Project Charter

Team 14

Catalog

1. Project Name or Identification

Project Name: Forgery Detection Platform for News Scenarios

Technology:

Large Language Models (LLMs)

Multimodal Fusion (CLIP, BLIP)

Explainable AI (XAI)

Deepfake Detection

Version: 2.0

Date: March 23, 2025

2. Project Stakeholders

Name/Organization	Role	Phone Number	Email
君同未来科技有限责任公司	Project Sponsor		

张茁含睿	Project Leader	18291665443	3590598688@qq.com
胡家豪	Team member	15727198260	2015294440@qq.com
黄俊哲	Team member	17767668098	2253640@tongji.edu.cn
陶子芾	Team member	18049037002	tzf9282003@163,com

3. Project Discription

Background

With the rapid proliferation of social media, short video platforms, and self-media, the dissemination of fake news has become a pressing global concern. The widespread availability of digital content and the ease of sharing information have significantly accelerated the spread of misinformation, often leading to severe consequences. False narratives can manipulate public opinion, influence political outcomes, disrupt financial markets, and even pose threats to national security. The dynamic nature of forgery techniques, including AI-generated deepfakes and manipulated multimedia content, has made it increasingly difficult to detect and counteract fake news using conventional methods.

Description of the Challenge or Opportunity

Despite various efforts to combat misinformation, current fake news detection methods face several critical limitations. Many existing systems struggle with accuracy, often failing to distinguish between authentic and fabricated content with high confidence. Additionally, these methods frequently lack interpretability, making it difficult for users and decision-makers to understand the rationale behind classification results. Furthermore, most detection techniques are optimized for a single modality, such as text-based or image-based forgery detection, and lack the adaptability needed to handle multimodal content effectively. Given the rapidly evolving nature of misinformation tactics, an advanced and comprehensive solution is urgently needed. This presents an opportunity to develop a more robust, scalable, and adaptable system that can keep pace with the sophistication of modern misinformation techniques.

Overview of the Desired Impact

This project aims to develop an AI-powered forgery detection platform that

significantly enhances detection accuracy, ensures adaptability across multiple content types, and provides clear and transparent interpretability. The platform will integrate large language models (LLMs) for deep semantic analysis of textual content while employing multimodal fusion techniques such as CLIP and BLIP to analyze and correlate information across text, images, and videos. This comprehensive approach will allow the system to identify forged content with greater precision and reliability. By leveraging state-of-the-art AI technologies, this system will improve fake news identification across diverse domains, including politics, economy, technology, and culture. The platform will be designed to support real-time detection, providing timely alerts and detailed reports to stakeholders such as news agencies, policymakers, and social media platforms. Through a combination of high detection accuracy, interpretability, and multimodal adaptability, this project seeks to establish a robust defense mechanism against the growing challenge of digital misinformation.

4. Project Objectives

MOV Element	Description	Measurement Criteria	How to Measure
Project Goal	Develop an AI-powered platform capable of efficiently detecting and analyzing fake news in various formats (text, images, and videos).	Successful creation and deployment of the platform with the ability to detect fake news in multiple formats.	Platform deployment, functionality tests, successful detection of fake news.
Value to the Organization	Enhance the ability to combat misinformation across various domains like politics, economy, technology, and culture, increasing reliability and trust	Accuracy in detecting fake news, system's effectiveness in handling diverse news formats.	Platform's performance metrics, accuracy rate above predefined benchmark.

	in information.		
Financial Impact	Reduce economic losses caused by misinformation, including the disruption of financial markets or consumer behavior.	Decrease in financial disruptions or loss of public trust linked to misinformation.	Comparative analysis of financial incidents before and after system deployment.
Impact on Stakeholders	Improve trust among stakeholders such as news agencies, social media platforms, and policymakers by providing reliable news verification.	Stakeholder engagement and trust in using the platform for news verification.	Stakeholder surveys, feedback on reliability and trust in platform.
Risk Reduction	Minimize the risks associated with the spread of fake news, particularly in sensitive areas like national security and politics.	Detection and mitigation of harmful fake news cases, particularly in critical domains.	Number of false narratives flagged and prevented, especially in critical domains (politics, security).
Adaptability and Scalability	Ensure that the platform can scale and adapt to new and evolving misinformation techniques and increasing amounts of data.	Platform's ability to process large-scale data efficiently while maintaining accuracy.	System performance under stress, scalability tests, and adaptation to emerging trends in forgery techniques.
Long-term	Establish a	Long-term system	Ongoing

Benefits	sustainable platform that continues to improve and remain effective in the long term, addressing the growing issue of digital misinformation.	usage, continuous updates, and platform evolution based on new data.	system usage statistics, feedback from users, and adaptation of the platform to new challenges.

5. Project Scope

The Forgery Detection Platform for News Scenarios will focus on the design and development of an AI-powered forgery detection system that enhances accuracy, interpretability, and adaptability across multiple media formats. The scope of this project is defined to ensure that the objectives align with the available resources, technology capabilities, and intended use cases.

In Scope:

The following areas are included in the project to ensure a robust and effective forgery detection platform:

- Design and implementation of an AI-driven forgery detection system.
  The project will involve the development of an advanced AI-based system capable of identifying and analyzing fake news. This system will integrate machine learning, deep learning, and natural language processing (NLP) techniques to ensure high-accuracy forgery detection across various content types.
- Integration of text, image, and video analysis models.
  To address the multimodal nature of fake news, the system will incorporate models for textual, visual, and video-based analysis. Techniques such as large language models (LLMs) for text processing, CLIP for image-text correlation, and deepfake detection models for video content will be implemented. This will enable cross-modal consistency checks to detect discrepancies between different media elements.
- Development of a real-time detection framework.
  Given the fast-paced nature of news dissemination, the system will be designed for real-time processing and forgery detection. Optimized AI pipelines and scalable cloud-based architectures will be implemented to ensure that the system analyzes

large datasets efficiently and provides instant results without significant computational delays.

- Creation of user-friendly analysis reports.
  The system will generate comprehensive, explainable reports that provide insights into the forgery detection process. These reports will include detection confidence scores, key indicators of manipulation, highlighted inconsistencies, and justification for classification decisions. The reports will be tailored for different stakeholders, including news agencies, policymakers, and fact-checking organizations, ensuring usability across various domains.

- Deployment of the platform for testing and evaluation.
  The system will undergo rigorous testing and validation before full deployment. This phase will include:

  - Internal testing to evaluate accuracy, performance, and robustness.
  - Pilot testing with selected users (e.g., news organizations and researchers).
  - User feedback collection to refine the system's functionality.
    The goal is to ensure that the platform performs reliably under real-world conditions and meets user expectations.

Out of Scope:

The following aspects are beyond the scope of this project and will not be covered:

- Legal actions against fake news sources.
  The platform will focus on detecting and analyzing fake news rather than enforcing legal consequences. Any legal actions or penalties related to misinformation will be left to government agencies, legal institutions, or regulatory bodies.

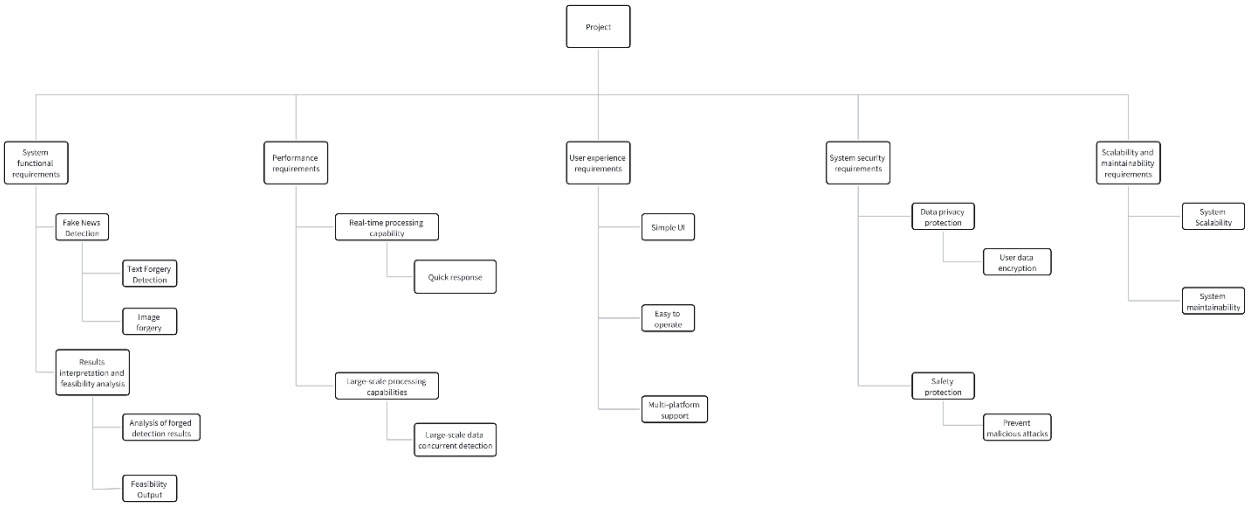
- Government regulations and policy-making.
  While the system may provide insights into fake news trends, it will not engage in creating or influencing legislation regarding misinformation control. The responsibility for establishing policy frameworks, content regulation laws, and compliance measures lies with government and legal entities.

- Manual content moderation beyond system-generated analysis.
  The project aims to develop an automated AI-powered detection system, meaning that human moderators will not manually review flagged content. Instead, the platform will provide detection results, evidence-based explanations, and probabilistic assessments, allowing users, fact-checking organizations, or policymakers to make

informed decisions.

By clearly defining the project's scope, this document ensures that all efforts remain focused on delivering an effective, AI-driven forgery detection platform while avoiding areas beyond the project's technological and operational boundaries.

RBS (requirement breakdown structure)

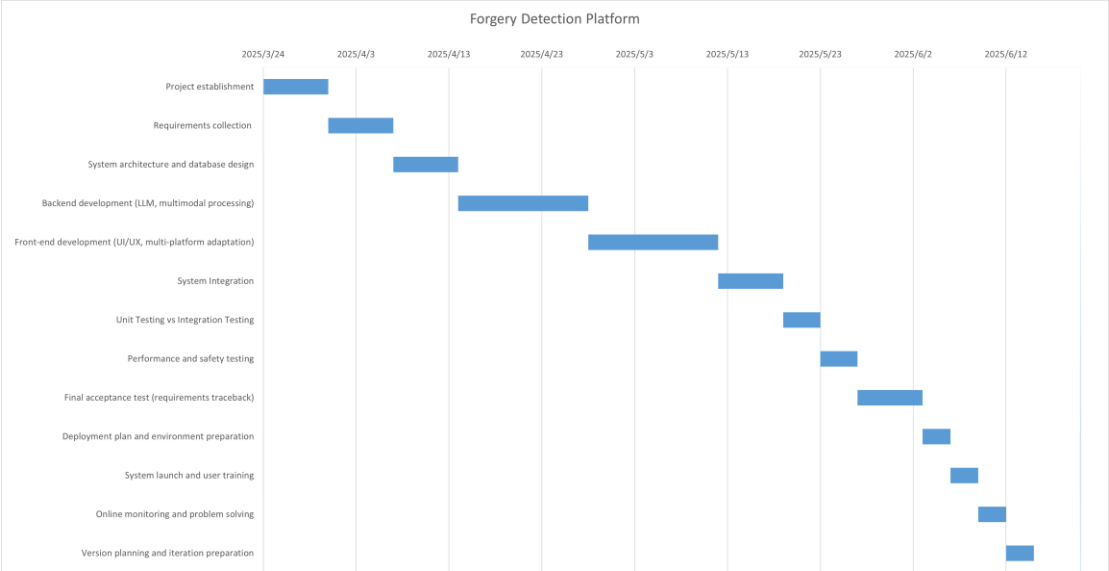


The clear version of RBS can be viewed through this link:

https://v1y1xrajrd9.feishu.cn/wiki/DfyGwtoZ3irlW7kTQpBc2qltnsb?from=from_copylink

Detailed explanation of RBS can also be viewed in the Appendices at the end of the Project Charter.

6. Project Scbedule Summary



The clear version of Gantt chart can be viewed through this link:

https://v1y1xrajrd9.feishu.cn/wiki/DfyGwtoZ3irlW7kTQpBc2qltnsb?from=from_copylink

Project start date:

March 24, 2025

Project end date:

June 12, 2025

Timeline of project phases and milestones

Phase	**Start Date**	**End Date**	**Key Deliverables**
1. Project Establishment	2025/3/24	2025/3/24	Project charter approval, team onboarding.
2. Requirements Collection	2025/3/25	2025/4/3	Finalized requirements document.
3. System Architecture Design	2025/4/4	2025/4/13	System architecture diagrams, DB schema.
4. Backend Development	2025/4/14	2025/5/3	Functional LLM & multimodal processing APIs.

5. Frontend Development	2025/5/4	2025/5/13	UI prototypes, multi-platform interfaces.
6. System Integration	2025/5/14	2025/5/23	Integrated backend-frontend system.
7. Testing Phase	2025/5/24	2025/6/2	Test reports (unit, integration, performance).
8. Deployment & Launch	2025/6/3	2025/6/12	Deployed platform, user training materials.

Project reviews and review dates

Review Stage	**Date**	**Focus Areas**
Kickoff Meeting	2025/3/24	Validate project scope, roles, and timeline.
Requirements Review	2025/4/3	Approve finalized requirements and priorities.
Design Review	2025/4/13	Evaluate system architecture and scalability.
Backend Demo	2025/5/3	Verify API functionality and processing accuracy.
Frontend UX Review	2025/5/13	Assess UI usability and cross-platform compatibility.
Integration Checkpoint	2025/5/23	Confirm end-to-end system functionality.
Pre-Launch Review	2025/6/2	Finalize test results and deployment readiness.
Post-Launch Retrospective	2025/6/12	Review user feedback and initial performance metrics.

Calculate the total project time

- Daily Working Hours: Assume 2 hours per day
- Workdays per Week: Set to 5 days.
- Project Deadline: It may be completed earlier than indicated in the Gantt chart (e.g., before early June).

Adjusted Estimation

1. Project Duration:

- March 24, 2025 – June 12, 2025 (81 days).
- 5 workdays per week, this results in 59 working days (excluding 22 weekend days).

2. Total Estimated Work Hours:
   - 2 hours per day.
   - 59 days × 2 hours/day = 118 hours.

7. Proiect Budget Summary

Total project budget

Considering that this project mainly relies on existing model processing and nighttime dataset resources, the main expenses may come from renting servers and other possible expenses, and the budget for this project is not currently calculated.*(Primary reliance on existing models and nighttime dataset resources. Major expenses limited to server rentals and incidental costs.)*

Budget broken down by phase

Phase	**Estimated Cost (RMB)**	**Key Expenses**
1. Project Establishment	0	Internal planning, no direct costs.
2. Requirements Collection	0	Stakeholder meetings, no external resources.
3. System Design	0	Cloud-based architecture tools (e.g., Lucidchart).
4. Backend Development	0	Server rentals (AWS/GCP), API testing.
5. Frontend Development	0	UI prototyping tools (Figma), cross-platform testing.
6. System Integration	0	Debugging tools, additional

		server capacity.
7. Testing Phase	0	Load testing tools (JMeter), security audits.
8. Deployment & Launch	0	User training, documentation hosting.
Contingency (10%)	0	Unplanned server scaling or tool licenses.
Total	0	*(Excluding existing model/dataset costs)*

8. Quality Issues

The Forgery Detection Platform must meet rigorous quality standards across all functional areas. Below are the key quality requirements organized by category:

Issues

Accuracy Requirements

- The system must achieve minimum precision of 90% in detecting forged content, with particular attention to minimizing false positives that could incorrectly flag legitimate news.
- Recall rates should maintain at least 85% to ensure comprehensive detection of fake news items.
- For multimodal content analysis, the system must demonstrate consistent performance across text (92% accuracy), images (88% accuracy), and videos (85% accuracy).
- Deepfake detection capabilities should correctly identify at least 80% of synthetic media in testing scenarios.

Performance Standards

- Response times must not exceed 1 second for text analysis and 5 seconds for image verification.
- Video processing should complete within 5 seconds per minute of footage.
- The system must maintain 99.9% uptime during operational hours, excluding

scheduled maintenance periods.

- Under peak loads, the platform should handle 10,000 concurrent analysis requests without significant degradation in performance.

Explainability Requirements

- All detection results must include clear, human-readable explanations highlighting:
    - Specific inconsistencies found in the content
    - Confidence levels for each modality analyzed
    - Visual indicators of manipulated regions in images/videos
- The system should generate comprehensive reports in multiple formats (PDF, HTML) that:
    - Present evidence in a logical, traceable manner
    - Include comparative analysis with known authentic content
    - Provide references to supporting data sources

Security and Privacy

- Implement end-to-end encryption for all data transmissions and storage.
- Establish robust access controls with multi-factor authentication for administrative functions.
- Conduct quarterly security audits to identify and address potential vulnerabilities.
- Maintain comprehensive activity logs with 180-day retention for audit purposes.
- Ensure all training datasets undergo rigorous bias testing before model incorporation.

User Experience

- The interface must support responsive design for desktop (4K resolution) and mobile devices.
- Implement accessibility features compliant with WCAG 2.1 AA standards.
- Provide multilingual support for at least English and Chinese interfaces.
- Include intuitive navigation with contextual help features.
- Maintain consistent performance across Chrome, Safari, and Edge browsers.

Reliability and Maintainability

- The codebase must maintain detailed documentation including:
    - API specifications with version control
    - Database schema diagrams
    - Deployment architecture documentation
- Implement automated testing frameworks covering:
    - 80% unit test coverage for critical modules
    - Integration test suites for all major components
    - Performance benchmarking tests
- Establish clear procedures for:
    - Hotfix deployments
    - Version upgrades
    - Data migration processes

Compliance Requirements

- Adhere to regional data protection regulations (GDPR for EU, PIPL for China).
- Maintain documentation for all algorithmic decision-making processes.
- Implement content moderation protocols that respect freedom of expression principles.
- Provide transparency reports detailing system performance and error rates.

Quality Assurance Measures

Testing Protocols

- Conduct adversarial testing with progressively sophisticated fake content
- Perform stress testing at 150% of expected maximum load
- Implement continuous monitoring for model drift detection

Validation Processes

- Monthly accuracy audits using newly collected test datasets
- Quarterly third-party security assessments

- Biannual usability studies with representative user groups

Improvement Mechanisms

- Establish feedback channels for end-users to report false positives/negatives
- Maintain a public issue tracker for transparency
- Implement automated retraining pipelines for model improvement

9. Resources Required

People

- Project Manager: Oversees project execution, timeline, and stakeholder communication.
- Backend Developers: Build API services, database integration, and cloud deployment.
- Frontend Developers: Design and implement the user interface (web/mobile).
- Data and AI developer: perform model training, and evaluate accuracy, Develop and optimize forgery detection models (LLMs, multimodal fusion).
- QA Testers: Conduct functional, performance, and security testing.
- DevOps Engineer: Manages CI/CD pipelines and cloud infrastructure.

Technology

- AI Models:
  - Pretrained LLMs (e.g., GPT-4, BERT) for text analysis.
  - Multimodal models (CLIP, BLIP) for image-text correlation.
  - Deepfake detection tools (e.g., Microsoft Video Authenticator).
- Development Tools:
  - Python (PyTorch, TensorFlow), JavaScript (React), Docker, Kubernetes.
- Cloud Services:
  - AWS/GCP for server rentals (EC2, Cloud TPUs), storage (S3), and databases (PostgreSQL).

- Security:
    - Encryption tools (AES-256), access control frameworks (IAM).

Facilities

- Development Environment:
    - High-performance workstations (GPU-enabled) for model training.
    - Secure office space for team collaboration (optional for remote teams).
- Testing Labs:
    - Access to controlled environments for stress/load testing.

Other

- Datasets:
    - Labeled fake/real news datasets (e.g., DGM4 on Hugging Face).
    - Deepfake video repositories (e.g., FaceForensics++).
- Legal Compliance:
    - Consultations with legal experts on data privacy (GDPR/PIPL).

Resources to be provided

Resource	**Name of Provider**	**Date to Be Provided**
Cloud Compute Credits	君同未来科技有限责任公司	2025/4/1
Labeled News Datasets	Hugging Face / Internal Team	2025/3/30 (Phase 1)
GPU Workstations	Company IT Department	2025/4/5
Legal Advisory	External Law Firm	2025/5/10 (Pre-launch)

10. Assumptions and Risks

Assumptions Used to Develop Estimates

- Technical Assumptions:
    - Existing pretrained models (LLMs, CLIP/BLIP) can be fine-tuned effectively

for forgery detection.

    - Cloud service providers (AWS/GCP) will maintain stable pricing and availability during the project.
    - Sufficient labeled datasets (e.g., DGM4) will remain accessible for training and validation.

- Operational Assumptions:
    - Team members will dedicate 2 hours/day consistently throughout the project.
    - No major organizational restructuring will disrupt the project team.
    - Legal frameworks for content moderation will not change significantly during development.
- Resource Assumptions:
    - The sponsor company will provide GPU workstations by April 2025.
    - Nighttime computational resources will be available for model training.

Key Risks, Probability, and Impact

Risk	**Probability**	**Impact**	**Mitigation Strategy**
Rapid evolution of forgery techniques	High	Critical	Continuous model retraining; adversarial testing
Bias in training datasets	Medium	High	Regular fairness audits; diverse data sourcing
Cloud service outages	Low	High	Multi-cloud backup strategy
Legal challenges around content moderation	Medium	Medium	Early consultation with legal experts
Team attrition	Low	Medium	Documentation standards; cross-training

Constraints

- Technical Constraints:
    - Limited to analyzing text, images, and videos (no audio-only content).

    - Must use existing company-approved cloud providers (AWS/GCP).
    - Real-time processing requirement limits model complexity.
- Resource Constraints:
    - Maximum 2 hours/day per team member.
    - No budget for manual content moderation.
    - GPU access restricted to nighttime hours for cost control.
- Regulatory Constraints:
    - Must comply with Chinese data privacy laws (PIPL).
    - Cannot store user-uploaded content beyond 30 days.

Dependencies

- Internal Dependencies:
    - IT department for workstation provisioning (due April 2025).
    - Legal team for compliance reviews (Q2 2025).
- External Dependencies:
    - Hugging Face dataset availability.
    - OpenAI's CLIP model updates.
    - Cloud provider API stability.

Project Impact Assessment

- Positive Impacts:
    - Establishes the company as a leader in AI-powered content verification.
    - Potential 30-40% reduction in fake news propagation for partner platforms.
    - Creates reusable AI infrastructure for future projects.
- Negative Impacts:
    - Temporary productivity dip during system integration (May 2025).

    - Increased cloud computing costs during peak training periods.

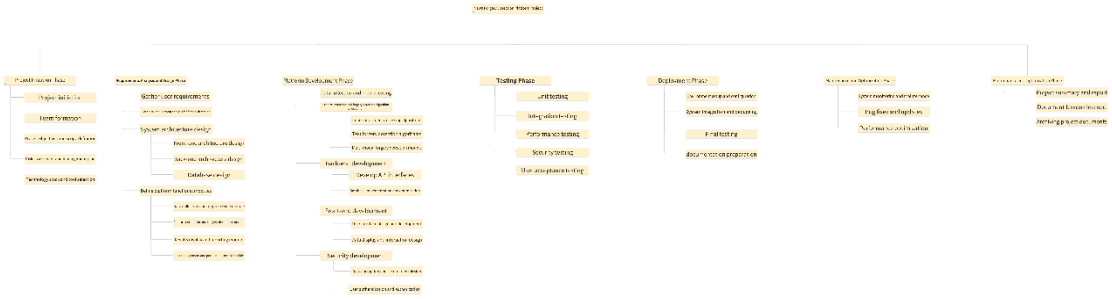
Outstanding Issues

- Unresolved Technical Issues:
    - Optimal threshold settings for multimodal consistency checks.
    - Handling of satirical content (may trigger false positives).
- Pending Decisions:
    - Final selection of deepfake detection model (pending April benchmarks).
    - UI language support beyond English/Chinese (budget-dependent).
- External Factors:
    - Potential changes to social media platform APIs.
    - Unclear regulatory stance on AI-generated content detection.

Risk Monitoring Plan

1. Monthly Risk Reviews: Evaluate top 3 risks and mitigation effectiveness.
2. Early Warning Indicators:
    - Model accuracy drops >5% in weekly tests.
    - Cloud costs exceed projections by 15%.
3. Contingency Reserves:
    - 10% buffer in project timeline.
    - $2,000 unallocated budget for emergency compute needs.

11. Work Breakdown Structure (WBS)



The clear version of WBS can be viewed through this link：

https://v1y1xrajrd9.feishu.cn/wiki/DfyGwtoZ3irlW7kTQpBc2qltnsb?from=from_copylink

Detailed explanation of WBS can also be viewed in the Appendices at the end of the Project Charter.

12. Project Administration

Communications Plan

- Stakeholder Communication
    - Frequency:
        - Weekly status updates (email) to core team
        - Biweekly progress reports (PPT) to sponsors
        - Monthly steering committee meetings
    - Channels:
        - Internal: Feishu for team collaboration
        - External: Encrypted emails for sensitive data
    - Escalation Path:
      Technical issues → Project Manager → CTO (within 24hrs)
- Meeting Structure

Meeting Type	Frequency	Participants
Daily Standup	Weekdays	Dev Team

Sprint Planning	Biweekly	PM + Tech Leads
Risk Review	Monthly	All Stakeholders

Scope Management Plan

- Control Process:
    1. All change requests must use CR-001 template
    2. Impact analysis by Tech Lead within 3 days
    3. Steering Committee approval for >8hr effort changes
- Scope Baseline:
    - Version 1.0 (approved 2025/3/23)
    - Frozen after Design Phase (2025/4/13)

Quality Management Plan

- Quality Gates:

Phase	Checkpoint	Sign-off Required
Requirements	100% testable criteria defined	Sponsor
Development	80% unit test coverage	QA Lead
Deployment	Penetration test passed	CISO

- Testing Strategy:
    - Automated: pytest (backend), Jest (frontend)
    - Manual:
        - Bias testing: Quarterly by Ethics Panel
        - UX testing: 20+ real users in pilot

Change Management Plan

1. Change Control Process
    - Submission: All changes require CRF-001 form (description, impact, priority).

- Evaluation:
    - *Technical*: Feasibility analysis by Tech Lead (48h).
    - *Business*: Sponsor reviews ROI using decision matrix.

2. Approval Authority

Change Type	Approver
Low-impact (≤5% cost)	Project Manager
High-impact	Change Control Board

3. Implementation

- Approved changes are scheduled in the next sprint.
- Document updates (requirements, WBS) within 24h.

4. Communication

- Notify all stakeholders via email/Feishu.
- Update Gantt chart and version log.

Tools: Jira for tracking, Confluence for documentation.

Human Resources Plan

1. Roles & Responsibilities

Role	Name	Key Tasks	Availability
AI Model Training	张茁含睿	Fine-tune LLMs, optimize detection models	Tue/Thu (4h/day)
Security Review	胡家豪	Penetration testing, compliance checks	Last Friday monthly
Frontend Dev	黄俊哲	UI/UX development, multi-platform support	Full-time
Backend Dev	陶子芾	API development, database integration	Full-time
Testing Team	All	Execute test cases, report bugs	As needed (see *Testing*

	members		*Phase* below)

2. Testing Phase Allocation

- Unit Testing: Backend/Frontend owners (陶子芾，黄俊哲)
- Integration Testing: All members (2h daily during Sprint 4)
- User Acceptance Testing: 张茁含睿 + 胡家豪 (final week)

3. Contingency

- Backup developer: Cross-trained on frontend/backend (documented in Appendix D).

Implementation & Closure Plan

- Deployment Phases:
    1. Soft Launch (2025/6/3): 3 news agency partners
    2. Full Release (2025/6/10): Public API access
- Closure Activities:
    - Final audit (2025/6/15)
    - Knowledge transfer docs due (2025/6/18)
    - Post-mortem meeting (2025/6/20)
- Success Metrics:
    - 95% of test cases passed
    - 80%+ user satisfaction in first month
    - Cloud costs within 10% of forecast

13. Approval

Name	**Position**	**Signature**	**Date**
[Project Sponsor]	君同未来科技有限责任公司		3.23.2025

[Project Manager]	张茁含睿		3.23.2025
[Team Lead]	张茁含睿		3.23.2025

References

This section lists all authoritative sources, research papers, tools, and frameworks referenced during the project's development.

Research Papers & Academic Sources

1. Radford, A., et al. (2021). "Learning Transferable Visual Models From Natural Language Supervision" (CLIP). *arXiv:2103.00020*.
    - https://arxiv.org/abs/2103.00020
2. Li, J., et al. (2022). "BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation". *arXiv:2201.12086*.
    - https://arxiv.org/abs/2201.12086
3. Shao, R., et al. (2023). "Detecting and Grounding Multi-Modal Media Manipulation" (DGM4 Dataset). *CVPR 2023*.
    - https://arxiv.org/abs/2304.02556

Technical Documentation

1. OpenAI CLIP Documentation:
    - https://openai.com/research/clip
2. Hugging Face Datasets:
    - DGM4 Dataset: https://huggingface.co/datasets/rshaojimmy/DGM4
3. PyTorch & TensorFlow:
    - Official documentation for model implementation.

Industry Standards & Regulations

1. General Data Protection Regulation (GDPR):
    - EU regulation on data privacy.
2. Personal Information Protection Law (PIPL):
    - China’s data privacy framework.
3. WCAG 2.1:
    - Web Content Accessibility Guidelines.

Project Management Frameworks

1. PMBOK® Guide (7th Edition):
    - Project Management Institute (PMI) standards.
2. ISO 9001:2015:
    - Quality management system requirements.

Terminology or Glossary

This section defines key terms and acronyms used throughout the project to ensure clear communication among stakeholders.

Technical Terms

- LLM (Large Language Model):
  AI models (e.g., GPT-4, BERT) trained on vast text data for natural language processing tasks.
  Usage: "LLMs will analyze news articles for semantic inconsistencies."
- Multimodal Fusion:
  Techniques combining multiple data types (text, images, videos) to improve detection accuracy.
  Example: "CLIP enables multimodal fusion by correlating images with captions."

- Deepfake:
  AI-generated synthetic media (videos/audio) that manipulate reality.
  Project Scope: "Our platform targets deepfakes with ≥80% detection accuracy."
- XAI (Explainable AI):
  Methods to make AI decisions interpretable to humans.
  Requirement: "All detections must include XAI reports highlighting evidence."

Project Management Terms

- MOV (Measurable Organizational Value):
  Quantifiable benefits the project delivers (e.g., "90% fake news detection rate").
- CR (Change Request):
  Formal proposal to modify project scope/requirements (Template: CR-001).
- Sprint:
  2-week development cycles with defined deliverables.

Domain-Specific Terms

- F1-Score:
  Balanced metric combining precision and recall (Target: ≥88%).
- Adversarial Testing:
  Intentional attacks on the system to evaluate robustness.
- PIPL (Personal Information Protection Law):
  Chinese data privacy regulation governing user data handling.

Acronyms

Acronym	**Full Form**
API	Application Programming Interface
AWS	Amazon Web Services
QA	Quality Assurance
UI/UX	User Interface/User Experience

Appendices

This section provides additional resources and references related to the Forgery Detection Platform for News Scenarios project. These resources include links to the project's code repository, research papers, and other relevant materials.

Project Resources

Git Repository:

The source code and development progress of the project can be accessed through the following Git repository:
https://github.com/turambar928/Softeware-management-course-project

Research Papers:

The following research papers provide the theoretical foundation and technical insights for the project:

[2304.02556] Detecting and Grounding Multi-Modal Media Manipulation

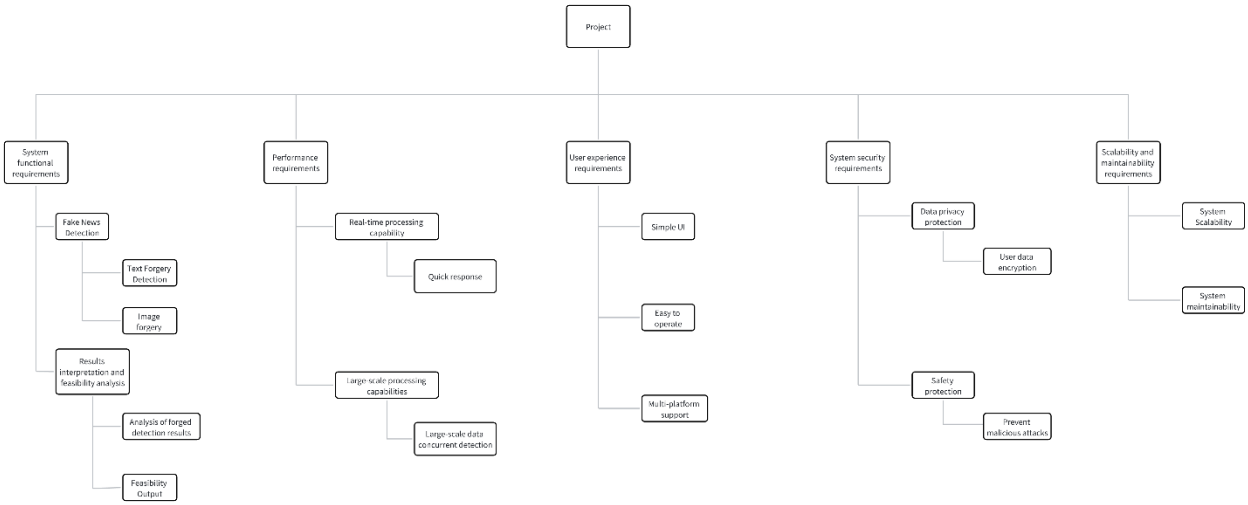
Dataset Repository:

The labeled datasets used for training and testing the AI models are available at:
rshaojimmy/DGM4 · Datasets at Hugging Face

Detailed Graphs Explainations：

https://v1y1xrajrd9.feishu.cn/wiki/DfyGwtoZ3irlW7kTQpBc2qltnsb?from=from_copylink

RBS (requirement breakdown structure)



The clear version of RBS can be viewed through this link：

https://v1y1xrajrd9.feishu.cn/wiki/DfyGwtoZ3irlW7kTQpBc2qltnsb?from=from_copylink

Detailed explanation of RBS:

a. System Functional Requirements

These are the essential functions the system must be able to perform.

- **Fake News Detection**: This is the core functionality, ensuring the system can detect false news content.
    - **Text Forgery Detection**: Detects alterations or falsifications in text.
        - **Image Forgery**: Detects forged or manipulated images within the content.
- **Results Interpretation and Feasibility Analysis**: This focuses on interpreting the results from the detection system and evaluating their feasibility.
    - **Analysis of Forged Detection Results**: A task that evaluates the accuracy and reliability of the forged content detection.
    - **Feasibility Output**: Analysis regarding the practicality and scalability of the results produced.

b. Performance Requirements

Performance defines how well the system should perform under certain conditions, such as speed and scale.

- **Real-time Processing Capability**: The system needs to process data in real-time.
    - **Quick Response**: The system must respond rapidly to queries and actions.
- **Large-scale Processing Capabilities**: The system must handle vast amounts of data and process them efficiently.
    - **Large-scale Data Concurrent Detection**: The system must support simultaneous detection of large datasets without compromising speed.

c. User Experience Requirements

These are the expectations for the user interface and user experience, focusing on making the system easy to use.

- **Simple UI**: The user interface should be straightforward and easy to navigate.

- **Easy to Operate**: The system should be intuitive, requiring minimal effort for users to operate.
- **Multi-platform Support**: The system should support multiple platforms to accommodate a wide range of devices and operating systems.

d. System Security Requirements

This section outlines the security measures required to protect user data and the integrity of the system.

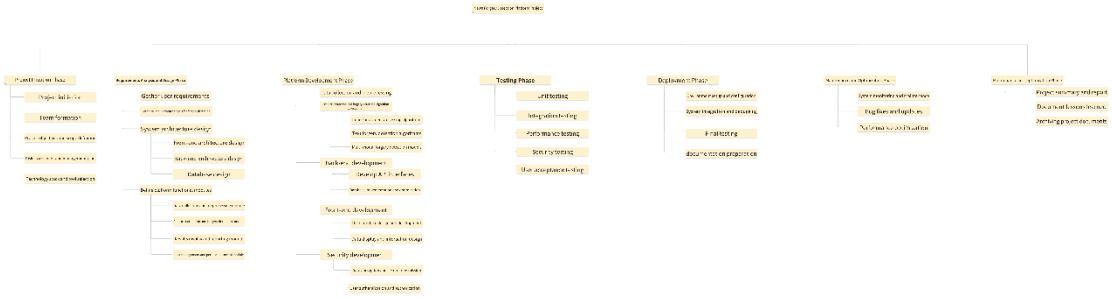
- **Data Privacy Protection**: Ensuring that user data is protected from unauthorized access or misuse.
    - **User Data Encryption**: Sensitive data should be encrypted to prevent data breaches.
- **Safety Protection**: The system must include mechanisms to safeguard against malicious threats.
    - **Prevent Malicious Attacks**: Measures to prevent attacks such as hacking or data corruption.

e. Scalability and Maintainability Requirements

These requirements focus on ensuring the system can grow and be maintained over time.

- **System Scalability**: The system should be able to scale efficiently as demand increases.
- **System Maintainability**: The system should be easy to maintain, with clear processes for updates and bug fixes.

Work Breakdown Structure (WBS)



The clear version of WBS can be viewed through this link：

https://v1y1xrajrd9.feishu.cn/wiki/DfyGwtoZ3irlW7kTQpBc2qltnsb?from=from_copylink

Detailed explanation of WBS:

1. **Project Initiation Phase**
    - **Project initiation**: Define the project's goals and scope to ensure all team members understand the project's core objectives.
    - **Team formation**: Select appropriate team members and allocate roles and responsibilities effectively.
    - **Risk assessment and management plan**: Evaluate potential risks during project execution and develop mitigation strategies.
    - **Technology stack and tool selection**: Choose appropriate technologies and tools to ensure smooth development throughout the project.
2. **Requirements Analysis and Design Phase**
    - **Gather user requirements**: Collect technical requirements and forgery data through communication and research with users to guide platform development.
    - **System architecture design**: Design the overall system architecture, including front-end, back-end, and database design, ensuring it supports the platform's core functionalities.
    - **Define platform functional modules**: Clarify the platform's functional modules, such as data collection and preprocessing, feature extraction and forgery detection, results display and reporting, and user management and access control.
3. **Platform Development Phase**
    - **Data collection and preprocessing**: Develop a module for collecting and preprocessing data, ensuring the platform can efficiently process forgery data.
    - **Feature extraction and forgery detection algorithm development**: Develop algorithms for detecting forgeries in images, videos, and text, using multi-modal data to enhance detection accuracy.
    - **Back-end development**: Develop back-end API interfaces and optimize the database design to ensure system efficiency and stability.
    - **Front-end development**: Design and implement the user interface to ensure user-friendly interaction and design.
    - **Security development**: Implement data encryption and secure transmission

mechanisms and design user authentication and authorization modules to protect user data and platform security.

4. Testing Phase

- **Unit testing**: Test each module independently to ensure its functionality is correct.
- **Integration testing**: Test the integration of all modules to ensure smooth cooperation between them.
- **Performance testing**: Test the platform's performance under heavy loads to ensure its stability and response time.
- **Security testing**: Identify and address any security vulnerabilities in the platform.
- **User acceptance testing**: Simulate real-world user scenarios to ensure the platform meets user requirements and offers a good user experience.

5. Deployment Phase

- **Environment setup and configuration**: Set up and configure the necessary runtime environment for the platform.
- **System integration and debugging**: Integrate the system and debug any issues that arise during the integration process, ensuring all parts work together smoothly.
- **Final testing**: Perform final functional and performance testing to ensure the platform is ready for deployment.
- **Documentation preparation**: Prepare project documentation to ensure a smooth handover for future support and maintenance.

6. Maintenance and Optimization Phase

- **System monitoring and maintenance**: Continuously monitor the platform to ensure its proper functioning and resolve any issues promptly.
- **Bug fixes and updates**: Regularly fix bugs in the system and release updates.
- **Performance optimization**: Conduct performance tuning to keep the platform operating at its optimal state.
- **Project summary and archiving**: Prepare a project summary report, document lessons learned, and archive project documents for future reference.