

사이버보안빅데이터센터

목 차

- I 사이버보안빅데이터센터 소개
- II 사이버보안빅데이터센터 이용 안내



I

사이버보안 빅데이터센터 소개

I 사이버보안빅데이터센터 소개



사이버보안 빅데이터 수집 현황 및 활용 체계

- ➡ KISA : 빅데이터 분석을 통해 침해사고 대응역량 강화 (7억건 위협정보)
- ➡ 산·학·연 : 위협 빅데이터와 분석 플랫폼을 제공하여 제품개발 및 연구 지원

빅데이터 수집

빅데이터 수집

- ☑ C-TAS 참여사 및 내부 시스템 연동 확대
- ☑ 국내외 위협 인텔리전스 및 OSINT 정보 수집

수집 데이터 확대

- ☑ IoT, ICS 등 4차 산업분야 위협정보 수집
- ☑ 해킹 관련 비정형 데이터 수집



사이버 위협정보
수집

〈현행 데이터〉

악성코드, 경유지, 유포지, 정보유출지,
C&C, 감염PCIP, 공격시도IP,
사고 정보 등 개별 위협정보

빅데이터 구축

위협 데이터셋 생성

- ☑ 악성코드, 공격시도IP 등 개별 위협정보
- ☑ 위협 간 연관정보 생성(프로파일링 등)

인공지능 학습데이터 관리

- ☑ 침해사고 데이터 및 분석 정보 관리
- ☑ 해킹 관련 실시간 이슈 정보 관리



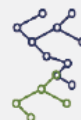
개별
위협정보



위협 간
연관정보



사고
정보



기타
정보

빅데이터 활용

대응 관점 빅데이터 분석

- ☑ 신규 위협 추이 및 사전 인지
- ☑ 공격 특징 추출·동일 공격자 추정

보안제품 개발 및 학술 연구

- ☑ 지능정보기술 기반의 보안장비 개발
- ☑ 분석알고리즘, 프로파일링기법 등 학술연구



위협 추이인지,
공격자 추정



제품개발,
학술연구

II

사이버보안 빅데이터센터 이용 안내

사이버보안빅데이터센터 이용절차

➡ 이용신청 방법



II 사이버보안빅데이터센터 이용



사이버보안빅데이터센터 이용절차

① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

1 - 1 보호나라 홈페이지 접속하기 (<https://www.boho.or.kr>)

① 빅데이터센터 선택



II 사이버보안빅데이터센터 이용



사이버보안빅데이터센터 이용절차

① 이용신청

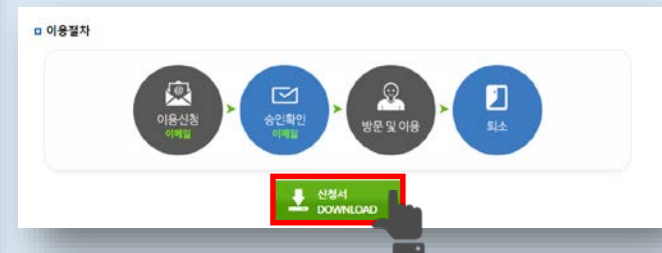
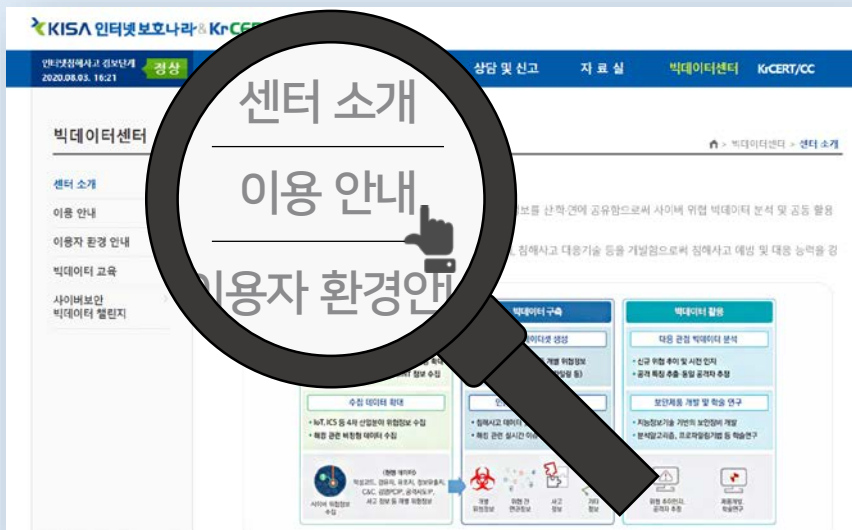
② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

1 - 1 보호나라 홈페이지 접속하기 (<https://www.boho.or.kr>)

② 이용안내 선택



II 사이버보안빅데이터센터 이용



사이버보안빅데이터센터 이용절차

① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

1 - 2 신청서 작성하기

이용신청 및 분석계획서[예시]	
기본 정보 ※ 방문신청 별 한 대의 패스워드PC가 제공되며, 모든 절차는 작성자를 통해 안내됩니다.	
사용자	성명 홍길동, 이메일 kildong@kisa.or.kr, 소속기관 한국인터넷진흥원, 추가 방문자 1명
신청 내역 정보	
사용일	2018-12-24
분석서버(VM) 자원 할당 요청 ※ 기본할당량으로 변경가능(단, 가용자원에 따라 요청방영여부 결정)	
CPU	4 Core, MEMORY 16 GB, STORAGE 500 GB
분석 계획 ※ 분석 목적을 명시하고, 분석 대상, 분석 방법, 분석 결과의 활용 방안 등을 상세히 기재합니다.	
분석과제명	미시관측을 이용한 비정상 행위탐색 자동추출 방안 연구
분석 기간	2018-12-24 ~ 2018-12-28
연구 배경 및 목적	○ 비정상 행위탐색을 자동화하기 위한 미시관측 기법 연구 필요 ※ KISA 패스워드 "미시관측"을 이용한 비정상 행위탐색 자동추출 방안 연구 수행에 필요한 비정상 행위탐색 및 분석환경 필요
연구 및 분석 계획	○ 인텔리전스 보고서 등 문서자료로부터 행위정보(IP, 도메인 등)를 자동으로 인식하고 추출하기 위한 미시관측 기술 연구 ○ 추출된 행위정보를 XML, JSON, CSV 등의 정형화된 형태의 행위정보로 자동생성하기 위한 자동화 기술 연구
기대효과 및 활용	○ 인텔리전스 보고서, OSINT 등으로부터 행위정보를 자동추출 할 수 있는 기술을 개발하여 사이버위협 식별이나 분석에 활용
개인정보 수집·이용 고지사항 ※ 개인정보 수집 항목 : 사이버보안빅데이터센터의 사용자 안전관리 및 서비스 제공을 위한 개인정보 항목 : 수집기간, 성명, 이메일, 소속기관, 추가 방문자 ※ 보유 및 이용기간 : 6개월(보통자료관리법상 제26조제2항) ※ 동의 거부 권리 및 동의 거부 시 불이익 : 개인정보의 수집·이용에 동의하지 않을 경우, 본 센터의 서비스 제공이 불가하며, 개인정보의 수집·이용에 동의한 후에도 언제든지 동의 철회가 가능합니다. ※ 본 센터는 개인정보보호법 제15조 조항에 근거하여 작성되었습니다. ※ 개인정보 수집·이용 동의 여부 : <input checked="" type="checkbox"/> 동의함 <input type="checkbox"/> 동의안함	

이용신청 및 분석계획서[예시]					
기본 정보 ※ 방문신청 별 한 대의 패스워드PC가 제공되며, 모든 절차는 작성자를 통해 안내됩니다.					
사용자	성명	이메일	소속기관	추가 방문자	
	홍길동	kildong@kisa.or.kr	한국인터넷진흥원	작성자 외 1명	
신청 내역 정보					
사용일	2020-08-24				
분석서버(VM) 자원 할당 요청 ※ 기본할당량으로 변경가능(단, 가용자원에 따라 요청방영여부 결정)					
CPU	4 Core	MEMORY	16 GB	STORAGE	500 GB

- 기본 정보 및 사용일, 분석서버 할당 요청 기입
- 추가 방문자 VM 희망 시 이메일 추가 기입
- 사용일은 실제 방문일

II 사이버보안빅데이터센터 이용



사이버보안빅데이터센터 이용절차

① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

1 - 2 신청서 작성하기

이용신청 및 분석계획서(예시)				
기본 정보				
사용자	성명	이메일	소속기관	승가 방문자
	홍성민	hlsbong@kisa.or.kr	한국인터넷진흥원	제1차 1월
신청 내역 정보				
신청일	2018-12-24			
분석사제(VM) 자원 할당 요청 ※ 가상화환경으로 분석가능한 기능에 한함				
CPU	4 Core	MEMORY	16 GB	STORAGE
				500 GB
분석 계획				
본 분석 계획은 제1차 1월 분석계획을 본 분석기간의 기준과 동일하게 기재				
분석과제명	머신러닝을 이용한 비정형 위협정보 자동추출 방안 연구			
분석 기간	2018-12-24 ~ 2018-12-28			
연구 배경 및 목적	<ul style="list-style-type: none"> ○ 비정형 위협정보를 자동으로 추출하기 위한 머신러닝 기법 연구 필요 ※ KISA 위탁과제 "머신러닝을 이용한 비정형 위협정보 자동추출 방안 연구" 수행에 필요한 비정형 위협정보 및 분석환경 필요 			
연구 및 분석 계획	<ul style="list-style-type: none"> ○ 인텔리전스 보고서 등 문서파일로부터 위협정보(IP, 도메인 등)를 자동으로 인식하고 추출하기 위한 머신러닝 기술 연구 ○ 추출된 위협정보를 XML, JSON, CSV 등의 정형화된 형태의 위협지표로 자동생성하기 위한 자동화 기술 연구 			
기대효과 및 활용	<ul style="list-style-type: none"> ○ 인텔리전스 보고서, OSINT 등으로부터 위협정보를 자동추출 할 수 있는 기술을 개발하여 사이버위협 빅데이터 분석에 활용 			
개인정보 수집·이용 고지사항				
<ul style="list-style-type: none"> 가) 개인정보 수집 항목 : 사이버보안빅데이터센터의 사용자 안전관리 및 서비스 제공 나) 수집하는 개인정보 항목 : 소속기관, 성명, 이메일 다) 보유 및 이용기간 : 5년(공공기관출력정보관리법 제26조제1항) 라) 동의 거부 권리 및 동의 거부 시 불이익 : 개인정보의 수집·이용에 동의할 수 없으며, 이 경우 사이버보안빅데이터센터에 요청한 개인정보는 삭제됩니다 마) 본 동의서는 개인정보보호법 제15조 조항에 근거하여 작성되었습니다 				
<div> <div>개인정보 수집·이용 동의</div> <div>동의함</div> <div>동의함</div> </div>				

분석 계획		※ 동일 건으로 재방문 시 분석과제명 및 분석기간만 기준과 동일하게 기재
분석과제명	머신러닝을 이용한 비정형 위협정보 자동추출 방안 연구	
분석 기간	2020-08-24 ~ 2020-08-28	
연구 배경 및 목적	○ 비정형 위협정보를 자동으로 추출하기 위한 머신러닝 기법 연구 필요 ※ KISA 위탁과제 “머신러닝을 이용한 비정형 위협정보 자동추출 방안 연구” 수행에 필요한 비정형 위협정보 및 분석환경 필요	
연구 및 분석 계획	○ 인텔리전스 보고서 등 문서파일로부터 위협정보(IP, 도메인 등)를 자동으로 인식하고 추출하기 위한 머신러닝 기술 연구 ○ 추출된 위협정보를 XML, JSON, CSV 등의 정형화된 형태의 위협지표로 자동생성하기 위한 자동화 기술 연구	
기대효과 및 활용	○ 인텔리전스 보고서, OSINT 등으로부터 위협정보를 자동추출 할 수 있는 기술을 개발하여 사이버위협 빅데이터 분석에 활용	

■ 분석 계획, 목적, 활용 내용 기입

II 사이버보안빅데이터센터 이용



사이버보안빅데이터센터 이용절차

① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

1 - 2 신청서 작성하기

이용신청 및 분석계획서(예시)			
기본 정보 본 센터를 통해 본 센터 회원이 신청한 분석을 위한 신청서입니다. 신청서는 본 센터 홈페이지에 게시되어 있습니다.			
사용자	성명	이메일	소속기관
	홍길동	hgilong@kisa.or.kr	한국인터넷진흥원
신청 내역 정보			
신청일	2018-12-24		
분석사제(VMO) 자원 할당 요청 (본 센터에서 제공되는 분석사제에 대한 상세정보는 본 센터 홈페이지에 게시되어 있습니다.)			
CPU	4 Core	MEMORY	16 GB
STORAGE	500 GB		
분석 계획 본 센터를 통해 본 센터 회원이 신청한 분석을 위한 신청서입니다. 신청서는 본 센터 홈페이지에 게시되어 있습니다.			
분석과제명	미시데이터를 이용한 비정형 위험정보 자동추출 방안 연구		
분석 기간	2018-12-24 ~ 2018-12-28		
연구 배경 및 목적	<ul style="list-style-type: none"> ○ 비정형 위험정보를 자동으로 추출하기 위한 미시데이터 자동 추출 방안 연구 ○ KISA 위키데이터 "미시데이터를 이용한 비정형 위험정보 자동추출 방안 연구" 수행에 필요한 위험정보 및 분석환경 정보 		
연구 및 분석 계획	<ul style="list-style-type: none"> ○ 위험정보를 보고서 등 문서자료로부터 위험정보(IP, 도메인 등)를 자동으로 인식하고 추출하기 위한 미시데이터 자동 추출 방안 연구 ○ 추출된 위험정보를 XML, JSON, CSV 등의 정형화된 형태의 위험정보로 자동 생성하기 위한 자동화 기술 연구 		
기타요구 사항	<ul style="list-style-type: none"> ○ 위험정보를 보고서, OSINT 등으로부터 위험정보를 자동추출 할 수 있는 기술 등 개발하여 사이버위협 억제에 대한 분석에 활용 		

《개인정보 수집·이용 고지사항》	
○ 개인정보 수집 이용 목적	: 사이버보안빅데이터센터의 사용자 본인확인 및 서비스 제공
○ 수집하는 개인정보 항목	: 소속기관, 성명, 이메일
○ 보유 및 이용기간	: 5년(공공기록물관리법시행령 제26조①항)
○ 동의 거부 권리 및 동의 거부에 따른 불이익	: 개인정보의 수집·이용에 동의를 거부할 수 있으며, 이 경우 사이버보안빅데이터센터 이용이 제한됩니다.
※ 본 동의서는 개인정보보호법 제15조 ②항에 근거하여 작성되었습니다.	
▷ 개인정보 수집·이용 동의여부	동의 <input type="checkbox"/> 동의안함 <input type="checkbox"/>

- 개인정보 수집 이용 고지사항 확인 후 동의여부 체크
- 신청서 작성 완료 후 bigdata@krcert.or.kr 메일 전송
- 신청서가 적합한지 심사 이후 승인여부를 이메일로 안내
- 신청서 당 하나의 분석서버 제공
- 최종 방문일 기준 한 달 경과 시 가상머신 삭제

II 사이버보안빅데이터센터 이용



사이버보안빅데이터센터 이용절차

① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

2 - 1 센터 방문



- 한국인터넷진흥원 서울청사 (서울시 송파구 중대로 135(가락동 78) IT벤처타워 서관 8층)
- 이용 시간 : 평일 10:00 - 17:00 (토요일, 공휴일 휴관)
- 문의 : 02-405-5396

④ 퇴실

- 보안준수 서약서 검토 후 서명

사이버보안빅데이터센터 이용절차

① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

2 - 3 분석인프라 사용

- 이용신청 시 이메일 아이디 부분 사용
- 기본비밀번호 제공
- 로그인 후 변경

- 로그인 후 상단바 VM > KISA > 아이디 VM 선택
- 가상머신 연결 후 접속 완료

사이버보안빅데이터센터 이용절차

① 이용신청

② 방문 및 이용

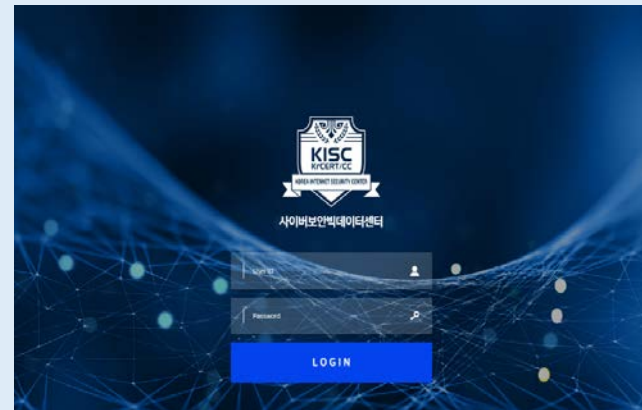
③ 분석결과 반출신청

④ 퇴실

2 - 3 분석인프라 사용



- Chrome 접속 > 사이버보안빅데이터센터 사이트 접속



- 가상머신 로그인과 동일 계정으로 로그인

사이버보안빅데이터센터 이용절차

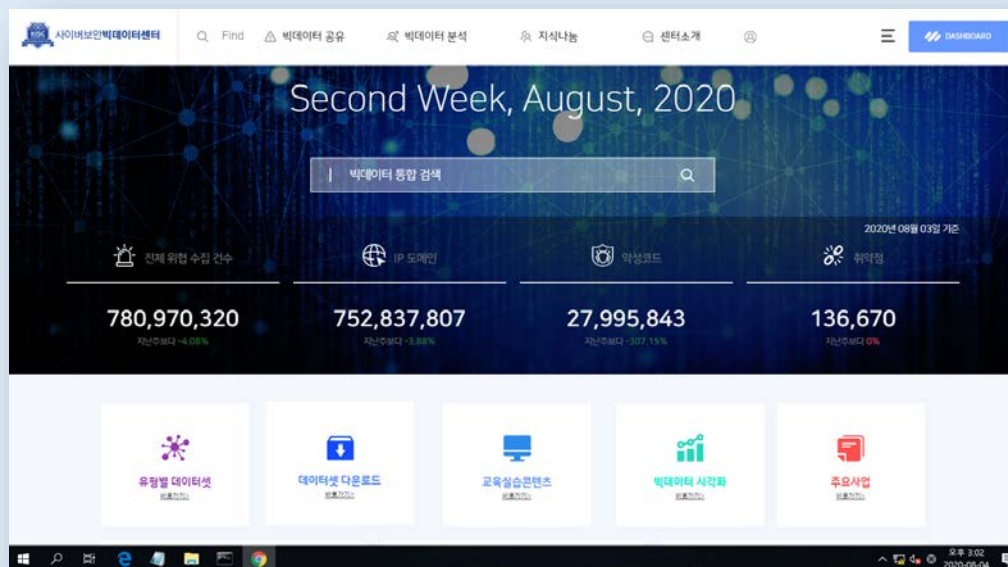
① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

2 - 3 분석인프라 사용



▪ 접속 완료

II 사이버보안빅데이터센터 이용



사이버보안빅데이터센터 이용절차

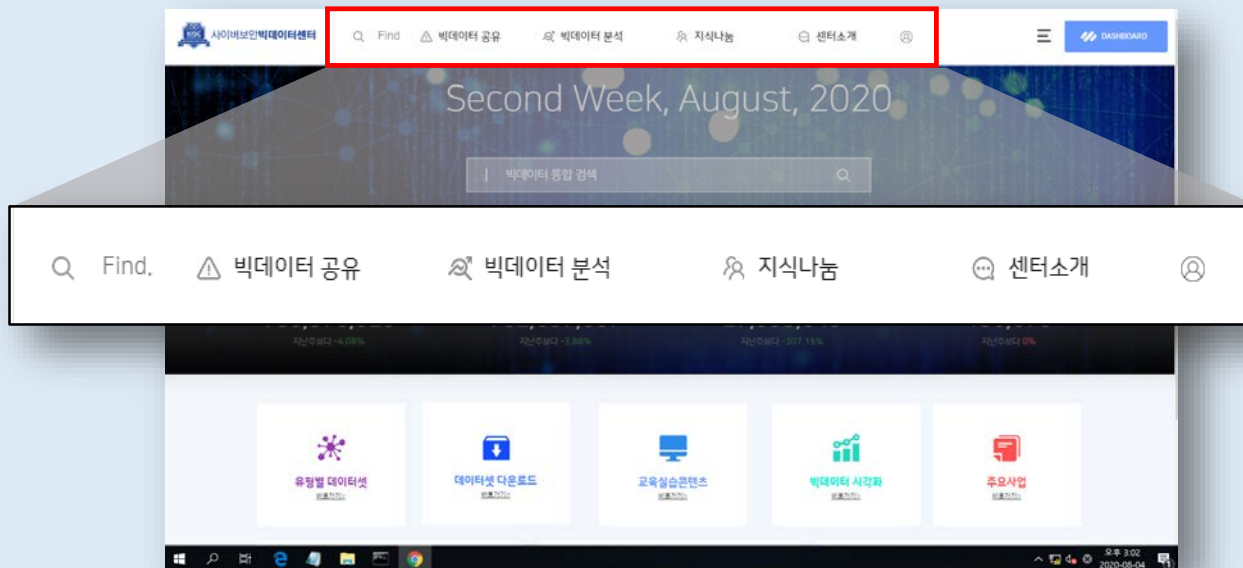
① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

2 - 4 사이버보안빅데이터사이트



사이버보안빅데이터센터 이용절차

① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

2 - 4 사이버보안빅데이터사이트 - 빅데이터 공유

① 전체 데이터

구분	설명
표준데이터	ASN, 정상도메인, 소프트웨어 최신버전정보, 기타 데이터셋
위협데이터	IP/ 도메인 - 경유지, 유포지, DDoS, C&C, 피싱, 파밍 등
	악성코드 - 랜섬웨어, 런처, 원격제어, 스팸 등
채널별 원본 데이터	수집채널(KISA) - 악성코드 유포 사이트, 스팸메일, 웹헬, 모바일 악성앱, 피싱사이트, 취약점 등
	수집채널(국내외) - 웹스크래핑, 포토스캔, 악성코드 등
	수집채널(OSINT) - 공격 IP, 명령제어서버, 피싱사이트 등
	수집채널(참여자) - 위협정보
K- 사이버 시큐리티 챌린지	웹서버 공격패턴, AI기반 악성 도메인 예측 데이터

II 사이버보안빅데이터센터 이용



사이버보안빅데이터센터 이용절차

① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

2 - 4 사이버보안빅데이터사이트 - 빅데이터 공유

② 위협 데이터

수집일시	채널	유형	IP	국가	도메인	DGA스코어	URL	포트
2020-08-05	virtuall	distroute			vmousconthangug	0.25	http://vmousconthangug.net/	
2020-08-05	virtuall	distroute			backgrounds.pk	0.00	http://backgrounds.pk/pic/0.0.0.0	
2020-08-05	virtuall	distroute			jamshed.pk	0.16	http://jamshed.pk/201.0.0.0	
2020-08-05	virtuall	distroute			www.advertad.net	0.00	http://www.advertad.net/advert/0.0.0.0	
2020-08-05	virtuall	distroute			www.advertad.net	0.00	http://www.advertad.net/advert/0.0.0.0	

II 사이버보안빅데이터센터 이용



사이버보안빅데이터센터 이용절차

① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

2 - 4 사이버보안빅데이터사이트 - 빅데이터 공유

③ 데이터셋 다운로드

번호	분류	현재상태	자료유형	등록일
456	vulnerability	일	STIX CSV JSON XML	2020-08
455	sample	일	STIX CSV JSON XML	2020-08
454	address	일	STIX CSV JSON XML	2020-08
453	vulnerability	일	STIX CSV JSON XML	2020-07
452	sample	일	STIX CSV JSON XML	2020-07
451	address	일	STIX CSV JSON XML	2020-07
450	vulnerability	일	STIX CSV JSON XML	2020-06
449	sample	일	STIX CSV JSON XML	2020-06
448	address	일	STIX CSV JSON XML	2020-06
447	vulnerability	일	STIX CSV JSON XML	2020-08

사이버보안빅데이터센터 이용절차

① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

2 - 5 사이버보안빅데이터센터 분석도구

- 빅데이터분석(13종), 인공지능(6종), 보안도구(11종), 시각화 도구(4종)

구분	용도	도구명
빅데이터 분석 도구	통계분석	R, Pandas/NumPy, SciPy, Spark SQL, Excel
	자연어처리	NLTK, KoNLP, spaCy, Gensim, Stanford CoreNLP
	그래프분석	Gephi, Spark GraphX, NodeXL
인공지능 도구	머신러닝	Spark MLlib, Scikit-learn, Weka
	딥러닝	TensorFlow, Keras, PyTorch
보안도구	시스템 보안	Sysinternals Suite, HxD, PEID, OllyDbg, Immunity Debugger, WinDbg, IDA Free, Yara
	네트워크 보안	WireShark, Fiddler, Snort
시각화 도구	데이터 시각화	Logstash, Kibana, Matplotlib/Seaborn, Tableau Public

사이버보안빅데이터센터 이용절차

① 이용신청

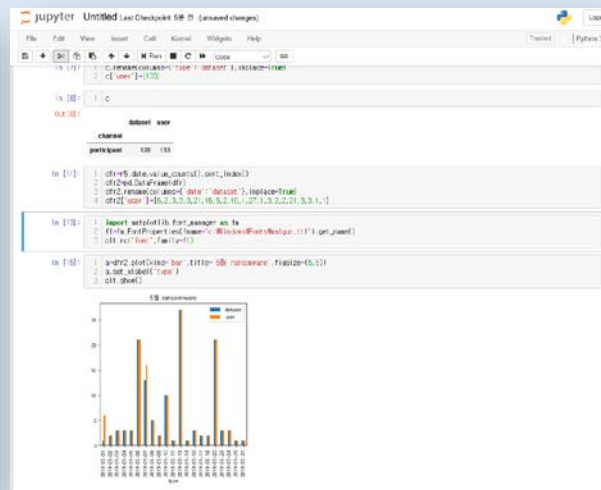
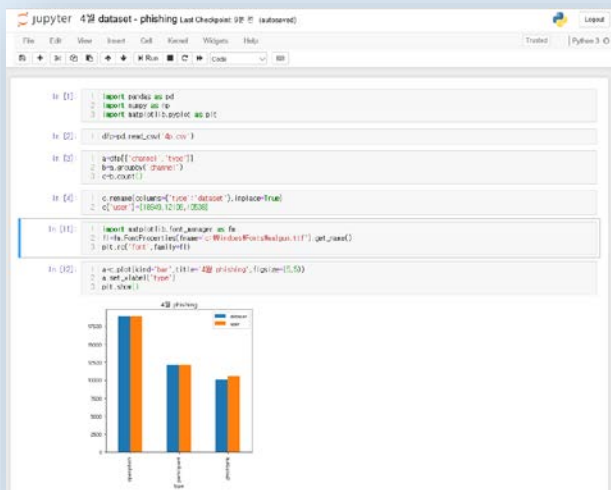
② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

2 - 6 분석 예시

- Jupyter notebook을 활용한 분석 예시



사이버보안빅데이터센터 이용절차

① 이용신청

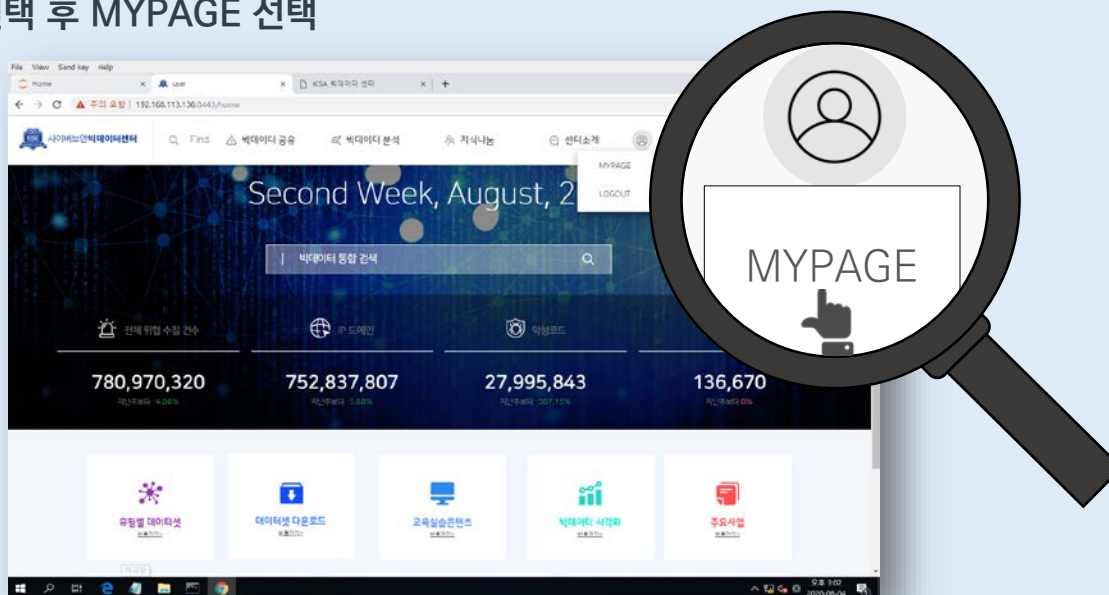
② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

③ - 1 반출 신청하기

① 아이콘 선택 후 MYPAGE 선택



사이버보안빅데이터센터 이용절차

① 이용신청

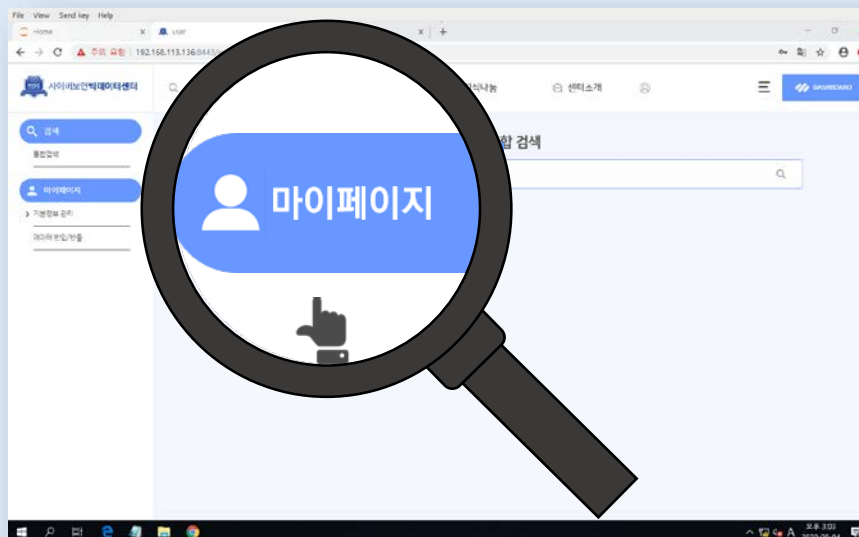
② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

③ - 1 반출 신청하기

② 아이콘 선택 후 마이페이지 선택



사이버보안빅데이터센터 이용절차

① 이용신청

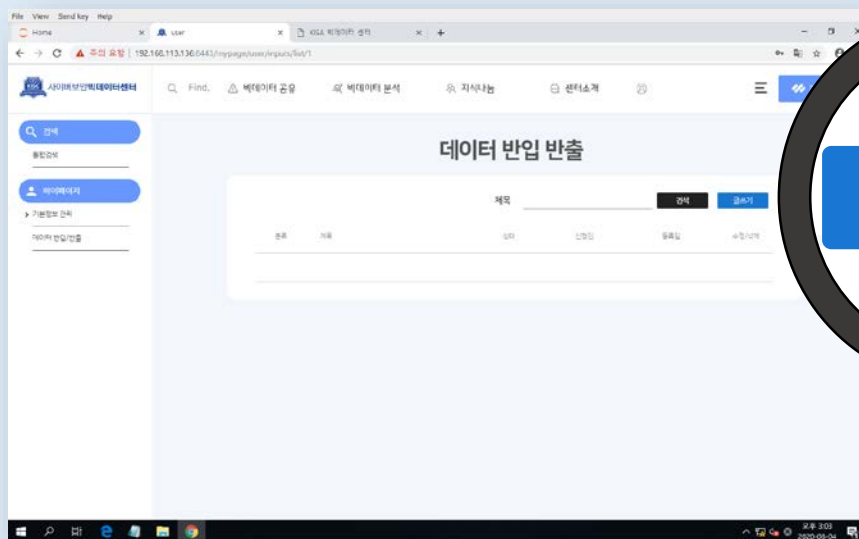
② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

③ - 1 반출 신청하기

③ 글쓰기 버튼 클릭



사이버보안빅데이터센터 이용절차

① 이용신청

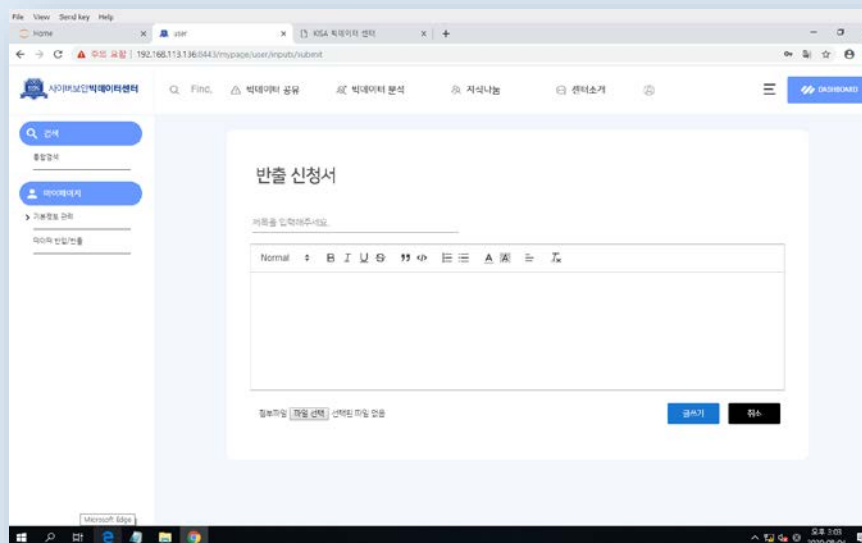
② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

③ - 1 반출 신청하기

④ 반출 신청서 작성 후 글쓰기 버튼 클릭 -> 신청 완료



사이버보안빅데이터센터 이용절차

① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

③ - 2 반출 심의

- 분석결과 반출여부 검토 (3-4일 소요)
- 반출 승인 시 이용신청서에 기입한 이메일로 자료 전송
- 원시데이터 사용유무 확인
- 데이터 유출, 보안위배사항 점검

사이버보안빅데이터센터 이용절차

① 이용신청

② 방문 및 이용

③ 분석결과 반출신청

④ 퇴실

4 방문의견서 제출

- 방문의견서 작성 후 제출

방문의견서 작성 후 제출

한국인터넷진흥원 KISA

「사이버보안빅데이터센터 방문의견서」

o 날짜 및 장소 : 년 월 일(요일), KISA 서울청사 8층 사이버보안빅데이터센터

이름		소속/직책	
휴대폰번호		서명	
의견			

<개인정보 수집이용>
o 개인정보 수집 이용 목적 : 의견수렴
o 수집하는 개인정보 항목 : 이름, 소속, 휴대전화번호
o 보유 및 이용기간 : 소독세입 등 관련업무에 따른 보유기간

※ 근거법령 : 개인정보보호법 제15조, 제24조의2

동의 ☐ 동맥 ☐ 동맥안함 ☐

사이버보안 빅데이터 활용 강화 활동



사이버보안 빅데이터 활용 강화 안내

2020년 교육 · 챌린지

1

빅데이터 교육

- ☑ 공통·활용·심화과정
- ☑ 보안 빅데이터 활용 경험

2

아이디어 챌린지

- ☑ AI 데이터셋 아이디어 공모
- ☑ 이메일, 웹로그 분야 빅데이터 분석 챌린지

2020년 사이버보안 빅데이터 활용 강화

CYBER SECURITY

BIGDATA

사이버보안 분야 빅데이터 활용 활성화를 위한 교육·챌린지 프로그램을 소개합니다. 사이버보안 분야 빅데이터 활용 활성화를 위한 교육·챌린지 프로그램을 소개합니다. 사이버보안 분야 빅데이터 활용 활성화를 위한 교육·챌린지 프로그램을 소개합니다.

교육개요

교육대상: 사이버보안 분야 빅데이터 활용 활성화를 위한 교육·챌린지 프로그램을 소개합니다. 사이버보안 분야 빅데이터 활용 활성화를 위한 교육·챌린지 프로그램을 소개합니다. 사이버보안 분야 빅데이터 활용 활성화를 위한 교육·챌린지 프로그램을 소개합니다.

교육구분

구분	대상	교육일	교육시간	교육장소	비고
공통	초·중·고	9/14(수) - 9/15(목)	9:00 - 12:00	서울시교육청	온라인
	고1	9/14(수) - 9/15(목)	9:00 - 12:00	서울시교육청	온라인
	고2	9/14(수) - 9/15(목)	9:00 - 12:00	서울시교육청	온라인
	고3	9/14(수) - 9/15(목)	9:00 - 12:00	서울시교육청	온라인
활용	초·중·고	9/14(수) - 9/15(목)	9:00 - 12:00	서울시교육청	온라인
	고1	9/14(수) - 9/15(목)	9:00 - 12:00	서울시교육청	온라인
	고2	9/14(수) - 9/15(목)	9:00 - 12:00	서울시교육청	온라인
	고3	9/14(수) - 9/15(목)	9:00 - 12:00	서울시교육청	온라인
심화	초·중·고	9/14(수) - 9/15(목)	9:00 - 12:00	서울시교육청	온라인
	고1	9/14(수) - 9/15(목)	9:00 - 12:00	서울시교육청	온라인
	고2	9/14(수) - 9/15(목)	9:00 - 12:00	서울시교육청	온라인
	고3	9/14(수) - 9/15(목)	9:00 - 12:00	서울시교육청	온라인

주최: 한국인터넷진흥원, 후원: 한국인터넷진흥원, 후원: 한국인터넷진흥원

감사합니다.