E2E Test Comprehensive Fix Plan - January 8, 2026

Executive Summary

Current Status: 161/205 tests passing (78.5%) **Target**: 210/210 tests passing (100%) **Gap**: 49 tests need fixing (39 failures + 5 skipped + restore to 210 total)

Test Results After Rollback

- **Passed**: 161 (78.5%)
- **Failed**: 39 (19.0%)
- **Skipped**: 5 (2.4%)
- **Total**: 205 tests

Note: Rollback performed worse than Phase 3 (173/210 = 82%). This requires investigation.

Failure Categories

---

1. Agent Performance - Parallel Execution (5 tests)

Failed Tests: - `[chromium] › ai-agents.spec.ts:145:7 › Should support parallel agent execution` - `[firefox] › ai-agents.spec.ts:145:7 › Should support parallel agent execution` - `[webkit] › ai-agents.spec.ts:145:7 › Should support parallel agent execution` - `[Mobile Chrome] › ai-agents.spec.ts:145:7 › Should support parallel agent execution` - (Mobile Safari not shown but likely failed)

Root Cause: TBD - needs research into parallel async operations in Playwright

2. Error Handling - Timeout Scenarios (5 tests)

Failed Tests: - `[chromium] › ai-agents.spec.ts:243:7 › Should handle timeout scenarios` - `[firefox] › ai-agents.spec.ts:243:7 › Should handle timeout scenarios` - `[webkit] › ai-agents.spec.ts:243:7 › Should handle timeout scenarios` - `[Mobile Chrome] › ai-agents.spec.ts:243:7 › Should handle timeout scenarios` - (Mobile Safari not shown but likely failed)

Root Cause: TBD - needs research into timeout testing patterns

3. Authentication - Validation Errors (5 tests)

Failed Tests: - `[chromium] › authentication.spec.ts:23:7 › Should show validation errors for invalid login` - `[firefox] › authentication.spec.ts:23:7 › Should show validation errors for invalid login` - `[webkit] › authentication.spec.ts:23:7 › Should show validation errors for invalid login` - `[Mobile Chrome] › authentication.spec.ts:23:7 › Should show validation errors for invalid login` - (Mobile Safari not shown but likely failed)

Root Cause: Error messages not appearing in UI - needs form error handling research

4. Script Management - List Display (5+ tests)

Failed Tests: - `[chromium] › script-management.spec.ts:194:7 › Should display list of uploaded scripts` - `[firefox] › script-management.spec.ts:194:7 › Should display list of uploaded scripts` - `[webkit] › script-management.spec.ts:194:7 › Should display list of uploaded scripts` - (More browsers likely affected)

Root Cause: React Query data loading after authentication redirect - needs research

5. Firefox-Specific Issues (6 tests total)

Failed Tests: - `[firefox] › ai-analytics.spec.ts:113:7 › Should display analytics dashboard page` - `[firefox] › ai-analytics.spec.ts:132:7 › Should display cost metrics` - `[firefox] › script-management.spec.ts:35:7 › Should display upload button` - `[firefox] › script-management.spec.ts:208:7 › Should allow searching scripts` - (Plus 2 from other categories)

Root Cause: Firefox-specific rendering or timing - needs browser-specific research

6. Mobile Chrome Analytics (1 test)

Failed Tests: - `[Mobile Chrome] › ai-analytics.spec.ts:161:7 › Should display model performance metrics`

Root Cause: Mobile viewport rendering issue - needs mobile testing research

7. Skipped Tests (5 tests)

Reason: Require environment variables (API keys, external services)

Action Needed: Document how to enable these tests

Research Topics for 2026 Best Practices

1. **Playwright + React Query v5 Integration**
2. Query state management in E2E tests
3. Waiting for cache hydration

4. Testing refetch behavior

5. **Form Validation Error Display**

6. React Hook Form error rendering
7. Accessibility-compliant error messages

8. Testing form validation states

9. **Parallel Async Operations Testing**

10. Testing concurrent API calls
11. Race condition detection

12. Promise.all() testing patterns

13. **Timeout Scenario Testing**

14. Simulating slow network conditions
15. Testing error boundaries

16. Graceful degradation patterns

17. **Firefox-Specific E2E Testing**

18. Gecko engine timing differences
19. Firefox DevTools integration

20. Browser-specific selectors

21. **Mobile Viewport Testing**

22. Touch event simulation
23. Viewport-specific rendering
24. Mobile performance testing

Comprehensive Fix Approach

Phase 1: Research (Current)

- [ ] Internet research for each failure category
- [ ] MCP tools analysis of failing test code
- [ ] Review 2026 Playwright documentation
- [ ] Study React Query v5 + testing patterns

Phase 2: Validation & Authentication Fixes

- [ ] Fix form error display mechanism
- [ ] Ensure error messages visible in all browsers
- [ ] Add proper ARIA labels for accessibility
- [ ] Test across all 5 browser configurations

Phase 3: Data Loading & React Query Fixes

- [ ] Implement proper wait strategies for React Query
- [ ] Add data-testid attributes for loaded states
- [ ] Handle authentication redirect → data fetch flow
- [ ] Test list display after navigation

Phase 4: Agent & Async Operation Fixes

- [ ] Fix parallel agent execution tests
- [ ] Implement timeout scenario testing
- [ ] Add proper error boundaries
- [ ] Test async operation completion

Phase 5: Browser-Specific Fixes

- [ ] Address Firefox rendering issues
- [ ] Fix Mobile Chrome analytics display
- [ ] Ensure consistent behavior across browsers
- [ ] Test mobile viewport edge cases

Phase 6: Verification & Documentation

- [ ] Run complete test suite
- [ ] Verify 210/210 tests passing
- [ ] Document all changes made
- [ ] Create maintenance guide

Resources to Utilize

1. **Internet Research (January 2026)**
2. Latest Playwright documentation
3. React Query v5 testing guides
4. Browser compatibility databases

5. Stack Overflow recent solutions

6. **MCP Tools**

7. Serena for code analysis
8. Symbol navigation for test structure

9. File search for related patterns

10. **All Available Agents**

11. Research agents for documentation
12. Code analysis agents for patterns
13. Testing agents for strategy

Success Criteria

- ✅ 210/210 tests passing (100%)
- ✅ All browsers working (Chromium, Firefox, Webkit, Mobile Chrome, Mobile Safari)
- ✅ Zero skipped tests (unless environment-dependent)
- ✅ Stable test suite (no flakiness)
- ✅ Comprehensive documentation

Timeline

Target: Complete fix within this session

Document created: 2026-01-08 Status: Research phase initiated