

범죄수사 데이터 기반 스토리라인 시각화 인터페이스에 관한 연구

이지현, 정종진
한국전자기술연구원 정보미디어연구센터
leejihyun@keti.re.kr, mozzalt@keti.re.kr

A Study on the Visualization Interface of Storyline Based on Criminal Investigation Data

Ji Hyun Lee, Jong Jin Jung
Information Media Research Center, Korea Electronics Technology Institute

요 약

본 연구는 입력된 수사문건에 대해 시간 속성을 기반으로 범죄사건의 흐름을 나타내는 스토리라인을 자동으로 시각화하여 사건의 전체적인 흐름에 대한 이해도를 향상하는 데 도움을 주는 범죄사실에 특화된 스토리라인 시각화 인터페이스에 관한 연구이다. 이를 위해 객관적인 사실 관계 파악과 논증 검증이 중요한 범죄수사에서 중요한 자료인 판결문 데이터를 활용하였으며, 판결문 중에서도 육하원칙에 따라 해당 범죄의 구성요건적인 특징이 명확히 기재되어 있는 범죄사실을 대상으로 하여 사건의 흐름에 따라 추출된 데이터를 자동으로 시각화하였다. 이는 현장에서 실무자들이 범죄사건의 흐름을 쉽고 빠르게 파악하고 더 나아가 종합적이고 입체적으로 파악하는 데 도움이 될 것으로 기대된다.

1. 서론

현장 실무자들이 범죄사건의 전체적인 흐름을 파악하기 위해 확보된 자료를 일일이 분석하여 시간 흐름의 순서대로 사건의 주요 인물, 주요 행위가 일어난 시점, 수사의 쟁점 등 사건의 전반적인 구성을 파악하는데[1] 이는 인력과 시간이 많이 소요될 수밖에 없다. 이를 개선하고자 자동으로 구성되는 스토리라인 시각화 인터페이스를 제공하고자 한다. 이를 위해 본 연구에서 활용한 학습데이터는 객관적인 사실 관계 파악과 논증 검증이 중요한 범죄수사에서 유사 범죄 검색, 쟁점 분석 등 수사 지원 작업 시 중요한 역할을 하여 법률 및 수사 분야 인공지능 영역에서 학습데이터로서의 가치가 높아지고 있는 판결문을 활용하였다. 특히 판결문 중 범죄사실은 육하원칙에 따라 해당 범죄의 구성요건적인 특징이 명확히 기재되어 있기 때문에 사건의 전반적인 흐름을 파악하거나 쟁점을 유추하는 등에 도움이 될 수 있다.[2] 따라서 본 논문에서는 판결문의 범죄사실을 중심으로 사건의 흐름에 따라 추출된 데이터를 자동으로 시각화하여 사건의 전체적인 흐름에 대한 이해도를 향상할 수 있는 범죄사실에 특화된 스토리라인 시각화 인터페이스를 제안하고자 한다.

2. 본론

본 연구에서는 한림대학교 정보법과학연구소의 범죄사실 정보 추출 가이드라인을 참고하여 추출한 데이터를 스토리라인 시각화 데이터로 활용하였다. 분석대상이 되는 판결문 내 범죄사실과 범죄사실 정보 추출 작업에 대한 정의는 아래의 표와 같다.[3]

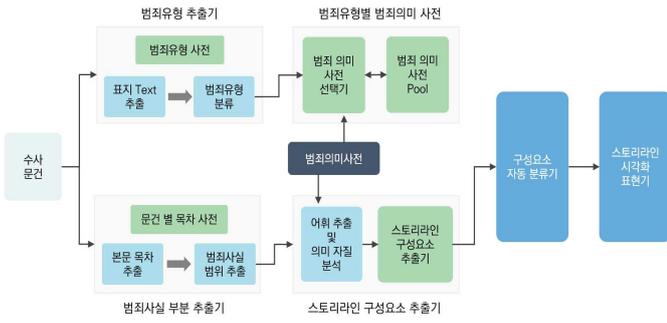
<표 1> 용어 정의

범죄사실	범죄 구성요건에 맞추어 법률적으로 재구성한 사실
범죄사실 정보 추출	판결문 데이터의 범죄사실 부분에서 사건 규명에 필수적인 정보를 추출하는 작업

수사문건을 입력하여 스토리라인 시각화를 하기까지의 시스템 동작 시나리오는 하단 그림 1과 같다. 이 중, 스토리라인 시각화에 필요한 데이터를 추출하는 모듈과 시각화 모듈을 살펴본다.[1]

2.1. 범죄사실 범위 추출

범죄사실 정보를 추출하기 위해서는 입력된 수사문건(본 연구에서는 판결문에 한함) 내에서 범죄사실 범위가 어디인지 파악해야 한다. 이를 위해 입력된 수사문건의 목차를 활용하여 범죄사실 범위를 파



(그림 1) 스토리라인 시각화 시스템 흐름도

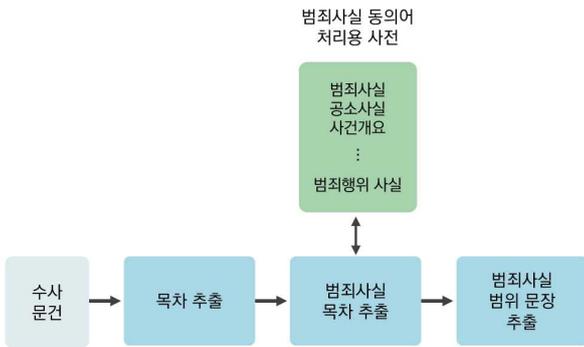
id	범죄사실	도구	제안방법	공격범위 및 횟수	공격행위 및 수단	의의	상해부위
CA58000511201207500132	범죄사실	인	인	인	인	인	인
CA58000511201307500008	범죄사실	인	인	인	인	인	인
CA58000511201707500109	범죄사실	인	인	인	인	인	인
CA58000511201507500019	범죄사실	인	인	인	인	인	인
CA58000511201907500005	범죄사실	인	인	인	인	인	인
CA580005112021075000126	범죄사실	인	인	인	인	인	인
CA580005112004075000107	범죄사실	인	인	인	인	인	인
CA580005112013075000154	범죄사실	인	인	인	인	인	인
CA580005112004075000095	범죄사실	인	인	인	인	인	인
CA580005112019075000082	범죄사실	인	인	인	인	인	인
CA580005112012075000181	범죄사실	인	인	인	인	인	인
CA580005112013075000090	범죄사실	인	인	인	인	인	인
CA5800051120017075000115	범죄사실	인	인	인	인	인	인
CA580005112015075000057	범죄사실	인	인	인	인	인	인
CA580005112019075000126	범죄사실	인	인	인	인	인	인
CA5800051120012075000145	범죄사실	인	인	인	인	인	인
CA580005112004075000109	범죄사실	인	인	인	인	인	인
CA580005112013075000156	범죄사실	인	인	인	인	인	인
CA580005112004075000009	범죄사실	인	인	인	인	인	인
CA580005112015075000181	범죄사실	인	인	인	인	인	인
CA580005112012075000047	범죄사실	인	인	인	인	인	인
CA580005112014075000133	범죄사실	인	인	인	인	인	인
CA580005112017075000117	범죄사실	인	인	인	인	인	인
CA580005112015075000082	범죄사실	인	인	인	인	인	인
CA580005112019075000012	범죄사실	인	인	인	인	인	인
CA5800051120012075000189	범죄사실	인	인	인	인	인	인
CA580005112004075000127	범죄사실	인	인	인	인	인	인
CA580005112013075000016	범죄사실	인	인	인	인	인	인

(그림 3) 범죄사실 정보 추출 예시

악한다. 수사문건의 목차는 정해진 규칙이 없이 범죄유형이나 작성자마다 조금씩 상이하지만 대표적인 몇 가지의 유형으로 분류할 수 있으며, 범죄사실을 기술하기 위해 주로 사용하는 목차 제목을 사전에 정의하여 맵핑테이블로 구성한 뒤 새로 입력된 문서에서 도출한 목차 제목을 동의어 처리하여 범죄사실 범위를 파악하여 추출한다.[1]

2.3. 범죄수사 데이터 기반 스토리라인 시각화

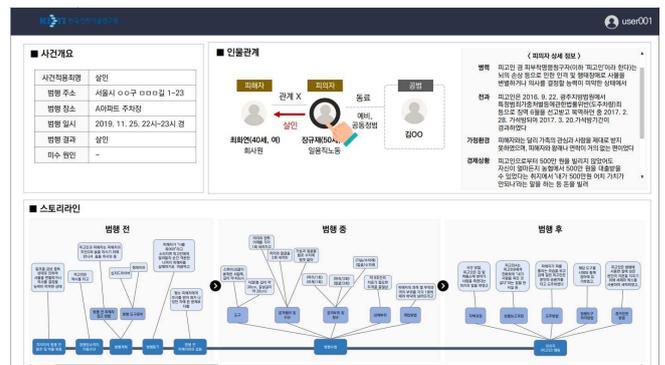
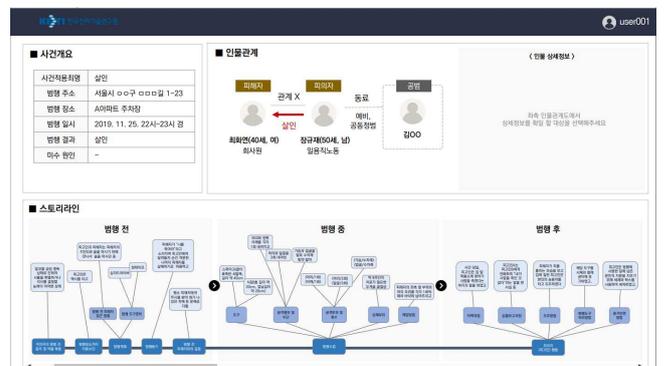
본 연구에서 제안하는 범죄수사 데이터 기반 스토리라인 시각화 인터페이스는 한림대학교 정보법과학연구소에서 정의한 40여 개의 범죄정보를 토대로 기본 정보, 인물 관계도, 인물별 상세정보, 범행 전/중/후의 정보를 시각화하여 범주의 구성요건적인 특징을 한눈에 쉽고 빠르게 확인할 수 있도록 설계하였다. 특히 범행 전/중/후 정보는 지식 요소들을 정의, 분석, 분류하여 그 지식 간의 관계를 표현하는[4] 지식맵(Knowledge Map) 구조를 응용하여 설계하였다. 이때 사용자의 인터랙션에 따라 그에 알맞은 정보나 화면을 제공하여 사용자의 의도에 따른 정보를 제공한다.



(그림 2) 범죄사실 범위 추출 과정 개념도

2.2. 범죄사실 정보 추출

앞서 언급한 것과 같이 본 연구에서는 범죄수사 데이터 스토리라인 시각화를 위해 한림대학교 정보법과학연구소의 범죄사실 정보 추출 가이드라인을 참고하여 판결문에서 추출한 데이터를 활용하였다. 한림대학교 정보법과학연구소의 범죄사실 정보 추출 가이드[3]는 판결문 내 범죄사실에서 의미 있는 정보를 탐색하여 유사사건 검색, 사건의 시간 구성 시각화 등에 활용 가능한 범죄정보를 정의하였다. 해당 가이드는 총 4개 영역(기본정보, 범행 전/중/후)에 대해 3개 계층(대분류, 중분류, 소분류)의 분류 체계를 가지고 40여 개의 범죄사실 정보 추출 분류 항목을 정의하고 있다.[3] 이를 토대로 그림 3과 같이 실제 살인 및 살인미수 판결문을 대상으로 범죄사실 정보를 추출한다.



(그림 4) 범죄사실 데이터 기반 스토리라인 시각화 인터페이스 예시

3. 결론

본 연구에서는 판결문 내에서 육하원칙에 따라 해당 범죄의 구성요건적인 특징이 명확히 기재되어 있는 범죄사실에서 추출된 데이터를 기반으로 사건 구성에 대한 이해도를 향상시킬 수 있는 범죄수사에 특화된 스토리라인 시각화 인터페이스를 제안하였다. 이는 현장에서 실무자들이 범죄사건의 구조 및 흐름을 쉽고 빠르게 파악하고 더 나아가 종합적이고 입체적으로 파악하는 데 도움이 될 것으로 기대된다. 아울러 향후 연구에서는 다양한 최종의 판결문 데이터와 많은 종류의 수사문건을 토대로 딥러닝 모델을 학습시켜 범죄수사 데이터 스토리라인 시각화 수준을 고도화 및 세분화시키는 연구가 필요하다.

ACKNOWLEDGMENT

이 논문은 2023년도 정부(경찰청)의 재원으로 지원받아 수행된 연구결과임 [내역사업명: AI 기반 범죄수사 지원 / 연구개발과제번호: PR10-02-000-21]

참고문헌

- [1] 정종진, 박종빈, 박성주, “주요 범죄사실 개요 분석을 통한 범죄사실 타임라인 자동 작성 시스템 연구”, ACK 2021년 추계학술대회, 2021, vol.28, no.2, pp. 622-625.
- [2] 이유나, 박성미, 박노섭, “개체명 인식과 이벤트 추출을 통한 판결문 범죄사실 구성요소 및 스토리라인 시각화방안 연구”, ACK 2022년 추계학술대회, 2022, vol.29, no.2, pp. 490-492.
- [3] 한림대학교 정보법과학연구소, “범죄사실 정보 추출 가이드 v.1.5”, 2023.
- [4] 이지수, 정겨운, 이경원, “스토리텔링의 시각화를 위한 개념적 맵들의 특성분석”, 한국HCI학회 학술대회, 2008, pp. 1214-1219.