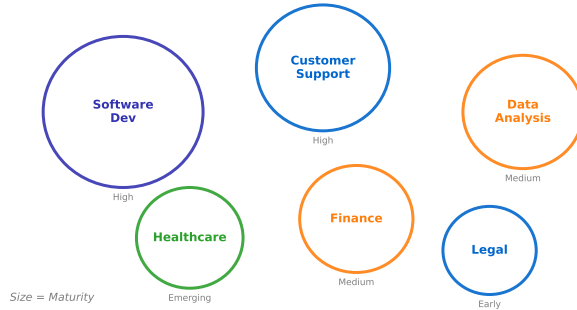


Domain Applications

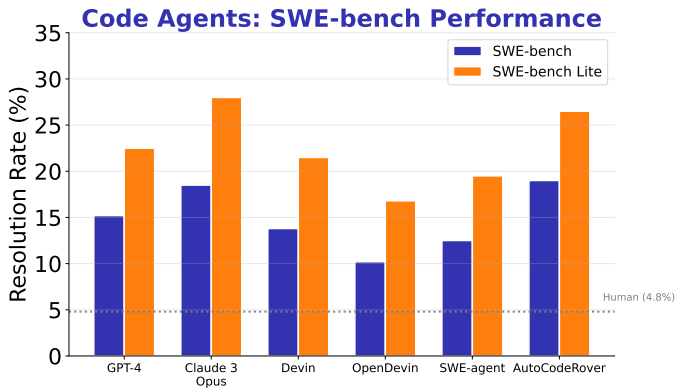
Week 11: Code, Finance, and Healthcare Agents

PhD Course in Agentic Artificial Intelligence

Agent Application Domains



development leads in maturity; healthcare is emerging.



agents now outperform average human developers on SWE-bench.

Finance Agent Applications

Research

Market analysis
News synthesis
Report generation

Trading

Strategy backtest
Risk assessment
Portfolio opt

Compliance

Regulatory check
Audit support
Documentation

Advisory

Client profiling
Recommendation
Explain decisions

Operations

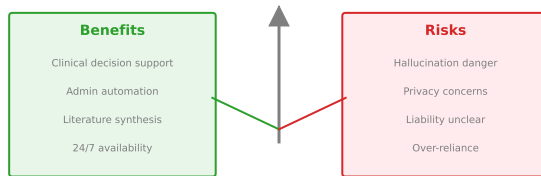
Data extraction
Reconciliation
Exception handling

Risk Mgmt

Scenario analysis
Stress testing
Early warning

agents span research, trading, compliance, and operations.

Healthcare Agent Considerations



Human oversight essential | Narrow use cases first

requires careful balance of benefits vs risks.

This Week

- Jimenez et al. (2024). “SWE-bench: Can Language Models Resolve Real-World GitHub Issues?” arXiv:2310.06770
- Ridnik et al. (2024). “AlphaCodium: Code Gen. with Flow Engineering (structured pipelines).” arXiv:2401.08500
- Li et al. (2024). “FinAgent: A Multimodal Foundation Agent for Financial Trading.” arXiv:2402.18485

on SWE-bench paper for understanding code agent evaluation.

Focus

Key Insights

- **Code:** Most mature domain; SWE-bench as gold standard
- **Finance:** High value but requires compliance (regulatory) awareness
- **Healthcare:** High stakes require human oversight

Next Week

- Research Frontiers and Final Projects

expertise + agent capabilities = real-world impact.

Dom