

BenchmarkReport

September 6, 2026

BenchmarkReport

Kind: Interface

Source: `atloria-monorepo/libs/agent-core/src/benchmark/types.ts`

Summary of a complete benchmark run.

Properties

Property	Type
<code>suite</code>	<code>BenchmarkTask['suite']</code>
<code>model</code>	<code>string</code>
<code>timestamp</code>	<code>string</code>
<code>totalTasks</code>	<code>number</code>
<code>passedTasks</code>	<code>number</code>
<code>passRate</code>	<code>number</code>
<code>totalDurationMs</code>	<code>number</code>
<code>totalTokens</code>	<code>{ input: number; output: number; total: number }</code>
<code>totalCostUsd</code>	<code>number</code>
<code>avgCostPerTask</code>	<code>number</code>
<code>byDifficulty</code>	<code>Record<string, { total: number; passed: number; rate: number }></code>
<code>byLanguage</code>	<code>Record<string, { total: number; passed: number; rate: number }></code>
<code>results</code>	<code>TaskResult[]</code>
<code>toolStats</code>	<code>Record<string, { count: number; avgDurationMs: number }></code>
<code>failureBreakdown</code>	<code>Record<string, number></code>

Property	Type
byRepo	Record<string, { total: number; passed: number; rate: number; topFailure?: string }>
cumulativePassAtK	Record<number, { passed: number; rate: number }>
marginalGains	Record<string, number>
passAtK	{ k: number; passed: number; rate: number }
passAtKResults	PassAtKResult[]
costEfficiency	`{

```
firstPassSuccess: { avgTokens: number; avgDurationMs: number; avgTurns: number; count: number }
recoveredSuccess: { avgTokens: number; avgDurationMs: number; avgTurns: number; count: number }
finalFailure: { avgTokens: number; avgDurationMs: number; avgTurns: number; count: number };
```

```
} | | fileLocalizationStats | {
avgRecall: number;
avgPrecision: number;
/** Tasks with >0 overlap that still failed /
correctAreaButFailed: number;
/* Tasks with 0 overlap */
zeroOverlap: number;
}` |
```