TTA Standard

정보통신단체표준(국문표준)

TTAS.KO-12.00XX 제정일: 2019년 12월 XX일

디지털 포렌식 조사를 위한 통합 정보 처리 규격 - 제1부 : 개요 및 요구사항 -

Data Expression Standard for Digital Forensic Investigation: Part 1. Overview and Requirements

TTA 한국정보통신기술협회
Telecommunications Technology Association

표준초안 검토 위원회 사이버보안 프로젝트그룹(PG503)

표준안 심의 위원회　정보보호 기술위원회(TC5)

	성명	소 속	직위	위원회 및 직위	표준번호
표준(과제) 제안	이상진	고려대학교	교수	사이버보안프로젝트그룹 특별위원	
	한재혁	고려대학교	연구원		
표준 초안 작성자	이상진	고려대학교	교수	사이버보안프로젝트그룹 특별위원	
	손태식	아주대학교	교수		
	이성주	㈜인정보	이사		
	박경해	㈜클루드인	이사		
	한재혁	고려대학교	연구원		
	윤우성	고려대학교	연구원		
	김지언	고려대학교	연구원		
사무국 담당	황예지	TTA			

본 표준 발간 이전에 접수된 지식재산권 확약서 정보는 본 표준의 '부록(지식재산권 확약서 정보)'에 명시하고 있으며, 이후 접수된 지식재산권 확약서는 TTA 웹사이트에서 확인할 수 있습니다.
본 표준과 관련하여 접수된 확약서 외의 지식재산권이 존재할 수 있습니다.

발행인 : 한국정보통신기술협회 회장
발행처 : 한국정보통신기술협회
13591, 경기도 성남시 분당구 분당로 47
Tel : 031-724-0114, Fax : 031-724-0109
발행일 : 2019.12

서 문

1 표준의 목적

디지털 포렌식 조사는 디지털 데이터로부터 사건의 단서나 증거와 같은 정보를 도출하기 위해 정보저장매체로부터 데이터를 수집하고 분석한다. 이러한 조사 과정에서 처리되는 데이터는 그 데이터를 생성한 응용프로그램이나 저장되는 환경에 따라 매번 다르게 표현되기 때문에 효율적인 분석이 어려우며, 일관된 분석을 위해서는 정형화된 형태로 가공해야 한다.
이 표준은 디지털 포렌식 조사에서 동일한 성격의 데이터를 일관된 형태로 분석하기 위해 분석용 데이터의 저장 규격을 정의하는 것이 목적이다. 총 3부로 구성되어 있으며, 1부에서는 데이터 처리 규격을 정의하기에 앞서 개요와 요구사항을 정의하며, 2부에서는 데이터 종류별로 처리 규격을 정의하고, 3부에서는 기존 환경(레거시)을 고려하여 상호 호환성을 위한 사용방법을 정의한다.

2 주요 내용 요약

이 표준은 디지털 포렌식 조사에서 사용되는 데이터 모델과 사용범위를 포함한다. 이 표준을 개발하는 과정에서 요구되는 사항을 정의하고 데이터를 표현하는 방법과 이를 위한 구성요소를 설명한다. 또한, 참고표준과 표현결과를 비교함으로써 차별성을 명확하게 한다.

3 인용 표준과의 비교

3.1 인용 표준과의 관련성

- 해당 사항 없음

3.2 인용 표준과 본 표준의 비교표

- 해당 사항 없음

Preface

1 Purpose

On the digital forensic investigation, it's need to collect and analyze digital data, which is usually recorded on the storage media, to derive information from digital evidence. In this process, data is expressed fairly differently depending on the analysis environment by applications, methods, or investigators. The reason why it is difficult to analyze efficiently, therefore it should be processed in a regular form.

The purpose of this standard is to define the specification of the data expression in order to analyze digital data of the same categorization in a digital forensic investigation. This standard consists of three parts. Part 1 describes the framework and defines requirements, Part 2 defines the data types and specification for each data type, and Part 3 describes the usage method for interoperability considering the existing environment (i.e. legacy)

2 Summary

This standard defines specifications of data expression for digital forensic investigation and suggests how to use it. This covers the data model and scope of usage in the digital forensic investigation. Part 1 includes the requirements of this standard, how the data expression is represented, and what components there are. This also clarifies a differentiation by comparing reference with other standards.

3 Relationship to Reference Standards

3.1 Relationship to Reference Standards

- None

3.2 Comparison between This Standard and Reference Standards

- None

목 차

디지털 포렌식 조사를 위한 통합 정보 처리 규격
- 제1부 : 개요 및 요구사항 -
(Data Expression Standard for Digital Forensic investigation: Part 1. Overview and Requirements)

1 적용 범위

디지털 포렌식 조사는 사이버 범죄 수사를 포함하여 디지털 기기와 관련된 모든 조사를 포괄하는 개념으로, 수사기관에서 용의자를 조사하거나 기업 감사와 같은 상황에서 디지털 포렌식 기술을 활용한 일련의 조사 과정을 의미한다.

이 표준은 디지털 포렌식 조사에 사용되는 데이터를 처리하기 위한 규격을 정의하고 표현하기 위해서 작성되었다. 이 규격은 디지털 포렌식 조사에서 그 필요성이 대두됨에 따라 개발되었으며, 사용범위를 포괄적으로 수용할 수 있고 시기적절하게 효율적인 데이터 교환 및 공유 등의 처리를 지원하다. 따라서 사이버 보안을 강화할 수 있으며, 정형화된 데이터 기반의 디지털 포렌식 분석 도구의 개발을 촉진시킬 수 있다.

이 표준은 데이터를 일관성 있게 처리하기 위한 규격을 제공하고 있을 뿐, 어떠한 의무를 강요하지 않으며 권고안으로 제안한다. 다만, 기존에 사용하는 규격과의 차이를 포괄하고 호환성을 보장하기 위해 사용방법과 유스케이스를 포함한다.

2 인용 표준

해당 사항 없음

3 용어 정의

3.1 디지털 증거 (digital evidence)

디지털 형태로 저장되거나 전송되는 데이터로 쟁점이 되는 사건을 입증하거나 반박하는 데 도움이 되는 정보

3.2 디지털 증거 정보

한 사건에서 수집한 각 정보저장매체 및 디지털 데이터에 대한 채증 자료, 증거 파일, 증거 파일 생성 정보

3.3 레거시 (legacy)

과거에 개발되어 현재에도 사용 중인 낡은 하드웨어나 소프트웨어 [출처: TTA 정보통신 용어사전]

3.4 분석 정보

수집한 정보저장매체 및 디지털 데이터에 대한 분석 결과물로서, 추출 파일, 분석 보고서 파일, 촬영 정보 등을 의미함

3.5 사건 정보

디지털 증거의 사건 정보로 사건 번호, 기관명, 기관 부서명, 이송 받는 기관명을 포함

3.6 아티팩트 (artifact)

운영체제나 응용프로그램을 사용하면서 자동으로 생성된 기록 또는 흔적. 예를 들어, 윈도우즈 운영체제가 설치된 시스템에는 레지스트리, 프리패치, 이벤트로그 등이 있다.

3.7 유스케이스 (use case)

시스템이 수행하는 일련의 조치사항(sequences of actions)들을 설명한 것. 시스템이 이들 조치사항을 수행함으로써 행위자에게 가시적인 결과를 제공한다. [출처: UML Guide]

3.8 이미지 파일 (image file)

특정 매체에 저장된 데이터 전체를 하나의 파일로 만든 것. 디스크, ROM, DVD 등 기록이 가능한 매체 내에 전체나 일부 데이터를 파일 형태로 저장된 것이며, 압축이나 암호화될 수 있다.

3.9 정보저장매체(storage media)

데이터를 저장 및 보관할 수 있는 모든 형태의 매체를 의미하며, 하드디스크, CD-ROM, USB 메모리, 백업테이프 등이 있음

3.9 채증

법 집행에 있어서 어떠한 사실을 입증하기 위해 증거를 수집하는 행위 또는 과정. 채증은 채집된 증거라는 말의 준말임

3.10 파싱 (parsing)

파서(parser) 역할을 하는 컴퓨터가 문장 단위의 문자열을 토큰(token) 단위로 분류하고 이를 문법에 따라 재구성하는 구문 분석 과정

3.11 파일시스템 (filesystem)

파일이나 자료를 쉽게 발견하고 접근할 수 있도록 보관 또는 조직하는 체계. 대부분의 운영체제는 파일시스템을 가지고 있으며 주로 파일 관리를 목적으로 함. 많이 사용되는 파일시스템으로는 윈도우즈의 FAT, NTFS가 있으며 유닉스 계열의 운영체제에서는 EXT, ZFS, HFS 등이 있음

4 약어

CAPEC	Common Attack Pattern Enumeration and Classification
CASE	Cyber-investigation Analysis Standard Expression
CD	Compact Disc
CTI	Cyber Threat Intelligence
CYBOX	Cyber Observable Expression
DVD	Digital Versatile Disc
EXIF	EXchangeable Image File format
JSON	Java Script Object Notation
PID	Process Identifier
ROM	Read-Only Memory
SNS	Social Networking Service
SSD	Solid State Drive
STIX	Structured Threat Information Expression
TAXII	Trusted Automated eXchange of Indicator Information
XML	eXtensible Markup Language
YAML	Yet Another Multicolumn Layout

5 디지털 포렌식 조사를 위한 통합 정보 처리 규격

이 표준은 디지털 포렌식 조사에서 분석에 필요한 정보를 표현하기 위한 규격이며, 표준의 제정 목적은 수집된 데이터를 사람이 쉽게 읽을 수 있는 수준으로 정규화 함으로써 포렌식 분석 도구의 개발을 촉진하는 것이다.

디지털 포렌식 조사에서 다뤄지는 데이터들은 단계별로 처리하는 주체나 방법에 따라 표현하는 방법이 다르기 때문에 수집된 정보를 통합하여 살펴보기가 어렵고 추가적인 단서를 확보하는 과정이 효과적이지 않다.

따라서 이 표준은 디지털 포렌식 조사에서 발생하는 어려움을 개선하기 위해서 정보를 통합하여 처리할 수 있는 규격을 정의하고, 이 규격이 활용되기 위한 사용법과 유스케이스를 포함한다.

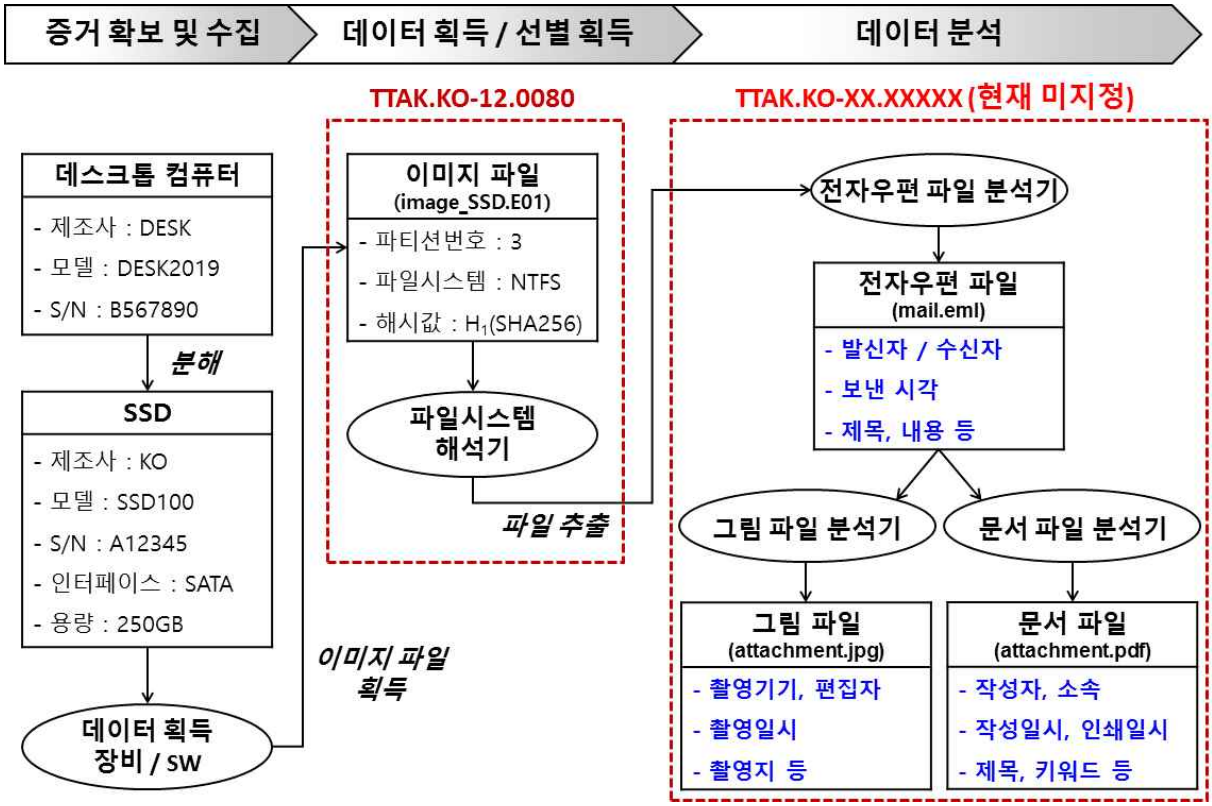
이 표준은 총 3부로 구성되며, 제1부는 표준에서 표현할 수 있는 사용범위와 기본적인 개념, 참고표준과의 비교를 통해 차별되는 특징을 살펴본다. 제2부는 정보를 표현하기 위한 속성을 정의하고, 제3부는 본 표준을 활용한 예제를 유스케이스로 제시하고 레거시와 상호 호환을 위한 참조 모델을 제공한다.

5.1 데이터 모델과 사용범위

디지털 포렌식 조사는 디지털 기기 내부의 데이터가 법적 증거 능력을 갖도록 수집, 보관, 분석, 보고하는 일련의 과정이다. 먼저 확보된 정보저장매체로부터 시스템 동작이나 사용자 행위에 의해 생성된 기록과 관련된 데이터를 수집해야 한다. 이를 위해 데이터 획득 장비나 소프트웨어를 이용해서 데이터가 저장되어 있는 정보저장매체로부터 조사에 필요한 부분을 복제한다. 그 이후 획득한 데이터를 파일 단위로 접근하고, 특정 파일의 구조를 분석함으로써 사건의 단서나 사람의 행위, 데이터의 의미를 해석하여 사건의 실체를 파악한다.

이 과정을 단계별로 도식화하여 예로 들면 (그림 5-1)과 같다. 데스크톱 컴퓨터에서 장착된 정보저장매체인 SSD를 분해하여 수집하고, 데이터 획득 장비를 이용하여 이미지 파일을 생성한다. 여기서 생성된 이미지 파일은 디지털 증거 파일 교환포맷(TTAK.KO-12.0080)으로 저장시킬 수 있다. 이후 파일시스템을 해석하여 특정 파일에 접근할 수 있으며, 전자우편 파일에 포함된 발·수신자, 보낸 시각, 제목, 내용 등의 정보를 확인한다. 또한, 전자우편에 첨부된 그림 파일에서 EXIF 정보 분석을 통해 사진을 촬영한 기기, 편집한 사람, 사진이 촬영된 시각이나 촬영된 장소를 알 수 있고, 문서 파일에서 작성자 등의 정보를 알 수 있다.

디지털 포렌식 조사에서 다루는 파일의 종류는 그 범위가 넓고 다양하지만, 분석결과를 표현하는 방법이 없으므로, 이 표준은 파일 단위로 분석한 결과를 통합된 형태로 관리할 수 있도록 정보를 표현하는 방법을 정의한다.



(그림 5-1) 디지털 포렌식 조사에서 요구되는 규격의 사용범위 (예: 전자우편 파일 등)

디지털 데이터로 저장되어 있는 내용을 분석한 결과는 객체(object) 단위로 표현되며, 크게 '기록(trace)'과 '이벤트(event)'로 구분할 수 있다.

기록은 특정 대상의 상태나 설정정보와 관련하여 윈도우 레지스트리나 응용 프로그램에서 설정된 값과 같은 내용을 표현하기 위한 모델이다. 이벤트는 시스템 동작이나 사용자 행위에 의해 발생되었던 내용을 표현하기 위해 설계된 모델이며, 하나 이상의 서로 다른 객체(기록, 이벤트)를 연관시키기 위한 표현을 지원한다. 예를 들어, 문서 파일에서 파일 확장자, 파일 크기, 글자 수와 같은 정보는 기록 객체로 표현하고, 작성한 사람, 작성한 일시, 편집 프로그램과 같은 정보는 이벤트 객체로 표현한다.

5.2 요구사항 정의

정보를 처리하기 위한 규격은 분석을 통해 확인된 정보가 정확하고(correctness), 명료하며(conciseness), 간결한(clearness) 형태를 가지고 있어야 한다는 것이 기본적으로 본 표준을 활용하여 데이터를 표현하는 데 요구되는 사항이다.

5.2.1 정확성(correctness)

데이터를 규격에 맞추어 표현하는 목적은 특정 사실을 조사관 등이 서로 전달하고 공유하는 것이다. 따라서 분석을 통해 확인된 정보들은 정확해야 한다. 여러 출처로부터 수집된 데이터의 분석 결과를 정확하게 표현하기 위해서 <표 5-1>과 같이 '누가', '언제', '어디서', '무엇을', '어떻게', '그 외'로 구성할 수 있는 육하원칙에 기반을 두어 표현한다.

<표 5-1> 정확한 표현을 위해 사용하는 육하원칙

구분		내용
who	누가, 누가 누구에게	행위 또는 사건의 주체, 대상
when	언제, 언제부터 언제까지	행위 또는 사건이 일어난 시간, 기간
where	어디서, 어디에서 어디로	행위 또는 사건이 일어난 장소
what	무엇을	객체, 대상, 행위
how	어떻게	방법
etc	그 외	위의 구성요소로 표현되지 않으나, 디지털 포렌식 조사에 필요한 내용 일체

5.2.2 명료성(conciseness)

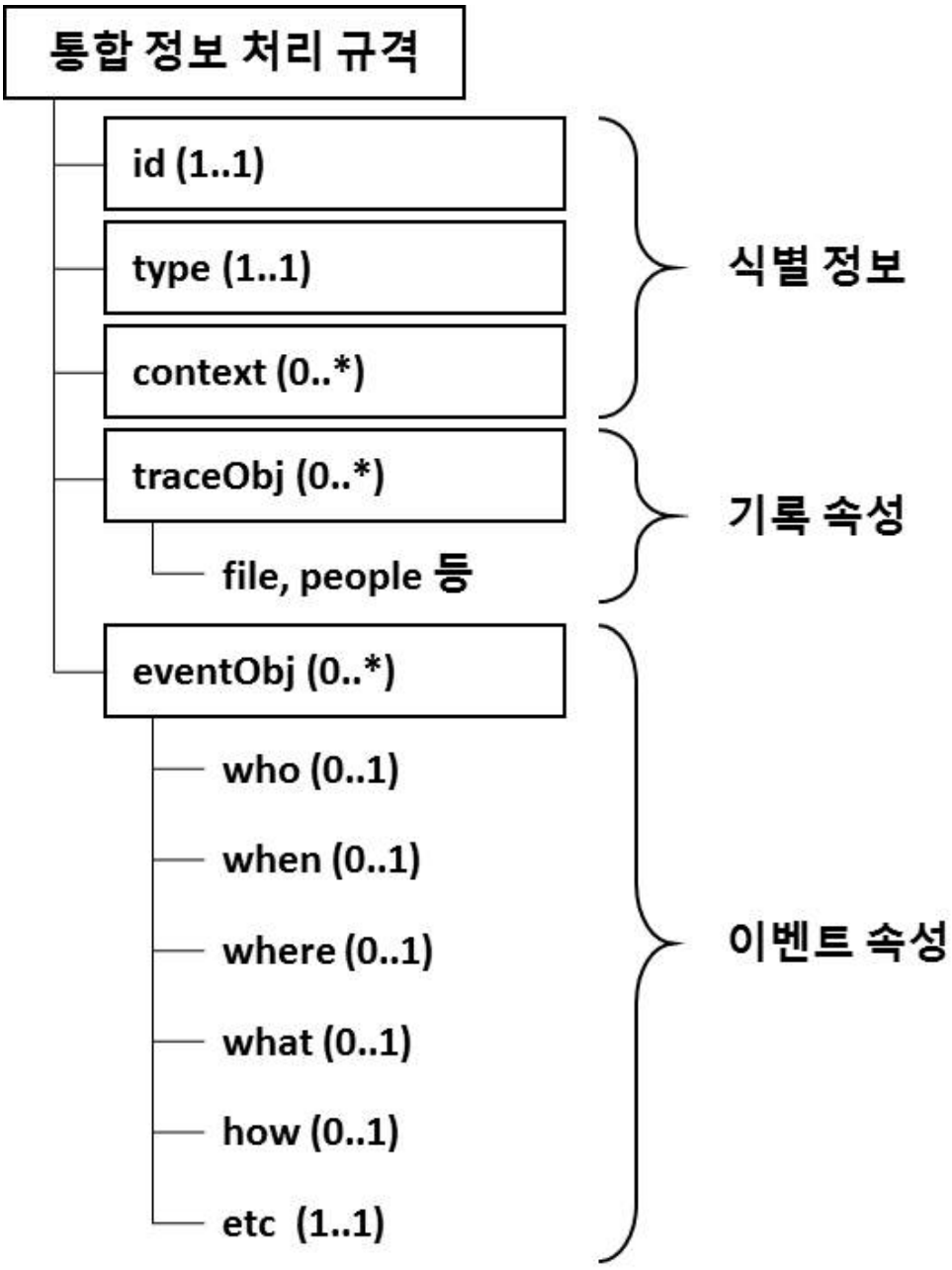
하나의 표현은 하나의 정보로 해석되어야 한다. 데이터 처리 규격으로 표현되는 정보는 단순히 파싱되어 출력되는 결과와 다르고, 둘 이상의 정보가 하나의 정보로 표현이 가능하다. 따라서 다르게 해석될 수 있는 표현을 피해야 하며, 정의되지 않은 내용을 규격에 포함시킬 경우에도 분명한 표현과 설명이 동반되어야 한다.

5.2.3 간결성(clearness)

정보를 간결하게 표현한다는 것은 자연스럽게 문장으로 읽을 수 있어야 한다는 뜻이다. 누구나 이해할 수 있게 작성되어야 하며, 불필요하거나 중복되는 데이터는 지양해야 한다. 데이터를 처리하는 과정에서 수집되는 모든 정보가 디지털 포렌식 조사에서 필요한 것은 아니다.

5.3 데이터 표현 구조

통합 정보 처리 규격에서 사용하는 데이터 모델은 자료 표현 방법 중 하나인 JSON 문법을 기준으로 개발하였으며, 특정 문법(XML, YAML 등)에 종속적인 문법이 필요할 경우에는 명시적으로 추가되었음을 표시해야 한다. 이 구조의 구성요소별 자세한 내용은 본 표준의 6장에서 기술한다.



(그림 5-2) 디지털 포렌식 조사를 위한 통합 정보 처리 규격에서 데이터 표현 구조

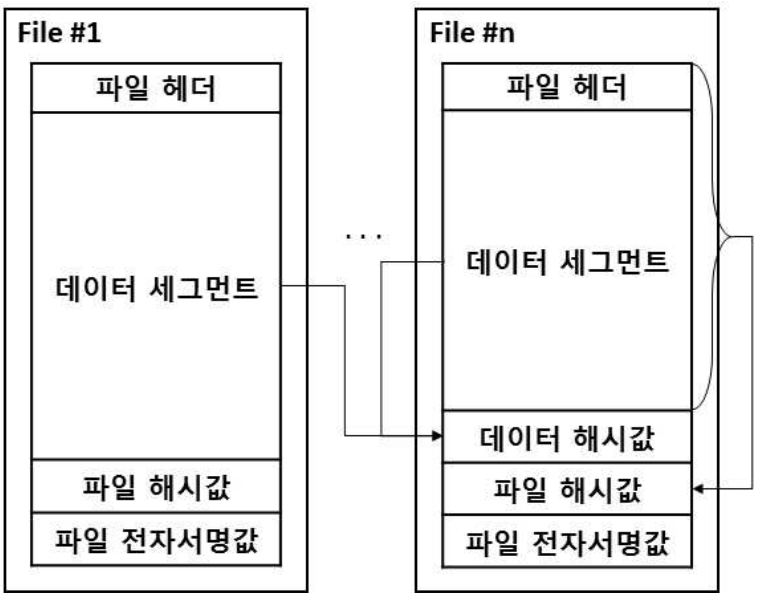
5.4 참고표준과의 차별성

통합 정보 처리 규격의 표현 범위를 참고표준과 비교하여 이 표준의 목적을 제고하고 차별성을 확인한다. 참고표준으로는 '디지털 증거 파일 교환포맷 (TTAK.KO-12.0080)'과 '구조화된 위협 정보 표현 규격(STIX) 버전 2.0 (TTAE.OT-12.0019)'이 있으며, 그 밖의 참고문헌은 부록을 참조한다.

5.4.1 디지털 증거 파일 교환포맷 (TTAK.KO-12.0080)

디지털 증거 파일 교환포맷은 디지털 포렌식 도구(하드웨어 장비, 소프트웨어)에서 획득한 데이터를 원활하게 교환하기 위해서 이미지 파일에 대한 포맷을 정의하였다. 이미지 파일은 데이터 획득 도구를 개발하는 업체에 따라 상이하므로 특정 도구에 의존적으로 진행되던 제한적인 상황에 대한 해결책으로 제시되었다.

이미지 파일과 개별 파일을 저장할 수 있는 범용적인 교환포맷을 (그림 5-3)과 같이 정의하였으며, 파일 헤더에 데이터의 크기, 증거 파일을 작성한 기관, 작성자 이름, 식별번호, 작성된 시간을 필수적으로 포함하여 저장하므로 데이터의 무결성, 원본성, 진정성을 보장할 수 있도록 정의되었다.



(그림 5-3) 디지털 증거 파일 교환포맷

즉, 디지털 증거 파일 교환포맷은 디지털 포렌식 절차(그림 5-1)에서 데이터 획득(또는 선별 획득)을 통해 생성된 이미지 파일을 저장하기 위한 포맷을 정의한다. 하지만 데이터 분석을 통해 출력된 결과를 처리하는 기능을 지원하지 못하므로, 디지털 포렌식 조사에서 그 사용범위가 제한적이다. 하지만 이 표준(TTAK.KO-12.00XX)은 데이터 분석 단계에서 적용할 수 있는 내용으로써, 획득된 데이터를 분석하여 출력된 내용을 처리할 수 있도록 정규화된 표현으로 작성하기 위한 포맷을 정의한다.

만약, 하나의 사건에서 수집한 다수의 정보저장매체와 디지털 데이터를 연속적이고 통합적으로 관리하기 위해 사건 정보, 디지털 증거 정보(현장 채증 정보, 증거 파일 작성 정보를 포함), 분석 정보(추출 파일 정보, 분석 보고서 정보, 촬영정보)를 구조화하여 저장하는 포맷이 필요할 경우에는 디지털 증거 데이터 패키지(KS X 1220)를 참고한다.

5.4.2 구조화된 위협 정보 표현 규격 (STIX 버전 2.0, TTAE.OT-12.0019)

STIX는 사이버 위협 인텔리전스(CTI) 정보를 교환하는데 사용하는 언어로써, 악성코드 특성, 침입 탐지, 사고 대응 및 관리, 디지털 포렌식 등에서 활용할 수 있도록 사이버 공간에서 관측되는 객체와 그 속성을 구조화된 표현으로 정의한다. STIX를 이용하면 사이버 공격에 대해 효과적으로 대응할 수 있도록 정보를 공유할 수 있다. (STIX 버전 2.0은 CAPEC이나 CYBOX에서 정의한 속성을 포함하여 표현한다.)

STIX와 본 표준은 정보를 구조화된 표현으로 정의하는 것을 목적으로 한다는 공통점이 있으나, 근본적으로 활용목적이 다르므로 표현범위가 상이하다. STIX는 사이버 공간상의 네트워크 또는 호스트에서 일어난 일에 관한 사실을 표현하지만, 본 표준은 시스템 자체에서 일어난 흔적에 관한 사실(아티팩트, 특정 파일 등을 분석한 결과)을 표현하고 육하원칙에 기반을 둔다. 다만, 본 표준은 STIX에서 정의한 속성과 연계하여 표현이 가능하도록 개발하였다.

<표 5-2> STIX와 통합 정보 처리 규격 비교

구분	STIX 버전 2.0	통합 정보 처리 규격
대상	시스템, 네트워크 패킷	시스템, 이미지 파일
목적	사이버 위협 정보 공유	디지털 포렌식 조사
공통점	- 정보의 개념을 표준화하고 구조화하여 일관된 분석과 자동화된 해석이 가능하게 만들기 위한 표현 규격 - JSON 포맷으로 작성	
차이점	- 표현범위 : 네트워크에서 일어난 일 중심 (통신, 서비스, 프로세스 등)	- 표현범위 : 시스템에서 일어난 흔적 중심 (아티팩트, 파일 등) - 육하원칙에 기반을 둔 표현
기타	- 미국 MITRE 주체, 국제 표준 - 정보 전송 규격(TAXII) 지원	- CASE* 커뮤니티 주체, 개발진행

* CASE(Cyber-investigation Analysis Standard Expression), https://caseontology.org/

본 표준은 STIX를 활용하여 정보를 표현하는데 있어 목적에 부합되는 표현이 제한되므로 디지털 포렌식 조사 관점에서 차별성을 가진다. (그림 5-4)는 STIX 버전 2.0과 본 표준으로 표현한 결과의 차이를 보여준다. 사진을 저장하는 파일 포맷 중 하나인 JPEG 파일은 EXIF에 해당 파일과 관련된 메타데이터를 추가로 포함하며, STIX로는 이러한 표현이 제한된다.

STIX 버전 2.0 표현

```
{
    "type": "bundle",
    "id": "7b8c44ce1d15",
    "Description": "An example",
    "Objects": [
        {
            "File_Name": "attachment.jpg",
            "Size_In_Bytes": 25523
        },
        {
            "image_is_compressed": True,
            "Image_File_Format": "JPEG/JFIF",
            "Image_Height": 1932,
            "Image_Width": 2576,
            "Bits_Per_Pixel": 256
        }
    ]
}
```

통합 정보 처리 규격 표현

```
{
    "id": "7b8c44ce1d15",
    "type": "picture",
    "Context": "An example ",
    "traceObj": {
        "id": "jpg_file1",
        "File_Name": "attachment.jpg",
        "File_Size": 25523,
        "Created_Time":
          "2019-12-10T15:25:37+09:00"
    },
    "eventObj": {
        "who": {
            "Manufacturer": "Apple",
            "Model": "iPhone 6s",
            "Software": "12.0.1",
        },
        "when": {
            "DigitizedTime":
              "2018-11-10T15:25:37-05:00"
        },
        "where": {
            "Location": {
                "latitude": "37.588672",
                "longitude": "127.046808"
            }
        },
        "what": {
            "Orientation": "top-left",
            "ExposureTime": "1/15s",
            "F_number": "f/2.2",
            "Pixel_X_dimension": 2576,
            "Pixel_Y_dimension": 1932,
            "ColorSpace": "sRGB",
            "CompressionMethod":
              "JPEG/JFIF",
        },
        "etc": {
            "Source": "jpg_file1"
        }
    }
}
```

(그림 5-4) STIX와 통합 정보 처리 규격의 표현 비교(예): 그림 파일

6 데이터 표현을 위한 구성요소

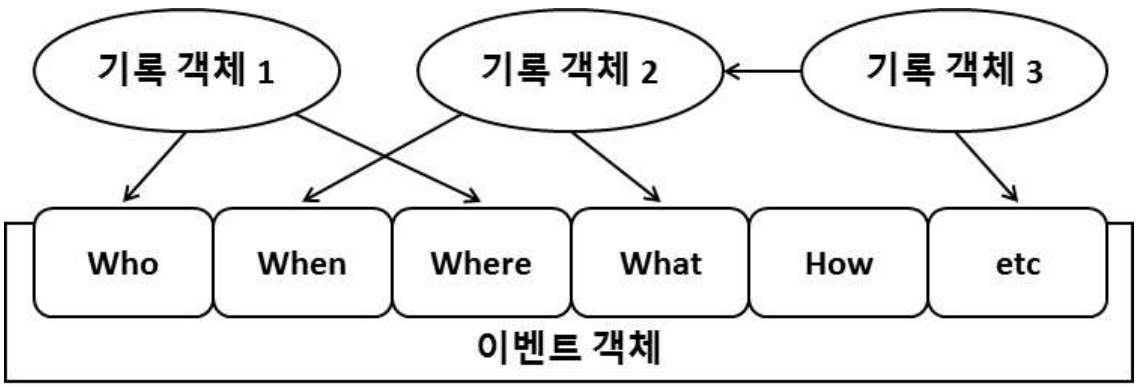
이 표준에서 정의하는 규격은 디지털 포렌식 조사에서 유용하게 활용될 수 있도록 조사 대상으로부터 확인된 정보를 유연한 표현으로 작성하는 것을 목표로하며, 정보를 통합하여 처리하여 추가적인 단서를 도출하기 위한 처리를 지향한다.

규격은 JSON 형식을 기본으로 객체 단위로 표현하며, 객체는 키(key)-값(value) 형태로 속성을 표현한다. 이 장은 객체와 속성에 대한 설명을 기술하고, 정보의 종류에 따라 객체를 종류별로 정의하였다.

6.1 객체 (object)

객체는 데이터의 성격에 맞게 상태나 발생되었던 이벤트를 표현하기 위한 모델이다. 각 객체는 데이터의 종류에 따라 표현하고자 하는 내용을 정의된 속성으로 작성되며, 모든 객체는 공통적으로 객체를 구분할 수 있는 속성의 성격에 따라 기록 속성과 이벤트 속성으로 구분된다. 객체는 기록 객체와 이벤트 객체로 구분되며, 식별 정보와 기록 속성만 가지고 있는 객체를 기록 객체라고 부르고, 이벤트 속성을 포함한 객체를 이벤트 객체로 부른다.

기록 객체는 조사 대상으로부터 관찰된 사실을 기술하며, 특정 대상에 대한 상태를 표현하기 위해 사용한다. 이벤트 객체는 기록 객체로부터 확인된 내용으로 추론이 가능한 정보를 표현하기 위해 설계된 모델이다. 기록 객체로 표현된 사실들을 서로 연관시키고 관련 객체에 이벤트를 연관시키는 기능을 지원한다. 이 모델은 사용자 행위나 시스템에서 발생된 동작을 포함하여 대상에 대한 관계를 표현할 수 있도록 하며 육하원칙에 기반을 둔 형태로 표현한다. (그림 6-1)은 기록 객체와 이벤트 객체의 활용을 설명하기 위한 예제로 정보들의 관계를 표현한 그림이다. 기록 객체와 이벤트 객체에서 사용하는 속성의 종류는 <표 6-1>과 같다. (그림 6-2)는 통화기록을 규격에 맞추어 작성한 이벤트 객체의 예시이다.



(그림 6-1) 정보를 표현하기 위한 기록 객체와 이벤트 객체의 관계

<표 6-1> 속성의 종류

No.	속성 이름	속성 내용
1	file	파일의 메타데이터 정보. 기본적으로 파일시스템을 분석한 결과를 저장할 수 있으며, 파일 형식마다 파일 내부에 포함되어 있는 메타데이터를 표현하도록 다음 파일 형식을 고려하여 세분화하여 정의한다. - 압축파일(archive) : ZIP, 7z, RAR 등 압축된 파일 - 실행파일(executable) : 운영체제에서 실행되는 파일 - 문서파일(document) : 열람 및 편집할 수 있는 파일 - 글꼴파일(font) : 글꼴 형태가 정의되어 있는 파일 - 미디어파일(media) : 사진, 동영상 파일 - 메모리파일(memory) : 메모리의 데이터를 복제한 파일
2	location	GPS 기록, 주소지 등 장소의 정보를 표현하거나 출발지, 경유지, 목적지를 포함하여 특정 대상의 이동한 기록을 표현할 수 있는 속성을 정의한다.
3	string	특정 파일에서 파싱된 문자열을 표현하기 위해서 시스템마다 사용하는 문법이나 인코딩 방식이 다양한데 이러한 정보를 표현할 수 있는 속성을 정의한다.
4	people	특정 대상과 관련된 정보를 표현할 수 있는 속성이다. 이름, 연락처(전화번호, 전자우편), 사진, 계좌정보, 인터넷을 이용하여 사용하는 서비스의 계정 정보 등이 정의되어 있다.
5	calender	특정 대상의 일정 기록을 나타내거나 시스템에서 발생된 알림이나 사건을 표현할 수 있는 속성이다. 디지털 포렌식 조사에서 시계열 분석결과를 공유하는 목적으로도 사용할 수 있다.
6	call	디지털 기기 내 저장되어 있던 전화기, 컴퓨터 등의 기기에서 음성이나 영상 전화로 주고받은 기록이나 통신사로부터 확보한 통화내역을 표현할 수 있도록 정의한 속성이다.
7	cloud	클라우드를 기반으로 하는 서비스가 많아지고 있는데, 특히 파일을 저장하기 위한 공간을 대상으로 조사가 진행되는 경우 여기에 정의된 속성을 참고할 수 있다.
8	connection	기기 또는 매체 간 연결된 기록을 표현할 수 있는 속성이다. 시스템과 외부저장매체와의 연결, 무선 인터넷 연결(예: 와이파이), 블루투스를 이용한 연결 등 다양한 방식과 상황을 속성으로 정의한다.

No.	속성 이름	속성 내용
9	email	전자우편을 통해 주고받은 기록을 표현할 수 있는 속성이다. 전자우편은 파일 형태로 내용과 이력 정보(IP, 도메인 등)가 저장되고 업무에서 많이 사용되므로, 파일을 파싱한 결과는 디지털 포렌식 조사에서 유용하다. 전자우편에 첨부된 파일(예: 문서, 압축 등)은 파일 속성을 함께 사용할 수 있다.
10	exchange	특정 대상과의 거래한 기록을 표현할 수 있는 속성이다. 이 속성을 사용할 수 있는 예제로는 계좌이체, 신용카드 결제, 현금이나 현물의 주고받은 상황이 있다.
11	message	메시지 속성은 특정 대상(들)과 주고받은 메시지를 표현할 수 있다. 휴대전화의 문자내역, PC나 스마트폰에 설치된 메신저 애플리케이션을 이용하여 주고받은 내용을 나타낼 수 있다. 사진이나 동영상 같은 메시지는 파일 속성을 이용하여 보다 정확하게 표현할 수 있다.
12	social	소셜네트워킹서비스(SNS) 등을 이용하여 인터넷 상으로 게시되어 있는 내용이나 사용자가 관련 서비스를 이용한 기록을 나타낸다. 웹 기록 속성으로도 일부 표현이 가능하나, 관련 모바일 애플리케이션의 사용기록을 웹 기록과 구분하여 표현할 수 있다.
13	web	웹 브라우저를 이용하여 인터넷에 접속한 기록은 사용자의 행위를 파악할 수 있는 유용한 정보이다. 웹페이지 방문기록, 다운로드한 파일 목록, 캐시, 쿠키를 분석한 정보를 표현할 수 있는 속성을 정의한다.
14	system	시스템에서 발생한 이벤트가 기록된 로그(아티팩트)를 파싱한 결과를 표현하기 위한 속성을 정의한다. 윈도우즈가 설치된 시스템의 경우, 레지스트리, 프리패치, 이벤트로그와 같은 파일이 포함하는 정보를 나타낼 수 있다.
15	network	네트워크 통신에서 발생된 이벤트를 표현할 수 있다. 해당 속성은 STIX에서 정의한 속성과 중복되므로, 본 표준에서는 최소로 정의하고 STIX에서 정의한 속성을 이용하여 표현한다. STIX와 연동한 표현은 3부의 5.3 내용을 참조한다.

6.1.1 식별 정보

식별 정보는 객체를 구분하고 예외사항을 전달하기 위한 용도로 사용되는 속성이다. id, type 속성은 반드시 포함되어야 하며, context는 필요시 사용할 수 있다.

<표 6-2> 식별 정보의 종류

No.	속성 이름	속성 내용
1	id	id 속성은 객체를 고유하게 식별할 수 있게 하는 보편적인 식별자이다. 사용자에 의해 생성되었거나 정보를 공유하는 관계에서의 객체는 모두 다른 식별자를 가지고 있어야 한다. 동일한 식별자를 갖는 경우에는 중복객체를 제거하거나 식별자 재생성 등의 조치를 취한다.
2	type	type 속성은 객체의 유형을 식별한다. type 속성의 값은 본 표준의 2부에서 정의되고, 정의된 유형 중 하나의 이름을 가져야 한다. * STIX와 연계하여 표현할 경우에는 STIX 버전 2.0에서 정의한 type을 사용할 수 있다.
3	context	해당 객체를 전달받는 사용자에게 객체의 생성 이력이나 특이사항 등을 추가적으로 설명하기 위한 속성

6.1.2 기록 속성

기록 속성은 특정 대상의 상태나 시스템 설정 정보와 같은 내용을 표현하기 위한 속성이다. 예를 들어, 파일시스템에 기록되어 있는 파일의 메타데이터(파일 이름, 저장위치, 크기 등)와 파일 내부에 저장되어 있는 정보, 장소나 경로에 대한 정보, 연락처에 저장되어 있거나 특정 서비스 사용을 위한 사용자 계정 정보, 그리고 인코딩 정보 등이 표현되도록 지원한다.

6.1.3 이벤트 속성

이벤트 속성은 데이터를 처리한 결과가 사용자에게 정확하고 명확하게 정보를 제공할 수 있도록 본 표준에서 개발한 속성이다. <표 5-1>과 같이 육하원칙에 기반을 두었으며, 디지털 포렌식 조사에서 확인되어야 할 정보의 범위를 '누가', '언제', '어디서'만으로 표현하기에는 제한되므로 그 범위를 확대하여 각각 '누가 누구에게', '언제부터 언제까지', '어디에서 어디로' 개념을 포함할 수 있는 범위로 확대하였으며, '왜' 대신에 '그 외'로 대체하여 사용한다. '그 외'에는 육하원칙으로 구분하기 모호한 정보를 가지는 속성을 처리한다.

예시

```
{
  "id": "UserActivity",
  "type": "call",
  "traceObj": [
    {
      "id": "Source1",
      "type": "file",
      "traceObj": {
        "FilePath": "/data/data/com.provider.telephone/calllog.db",
        "ModifiedTime": "2019-04-01T09:12:11+09:00",
        "ByteRuns": [{offset=503808, len=8192}],
        "HashValue": "40b3015192aea60377986a24ea238af50d711efb",
        "HashType": "SHA1"
      }
    },
    {
      "id": "Source2",
      "type": "file",
      "traceObj": {
        "FilePath": "/data/data/com.provider.contact/contact.db",
        "ByteRuns": [(offset=122880 len=4096)]
      }
    },
    {
      "id": "Person1",
      "type": "people",
      "traceObj": {
        "FirstName": "Chulsoo",
        "LastName": "Kim",
        "PhoneNumber": "010-1234-5678",
        "Source": "Source1"
      }
    },
    {
      "id": "Person2",
      "type": "people",
      "traceObj": {
        "FirstName": "Younghee",
        "LastName": "Kim",
        "PhoneNumber": "010-8765-4321",
        "Source": "Source1"
      }
```

```
    {
      "id": "Location1",
      "type": "location",
      "traceObj": {
        "StartPoint": "37.5665 N, 126.9780 E"
        "Source": "Source2"
      }
    }
  ],
  "eventObj": [
    {
      "who": {
        "From": "Person1",
        "To": "Person2"
      },
      "when": {
        "StartTime": "2019-03-23T12:34:56+09:00",
        "EndTime": "2019-03-23T12:43:12+09:00"
      },
      "where": {
        "Location": "Location1"
      },
      "what": {},
      "How": {
        "Application": "default",
        "CallType": "Incomming"
      },
      "etc": {
        "Source": ["Source1", "Source2"]
      }
    }
  ]
}
```

(그림 6-2) 통합 정보 처리 규격을 이용한 표현: 통화기록

6.2 속성 (property)

속성은 객체의 특징이나 성질을 나타내며 데이터 처리(해석)의 결과로 확인될 수 있는 값이다. 시스템에서 사용하는 운영체제나 응용프로그램의 종류가 다양하므로 데이터가 저장되어 있는 형태는 특정한 규칙을 가지고 있지 않은 경우가 많다. 따라서 객체의 종류에 따라 공통적으로 포함되는 속성을 규정함으로써 통합하여 데이터 처리가 가능하고 정보의 교환이 용이할 것이다.

6.2.1 자료형 (data type)

디지털 포렌식 조사에서 다루어지는 데이터의 종류가 다양하고 방대하므로 처리된 결과의 효율적인 관리를 위해 정보를 저장하는 방식을 <표 6-3>의 범위 내에서 사용한다.

<표 6-3> 자료형 정의

자료형	설명
boolean	참과 거짓을 표현하기 위한 자료형. True 또는 False로 표현한다. 예) True
integer	정수를 표현하기 위한 자료형 예) 1027
float	실수를 표현하기 위한 자료형 부동소수점을 표현하는 숫자이며, IEEE 754을 준수한다. 예) 123.456
string	문자열을 표현하기 위한 자료형 (UTF-8, UTF-16 등 유니코드를 인코딩)
datetime	날짜와 시간을 표현하는 자료형 날짜와 시간 표현 형식에 관한 프로파일(TTAE.IF-RFC3339)을 준수하여 표현한다. 예) 2019-12-10Z, 2019-12-10T15:25:37.456+09:00
list	여러 개의 값이나 자료형을 하나로 관리하기 위한 자료형
object	객체를 활용하여 표현하는 자료형 (객체 내부에 객체를 표현)

6.2.2 반복수 (multiplicity)

이 표준에서의 반복수 표기는 <표 6-4>와 같은 의미를 가진다. 필수로 작성해야 하는 키임에도 불구하고 해당하는 값이 없는 경우에는 “null” 또는 공백(“”)으로 작성한다.

<표 6-4> 반복수 표현 설명

반복수	필수/선택	설명
0..1	선택	해당 항목이 선택적으로 1개 존재할 수 있다
0..*	선택	해당 항목이 선택적으로 1개 이상 존재할 수 있다 (복수 개 존재 가능)
1..1	필수	해당 항목이 필수적으로 1개 존재해야한다
1..*	필수	해당 항목이 필수적으로 1개 이상 존재해야한다 (복수 개 존재 가능)

6.2.3 명명 규칙 (naming rule)

통합 정보 처리 규격에서 속성은 디지털 포렌식 조사에서 발생될 수 있는 다양한 상황을 모두 지원하기 위해 관련된 모든 유형의 정보를 처리하는 것을 목표로 하지만, 표준에서 제공하는 유스케이스가 제한적이고 환경 변화를 빠르게 반영되지 못하는 상황을 고려하여 규격의 지속적인 개선을 위해 확장이 가능하도록 지원한다. 즉, 객체와 속성을 사용자가 추가할 수 있도록 규칙을 제공한다.

이 표준에서 제공하는 표현을 기본으로 작성하고 동일한 내용을 표현할 수 있는 별도의 키를 다시 생성하지 않도록 한다(중복의 최소화). 사용자가 주어진 상황에서 필요한 정보만 표현이 가능하도록 필수적으로 포함시켜야 하는 키를 최소화한다(사용의 유연성). 또한, 자동으로 작성하는 기능을 지원하기 위해 객체의 구조와 속성에 대한 표현은 이 표준에서 정의한 규칙을 준수하여 작성하도록 한다(표현의 일관성).

① 키 이름은 영어 대소문자와 숫자만 사용하며, 첫 문자는 대문자로 작성한다.
예) Source

② 숫자는 동일한 키를 반복하여 사용할 때 각각을 구분하기 위해 사용한다.
예) Source1, Source2, Source3

③ 복합어 형태로 키 이름을 생성할 경우에는 중간에 시작하는 새로운 단어의 첫 문자는 대문자로 한다.
예) SourceType, SourceOwner

부 록 Ⅰ-1

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

지식재산권 확약서 정보

Ⅰ-1.1 지식재산권 확약서

해당 사항 없음

※ 상기 기재된 지식재산권 확약서 이외에도 본 표준이 발간된 후 접수된 확약서가 있을 수 있으니, TTA 웹사이트에서 확인하시기 바랍니다.

부 록 Ⅰ-2

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

시험인증 관련 사항

Ⅰ-2.1 시험인증 대상 여부

해당 사항 없음

Ⅰ-2.2 시험표준 제정 현황

해당 사항 없음

부 록 Ⅰ-3

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

본 표준의 연계(family) 표준

Ⅰ-3.1 디지털 포렌식 조사를 위한 통합 정보 처리 규격 - 제2부 : 데이터 종류별 규격 정의

정보를 통합하여 표현하기 위한 속성을 정의하고 예제로 작성한 표현식을 제공

Ⅰ-3.2 디지털 포렌식 조사를 위한 통합 정보 처리 규격 - 제3부 : 데이터 처리 상호 호환을 위한 참조 모델

규격의 활용을 유스케이스로 제시하고 레거시와 상호 호환을 위한 참조 모델을 제공

부 록 Ⅰ-4

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

참고 문헌

[1] TTAK.KO-12.0080, "디지털 증거 파일 교환포맷", 2008.12.19.

[2] TTAE.OT-12.0019, "구조화된 위협 정보 표현 규격(STIX) 버전 2.0", 2018.12.19.

[3] KS X 1220, "디지털 증거 데이터 패키지", 2014.11.28.

[4] TTAS.KO-12.0058/R1, "디지털 증거 수집 보존 가이드라인", 2017.12.13.

[5] OASIS Cyber Threat Intelligence (CTI) TC, STIX(Structured Threat Information Expression) 2.0, 2018.05.10.

[6] Mirtre, Cyber Observable eXpression(CybOX) 2.1, 2014.01.23.

[7] Rutkowski A, Kadobayashi Y, Furey I, Rajnovic D, Martin R, Takahashi T, Schultz C, Reid G, Schudel G, Hird M, Adegbite S., "CYBEX – The Cybersecurity Information Exchange Framework (X.1500)", ACM SIGCOMM Computer Communication Review. 40(5), pp. 59-64. 2010.10.22.

[8] Casey E, Barnum S, Griffith R, Snyder J, van Beek H, Nelson A., "Advancing coordinated cyber-investigations and tool interoperability using a community developed specification language", Digital Investigation. 22, pp. 14-45. 2017.09.01.

[9] Biasiotti, M.A., Mifsud Bonnici, J.P., Cannataci, J., Turchi, F. "Handling and Exchanging Electronic Evidence Across Europe", Vol. 39. Springer, 2018.

[10] Casey E, Back G, Barnum S., "Leveraging CybOX™ to standardize representation and exchange of digital forensic information", Digital Investigation. 12, pp. S102-110. 2015.05.01.

[11] Garfinkel, S., "Digital forensics XML and the DFXML toolset", Digital Investigation, 8(3-4), pp.161-174. 2012.

[12] TTAE.IF-RFC3339, “날짜와 시간 표현 형식에 관한 프로파일”, 2004.12.23.

※ 상기 기재된 참고 문헌의 발간일이 기재된 경우, 해당 표준(문서)의 해당 버전에 대해서만 유효하며, 연도를 표시하지 않은 경우에는 해당 표준(권고)의 최신 버전을 따름

부 록 Ⅰ-5

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

영문표준 해설서

해당 사항 없음

부 록 Ⅰ-6

(본 부록은 표준을 보충하기 위한 내용으로 표준의 일부는 아님)

표준의 이력

판수	채택일	표준번호	내용	담당 위원회
제1판	2019.12.XX	제정 TTAK.KO-12.00xx	-	사이버보안 프로젝트그룹 (PG503)