

Langchain의 출력파서(Output Parser)는 LLM의 출력을 더 유용하고 구조화된 형태로 변환하는 컴포넌트입니다.

출력파서의 역할

- LLM의 출력을 받아 더 적합한 형식으로 변환
 - LLM의 출력을 원하는 형태 변환, 원하는 정보만 추출
- 구조화된 데이터 생성에 매우 유용
- LlangChain 프레임워크에서 다양한 종류의 출력 데이터를 파싱하고 처리
- 애플리케이션 개발할때 자동화를 위해 정형화된 데이터가 필요한데, 스키마에 맞는 데이터를 맞춰주는 역할

출력파서의 이점

- 구조화: LLM의 자유 형식 텍스트 출력을 구조화된 데이터로 변환, 자동화에 이점이 있습니다.
- 일관성: 출력 형식을 일관되게 유지하여 후속처리를 용이하게 합니다.
- 유연성: 다양한 출력 형식(JSON, LIST, DICTIONARY)으로 변환이 가능합니다.

출력파서를 사용했을때와 안했을때 비교

이메일 내용

date : 2024-11-19

안녕하세요. 이번 블록체인 컨트랙트 오딧 결과, 여러 보안 취약점이 발견되었습니다. 첫 번째로, 특정 함수가 관리자 권한 없이 호출될 수 있는 접근 제어 문제가 존재하여 시스템이 권한 없는 접근에 노출될 위험이 있습니다. 두 번째로, 사용자 민감 데이터가 로그에 기록되어 외부에서 해당 데이터에 접근할 가능성이 있는 민감한 데이터 노출 문제가 확인되었습니다. 마지막으로, 특정 스마트 컨트랙트가 재진입 공격에 취약하여 자금 손실과 같은 심각한 보안 사고가 발생할 수 있는 위험이 있습니다. 이러한 문제들은 모두 보안에 심각한 영향을 미칠 수 있으므로 빠른 수정을 권장합니다. 추가 질문이 있으신 경우 언제든지 연락 주시기 바랍니다. 감사합니다.

"중요한 내용을 정리해주세요." LLM에게 일반 요청

발견된 취약점

1. 접근 제어 문제

- ****설명****: 특정 함수가 관리자 권한 없이 호출될 수 있는 취약점.
- ****영향****: 권한 없는 접근으로 시스템이 악용될 위험이 있음.

- ****권장 조치****: 접근 제어를 강화하고 권한 검증 로직을 추가.

2. 민감한 데이터 노출

- ****설명****: 사용자 민감 데이터가 로그에 기록되어 외부에서 접근 가능성이 있음.
- ****영향****: 민감 정보 유출 가능성으로 인해 사용자 신뢰 손상 및 법적 문제가 발생할 위험이 있음.
- ****권장 조치****: 민감 데이터 로그 기록을 제거하거나 암호화 처리.

3. 재진입 공격 가능성

- ****설명****: 특정 스마트 컨트랙트가 재진입 공격에 취약.
- ****영향****: 자금 손실 등 심각한 보안 사고 발생 가능성.
- ****권장 조치****: 재진입 방지 로직 추가 및 보안 강화.

요약

- 발견된 주요 이슈: 접근 제어 문제, 민감 데이터 노출, 재진입 공격 가능성.
- 영향: 시스템 악용, 데이터 유출, 자금 손실 가능성 등 보안에 심각한 위험 초래.
- 권장 조치: 문제별 빠른 수정 및 보안 강화.

추가 정보

- 발신자: 오딧 팀
- 연락처: 이메일(june@audit.com)

LLM에게 JSON 스키마를 사용한 요청

JSON 스키마 예시:

```
{  
    "subject": "이메일 제목",  
    "email": "발신자 이메일",  
    "summary": "이슈 내용 요약",  
    "date": "이슈 전달 날짜 (YYYY-MM-DD)"  
}
```

LLM 응답:

```
{
  "subject": "블록체인 컨트랙트 오딧 결과 보고",
  "email": "june@audit.com",
  "summary": "이번 오딧에서 주요 취약점으로 접근 제어 문제, 민감한 데이터 노출, 재진입 공격 취약성이 발견되었습니다. 빠른 수정이 요구됩니다.",
  "date": "2024-11-19"
}
```

차이점 분석

항목	일반 요청	JSON 요청
구조화 여부	자유 형식의 비구조화된 텍스트로 제공. 사람이 읽기 쉽지만 자동화에 부적합.	사전 정의된 스키마를 따르는 구조화된 데이터로 제공. 기계적 처리 및 데이터베이스 저장에 적합.
요약 정보의 명확성	이슈 내용은 포함되지만 요약의 형식과 표현이 일관되지 않을 수 있음.	스키마 필드별로 정보가 구분되므로 요약의 명확성이 보장됨.
활용 용도	이메일 회신, 보고서 작성 등 수작업이 필요한 상황에 적합.	시스템 통합, 자동화된 보고서 생성, 데이터 분석 등 기계적 처리에 적합.
추가 정보의 표현 가능성	발신자 정보나 날짜가 명시적으로 표시되지 않음.	발신자 이메일, 이슈 전달 날짜 등 부가 정보를 스키마 필드로 추가하여 명시적으로 표현.
사용 사례	사람이 직접 읽고 이해하거나 회의 등에서 발표하는 데 적합.	데이터베이스 저장, 자동화된 알림 시스템, 분석 및 리포트 생성에 적합.

결론

- 일반 요청은 빠르게 내용을 요약하고 사람이 읽는 데 적합한 결과를 제공합니다.
- JSON 요청은 구조화된 결과를 생성하여 자동화 시스템 및 데이터 분석 워크플로우에 최적화된 형식을 제공합니다.