# LLM 自动化评估结果

### 功能需求评估表

根据您提供的系统概述和需求文档，我首先执行了以下步骤：

- \*\*统计数据实体\*\*：从功能需求和系统概述中提取系统内需要管理的核心数据实体（不包括外部接口）。数据实体包括：Email（邮件）、Event（事件）、Task（任务）、Custom Phrase（自定义短语）、Folder/Label（文件夹/标签）、Category（分类）、Configuration Data（配置数据）。\*\*总计7个数据实体\*\*。

- \*\*合并功能需求\*\*：按照您的规则（合并同一数据实体的相同功能点，剔除系统通用设置类需求）进行合并。原始功能需求共12条（FR-1至FR-12），合并后剔除0条（无通用设置类需求），最终功能需求为9条。合并细节：

- FR-3、FR-4、FR-5 合并为“管理自定义短语”（涉及同一实体 Custom Phrase）。

- FR-11 和 FR-12 合并为“管理配置数据”（涉及同一实体 Configuration Data）。

- 其他需求未合并（功能点不同或涉及不同实体）。

- \*\*判断幻觉需求\*\*：基于定义（需求引入新操作涉及的数据实体无法与现有实体关联，导致E-R图不弱连接）。所有需求均关联现有实体（E-R图弱连接：Email ↔ Event ↔ Task ↔ Custom Phrase ↔ Folder/Label ↔ Category ↔ Configuration Data），因此无幻觉需求。

- \*\*判断完整引用需求\*\*：基于定义（需求未引用未在文档中定义的功能、输入或输出）。评判不严格，只要在功能需求或外部接口中提及即视为完整引用。所有需求在文档中有明确定义或引用。

- \*\*判断可测试需求\*\*：基于定义（输入能达成所需输出，依赖外部接口或用户交互也视为可测试）。评判不严格，所有需求均可通过测试用例验证（如模拟输入/输出、UI操作）。

以下是完整的评估表。每条需求基于合并后的列表进行评判（幻觉需求为“否”时，才评估完整引用和可测试性）。评判理由基于需求描述、系统概述和约束。

| 功能需求描述 (基于合并后) | 幻觉需求? | 完整引用? | 可测试? | 评判理由 |

|-------------------------|-----------|-----------|---------|----------|

| FR-1: 检测邮件内容中的事件类型（如会议请求、截止日期） | 否 | 是 | 是 | 涉及现有实体 Email 和 Event（系统概述提及邮件和事件）。输入（邮件内容）和输出（事件类型）在 FR-1 定义，无未引用元素。输入可通过测试邮件验证输出。 |

| FR-2: 从邮件内容提取事件细节（日期、时间、参与者等） | 否 | 是 | 是 | 涉及现有实体 Email 和 Event（概述中事件管理）。输入（邮件内容）和输出（细节）在 FR-2 定义。输入可通过测试邮件提取细节。 |

| Combined FR-3,4,5: 用户定义和管理自定义短语（最多50条/用户，最长50字符），通过配置模块编辑/删除 | 否 | 是 | 是 | 涉及现有实体 Custom Phrase（FR-3定义）。输入（用户操作）和输出（短语管理）在 FR-3,4,5 定义。可通过UI测试定义、编辑操作。 |

| FR-6: 用户创建、编辑和删除任务和事件 | 否 | 是 | 是 | 涉及现有实体 Task 和 Event（概述中任务和日历管理）。输入（用户操作）和输出（任务/事件操作）在 FR-6 定义。可通过界面测试创建、删除功能。 |

| FR-7: 提供统一仪表板，按日期和优先级显示任务和事件 | 否 | 是 | 是 | 涉及现有实体 Task 和 Event。输入（任务/事件数据）和输出（显示视图）在 FR-7 定义。可通过模拟数据验证显示逻辑。依赖用户交互但可测试。 |

| FR-8: 自动创建并应用至少三个文件夹/标签（基于事件类型，2秒内） | 否 | 是 | 是 | 涉及现有实体 Event 和 Folder/Label（FR-8 引入 Folder/Label，但关联 Event）。输入（事件检测）和输出（文件夹应用）在 FR-8 定义。可通过性能测试验证2秒内创建。约束中邮件客户端定义，完整引用。 |

| FR-9: 用户手动分配邮件到文件夹/标签 | 否 | 是 | 是 | 涉及现有实体 Email 和 Folder/Label。输入（用户操作）和输出（分配结果）在 FR-9 定义。可通过UI测试分配功能。 |

| FR-10: 用户标记邮件为预定义分类（"Urgent"等）或自定义分类（最多5个） | 否 | 是 | 是 | 涉及现有实体 Email 和 Category（FR-10 引入 Category，但关联 Email）。输入（用户操作）和输出（分类标记）在 FR-10 定义。可通过测试标记和验证分类应用。 |

| Combined FR-11,12: 提供配置模块管理自定义规则，数据以JSON存储并支持导入/导出 | 否 | 是 | 是 | 涉及现有实体 Configuration Data。输入（用户配置）和输出（JSON存储）在 FR-11,12 定义。可通过测试导入/导出功能验证数据完整性。 |

\*\*总结\*\*：

- 整个文档涉及的数据实体有\*\*7个\*\*（Email, Event, Task, Custom Phrase, Folder/Label, Category, Configuration Data）。

- 经过相同功能点合并后，功能需求共\*\*9条\*\*（合并前12条，合并剔除0条通用类需求）。

- 其中，幻觉需求有\*\*0条\*\*，非幻觉需求有\*\*9条\*\*。

- 非幻觉需求中，完整引用的功能需求有\*\*9条\*\*。

- 非幻觉需求中，可测试的功能需求有\*\*9条\*\*。

\*\*评估说明\*\*：

- \*\*幻觉需求判断\*\*：所有需求均关联核心系统实体（E-R图弱连接）。例如，FR-8 引入 Folder/Label，但基于 Event 操作；FR-10 引入 Category，但用于 Email。无孤立实体。

- \*\*完整引用判断\*\*：所有需求的输入/输出在文档中定义（功能需求或约束）。例如，外部接口（如邮件客户端）在约束中定义，视为完整引用。

- \*\*可测试判断\*\*：所有需求可设计测试用例（如模拟输入、自动化测试）。例如，FR-8 的“2秒内”可性能测试；用户交互需求（如FR-9）可通过UI测试。

- \*\*严格评判\*\*：我稍严格评估（如FR-7输出是视图，但可测试显示逻辑；FR-8依赖外部客户端，但约束支持测试）。所有需求均符合标准，无问题需求。