University of Essex

Online

"Design Proposal for Academic Research Online Agent System"

Group F – Weekly Meeting Notes

Week 1 – Project Kickoff & Team Contract

Date: 09 August 2025

- All team members reviewed and signed the Team Contract with assigned roles:

  ✓ Fahad Saleh – Project Lead (oversees timelines, stakeholder communication)

  ✓ Rayyan Alnaqbi – Lead Developer (manages codebase, agent logic, API integration)

  ✓ Ali Alzahmi – Researcher/QA (validates algorithms, testing coverage)

- Finalised project title: Academic Research Online Agent System.

- Established communication via Microsoft Teams and WhatsApp ; weekly meetings scheduled on Mondays.

- Set up GitHub repository for version control and Trello board for task tracking.

- Agreed on PEP8/Pylint coding standards and Agile sprints (6-week timeline).

Name	Attendance Status
Fahad Saleh	Attended
Rayyan Alnaqbi	Attended
Ali Alzahmi	Attended

Week 2 – System Requirements & Architecture Planning

Date: 13 August 2025

Rayyan presented software libraries:

- Scrapy for crawling (asynchronous architecture, proxy rotation).

- Selenium for dynamic content (CAPTCHA handling via 2Captcha API).

- SQLAlchemy and AWS S3 for storage.

Ali proposed hardware specs:

- Local deployment (Intel i5 + 8GB RAM) vs. cloud (AWS EC2 t3.medium).

Fahad outlined the Agile sprint breakdown:

- Sprint 1: Crawler Agent development (Weeks 1–2).

- Sprint 2: Processor Agent pipelines (Weeks 3–4).

- Sprint 3: Storage Agent integration (Weeks 5–6).

- Team began drafting the System Requirements section of the report.

Name	Attendance Status
Fahad Saleh	Attended
Rayyan Alnaqbi	Attended
Ali Alzahmi	Attended

Week 3 – Agent Design & Methodology Finalization

Date: 17 August 2025

- Rayyan developed the Crawler Agent prototype with proxy rotation (BrightData API).

- Ali tested ethical compliance measures:

- Implemented `robotexclusionrulesparser` for robots.txt adherence.

- Anonymised IPs using Tor networks.

- Team debated Python vs. Rust/Node.js:

- Chose Python for NLP ecosystem (spaCy, NLTK) and Docker compatibility.

- Fahad finalized the UML sequence diagram for agent communication (Redis Pub/Sub).

- Drafted Methodology section (Agile implementation, JSON schema validation).

Name	Attendance Status
Fahad Saleh	Attended
Rayyan Alnaqbi	Attended
Ali Alzahmi	Attended

- Addressed dynamic content rendering:

- Rayyan integrated Selenium's `WebDriverWait` with 20s timeout for AJAX-heavy sites.

- Resolved data volume issues:

- Ali implemented zlib compression and LRU caching for SQLite.

- Fahad reviewed system architecture diagrams (class/sequence diagrams).

- Team peer-reviewed code for GDPR compliance and performance optimisations.

- Began drafting Challenges and Mitigations section.

Name	Attendance Status
Fahad Saleh	Attended
Rayyan Alnaqbi	Attended
Ali Alzahmi	Attended

Week 5 – Report Drafting & Presentation Prep

Date: 31 August 2025

- Ali finalised the Results section, including TF-IDF relevance scoring and Parquet storage benchmarks.

- Rayyan completed Storage Agent integration with AWS S3 (multipart uploads, KMS encryption).

- Fahad consolidated the Conclusion and Justification of Key Choices (e.g., multi-agent vs. monolithic).

- Team created presentation slides highlighting agent workflows and ethical compliance.

- Conducted a dry run of the final presentation for internal feedback.

Name	Attendance Status
Fahad Saleh	Attended
Rayyan Alnaqbi	Attended
Ali Alzahmi	Attended

Week 6 – Final Review & Submission

Date: 06 September 2025

- Performed final code review:

- Validated Crawler Agent's 95% link accuracy on PubMed (pytest).

- Ensured OAuth 2.0 security for AWS access.

- Ali formatted references and cross-checked citations (APA 7th edition).

- Fahad submitted the report and codebase to the portal ahead of the deadline.

- Team rehearsed the presentation, focusing on GDPR compliance and system scalability.

Name	Attendance Status
Fahad Saleh	Attended
Rayyan Alnaqbi	Attended
Ali Alzahmi	Attended