

10개 시나리오 전체 요약

#	이름	Gap	예상	Test Input
01	자동 신뢰도 해결	0.20	자동 해결	프로젝트 문서 작성
02	동등 신뢰도 충돌	0.00	충돌 감지	회의 준비 및 안건 정리
03	Keyword 우세	0.23	Keyword 승	팀과 월간 회의 진행
04	PARA 우세	0.28	PARA 승	프로젝트 마감일 계획
05	신뢰도 차이 큼	0.50	명확한 승자	연도별 전략 수립 및 검토
06	신뢰도 낮음	-	Review 필요	기타 활동
07	Review 필요	-	Review 플래그	프로젝트 회의 및 검토
08	메타데이터 충돌	-	충돌 감지	회사 정책 변경 검토
09	경계값 (0.2)	0.20	경계값 테스트	분기별 재정 검토 보고서
10	극단값	-	극단값 처리	기타

실행 명령어

```
python -m tests.test_conflict_cases_scenarios.runner
✓ ModelConfig loaded from backend.config
```

실행 결과물

```
=====
=====
🚀 충돌 해결 테스트 스위트 시작!
=====

[1/10] Scenario 01: 자동 신뢰도 해결 실행...
=====
=====
📝 Scenario 1: 자동 신뢰도 해결 (명확한 승자)
=====
=====
Input: 프로젝트 문서 작성
=====

=====
🔍 원본 LLM 응답:
=====
```

```
=====
{
  "tags": ["업무"],
  "confidence": 0.70,
  "matched_keywords": {
    "업무": ["프로젝트"]
  },
  "reasoning": "프로젝트 문서 작성은 업무 관련 활동으로 명확히 분류됨",
  "para_hints": {
    "업무": ["Projects"]
  }
}
=====
```

추출된 JSON:

```
{
  "tags": ["업무"],
  "confidence": 0.70,
  "matched_keywords": {
    "업무": ["프로젝트"]
  },
  "reasoning": "프로젝트 문서 작성은 업무 관련 활동으로 명확히 분류됨",
  "para_hints": {
    "업무": ["Projects"]
  }
}
```

분류 결과:

- Category: Projects
- Confidence: 0.9
- Conflict: False
- Snapshot ID: Snapshot(id='snap_20251103_221755', timestamp=datetime.datetime(2025, 11, 3, 22, 17, 55, 287569), text='프로젝트 문서 작성', para_result={'category': 'Projects', 'confidence': 0.9, 'reasoning': '프로젝트 문서 작성은 명확한 목표(문서 작성)와 함께 진행될 것으로 보이며, 특정 기한이 암시될 수 있어 Projects로 분류.', 'detected_cues': ['프로젝트', '문서', '작성'], 'source': 'langchain', 'has_metadata': False}, keyword_result={'tags': ['업무'], 'confidence': 0.7, 'matched_keywords': {'업무': ['프로젝트']}}, 'reasoning': '프로젝트 문서 작성은 업무 관련 활동으로 명확히 분류됨', 'para_hints': {'업무': ['Projects']}}, conflict_result={'final_category': 'Projects', 'para_category': 'Projects', 'keyword_tags': ['업무'], 'confidence': 0.9, 'confidence_gap': 0.2, 'conflict_detected': False, 'resolution_method': 'auto_by_confidence', 'requires_review': False, 'winner_source': 'para', 'para_reasoning': '프로젝트 문서 작성은 명확한 목표(문서 작성)와 함께 진행될 것으로 보이며, 특정 기한이 암시될 수 있어 Projects로 분류.', 'reason': '명확한 승자 선택됨 (Gap: 0.20)'}, metadata={'confidence': 0, 'is_conflict': False, 'final_category': 'Projects'})

분류 결과 저장 완료: file_id=3, snapshot_id=snap_20251103_221755

Scenario 1 PASS!

- Category: Projects
- Confidence: 0.9
- No Conflict Detected

- DB Saved: file_id=3

[2/10] Scenario 02: 동등 신뢰도 충돌 실행...

Scenario 2: 동등 신뢰도 충돌 (User Review 필요)

Input: 회의 준비 및 안건 정리

원본 LLM 응답:

```
{  
    "tags": ["업무"],  
    "confidence": 0.75,  
    "matched_keywords": {  
        "업무": ["회의", "안건"]  
    },  
    "reasoning": "업무 관련 회의 준비와 안건 정리에 해당하는 키워드가 감지됨",  
    "para_hints": {  
        "업무": ["Projects"]  
    }  
}
```

추출된 JSON:

```
{  
    "tags": ["업무"],  
    "confidence": 0.75,  
    "matched_keywords": {  
        "업무": ["회의", "안건"]  
    },  
    "reasoning": "업무 관련 회의 준비와 안건 정리에 해당하는 키워드가 감지됨",  
    "para_hints": {  
        "업무": ["Projects"]  
    }  
}
```

분류 결과:

- Category: Areas
- Confidence: 0.8
- Conflict Detected: True
- Requires Review: True

분류 결과 저장 완료: file_id=4, snapshot_id=snap_20251103_221800

Scenario 2 PASS!

- Conflict Detected: True
- Requires Review: True

- DB Saved: file_id=4

[3/10] Scenario 03: Keyword 우세 실행...

=====

=====

✓ Scenario 3: Keyword 우세 분류 (Gap > 0.2)

=====

=====

Input: 팀과 월간 회의 진행

=====

=====

🔍 원본 LLM 응답:

=====

=====

```
{  
    "tags": ["업무"],  
    "confidence": 0.75,  
    "matched_keywords": {  
        "업무": ["회의"]  
    },  
    "reasoning": "팀과 월간 회의 진행이라는 업무 관련 키워드가 감지됨",  
    "para_hints": {  
        "업무": ["Projects"]  
    }  
}
```

=====

=====

📄 추출된 JSON:

```
{  
    "tags": ["업무"],  
    "confidence": 0.75,  
    "matched_keywords": {  
        "업무": ["회의"]  
    },  
    "reasoning": "팀과 월간 회의 진행이라는 업무 관련 키워드가 감지됨",  
    "para_hints": {  
        "업무": ["Projects"]  
    }  
}
```

📊 분류 결과:

- Category: Areas
- Confidence: 0.8
- Conflict: True

✖ Scenario 03 FAILED: 충돌이 감지되면 안 됨

[4/10] Scenario 04: PARA 우세 실행...

=====

💡 Scenario 4: PARA 우세 분류 (Gap > 0.2)

=====

=====

Input: 프로젝트 마감일 계획

=====

=====

🔍 원본 LLM 응답:

=====

```
{  
    "tags": ["업무"],  
    "confidence": 0.70,  
    "matched_keywords": {  
        "업무": ["프로젝트", "마감일", "계획"]  
    },  
    "reasoning": "프로젝트와 관련된 마감일 계획이 명확하게 업무 카테고리에 해당됨",  
    "para_hints": {  
        "업무": ["Projects"]  
    }  
}
```

=====

=====

📄 추출된 JSON:

```
{  
    "tags": ["업무"],  
    "confidence": 0.70,  
    "matched_keywords": {  
        "업무": ["프로젝트", "마감일", "계획"]  
    },  
    "reasoning": "프로젝트와 관련된 마감일 계획이 명확하게 업무 카테고리에 해당됨",  
    "para_hints": {  
        "업무": ["Projects"]  
    }  
}
```

=====

📊 분류 결과:

- Category: Projects
- Confidence: 0.9
- Conflict: False

✅ 분류 결과 저장 완료: file_id=5, snapshot_id=snap_20251103_221810

✅ Scenario 4 PASS!

- PARA Dominant: Projects ✓
- Confidence: 0.9 ✓
- DB Saved: file_id=5

[5/10] Scenario 05: 신뢰도 차이 큼 실행...

=====

📝 Scenario 5: 신뢰도 차이 큼 (Clear Winner)

=====

=====

Input: 연도별 전략 수립 및 검토

=====

=====

🔍 원본 LLM 응답:

=====

=====

```
{  
    "tags": ["업무"],  
    "confidence": 0.70,  
    "matched_keywords": {  
        "업무": ["전략"]  
    },  
    "reasoning": "전략 수립 및 검토는 업무 관련 활동으로 분류됨",  
    "para_hints": {  
        "업무": ["Projects"]  
    }  
}
```

=====

=====

📄 추출된 JSON:

```
{  
    "tags": ["업무"],  
    "confidence": 0.70,  
    "matched_keywords": {  
        "업무": ["전략"]  
    },  
    "reasoning": "전략 수립 및 검토는 업무 관련 활동으로 분류됨",  
    "para_hints": {  
        "업무": ["Projects"]  
    }  
}
```

📊 분류 결과:

- Category: Areas
- Confidence: 0.8
- Conflict: True

✖ Scenario 05 FAILED: 신뢰도가 매우 높아야 함: 0.8

[6/10] Scenario 06: 신뢰도 낮음 실행...

=====

=====

📝 Scenario 6: 신뢰도 낮음 (판단 어려움)

=====

=====

Input: 기타 활동

=====

=====

🔍 원본 LLM 응답:

=====

=====

{

 "tags": ["기타"],
 "confidence": 0.30,
 "matched_keywords": {},
 "reasoning": "명확한 키워드가 감지되지 않음",
 "para_hints": {
 "기타": ["Resources"]
 }
}

=====

=====

📄 추출된 JSON:

{

 "tags": ["기타"],
 "confidence": 0.30,
 "matched_keywords": {},
 "reasoning": "명확한 키워드가 감지되지 않음",
 "para_hints": {
 "기타": ["Resources"]
 }
}

📊 분류 결과:

- Category: Areas
- Confidence: 0.8
- Requires Review: False

✖ Scenario 06 FAILED: 신뢰도가 높아야 함: 0.8

[7/10] Scenario 07: Review 필요 실행...

=====

=====

✍ Scenario 7: Review 필요 (requires_review=True)

=====

=====

Input: 프로젝트 회의 및 검토

=====

=====

=====

🔍 원본 LLM 응답:

=====

=====

{

 "tags": ["업무"],
 "confidence": 0.70,

```

"matched_keywords": {
    "업무": ["회의", "프로젝트"]
},
"reasoning": "프로젝트 회의와 검토 관련 키워드가 감지되어 업무 카테고리에 해당됨",
"para_hints": {
    "업무": ["Projects"]
}
}
=====
```

추출된 JSON:

```
{
    "tags": ["업무"],
    "confidence": 0.70,
    "matched_keywords": {
        "업무": ["회의", "프로젝트"]
    },
    "reasoning": "프로젝트 회의와 검토 관련 키워드가 감지되어 업무 카테고리에 해당됨",
    "para_hints": {
        "업무": ["Projects"]
    }
}
```

분류 결과:

- Category: Areas
- Confidence: 0.8
- Requires Review: True

분류 결과 저장 완료: file_id=6, snapshot_id=snap_20251103_221828

Scenario 7 PASS!

- Requires Review: True
- DB Saved: file_id=6

[8/10] Scenario 08: 메타데이터 충돌 실행...

Scenario 8: 메타데이터 충돌 (다중 정보 상충)

Input: 회사 정책 변경 검토

원본 LLM 응답:

```
{
    "tags": ["업무"],
    "confidence": 0.70,
    "matched_keywords": {
        "업무": ["정책", "검토"]
    }
}
```

```
{
  "reasoning": "회사 정책 변경과 관련된 업무 활동이 명확하게 나타남",
  "para_hints": {
    "업무": ["Projects"]
  }
}
=====
=====
```

추출된 JSON:

```
{
  "tags": ["업무"],
  "confidence": 0.70,
  "matched_keywords": {
    "업무": ["정책", "검토"]
  },
  "reasoning": "회사 정책 변경과 관련된 업무 활동이 명확하게 나타남",
  "para_hints": {
    "업무": ["Projects"]
  }
}
```

분류 결과:

- Category: Areas
- Confidence: 0.8
- Conflict Detected: True

분류 결과 저장 완료: file_id=7, snapshot_id=snap_20251103_221834

Scenario 8 PASS!

- Category: Areas
- DB Saved: file_id=7

[9/10] Scenario 09: 경계값 테스트 실행...

Scenario 9: 경계값 테스트 (Gap = 0.2)

Input: 분기별 재정 검토 보고서

원본 LLM 응답:

```
{
  "tags": ["재무"],
  "confidence": 0.70,
  "matched_keywords": {
    "재무": ["재정"]
  },
  "reasoning": "재정 관련 키워드가 감지되어 재무 카테고리에 해당됨",
```

```
"para_hints": {  
    "재무": ["Archives"]  
}  
}  
=====
```

📄 추출된 JSON:

```
{  
    "tags": ["재무"],  
    "confidence": 0.70,  
    "matched_keywords": {  
        "재무": ["재정"]  
    },  
    "reasoning": "재정 관련 키워드가 감지되어 재무 카테고리에 해당됨",  
    "para_hints": {  
        "재무": ["Archives"]  
    }  
}
```

📊 분류 결과:

- Category: Archives
- Confidence: 0.9
- Conflict: False

✓ 분류 결과 저장 완료: file_id=8, snapshot_id=snap_20251103_221844

✓ Scenario 9 PASS!

- Boundary Test: Success ✓
- DB Saved: file_id=8

[10/10] Scenario 10: 극단값 테스트 실행...

=====

=====

✍ Scenario 10: 극단값 테스트 (매우 모호한 입력)

=====

Input: 기타

=====

=====

🔍 원본 LLM 응답:

=====

=====

```
{  
    "tags": ["기타"],  
    "confidence": 0.30,  
    "matched_keywords": {},  
    "reasoning": "명확한 키워드가 감지되지 않음",  
    "para_hints": {  
        "기타": ["Resources"]  
    }  
}
```

```
=====
```

```
=====
```

📄 추출된 JSON:

```
{  
  "tags": ["기타"],  
  "confidence": 0.30,  
  "matched_keywords": {},  
  "reasoning": "명확한 키워드가 감지되지 않음",  
  "para_hints": {  
    "기타": ["Resources"]  
  }  
}
```

📊 분류 결과:

- Category: Archives
- Confidence: 0.9
- Requires Review: False

✓ 분류 결과 저장 완료: file_id=9, snapshot_id=snap_20251103_221849

✓ Scenario 10 PASS!

- Extreme Case Handled: Success ✓
- DB Saved: file_id=9

```
=====
```

```
=====
```

📊 테스트 결과:

- ✓ Passed: 7/10
- ✗ Failed: 3/10

```
=====
```

```
=====
```

⚠️ 3개 시나리오 확인 필요