



# Multi-Agent System Evaluation Report



## testcase1 - PASS

**Input Goal:** Find the next SpaceX launch, check weather at that location, then summarize if it may be delayed. **Timestamp:** 2025-06-14 17:01:44 **Runtime:** 1.27 seconds **Result:** json

```
{ "launch": { "name": "USSF-44", "launchpad_id":  
"5e9e4502f509094188566f88", "launch_time":  
"2022-11-01T13:41:00.000Z" }, "weather": { "coord": { "lon":  
-80.64, "lat": 28.57 }, "weather": [ { "id": 803, "main":  
"Clouds", "description": "broken clouds", "icon": "04d" } ],  
"base": "stations", "main": { "temp": 24.66, "feels_like": 25.56,  
"temp_min": 24.66, "temp_max": 24.66, "pressure": 1020,  
"humidity": 91, "sea_level": 1020, "grnd_level": 1019 },  
"visibility": 10000, "wind": { "speed": 3.55, "deg": 175, "gust":  
7.87 }, "clouds": { "all": 52 }, "dt": 1749900346, "sys":  
{ "country": "US", "sunrise": 1749896670, "sunset": 1749946884 },  
"timezone": -14400, "id": 4164092, "name": "Merritt Island",  
"cod": 200 }, "assessment": "\u2705 Launch is likely to proceed  
as planned." } Details: All checks passed.
```

---



## testcase2 - FAIL

**Input Goal:** Get next launch and check delay conditions using weather **Timestamp:** 2025-06-14 17:01:46 **Runtime:** 0.0 seconds **Result:** json {} **Details:** Missing key: launch, Missing key: weather, Missing key: assessment

---



## spacex-weather-delay-test - FAIL

**Input Goal:** Find the next SpaceX launch, check weather at that location, then summarize if it may be delayed. **Timestamp:** 2025-06-14 17:01:46 **Runtime:** 1.23 seconds **Result:** json

```
{ "launch": { "name": "USSF-44", "launchpad_id":  
"5e9e4502f509094188566f88", "launch_time":  
"2022-11-01T13:41:00.000Z" }, "weather": { "coord": { "lon":  
-80.64, "lat": 28.57 }, "weather": [ { "id": 803, "main":  
"Clouds", "description": "broken clouds", "icon": "04d" } ],  
"base": "stations", "main": { "temp": 24.66, "feels_like": 25.56,  
"temp_min": 24.66, "temp_max": 24.66, "pressure": 1020,  
"humidity": 91, "sea_level": 1020, "grnd_level": 1019 },  
"visibility": 10000, "wind": { "speed": 3.55, "deg": 175, "gust":  
7.87 }, "clouds": { "all": 52 }, "dt": 1749900346, "sys":  
{ "country": "US", "sunrise": 1749896670, "sunset": 1749946884 },  
"timezone": -14400, "id": 4164092, "name": "Merritt Island",  
"cod": 200 }, "assessment": "\u2705 Launch is likely to proceed
```

as planned." } **Details:** Missing key: launchlocation, *Missing key: weatherforecast*,  
Unexpected assessment: ☒ Launch is likely to proceed as planned.

---