모바일 포렌식 시각화 툴 개발

K-shield 주니어 보안사고 분석대응반 월목반 김석진 | 월목반 문재식 월목반 민서정 | 월목반 박귀수 월목반 송어진 | 월목반 이용하 월목반 이재훈 | 월목반 이지호 월목반 한택승

목차

1.	프로	젝트 개요	3
	1.1	프로젝트 명 및 기간	3
	1.2	프로젝트 목적	3
	1.3	기대 효과	3
	1.4	프로젝트 배경	3
	1.5	프로젝트 목표	4
2.	프로젝트 조직		
	2.1	프로젝트 조직도	5
	2.2	프로젝트 책임 및 역할	5
3.	프로젝트 수행 일정		
	3.1	프로젝트 추진 일정	6
	3.2	프로젝트 세부 절차	6
	3.3	프로젝트 개발 시 사용한 라이브러리 및 공유자원	7
4.	분석	용 데이터 생성 및 추출	8
	4.1	사용 기기 루팅	8
	4.2	추출한 데이터	9
5.	GUI	화면 구성	12
6.	시각화1		
7.	추후 연구 예정2		
8	과려	자료	27

- 1. 프로젝트 개요
- 1.1 프로젝트 명 및 기간

프로젝트 명: 모바일 포렌식 시각화 툴 개발

프로젝트 기간: 2021.03. ~ 2021.06. (2.5개월)

프로젝트 수행 팀: 티미룸

1.2 프로젝트 목적

모바일 포렌식을 이용하는 사람들에게 기존에 있는 모바일 포렌식 도구들 보다 더 보기편하고 무료로 사용할 수 있는 툴을 제공하는 것이다. 또한 입문자가 쉽게 모바일 포렌식을 할수 있도록 도와주는 것이다.

1.3 기대 효과

- 고도의 경험을 가진 숙련자 외에 입문자 또한 쉽게 모바일 포렌식이 가능
- 모바일 포렌식 업을 하는 사람들이 보기 편한 모바일 포렌식 도구를 이용 가능

1.4 프로젝트 배경

현재 스마트폰 시장은 급속도로 팽창하는 중이다. 그 중에서도 안드로이드 스마트폰의 성장은 더욱 더 두드러진다. 모바일 기기의 사용이 증가하면서 모바일 포렌식에 대한 수요도 증가하고 있다. 최근 다양한 뉴스 기사에서 확인할 수 있다.

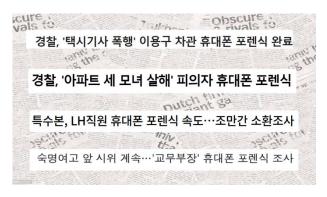


그림 1 모바일 포렌식 관련 기사

우리는 2021 년 기준 4 월 안드로이드 시장 점유율이 72.2%로 가장 높기에 우리는 안드로이드 운영체제를 타깃으로 잡았다. 오픈 소스의 장점을 기반으로 여러 제조사들의 다양한 기기 출시와 꾸준한 안드로이드 OS 업데이트는 사용자들에게 장점으로 부각되지만, 디지털 포렌식 관점에서는 어려움으로 작용하고 있다. 이는 논리적 포렌식은 대부분 루팅을 전제로 하는데, 다양한 기종과 지속적인 업데이트는 루팅을 어렵게 하기 때문이다.



그림 2 2021년 4월 세계 모바일 os 이용 비중

Android	ios	Samsung	Unknown	Windows	Tizen
73.64%	26.01%	0.34%	0.01%	0%	0%
Mobile Operating System Market Share in Republic Of Korea - April 2021					

그림 3 2021년 4월 국내 모바일 os 이용 비중

현재 상용화 되어있는 도구들은 비용적인 측면에서 개인이 쉽게 사용이 불가능하며, 기존에 있는 오픈소스에 의존할 수밖에 없다. 하지만 기존에 있는 오픈 소스들은 최신의 기기들에 적용하기 어려운 부분들이 존재하고, 시각화 부분에 아쉬움을 보인다. 우리는 이러한 현실태에서 모바일 논리적 이미징을 통해 사용자 위주의 행위에 대한 데이터를 추출하고 이를 시각화 작업을 통해서 더욱 분석하기 편리한 모바일 포렌식 시각화 도구를 제작할 것이다.

1.5 프로젝트 목표

- 휴대폰에서 추출한 데이터를 이해하기 쉽게 시각화하는 툴 만들기
- 위치정보와 시간정보를 바탕으로 사용자의 행위를 파악하기 쉽게 시각화

2. 프로젝트 조직

2.1 프로젝트 조직도

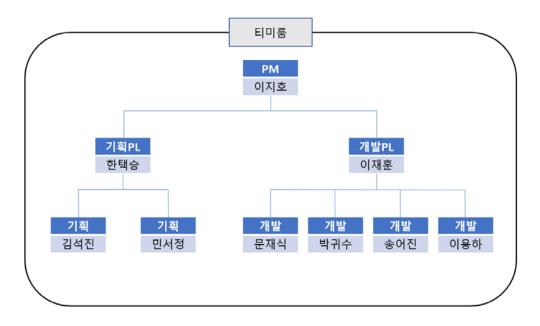


그림 4 프로젝트 팀 도식화

2.2 프로젝트 책임 및 역할

책임	역할	담당자명
PM	프로젝트 총괄 책임자	이지호
개발 PL	개발 책임자	이재훈
기획 PL	기획 및 조사 책임자	한택승
개발	장고를 이용한 대시보드 생성 및 시각화	문재식
개발	GUI 프로그램 개발, 시각화	박귀수
개발	장고를 이용한 대시보드 생성 및 시각화	송어진
개발	장고를 이용한 대시보드 생성 및 시각화	이용하
기획	기획 및 자료 조사, 케이스 생성	김석진
기획	기획 및 자료 조사, 케이스 생성	민서정

표 1 책임 및 역할

3. 프로젝트 수행 일정

3.1 프로젝트 추진 일정



그림 5 WBS

3.2 프로젝트 세부 절차

구분	수행업무	설명	담당자
착수	프로젝트 착수	프로젝트 수행 절차 및 일정 계획 수립	모두
준비 기초자료 수집		모바일 포렌식 관련 논문, 문서 등 자료수집	모두
	프로젝트 방향 설정	프로젝트 진행 방향 및 방법 설정	모두
	휴대폰 확보	안드로이드 기기 확보 및 버전 확인	이지호
	데이터 생성	확보된 기기에 데이터 쌓기	김석진
분석	데이터 파악	추출할 수 있는 데이터 분석	모두
개발	GUI프로그램 개발	Qt5를 이용하여 프로그램을 GUI로 사용할 수	박귀수
		있도록 개발	
	웹 서버 구축	장고를 이용한 웹 서버 구축 및 모델 구축	이재훈
	데이터 추출 및 분류	tar 데이터 추출 및 분류 자동화	이지호
	데이터 시각화	와이파이, 캘린더 데이터 시각화	박귀수
		타임라인 개발	이용하
		전화, 메시지 시각화 및 페이지 이동 개발	송어진
		미디어 시각화, 브라우저 및 메시지 기록 시각화	이재훈
		및 키워드 추천 개발	
		위치 정보 시각화	문재식,
			이재훈

표 2 프로젝트 인원 별 역할

3.3 프로젝트 개발 시 사용한 라이브러리 및 공유자원

구분	TYPE	Spec
개발 언어	Python	3.6.0
웹 프레임워크	Django	3.2.2
오픈 API	Google map API	

- 4. 분석용 데이터 생성 및 추출
- 4.1 사용 기기 루팅



그림 6 A8star 이미지

사용 기기 Galaxy A8 Star - 버전 안드로이드 10(최신버전 안드로이드 11) 루팅 과정

Frija 에서 순정 펌웨어를 찾은 후 오딘을 이용하여 순정 펌웨어 설치

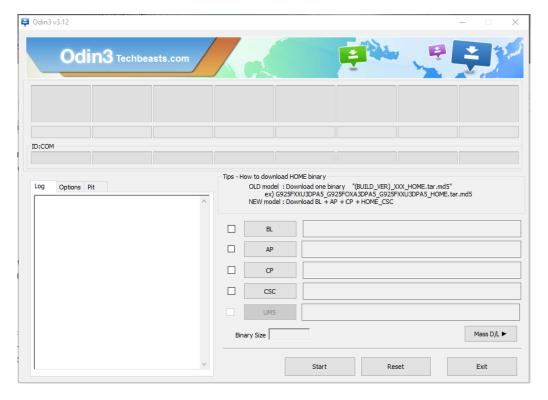


그림 7 odin3 실행창

Magisk 를 이용한 루팅



그림 8 Magisk 실행 화면

4.2 추출한 데이터

루팅을 하여 뽑아낸 데이터는 다음과 같다.

←Android 의 /data 안에 있는 폴더

그림 9 /data/



그림 10 /data/data/

↑추출한 데이터 안 userdata

userdata 안에는 db 파일, cache 등 관련 데이터가 들어 있다.

경로	설명
/data/data/[패키지	어플리케이션의 공유 설정 파일이 저장되는 영역
이름]/shared_prefs	

/data/data/[패키지	어플리케이션이 필요한 임시파일이 저장되는 영역
이름]/cache	
/data/data/[패키지	어플리케이션이 필요한 데이터베이스 파일이 저장되는 영역
이름]/database	
/data/data/[패키지	데이터베이스와 캐시를 제외한 어플리케이션에서 사용하는 일반
이름]/files	파일이 저장되는 영역
/data/data/[패키지 이름/lib	어플리케이션이 필요한 라이브러리 파일이 저장되는 영역

표 3 어플리케이션 데이터 설명

5. GUI 화면 구성

Timmy Room	- □ ×
분석할 모바일의 tar,build,prop 파일을 넣어주세요	tar 파일선택
	build 파일선택
	해제 경로
OK	Cancel

그림 11 실행 창

분석할 모바일의 /data/폴더는 모바일 기기의 userdata 가 들어있는 폴더로 이를 tar 로 압축한 압축파일을 경로에 넣는다. 모바일 기기의 정보가 담겨있는 build.prop 는 /system/의 밑에 있는 파일이다. 위의 두개의 파일은 루팅된 기기에서 추출해야 한다. 이후 파일위치를 선택하여 넣어준다. 이후 tar 파일에서 필요한 파일들을 추출하여 둘 위치경로를 선택한다.

pyQt5를 이용하여 파일을 선택할 수 있는 코드를 작성하였다.

```
def file_selecti(self): #파일 선택
FileOpen = QfileDialog.getOpenFileName(self, 'Open file', './', "datafile(*.dd *.tar )")#해당 확장자만 파일만 고를수 있도록
if(FileOpen[0]!=''):
    self.labeli.setText(f"tar 경로:{FileOpen[0]}")
    self.tarfilepath=FileOpen[0]
```

그림 12 파일 선택 코드

Tar 로 압축되어 있는 파일들을 압축 푸는 코드를 작성하였다.

```
def decompression(file_path, output_path):

tar = tarfile.open(file_path)
for folder in target_folder:
    if folder == './media/';
    for img in image_folder:
        image_name = set(tarinfo for tarinfo in tar.getmembers() if (tarinfo.name.startswith('./media/') and tarinfo.name.endswith(img)))
        tar.extractall(members=image_name, path=f"{image_path}/../mas/static/assets/images/")
        continue
    target_name = set(tarinfo for tarinfo in tar.getmembers() if tarinfo.name.startswith(folder)) # target_folder에 있는 폴더의 이름을 가진 모든 tarinfo를 target_name으로 둔다.
    tar.extractall(members=target_name, path=output_path)

tar.close()
```

그림 13 tar 압축 해제 코드

확인 버튼을 누르고 대시보드를 열어주는 코드를 작성하였다.

```
def ok(self):
    if self.tarfilepath!='' and self.buildfilepath!=''and self.outputpath!='':#파일 경로를 지정하였다면 시스템 명령으로 대시보드를 띄운다.
    try:
        path=os.path.dirname(os.path.abspath(_file__))
        self.progress(self)
        f=open(f"(path)/../경로.txt",'w')
        self.progress(self)
        factore for filepath)\n{self.outputpath}" #파일 경로를 기록한다. 추후에 was에서 읽음
    f.write(data)
    f.close()
        self.progress(self)
        shutil.copy(f"(self.buildfilepath)", f"{path}/../was/app/")#build.prop 파일 복사
    self.lapol1.setText("127.0.0.1:8000 으로 접속중입니다...기다려 주세요")
        self.label2.clear()
        self.label3.clear()
        self.label1.repaint() #객체의 label1을 다시 repaint 해준다.
        self.label3.repaint()
        read_tar.decompression(self.tarfilepath,self.outputpath)
        self.progress(self)
        subprocess.run(f'python {path}/../was/manage.py runserver',shell=True,timeout=0.5)
    except Exception as e:
        print(e)
        webbrovser.open("http://127.0.0.1:8000",1)#해당 url을 새 창으로 연다.
        sys.exit()
```

그림 14 대시보드 및 웹 시작 코드

이를 통하여 GUI 로 데이터를 선택하고 압축해제한 데이터들을 다음과 같은 과정을 거쳐서 대시보드에 보여준다. (example) contacts2.db

Db 의 각 테이블을 객체화 한다.

```
class contacts_model(models.Model): #이전

class Meta:
    managed = False
    app_label = "contacts"
    db_table = 'search_index'

id = models.AutoField(db_column='contact_id',primary_key=True)
name = models.TextField(db_column='sec_name',blank=True,null=True)
number = models.TextField(db_column='tokens',blank=True,null=True)
```

그림 15 DB 테이블 객체화 코드

DB의 경로를 선택한다.

그림 16 DB 경로 설정 코드

DB 라우터를 통해 정의한 모델이 각 데이터베이스를 찾아갈 수 있게한다.

```
class MultiDBRouter(object):
    def __init__(self):
        self.model_list =
        ['default', 'contacts', 'calllog', 'message', 'mms', 'map', 'chrome2', 'SAMINT', 'WebDowndata', 'Webext', 'Appinslog',
        'Media', 'Calendar', 'Kakao1', 'Kakao2']
    def db_for_read(self, model, **hints):
        if model._meta.app_label in self.model_list:
            return model._meta.app_label

        return None
    def db_for_write(self, model, **hints):
        return None
    def allow_relation(self, obj1, obj2, **hints):
        return None

    def allow_relation(self, db, app_label, model_name=None, **hints):
        return None
```

그림 17 DB 라우팅 코드

객체화한 모델을 통하여 필요한 데이터를 불러와 웹에서 시각화 해준다.

```
if context['url']=="contact.html": ##어진
page = request.GET.get('page', '1')

c=contacts_model.objects.values()
paginator = Paginator(c, 10)
page_obj = paginator.get_page(page)

context['calllog']=calllog_model.objects.raw("SELECT_id,datetime((date / 1000), 'unixepoch','localtime') AS date FROM calls where m_content IS NULL ORDER BY date ASC ") #메시지 아닌것만
context['contact']=c
context['page']=page_obj
```

그림 18 DB 데이터

6. 시각화

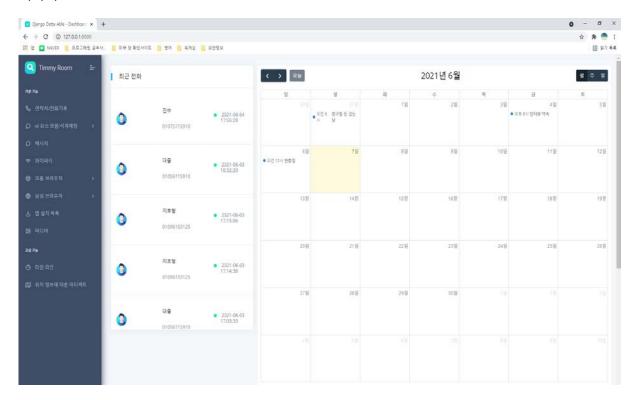


그림 19 메인화면1



그림 20 메인화면2

탭 안에는 연락처/전화기록, 메시지, 브라우저, 와이파이, 미디어, 앱 설치목록, 타임라인, 위치정보에 따른 아티팩트, 카카오톡 등이 있다. 메인 화면 안에는 캘린더와, 최근 전화, 기기 정보 등이 존재한다. 기기정보는 크롤링을 사용하여 기기의 사진을 불러온다.

크롤링 코드 (현재는 구글에 robots.txt 에서 disallow 되어있는 곳에서 하고 있는 중입니다. 추후 허용된 사이트로 변경 혹은 이미지를 추가하여 진행할 예정입니다.)

```
### Second Sec
```

그림 21 크롤링 코드

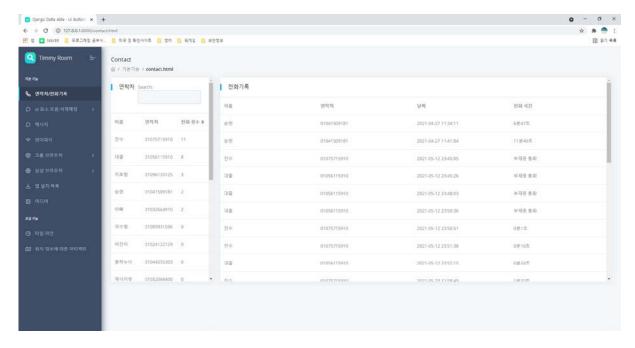


그림 22 연락처/ 전화기록 화면

화면은 연락처와 전화기록으로 되어있으며, 검색으로 검색도 가능하고 연락처의 이름을 클릭하여 전화기록에 그 사람과 연락한 내용만 보게 할 수 있다. 연락처의 원하는 사람의 전화기록을 보는 기능을 활용하면 다음과 같이 된다.

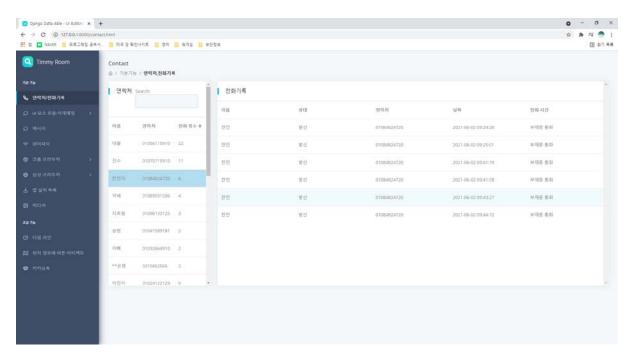


그림 23 원하는 연락처 번호 전화기록 표시

원하는 전화기록만 볼 수 있게 하는 코드를 작성하였다.

```
var ctl="((contact|length)]" //업략처 환경이
var cll="((colling|length)]" //업략기를 팬함이
f(document)-(ready(inputch())() (건략기를 팬함이
f(document)-(ready(inputch()))() (건략기능(in)-ind()
var count-s(*scalling)-find(*tr)-(su())-itsal()
);
i(*tr)-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-(su())-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-itsal()-it
```

그림 24 원하는 전화기록만 선택하는 코드

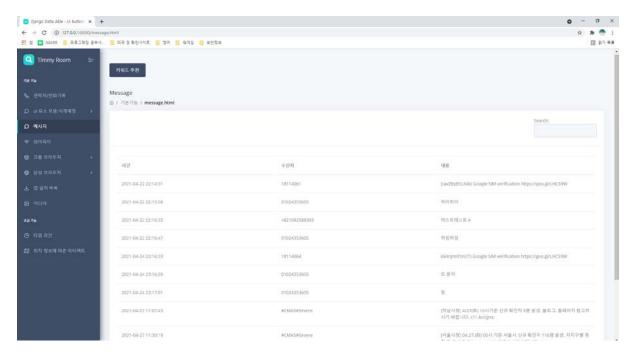


그림 25 메시지 화면

메시지, 브라우저의 검색 기록 등은 키워드 추천기능이 있어서 어떤 키워드를 많이 쳤었는지 보여줄 수 있다. 또한 메시지로 도착한 사진을 화면에 보여줄 수 있다.

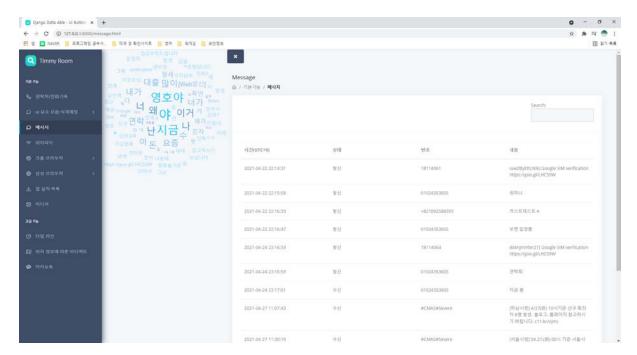


그림 26 키워드 추천 기능 화면

키워드를 추천하는 코드를 다음과 같다.

그림 27 키워드 추천 코드

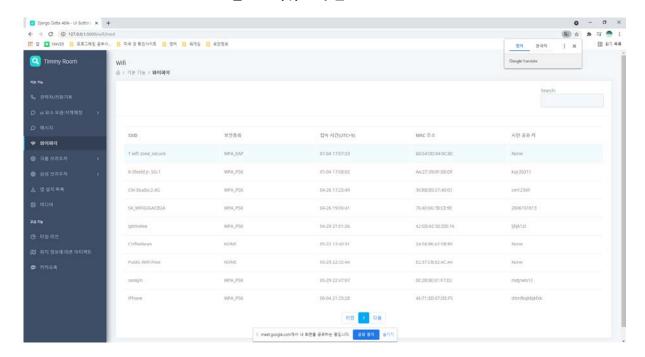


그림 28 와이파이 화면

AP name, 최초 접속시간, MAC 주소, 비밀번호 등의 정보들이 나타난다.

와이파이는 .xml 파일에 들어있어 파싱하여 데이터를 보여준다. 코드는 다음과 같다.

그림 29 xml파일 시각화 코드

브라우저 탭에는 브라우저 히스토리, 브라우저 다운로드, 브라우저 검색 기록 등이 있으며 각 화면은 다음과 같다.

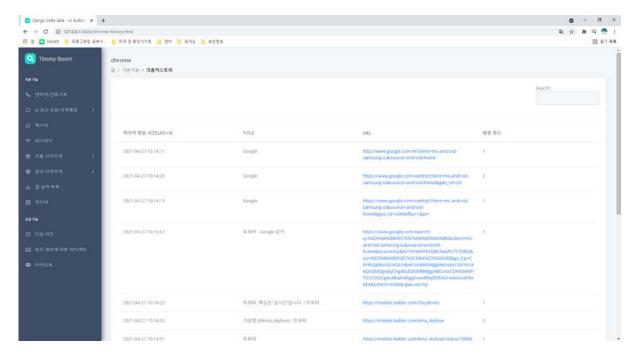


그림 30 크롬 히스토리 화면

마지막 방문 시간, TITLE, URL, 방문 회수 등이 나오며 URL을 클릭할 경우 인터넷과 연결하여 페이지를 보여준다.

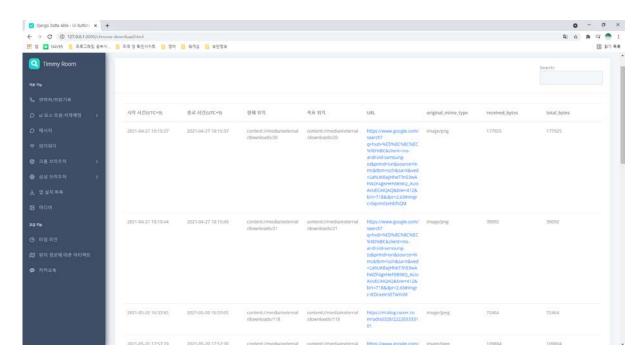


그림 31 브라우저 다운로드 화면

시작 시간, 종료 시간, 현재 위치, 목표위치, URL, orinial_mime_type, received_bytes, total_bytes 등 의 컬럼이 존재하며 URL을 누르면 받았던 페이지로 이동한다.

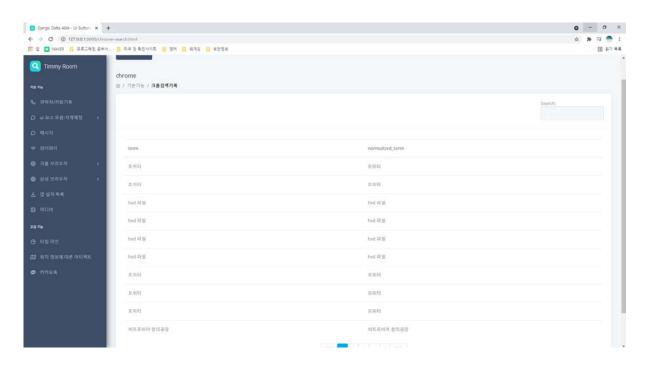


그림 32 브라우저 검색 기록 화면

브라우저 검색기록은 구글에서 검색한 내용 등을 기록하고 있으며 이를 통하여 키워드 추천으로 자주 검색했던 내용을 볼 수 있다.

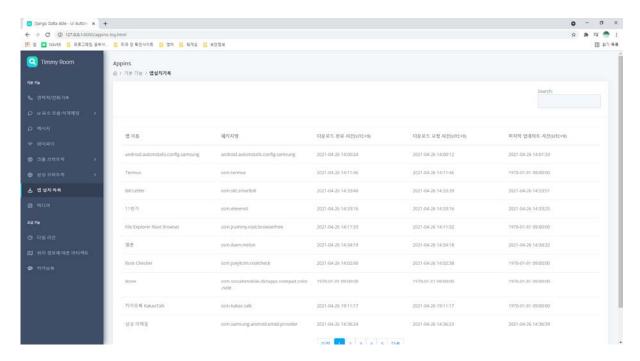


그림 33 앱 설치 목록 화면

앱을 설치한 목록들을 확인할 수 있으며, 앱 이름, 패키지명, 다운로드 완료 시간, 다운로드 요청 시간, 마지막 업데이트 등의 기록을 알 수 있다.

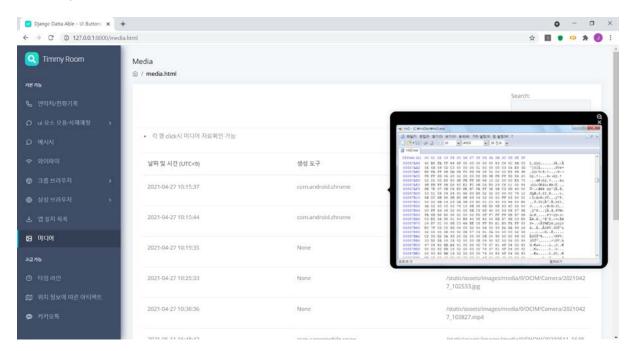


그림 34 미디어 화면

사진들의 시간, 생성 도구 이미지 경로 등이 나타나고 원하는 행 클릭 시 해당 이미지가 보인다.

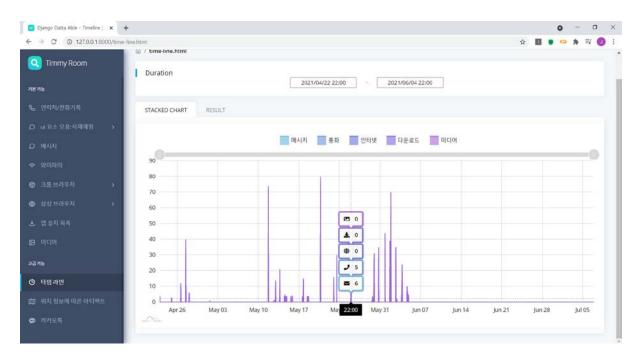


그림 35 타임라인 화면

STACKED CHART 탭에서는 메시지, 통화, 인터넷, 다운로드, 미디어 등의 시간기록을 바탕으로 하여 빈도수를 각각 표시하여 그래프로 보여준다.

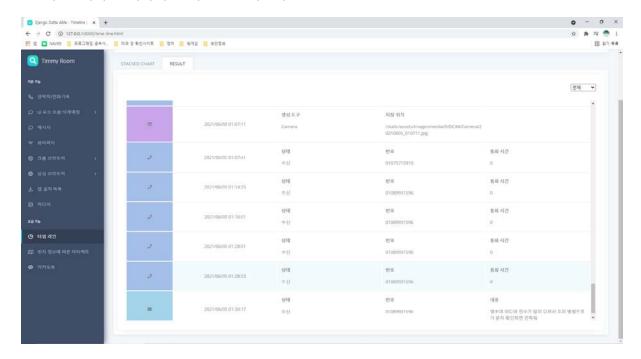


그림 36 타임라인 화면2

RESULT 탭에는 문자, 통화, 인터넷, 다운로드, 미디어 등의 것들로 시간안에 있었던 데이터들을 상세하게 볼 수 있다.

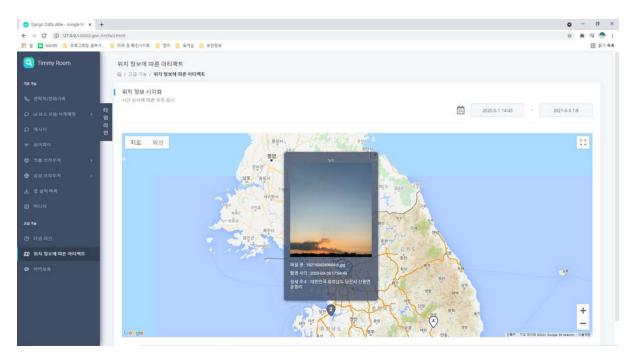


그림 37 위치 정보에 따른 아티팩트 화면

오른쪽에 있는 date picker 의 기간을 조정하여 원하는 시간대의 사진을 화면의 마크로 순서대로 보여준다. 이를 통해 동선을 파악할 수 있다. 마크를 클릭하면 사진을 볼 수 있다.

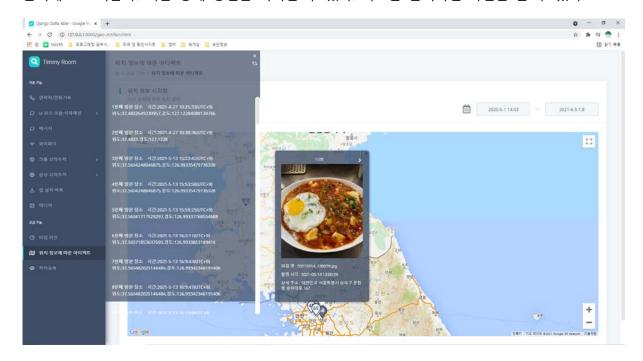


그림 38 위치 정보에 따른 아티팩트 화면2

타임라인 버튼을 이용하여 중복된 사진들을 하나하나 확인할 수 있다. 사진들의 시간을 보여주면서 위치정보가 같아 중복된 마크로 헷갈리게 되는 경우를 줄인다.

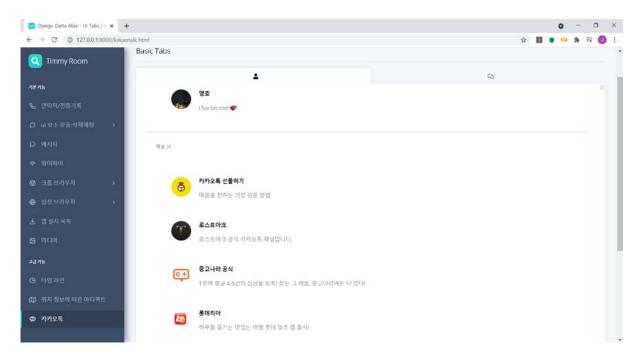


그림 39 카카오톡 화면

- 카카오톡 화면

카카오톡의 친구와 채팅 내역을 볼 수 있다.

카카오톡 복호화 코드가 존재하여 카카오톡 내용이나, 친구들을 볼 수 있으나, 개인정보 문제로 코드는 따로 올리지 않았습니다.(시나리오 데이터 포함)

- 7. 추후 연구 예정
- 사진데이터를 바탕으로 인물 식별 기능
- 키워드 분석 기능 -> 단순 추천이 아닌 연관할 수 있는 카테고리 분석
- 영상 데이터, 음성 데이터의 소리를 텍스트로 변환하여 보여주는 기능
- 삼성기기가 아닌 다른 기기들도 실행할 수 있도록 개발 예정
- Dd 파일에서 추출하거나, 더 나아가 휴대폰을 연결하면 바로 시각화 가능하도록 진행
- S노트 시각화 기능
- 원하는 데이터 PDF로 출력 기능

8. 관련 자료

https://gs.statcounter.com/os-market-share/mobile/worldwide/#monthly-202103-202103-bar

https://gs.statcounter.com/os-market-share/mobile/south-korea#monthly-202103-202103-bar

https://www.w3schools.com/

https://developer.mozilla.org/ko/docs/Web/JavaScript

https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artild=ART002550554

https://digitalis.postype.com/post/2290617

https://scienceon.kisti.re.kr/srch/selectPORSrchArticle.do?cn=JAKO201633056056055

https://www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE07480412