# 유용한 데이터를 가진 SNS정리.



<그림 1. 연령대별 SNS 사용 현황>

## Facebook

* API로 정보를 가져오는 것은 공개 페이지에 관해서만 가능.
* 가장 양질의 데이터 이지만, 크롤링이나 새로운 API를 짜는 방법을 생각해야 한다.

## 트위터

* 키워드를 포함한 글을 검색 가능. 글의 내용, 리트윗 수를 알 수 있다.
* 스트리밍 서비스도 가능한 것으로 보인다.
* <https://docs.microsoft.com/ko-kr/azure/hdinsight/hdinsight-hbase-analyze-twitter-sentiment>
* 데이터의 질은 별로지만 뽑아낼 수 있는 데이터는 가장 많음

## 인스타그램

* API가 대폭 변경되어 거의 대부분의 작업이 불가능해짐.
* API를 제작하여 사용하는 수 밖에 없음.
* 보통 사진이 주 내용이지만, 태그들이 있기때문에 관심사 분석은 가능할 것으로 보임

## 익명 SNS

* 블라인드, 모씨, 어라운드 등이 있다.
* 연령대를 추정할 수 없으므로, 이 정보를 모으게 된다면 전체 연령대에서의 공통 관심사, 고민거리 등으로 이용할 수 있을 것이다.
* 데이터를 사용하려면 API를 만들어 써야한다.

## 관심사가 정해져 있는 SNS

* 관심사가 정해져 있기 때문에 어떤 한가지의 정보에 대해서 양질의 데이터를 뽑아 낼 수 있을 것이라고 생각.
* 하지만 연령대 추측이 힘들기 때문에 이 또한 전체 연령대에서의 데이터로 생각해야한다.

## 카카오스토리

* 기존 API는 포스팅만 가능.
* 크롤링도 주변 사람들만 가능.

## 싸이월드 미니홈피

* 기존 API에서 자신의 정보 위주로만 조회가 가능.
* 홈페이지 기반으로 크롤링이 가능할 것이라고 보이나, 사이트의 실제 사용량이 얼마 되지 않는 것 같음.