As part of Research Paper:

Behavioral Analytics and Forensic Accounting: Understanding the Human Element in Fraud

Samuel F. Johnson-Rokosu

Appendix C: Detailed Data Analysis Plan: Coding Frameworks Aligned with RQs

The following coding frameworks will be structured around study research questions to analyse the qualitative data from interviews with forensic accountants and behavioural psychologists. Each framework contains primary codes (aligned to RQs), sub-codes (specific topics), and case study examples (Enron, Wirecard, FTX).

---

1. Coding Framework for *Behavioral Indicators* (RQ1)

Objective: Identify patterns in human behavior preceding fraud.

Primary Code	Sub-Codes	Examples from Case Studies
Communication Patterns	Evasive language, tone shifts	Emails from Enron: phrases like hiding losses, avoiding audits (negative sentiment score = -0.8).
Leadership Traits	Overconfidence, autocratic style	FTX CEO tweets: Overconfidence score = 0.92; Enron's Skilling centralization score = 0.92.
Psychological Biases	Optimism bias, groupthink	FTX investors' herd behavior (modularity = 0.17); Enron's "win at all costs" rationalization.
Lifestyle Anomalies	Sudden wealth, secrecy	Wirecard executives' unexplained asset accumulation before collapse.

Analysis Method:

- Use **NVivo** or **Atlas.ti** to tag interview excerpts with sub-codes.
- Quantify frequency (e.g., "evasive language" mentioned in 8/10 accountant interviews).

---

2. Coding Framework for *Ethical Trade-Offs* (RQ4)

Objective: Explore tensions between surveillance and privacy.

Primary Code	Sub-Codes	Examples from Case Studies
Privacy Violations	GDPR non-compliance, PII exposure	Wirecard: Post-GDPR monitoring efficacy dropped 40%; Enron's unflagged personal complaints.
Surveillance Efficacy	Anonymization challenges	BioCatch's keystroke dynamics vs. GDPR's "purpose limitation" principle.
Employee Trust	Fear of retaliation, transparency	Wirecard survey: 68% feared retaliation; Enron's lack of whistleblower protections.

Regulatory Gaps	Crypto opacity, decentralized risks	FTX: 0/10 exchanges complied with behavioral monitoring laws.

Analysis Method:

- Thematic analysis of ethical dilemmas (e.g., "How to balance AI monitoring with GDPR?").
- Map trade-offs using matrices (e.g., *Privacy vs. Fraud Detection Efficacy*).

3. Coding Framework for *Organizational Factors* (RQ3)

Objective: Uncover systemic vulnerabilities enabling fraud.

Primary Code	Sub-Codes	Examples from Case Studies
Leadership Style	Autocratic vs. ethical leadership	Enron's centralized hierarchy (Skilling's control score = 0.92).
Corporate Culture	Toxicity, "win at all costs"	Enron emails: 78% contained toxic terms like "avoid scrutiny."
Whistleblower Systems	Protections, reporting channels	Wirecard's governance score = 2.1/5 for whistleblower safeguards.
Internal Controls	Audit weaknesses, oversight gaps	Wirecard's 20-F filings: Traditional audits missed €1.9B cash discrepancies.

Analysis Method:

- Cross-reference interview data with case study metrics (e.g., governance scores).

---

4. Coding Framework for *Machine Learning Integration* (RQ2)

Objective: Assess AI's role in fraud detection.

Primary Code	Sub-Codes	Examples from Case Studies
AI Effectiveness	Accuracy, anomaly detection	Isolation Forest: 93% accuracy in Wirecard cash flow anomalies.
Data Integration	Behavioral + financial metrics	Enron: Combining NLP sentiment scores with stock volatility (r = -0.89).
Implementation Barriers	GDPR constraints, technical limits	Post-GDPR anonymization reduced Wirecard's monitoring efficacy by 40%.
Human-AI Collaboration	Auditor reliance on AI tools	83% of Wirecard fraud cases detected by AI vs. 0% by traditional audits.

Analysis Method:

- Code positive/negative perceptions of AI (e.g., "AI accelerates detection" vs. "privacy risks").

Implementation Steps

1. **Codebook Development**: Define codes/sub-codes with examples from case studies.
2. **Inter-Coder Reliability**: Train analysts to ensure consistency (e.g., Cohen's $\kappa \geq 0.8$).
3. **Triangulation**: Compare qualitative codes with quantitative data (e.g., survey results, AUC scores).
4. **Visualization**:
    - **Heatmaps**: Show frequency of behavioral indicators across cases.
    - **Network Diagrams**: Map relationships between codes (e.g., "overconfidence → liquidity gaps").

Example of Coded Interview Excerpt

Forensic Accountant Quote:

"In Enron, Skilling's emails were riddled with urgency and secrecy. We saw phrases like 'hide losses' months before the collapse."

Codes Applied:

- **Primary**: *Behavioral Indicators*
- **Sub-Codes**: *Communication Patterns (evasive language), Leadership Traits (autocratic style)*
- **Case Link**: Enron's NLP sentiment score = -0.8.

This structured approach ensures consistency in linking interview insights to the research objectives of the study, and allows for actionable conclusions on behavioural analysis in fraud detection.