[Your data in Search](https://myaccount.google.com/yourdata/search?pli=1)

[개인정보처리방침 – 개인정보 보호 및 약관 – Google](https://policies.google.com/privacy?hl=ko#infocollect)

[검색 알고리즘의 원리](https://www.google.com/intl/ko/search/howsearchworks/algorithms/)

[구글 음성 정보 수집 끄는 방법 / 꺼야하는 이유](https://fluorite94.tistory.com/449)

사생활 정의 <https://www.dbpia.co.kr/pdf/pdfView.do?nodeId=NODE07524737>

한국 가정 법률 상담소에 자문을 구하자.

사생활의자유의의미<https://easylaw.go.kr/CSP/CnpClsMain.laf?popMenu=ov&csmSeq=574&ccfNo=2&cciNo=3&cnpClsNo=1&search_put=%EC%82%AC%EC%83%9D%ED%99%9C>



<https://docs.google.com/forms/d/1MzFL9WE--oKDojLoB0PuWc1pE-jmEp7HAV-tsVxukXo/edit#responses>

**제 12조**

어느 누구도 자신의 사생활, 가족관계, 가정, 또는 타인과의 연락에 대해 외부의 자의적인 간섭을 받지 않으며, 자신의 명예와 평판에 대해 침해를 받지 않는다. 모든 사람은 그러한 간섭과 침해에 대해 법의 보호를 받을 권리가 있다.

구글이 음성인식을 통해 데이터를 수집하고 있다.

표선고 대상으로 한 설문(설문조사)-> 이러한 경험이 있다는 사람이 유의미하게 나왔다.

여기서 구글 약관 중 음성인식을 통해 데이터를 수집하지 않는다고 명시되어 있음.

이 것이 만약 사실이라면 소비자의 사생활 침해, 개인정보 불법 수집 등의 윤리적 문제가 발생.

그렇기에 구글의 이러한 행위에 대해 의문을 품게 되었기에 이 실험? 탐구?를 계획, 진행하게 되었다.

**이론적 배경**

**알고리즘이란?**

알고리즘(algorithm)은 주어진 문제를 논리적으로 해결하기 위해 필요한 절차, 방법, 명령어들을 모아놓은 것입니다.

알고리즘의 5가지 조건

유한성

명확성

입력

출력

효과성

**구글 검색 알고리즘 원리**

웹페이지를 검색 색인에 정리해 두고 사용자가 원하는 정보를 찾을 수 있도록 제시하는 것입니다.

원리: 검색 알고리즘은 가장 유용한 정보를 제공하기 위해 검색어의 단어, 페이지의 관련성 및 유용성, 출처의 전문성, 사용자의 위치 및 설정과 같은 다양한 요소를 고려합니다.

구글이 알고리즘을 위해 수집하는 데이터(여기서 구글 정책 언급)

고유 식별자, 브라우저 유형 및 설정, 기기 유형 및 설정, 운영체제, 통신사명과 전화번호를 포함한 모바일 네트워크 정보, 애플리케이션 버전 번호가 포함됩니다. 또한 IP 주소, 비정상 종료 보고서, 시스템 활동, 요청 날짜와 시간, 리퍼러 URL 등 사용자의 앱, 브라우저, 기기와 Google 서비스의 상호작용에 대한 정보를 수집합니다.

여기 음성에 대한 데이터를 수집하지 않고 있다고 알 수 있음. 하지만 위 설문조사를 보았을 때 이러한 경험이 있다는 것은 윤리적 문제가 있을 수 있다고 생각함.

(여기서 설문조사 결과 언급)

알고리즘이란?

알고리즘(algorithm)은 주어진 문제를 논리적으로 해결하기 위해 필요한 절차, 방법, 명령어들을 모아놓은 것입니다.

구글 검색 알고리즘 원리

목적: 유용하면서도 관련성 높은 검색 결과를 신속하게 제공할 수 있도록 수 천 억 개의 웹페이지를 검색 색인에 정리해 두고 사용자가 원하는 정보를 찾을 수 있도록 제시하는 것

**사생활**

**사생활 침해**

사생활 침해란?

:일반적으로 개인에 관련된 여러 가지 정보(성별, 주소, 나이, 재산정도, 학력, 취미 등)들이 다른 사람에게 노출되거나 악용되는 것

대한민국 헌법 제 17조: 모든 국민은 사생활의 비밀과 자유를 침해받지 아니한다.

대한민국 헌법에서도 사생활 침해는 엄연히 금지되있다.

세계 인권 선언 30개조 중 12조: 어느 누구도 자신의 사생활, 가족관계, 가정, 또는 타인과의 연락에 대해 외부의 자의적인 간섭을 받지 않으며, 자신의 명예와 평판에 대해 침해를 받지 않는다. 모든 사람은 그러한 간섭과 침해에 대해 법의 보호를 받을 권리가 있다.

<https://www.dbpia.co.kr/pdf/pdfView.do?nodeId=NODE02066172>

사생활 침해 사례

구글 사생활 침해 논란

<https://www.etoday.co.kr/news/view/1495315>

**실험 방법**

준비물: 같은 조건의 구글 계정 3~5개, 음성인식이 탑재된 기기.

실험 장소: 조용한 공간.

실험 준비

같은 조건의 구글 계정을 준비. (3~5개)

여기서 같은 조건의 구글 계정이란: 활동기록이 같고, 같은 카테고리의 광고 혹은 추천 검색어가 나오고, 계정 생일, 위치, 국적, 전화번호가 같고, 권한이 최대로 낮추어져 있는 계정.

조용한 장소 준비.

실험 과정

구글계정 한개는 최소, 한개는 최대, 한개는 음성인식만 끈다.

같은 활동(검색)을 한다. 같은 시간, 같은 위치에서.

여기서 어떠한 주제에 대해 이야기한다.(대화주제와 검색주제가 겹치지 않아야한다.)

대화전 주제1에 관한 검색을 모든 기기에서 한뒤 주제2와 관련된 대화를 컴퓨터가 켜져 있는 상태에서 한다.

실험 결과

IF) 3개의 계정에서 각각 광고 혹은 추천 검색어가 다르게 나온다면,

구글은 음성인식을 통해 우리의 대화를 수집하고 있다.

IF) 3개의 계정에서 다 같은 광고 혹은 추천검색어가 같게 나온다면,

실험이 잘못 됐다. or 음성인식 기능을 꺼 놓아도 음성인식을 통해 데이터를 수집하고 있다.

IF) 위의 결과 모두에 해당되지 않는다면,

재실험(마이크 등 재설정 후)

반복

통계

| 실험 횟수 | 기기A(권한 최소) | 기기B(음성만 OFF) | 기기C(권한 최대) |
| --- | --- | --- | --- |
| 1 |  |  |  |
| 2 |  |  |  |
| 3 |  |  |  |
| … |  |  |  |

**결론 및 제언**

**법 or 윤리**

윤리가 좀더 큰 거 같기도 .

결론이 구글에 메일을 보내야 한다.

법 <https://www.law.go.kr/%EB%B2%95%EB%A0%B9/%EC%95%BD%EA%B4%80%EC%9D%98%EA%B7%9C%EC%A0%9C%EC%97%90%EA%B4%80%ED%95%9C%EB%B2%95%EB%A5%A0>

약관법이 이미 있음┐

│

하지만 구글의 음성인식 약관을 동의 한 경우 이 법이 효력을 잃을 수 있다. (여기서는 윤리와 역자). 아무리 동의를 했더라도 사생활 침해인가?에 대한 의문을 던질 수 있음.

IF) 결과가 유의미하게 나온다면 캠페인 및 피켓팅 진행.

**1 5 10 2**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**구글이 음성인식을 활용하여 사생활 침해의 위협이 있는 행위를 하고 있는가?(후보)**

**구글의 음성인식 데이터 수집을 이용한 검색 알고리즘 형성에 대한 윤리·법적 문제 탐구.**

컴퓨터 과학 융합탐구 팀프로젝트 조원: 김누리, 이태혁, 정휘준

**Ⅰ. 연구의 필요성 및 목적**

이 연구는 구글이 음성인식이라는 기능을 활용하여 사람들의 사생활을 침해, 그리고 그를 통한 수익 창출에 관하여 과연 구글은 위법 행위를 하고 있는가? 라는 질문을 생성하여 연구를 한뒤 그 결과를 사람들에게 알림으로써 인터넷으로부터 안전한 삶을 보장받게 하려는 목적을 가지고 있다.

이 연구를 통해 우리는 구글이 음성인식이라는 기능을 윤리적인 허용범위 외에 따로 사생활을 침해할수 있는 행위를 하고 있는지를 알아보고자 이 연구를 진행하고자 한다.

**Ⅱ. 이론적 배경**

**1. 알고리즘**

**가. 알고리즘의 정의**

어떤 문제의 해결을 위하여, 입력된 자료를 토대로 하여 원하는 출력을 유도하여 내는 규칙의 집합.(네이버 국어사전)

**나. 구글 검색 알고리즘의 정의 및 정책**

웹페이지를 검색 색인에 정리해 두고 사용자가 원하는 정보를 찾을 수 있도록 제시하는 것.

원리: 검색 알고리즘은 가장 유용한 정보를 제공하기 위해 검색어의 단어, 페이지의 관련성 및 유용성, 출처의 전문성, 사용자의 위치 및 설정과 같은 다양한 요소를 고려한다. (Google 검색의 원리 중 검색 알고리즘)

**2. 사생활**

**가. 사생활의 정의**

1. 사생활의 의미

개인의 사사로운 생활 (네이버 국어사전)

1. 사생활의 자유의 의미

사생활의 자유는 사회공동체의 일반적인 생활규범의 범위에서 사생활을 자유롭게 형성해 나가고 그 설계 및 내용에 대해 외부로부터 간섭을 받지 않을 권리를 말한다. (찾기 쉬운 대한민국 법령 정보 中 인권침해)

**나. 사생활 침해**

1. 사생활 침해의 정의

일반적으로 개인에 관련된 여러 가지 정보(성별, 주소, 나이, 재산정도, 학력, 취미 등)들이 다른 사람에게 노출되거나 악용되는 것으로 타인의 사생활을 침해하면 법적 처벌을 받을 수도 있다.

1. 사생활 침해의 법적 고지

‘대한민국 헌법 제 17조: 모든 국민은 사생활의 비밀과 자유를 침해받지 아니한다.’, ‘세계 인권 선언 30개조 중 12조: 어느 누구도 자신의 사생활, 가족관계, 가정, 또는 타인과의 연락에 대해 외부의 자의적인 간섭을 받지 않으며, 자신의 명예와 평판에 대해 침해를 받지 않는다. 모든 사람은 그러한 간섭과 침해에 대해 법의 보호를 받을 권리가 있다.’ 대한민국 헌법 17조, 세계 인권 선언 30개조 중 12조를 보더라도 사생활은 침해되어서는 안 된다고 명시되어 있다.

1. 사생활 침해의 윤리적 문제

사생활 침해 피해자의

**Ⅲ. 연구 방법**

**1. 실험준비 및 시작**

**가. 계정 만들기**

올바른 연구 결과를 얻기 위해서는 사용할 계정들의 정보가 같아야 한다. 따라서 계정들의 아이디는 ‘phsteampro(번호)@gmail.com’ 으로 이름, 생년월일, 전화번호를 같게한다.

**나. 계정 권한 설정하기**

웹 및 앱 활동, 유튜브 기록, 광고 설정, 구글 검색의 개인 검색 결과 등 계정들의 권한을 다르게 설정해 구글에서 얻을 수 있는 사용자의 정보의 제한 정도를 다르게 했다. 이 연구에서는 모든 권한을 허용한 계정, 음성인식만을 허용한 계정, 유튜브 기록만을 허용한 계정을 사용했다.

3차 실험에서는 음성인식, 유튜브 기록 설정만을 다르게 하였다.

**다. 대화 전, 알고리즘 체크하기**

대화 전, 아무런 방문기록, 음성인식이 되어있지 않은 계정의 유튜브 영상(1줄에 4개, 10줄)과 나무위키, 코딩도장의 광고를 확인한다.

**라. 특정 주제로 대화하기**

팀원 3명이 컴퓨터를 켜두고 특정 주제로 약 1시간 동안 대화한다. 본 실험에서는 ‘게임’을 대화의 주제로 정했다.

**마. 대화 후, 알고리즘 체크하기**

대화가 끝나고 유튜브, 나무위키, 코딩도장의 페이지를 새로고침해서 나타나는 영상과 광고를 확인한다.

**Ⅳ. 연구 결과**

**1. 대화 전, 후의 유튜브, 나무위키, 코딩도장의 대화 주제 관련 알고리즘의 개수 비교**

1차 실험은 데이터가 유의미하게 나오지 않았다. 구글 계정 중 하나는 정지를 당했고 구글 애드센스와 관련없는 광고를 확인했다. 이는 구글의 음성인식으로 인해 변화한 알고리즘이 아니기 때문에 1차 실험의 데이터를 제외 시켰다.

2차 실험.

| 실험 데이터 | 권한 다끔 | 음성인식만 킴 | 다킴 |
| --- | --- | --- | --- |
| 2차실험 유튜브 | 전:게임 3 기타 37  후:게임 6 기타 34 | 전:게임 0 기타 40  후:게임 1 기타 39 | 전 : 게임 5 기타 35  후 : 게임 10기타30 |
| 2차실험 나무위키 | 전:게임 2 기타 1  후: 게임1 기타 2 | 전:게임 1 기타 2  후:게임 1 기타 2 | 전: 게임 1 기타 2  후: 게임 1 기타 2 |
| 2차실험 코딩도장 | 전: 게임 1기타 6  후:게임 4 기타 3 | 전:게임 0 기타 4  후:게임 1 기타 3 | 전: 게임 1 기타 6  후: 게임 2 기타 5 |

3차 실험.

| 실험 데이터 | 유튜브 기록 킴  음성인식 킴 | 유튜브 기록 끔  음성인식 킴 | 유튜브 기록 킴  음성인식 끔 | 유튜브 기록 끔  음성인식 끔 |
| --- | --- | --- | --- | --- |
| 3차 실험 유튜브 | 전: 게임0 기타 40  후: 게임6 기타 34 | 전: 게임 0 기타 40  후: 게임 1 기타 39 | 전: 게임 0 기타 40  후: 게임 1 기타 39 | 전: 게임1 기타 39  후: 게임0 기타 38 |
| 3차 실험 나무위키 | 전: 게임 1 기타 2  후: 게임 2 기타 1 | 전: 게임 1 기타 2  후: 게임 1 기타 2 | 전: 게임 0 기타 3  후: 게임 1 기타 2 | 전: 게임1 기타 2  후: 게임0 기타 4 |
| 3차 실험 코딩도장 | 전: 게임 0 기타 6  후: 게임 1 기타 5 | 전: 게임 2 기타 2  후: 게임 1 기타 3 | 전: 게임 0 기타 6  후: 게임 1 기타 5 | 전: 게임0 기타 7  후: 게임0 기타 4 |

**Ⅴ. 결론 및 제언**

**1. 결론**

데이터가 유의미하게 나오지 않은 1차 실험을 제외하면 공통적으로 유튜브 기록과 음성인식이 전부 다 켜져있는 기기에서만 유튜브에서의 대화 주제인 게임과 관련된 영상의 양이 증가하였다. 하지만, 나무위키, 코딩도장에서의 차이는 아무런 규칙이 나오지 않았다. 그렇다는 것은 음성인식 기능이 구글 애드센스 보다는 유튜브에 더 많은 영향을 끼치거나 구글 애드센스가 구글 음성인식 기능에 영향을 끼치지 않는다는 해석이 가능하다. 또는 구글과 직접적인 연관이 있는 사이트인 유튜브에서만 그 영향이 있는다는 해석도 가능하다.

본 연구에서는 구글이 알고리즘 이라는 방식에 사용될 정보들이 어떻게 입력되는지를 파악하고 그에 따른 사생활 침해를 우려함에 실행하게 되었다. 그리고 실험을 해본 결과 1차 실험에서는 준비 및 과정에서의 문제에 따른 결과의 오류가 발생하여 2차, 3차 실험을 이어서 하게 되었으며 그 결과 2,3차 실험에서는 유튜브의 기록저장 및 음성인식 기능을 키고서 게임에 대한 대화를 진행 했는데, 실험 후 유튜브의 알고리즘에서 눈에 띄게 대화 주제였던 게임에 관한 영상이 나오는 그 횟수가 증가함을 보여주었다. 이에 구글의 음성인식 데이터 수집은 이루어지고 있다고 결론을 도출했다.

**2. 제언**

본 연구에는 연구 횟수가 터무니 없이 적다. 실험 구성과 준비가 늦어졌고 융합 탐구 프로젝트 기간이 너무나 한정적이었기에 3회 밖에 실행하지 못한데다 그 중 하나는 데이터가 유의미하게 나올 수 없게 실험이 설계가 되었고 결과 또한 구글 계정이 하나가 정지당하여 무의미한 결과가 나왔다. 그리고 광고의 변화를 조사한 사이트의 광고가 거의 고정적인 알고리즘이 포함되어있었다. 나무위키의 경우에는 서브 컬쳐를 주로 다루는 사이트인 만큼 게임, 애니, 웹 소설 광고가 많이 나올 수 밖에 없고 코딩도장 또한 프로그래밍 교육 사이트로, 프로그래밍, 컴퓨터 등의 광고가 나오는 게 자연스러운 현상이었다.

하지만, 구글의 기본 설정에서는 음성인식이 꺼져있다. 그렇기에 실험 결과에 따르면 음성인식을 킨다는 것은 그런 데이터 수집에 동의한다는 뜻이고 이러한 경우에는 대한민국의 약관법에 접촉이 힘들 것이다. 그리고 구글 사용자 또한 권한 설정에 동의했기에 충분한 숙지가 있을 가능성또한 존재한다. 하지만, 여기서 중요시해야 할 것은 구글이 이러한 기술을 가지고 있고 실제로 이용자들을 대상으로 사용하고 있을 수도 있다는 것이다. 지금은 아닐 지 몰라도 언제든 범위를 확대가 가능하다는 것이다. 이에 따라 결론적으로 구글은 음성인식 기능을 끄더라도 우리의 사생활을 침해할 위험이 있는 정보수집을 할 수 있다고 해석 할 수 있다.

**참 고 문 헌**

Youtube.(2005).Youtube.URL:<https://www.youtube.com/>

나무위키.(2015).나무위키. URL:<https://namu.wiki/w/%EB%82%98%EB%AC%B4%EC%9C%84%ED%82%A4:%EB%8C%80%EB%AC%B8>

코딩도장.(불명).코딩도장.URL:<https://dojang.io/>

사생활.(불명).네이버 국어사전. URL:<https://ko.dict.naver.com/#/search?query=%EC%82%AC%EC%83%9D%ED%99%9C>

대한민국 헌법 부칙 제17조