ICC 2025 Project Proposal

1 Synopsis

1.1 Project Description

NERDS (**N**ext-generation **E**mergency **R**esponse **D**rone **S**ystem) is an autonomous, AI-powered platform designed to transform emergency search-and-rescue (SAR) operations in challenging environments such as forests, mountainous regions, or disaster-struck areas. By combining state-of-the-art (SOTA) models for real-time perception, prediction, human-drone interaction, and intelligent decision-making, **NERDS** aims to save lives, optimize rescue resources, and enhance crew safety.

The platform seamlessly integrates natural user interfaces (tablet sketching, hand gestures, voice commands) with real-time aerial mapping, multi-modal object detection, injury assessment, and predictive survivor localization. It is designed to operate with minimal user training, allowing both professional SAR teams and community responders to benefit from powerful AI tools under extreme time pressure and stress.

The concept is aligned with several of Idiap's research programs:

1. **Human-AI Teaming**: **NERDS** enables collaborative mission planning and execution between human rescuers and intelligent drones. The system learns from human-designed rescue strategies and feedback to continuously improve its own decision-making.

2. **AI for Life**: The core mission of **NERDS** is to save lives. It provides vital first-aid resources, enables communication between victims and medical staff, and offers survival guidance in critical moments.

3. **AI for Everyone**: Emergency situations demand fast, error-resistant decisions. **NERDS** supports human operators by reducing cognitive load and minimizing errors under pressure. Its intuitive interaction modes (e.g.,

gestures, voice) ensure accessibility even for users with no technical background. This democratization of advanced AI technology helps empower local communities and non-specialist responders.

4. **Sustainable & Resilient Societies**: By optimizing drone routes, identifying the most urgent needs, and reducing unnecessary deployment of vehicles or human teams, **NERDS** minimizes fuel and energy use, and enhances the overall resilience and efficiency of SAR systems.

1.2 Application Context

SAR operations face extreme challenges – difficult terrain, rapidly changing conditions, limited resources, and severe time pressure. Traditional SAR methods rely on human coordination, visual scanning, and limited aerial tools, which are often insufficient in large-scale disasters or remote environments. Delays, miscommunication, and incomplete situational awareness can cost lives, especially during the critical first hours after an incident.

NERDS addresses these issues by introducing an integrated, AI-driven drone platform capable of intelligent mission planning, autonomous navigation, and real-time survivor detection. Its multimodal interface – manipulation via gesture, voice, and touchscreen – enables intuitive human-drone collaboration even under stress or in resource-constrained field conditions. The system interprets mission intent from simple interactions, translates them into flight paths, and continuously adapts based on real-time data.

What distinguishes **NERDS** from existing solutions is its end-to-end intelligence and real-world deployability:

1. **Automated decision-making (ADM)**: **NERDS** autonomously detects survivors, assesses injuries via pose estimation, and suggests or initiates actions like medical supply drops or hazard avoidance.

2. **Predictive modeling**: Bayesian inference and behavioral models infer likely survivor locations from indirect evidence (e.g., gear, footprints).

3. **On-device AI**: Optimized models run on edge devices (e.g., via NVIDIA Jetson), enabling fully autonomous operation even without stable connectivity.

4. **Learning over time**: The system self-improves via human feedback, accumulated experience, adapting to new terrains, rescue tactics, and environmental patterns.

5. **Human-centered AI**: **NERDS** blends cutting-edge research in computer vision, robotics, and human-AI teaming into a single deployable platform that can be used by anyone – from professional rescue units to local volunteers with minimal training.

2 The Team

Bringing **NERDS** to life requires a well-balanced, interdisciplinary team that should ensure full technical coverage – From CV and reasoning AI, through drone hardware and edge processing, to clean UI and real-world interaction, research-to-market transfer and cross-domain versatility. The ideal 3-person core tech team would include:

1. **AI/Backend Engineer** (Computer Vision & Probabilistic Inference)

Tasks:
(a) Implement real-time object/person detection (YOLOv8, OpenPose)
(b) Develop Bayesian optimization modules for survivor localization
(c) Design trajectory prediction models (e.g., Trajectron++, Social LSTM)

(d) Optimize models for edge computing (Jetson, etc.)

Skills: (a) Python, PyTorch/TensorFlow, OpenCV, NumPy, ROS (Robot Operating System), LangChain, LLMs

(b) Bayesian inference

(c) GPU optimization and edge deployment (NVIDIA Jetson, ONNX)

2. **Frontend & UX Developer** (Web/Mobile Command Interface)

Tasks: (a) Build tablet/mobile interface for mission planning (map, video, gesture/voice input)

(b) Integrate Leaflet.js for sketching zones, route planning

(c) Visualize live detections, pose estimates, and system feedback

(d) Ensure low-latency, intuitive UX under stress conditions

Skills: (a) JavaScript/TypeScript, React or Vue

(b) Leaflet.js, WebRTC, Flask or FastAPI backend integration

(c) Figma or similar for UI prototyping, mobile-first/responsive UI design, accessibility best practices

3. **Embedded Systems & Drone Integration Engineer**

Tasks: (a) Implement drone control pipelines (pulp-dronet, SDK APIs, gesture/voice mapping)

(b) Ensure communication between drone, edge device, and frontend

(c) Integrate RTSP video, sensor data, and autonomous execution

(d) Field-test flight autonomy, GPS routing, and obstacle handling

Skills: (a) C++/Python for embedded systems

(b) DJI SDK, MAVLink, ROS

(c) Android/Linux hardware interfacing

(d) Networking (RTSP, WebSocket, MQTT) and latency profiling

Since our team only has 2 members, the tasks are to be distributed as follows:

Yulia: AI/Backend (Computer Vision & Probabilistic Inference) + Implement drone control pipelines (pulp-dronet, SDK APIs, gesture/voice mapping) + Ensure communication between drone, edge device, and frontend + Project lead

Hamid: Frontend & UX Developer tasks + Integrate RTSP video, sensor data, and autonomous execution + Field-test flight autonomy, GPS routing, and obstacle handling.

3 NERDS Prototype Description

3.1 Prototype Objectives

The prototype goal is to enable mobile/tablet-based intuitive mission planning, autonomous drone flight along a sketched path, real-time detection of survivors or hazards, and AI-generated contextual recommendations – all running in a semi-automated prototype loop.

Core Minimum Viable Product (MVP) Functionality:

1. **Mobile-based mission planning:** draw search area or route via mobile canvas (gesture-based or simple touch), export path as structured JSON to guide drone navigation;
2. **Semi-autonomous drone flight:** basic flight mission execution using DJI/Tello SDK or simulator, waypoint following based on user-defined path, live drone camera stream to dashboard (RTSP/WebRTC stack);
3. **Real-time detection from drone feed:** use lightweight YOLO for people, gear, fire/smoke detection, overlay detections on live stream (bounding boxes, labels);
4. **Contextual LLM-based action suggestions:** combine detections with simple metadata (e.g., "person lying on ground", "fire nearby"), send structured prompt to GPT-4 or Gemini API via LangChain; return suggestions such as "Injury likely, drop medical kit", "Fire detected in proximity; suggest rerouting and alert team", display suggestions in real time in the dashboard as system alerts;
5. **Alert logging/ event history:** store event types, timestamps, and AI recommendations, allow manual tagging/confirmation from user for feedback loop.

Stretch Features (if time permits)

1. **Real-time injury detection and assessment:** Integrate OpenPose (YOLO11) for posture and injury detection; detect unconscious or injured survivors based on body pose, display alerts and flag priority rescue cases in UI.
2. **Predictive Survivor Localization:** Incorporate evidence-based search: ML algorithm correlates found items with likely survivor locations, movement pattern analysis for guiding search strategy – predict future movement of survivors with Trajectron++ or social LSTM models) + bayesian optimization to continuously refine search areas based on new evidence.
3. **LLM-driven item prediction → CV detection loop**: Prompt an LLM with the high-level scenario (e.g. "Parachutist lost on a snowy mountain ridge") to generate a concise list of likely survivor items (helmet, backpack, ice axe, bright-colored clothing, emergency blanket, etc.), feed the list into the onboard CV detector (YOLO, CLIP-based zero-shot detector), if one of the items is recognized in the drone footage, the GPS coordinates are communicated to the team.

3.2 9-Day Roadmap $\rightarrow$5.5

Day 1: System Setup & Component Testing

1. Assemble hardware: drone, onboard compute, control interface devices (tablet/laptop).
2. Establish development environment for drone control, video streaming, and AI model inference.
3. Test manual drone flight, camera feed (RTSP), and streaming to a local device.
4. Verify object detection model (YOLOv8) works with test video frames.

Day 2: Object Detection Integration

1. Integrate YOLOv8 (YOLO11) with real-time drone camera feed, display bounding boxes (people, gear, fire/smoke) on live video in local UI.
2. Begin tagging detected items with GPS/map overlays.

3. Begin building modular architecture for later model additions.

Day 3: Core UI Components & Modular Architecture

1. Continue building modular architecture.
2. Implement interactive tablet-based interface using Leaflet.js + Flask or React.

Day 4: Mission Planning UI & Video/Map Interface

1. Enable sketching mission zones, visualizing drone path, and assigning tasks.
2. Enable WebRTC or alternative for low-latency video to browser.

Day 5: Gesture and Voice Command Integration

1. Integrate hand gesture control module (telloCommander, tello-gesture-control).
2. Integrate voice command module (telloCommander), map basic commands (e.g., "scan area," "return to base," "zoom in") to drone functions.

Day 6: Autonomous Flight Execution

1. Convert sketched zone on UI into drone flight path (via pulp-dronet or similar).
2. Enable terrain-aware autonomous navigation (forest/mountain sim).
3. Test mission abort/return triggers (low battery, blocked path, etc.)

Day 7: Action Recommendation Module

1. Implement rule-based reasoning for action decisions (e.g., "drop medical kit").
2. Integrate basic LLM queries via LangChain or LangGraph.
3. Start logging decisions and explainability outputs.
4. Display recommendations in the interface for human confirmation.

Day 8: Field Testing & System Integration

1. Conduct end-to-end test in a controlled outdoor environment, test full flow: UI planning $\rightarrow$ drone command $\rightarrow$ live detection $\rightarrow$ action recommendation.
2. Record latency, reliability, and detection accuracy.
3. Iterate UI and model parameters based on test results.
4. Start preparing the project presentation.

Day 9: Final Demo & Evaluation

1. Polish UI, add explanatory overlays, improve UX for interaction simplicity.
2. Finalize documentation/ user quidelines.
3. Prepare final demo script with varied rescue scenarios.

3.3 Milestones

1. Development environment set up, drone control + camera stream verified.
2. YOLOv8 integrated with live drone feed (people/gear detection).
3. Minimal UI (map + annotated video + command inputs) designed.
4. Core architecture established, core UI operational.
5. Gesture/voice commands via tello-gesture-control and telloCommander integrated.
6. Sketch-based mission planning and autonomous path execution enabled.
7. Action suggestion logic (rule-based prototype + LLM interface) implemented.
8. Integration + field test in outdoor environment successful.
9. Final polish, demo scenarios, documentation, final presentation.

3.4 Technical Challenges

The individual components for building a multimodal autonomous rescue system such as **NERDS** already exist. The corresponding AI models and tools, which are capable of real-time drone data analysis, survivor detection, pose estimation, and action recommendation need to be combined into a single framework (see Section 5.1 to learn about the modules and the corresponding existing tools). Challenges of the integration include:

- Real-time integration: fusing drone control, video processing, UI, and AI inference in a low-latency, robust architecture.
- Edge deployment constraints: running compute-intensive models (YOLOv8, OpenPose) on lightweight edge devices like NVIDIA Jetson or Raspberry Pi.
- Natural human-drone communication: ensuring gesture and voice interfaces are reliable in noisy, outdoor conditions.
- Multimodal reasoning: designing a decision module that merges vision, voice, and mission context into actionable insights.
- Compatibility & interoperability: ensuring that diverse modules—ranging from drone control and sensor data processing to AI inference and user interface—operate seamlessly within a unified workflow (incl. managing communication protocols, synchronizing inputs/outputs, and maintaining modularity while enabling real-time integration across hardware and software components).
- Usability under stress: making the system intuitive for field operators with minimal tech background.

4 Beyond ICC

We see the **NERDS** prototype as a foundational step toward a fully deployable, mission-critical system. Our follow-up plans include both technical development and strategic outreach:

4.1 Further Prototype Development

Post-challenge, we aim to expand the system's capabilities:

- Improve the robustness and accuracy of AI models (e.g., injury detection, trajectory prediction, action recommendation).
- Enhance the seamless integration between components (UI, drone control, inference, communication layers).
- Explore edge deployment optimization and multi-drone coordination.
- Consider integration possibilities with the SLF's **White Risk** platform and similar tools.
- Build additional modules to make the system adaptable to other tasks, such as perimeter monitoring, wildlife monitoring, etc.

4.2 Toward Commercialization

We plan to pursue Innosuisse funding to support applied development, field testing, and market validation. The long-term goal is to establish a spin-off company offering a modular, AI-powered drone platform for SAR and security applications.

4.3 Scientific Dissemination

We intend to present the project at major conferences such as **CoRL** (Conference on Robot Learning), **AMLD** (Applied Machine Learning Days), and **ICML** (International Conference on Machine Learning). The interdisciplinary nature of **NERDS** (CV, robotics, decision-making) makes it highly relevant to both academic and applied AI communities.

4.4 Collaborations & Ecosystem Building

We plan to initiate collaborations with key Swiss institutions and stakeholders:

- UZH Robotics and Perception Group, Idiap, ETH AI Center, EPFL AI Center, and SLF (Institute for Snow and Avalanche Research) for joint research and technical guidance.
- Weisse Arena Gruppe and other outdoor resorts (hiking, skiing, paragliding) as potential testbeds and early adopters.
- The Ark Foundation – to find and connect with other potential industrial collaborators.
- Engage in discussions with military and security stakeholders, where similar technology could support perimeter surveillance, disaster response, and critical infrastructure monitoring.

Ultimately, we envision **NERDS** evolving into a versatile platform that not only saves lives in emergencies, but also supports resilience, safety, and autonomy in a wide range of domains.

4.5 Business Model Canvas

Business Model Canvas

Key Partners

Academic: ETH Zurich, EPFL, UZH Robotics & Perception Group, Idiap, Ark, SLF

Industry: drone manufacturers (DJI, Parrot, Andruino), Weisse Arena Gruppe, Swisspeak Resorts, etc.

Public/NGO: Rega, Swiss Red Cross, local civil protection units

Tech: NVIDIA (Jetson), telecom providers (for real-time uplink)

Key Activities

- Prototyping, model integration, UI/UX development
- Field testing with industrial and research partners
- Business development & partnership outreach
- AI system optimization for edge hardware

Key Resources

- Multimodal AI models (YOLOv8/11, OpenPose, DeepWriting, Trajectron++)
- Drone platform and integration pipeline
- Edge devices & backend infrastructure
- Expert team (AI, backend, frontend, drone UI, project management)

Value Propositions

Faster, safer, smarter rescue operations:
- Drastically reduced response times
- Real-time survivor detection & assessment
- Semi-autonomous drones reduce personnel risk

Easy-to-use interface:
- Intuitive mission planning (voice, gesture, touchscreen)
- Requires little to no technical training

Scalable & modular:
- Customizable models (from injury detection to localization logic)
- Can be adapted to different terrains, missions, and objectives

Customer Relationships

- Co-development & feedback from early partners
- Training & onboarding workshops
- Post-deployment support (updates, retraining AI modules)
- Optional B2B SaaS or licensing model with service-level agreements

Channels

- Sales to public institutions
- Pilot projects via partnerships (e.g., ETH, EPFL, UZH, SLF, NGOs)
- Presentations at conferences (e.g., ALMD, CoRL, ICML)
- Collaborations with industrial partners (e.g., drone manufacturers, resorts)

Customer Segments

Primary:
- Public rescue organizations (mountain rescue, civil protection)
- NGOs (Red Cross)
- Government emergency services

Secondary:
- Outdoor sports resorts (ski, hiking, biking, paragliding)
- Security and defense
- Environmental & wildlife monitoring agencies

Cost Structure

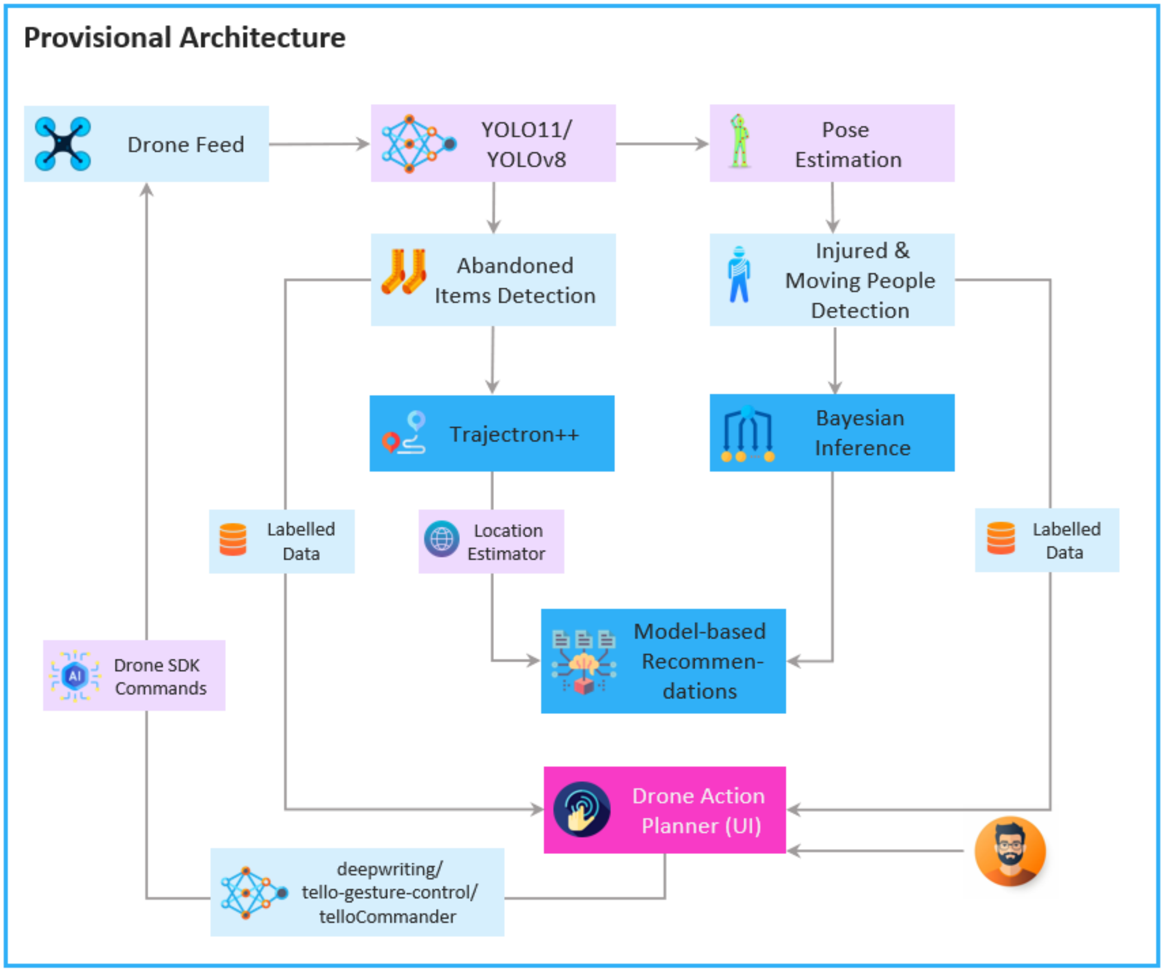
- R&D: salaries, hardware, data collection, testing
- AI model development & training (compute power, server/cloud services)
- Drone fleet integration & certification
- Field testing and validation
- Marketing, business development, compliance

Revenue Streams

- Innosuisse grant, further grants for scaling
- Initial hardware + software deployment packages
- Subscription for updates, AI improvements, or cloud services
- Custom integrations and consultancy, workshops
- Potential licensing for military/security use

5 Additional Input

5.1 Provisional NERDS Architecture



5.2 NERDS: Key Functions & Corresponding Models

1. Intelligent Mission Planning

    - Intuitive Interface: Draw search patterns directly on mobile devices (Model: deepwriting), hand gesture recognition (Model: tello-gesture-control), voice commands (Model: telloCommander)
    - Autonomous Execution: Drones navigate complex mountain/forest terrain independently (Model: pulp-dronet)
    - Dynamic Adaptation: Real-time route optimization based on conditions.
    - Real-Time Communication & Mapping (Tools: RTSP Streaming for drone camera feed, OpenCV + Flask + WebRTC for web/mobile real-time map and video, Leaflet.js + custom annotations to sketch zones on a phone or tablet)

2. Advanced Situational Analysis from Drone Footage

- Multi-Modal Detection: Simultaneous tracking of people, objects (backpacks, clothing, gear, fire/smoke), and environmental hazards in real time (Model: YOLOv8 (Ultralytics) or YOLO11)
- Injury Assessment: AI-powered pose estimation identifies medical emergencies (Model: OpenPose or YOLO11); detect unusual posture (e.g., unconscious, limping, collapsed).
- Risk Evaluation: Contextual analysis of terrain, weather, and rescue complexity
- Real-Time: GPU acceleration or edge computing (e.g., NVIDIA Jetson), lightweight versions working on edge devices.

3. Predictive Survivor Localization & Trajectory Prediction
   - Evidence-Based Search: Machine learning correlates found items with likely survivor locations
   - Behavioral Modeling: Movement pattern analysis for guiding search strategy – predict future movement of survivors, suggest optimal intercept routes or supply drop locations (Model: Trajectron++ or social LSTM models)
   - Bayesian Optimization: Continuously refines search areas based on new evidence.
4. AI for Action Recommendation
   - Situation-adaptive strategy suggestion: given detected posture, environmental hazards, and item types, infer: "This person is likely injured, drop medical kit" or "Fire nearby, reroute drone and notify crew" (Model: Rule-based + LLM hybrid, e.g., OpenAI GPT-4 or Google Gemini with LangChain or LangGraph).

5.3 Hardware Intended for Prototyping (in possession)

1. Drone: DJI Mini 4 Pro Fly More Combo
   - Flight time: up to 34 minutes
   - Weight: 249 g
   - Camera: 48 MP with advanced obstacle sensing and GPS stability
2. Edge Device (Mobile): Xiaomi Redmi 12C – for lightweight mobile interface for mission planning, command input
   - OS: Android 12
   - Memory: 2 GB RAM, 32 GB storage
3. Development Laptop: Lenovo Legion Pro 7i 16 – for local development, model training/inference, and high-performance testing
   - CPU: Intel Core i9
   - GPU: NVIDIA RTX 4080
   - Memory: 32 GB RAM, 1 TB SSD

5.4 Risk Analysis

Risk 1: Poor performance in unstructured environments, poor strategy suggestions. Mitigation Strategy: Train with synthetic + real SAR data; robustness testing in varied terrains; improvement via human-in-the-loop.

Risk 2: Edge compute bottlenecks. Mitigation Strategy: Use quantized, optimized models; hardware acceleration on Jetson devices.

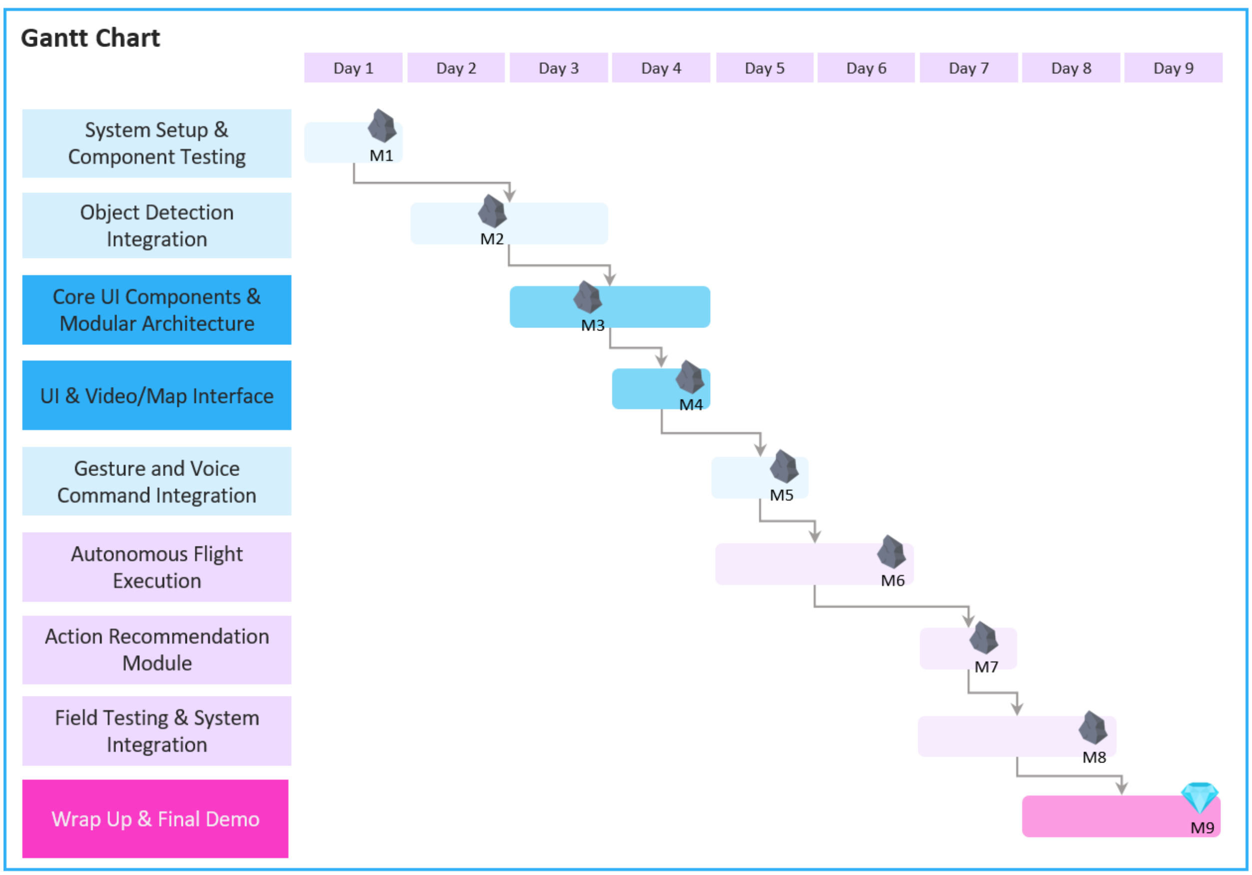
Risk 3: False detections. Mitigation Strategy: Ensemble methods + uncertainty-aware predictions.

Risk 4: Legal & ethical concerns (privacy, data use). Mitigation Strategy: Adhere to FAA/UAV guidelines; anonymized data; consult with ethics board.

Risk 5: Usability issues. Mitigation Strategy: Co-design with SAR experts; iterative UX testing.

Risk 6: Appearance of similar competing products. Mitigation Strategy: Continuous Innovation & R&D, agile scaling and refinements, strategic ecosystem partnerships, building a strong brand & customer loyalty.

5.5 Gantt Chart



6 Short CVs

References

1. Autonomous drone navigation: https://github.com/pulp-platform/pulp-dronet
2. Multiple object detection (YOLO11 or YOLOv8): https://github.com/ultralytics/ultralytics
3. Pose estimation: https://github.com/CMU-Perceptual-Computing-Lab/openpose
4. Handstrokes recognition for drone manipulation via phone/ tablet: https://github.com/emreaksan/deepwriting
5. Hand-gesture drone manipulation: https://github.com/kinivi/tello-gesture-control
6. Trajectory understanding: https://github.com/StanfordASL/Trajectron-plus-plus