing requirements.	High	Medium (11/19)	Enables stable and traceable requirement evolution
R3.6	C4	Inconsistent Requirements	Quality Control	Medium	ID2, ID3	Prevents misalignment and inconsistency.	High	High (16/19)	Reduces rework and improves model training quality
R3.6	C5.4	Tool/Process Limitations	Tooling and QA	Medium	ID2, ID3	Improves process compliance and traceability.	Moderate	High (15/19)	Enables scalable, auditable annotation workflows

Table 2. Criteria used to calculate Priority, Severity, and Frequency scores for practitioner recommendations and annotation challenges.

(The values are derived from cross-interview thematic analysis, considering the criticality of the challenge, its occurrence across interviews, and its impact on AI-enabled perception system (AIePS) performance.)

Dimension	Definition	How It's Measured	Heuristic Thresholds
Priority	Indicates the urgency and importance of implementing a recommendation to address a challenge	Based on the severity of the challenge it mitigates, how often it is mentioned, and whether it is critical for safe or effective AIePS function	- High: Addresses a critical challenge mentioned frequently and tied to safety/system failure - Medium: Mitigates an important but not system-critical issue - Low: Helpful but not urgent or high-impact
Severity	Reflects the negative impact of a challenge on data quality, model performance, or compliance	Inferred from interview narratives about the consequences of unresolved challenges (e.g., safety, legal, rework)	- Critical: Can cause unsafe AI behavior, compliance failure, or downstream harm - High: Major impact on annotation quality or consistency - Moderate: Impacts workflow but not safety or regulatory outcomes
Frequency	Represents how commonly a challenge was mentioned by interviewees	Counted via thematic coding across transcripts; can be expressed as ratios or categories	- High: Mentioned in ≥13 out of 19 interviews - Medium: Mentioned in 6–12 interviews - Low: Mentioned in ≤5 interviews