개인정보보호에서 비식별화의 기술적 문제점과 트렌드

이영환교수
건국대학교
경영대학 기술경영학과
정보통신기술대학원 금융IT 학과

Agenda

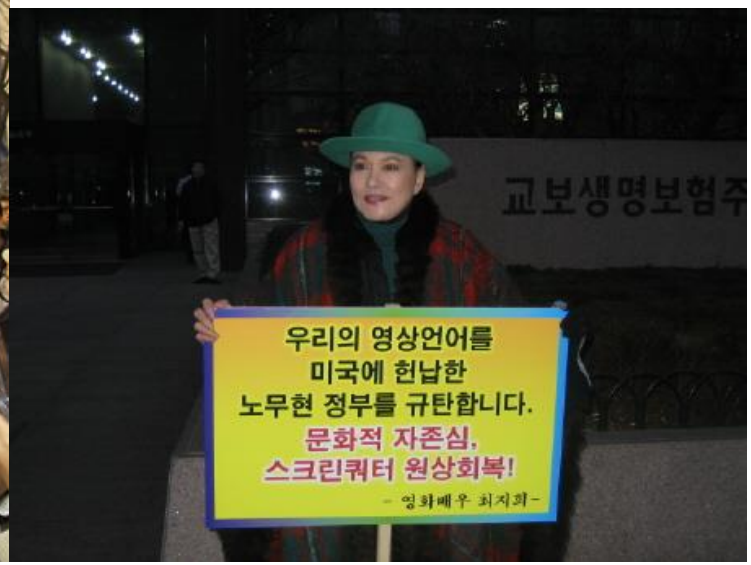
I. 비식별화와 법규제

II. 진화의 관점에서 본 데이터 산업

III. 결론

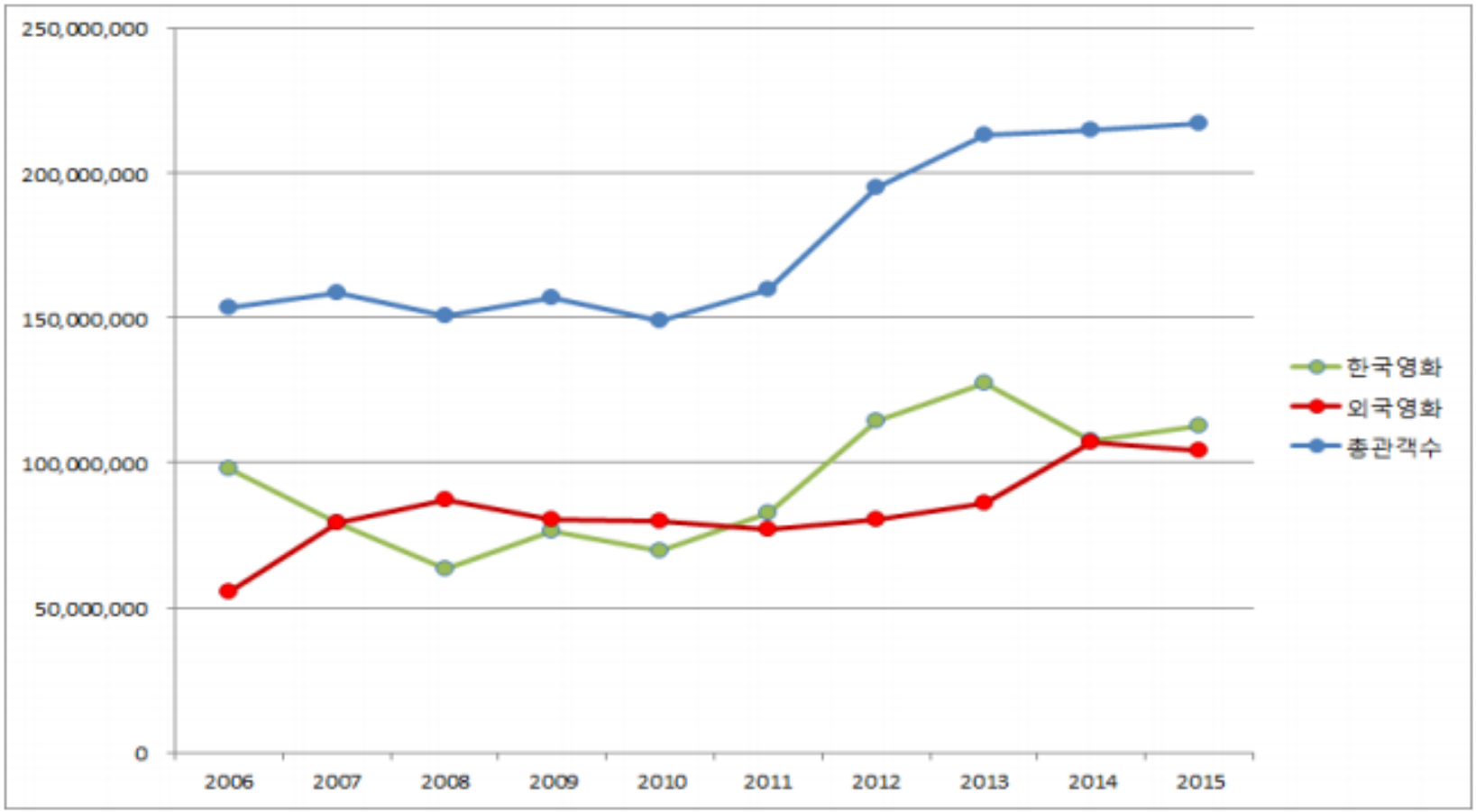
부록. 식별정보와 인증정보

영화 쿼터제



영화 쿼터제

<그림 1> 2006년-2015년 한국영화vs외국영화 극장 관객 수 추이



출처: 영화진흥위원회 산업정책연구팀. 2015 한국 영화산업 결산.

I. 비식별화와 법규제

● 개인정보보호: 식별정보와 인증정보 모두 보호

개인정보보호법이 보호하고 있는 것?

구분		근거	개인 식별 정보 항목
일반		· 개인정보보호법 제18조, 제23조, 제24조제1항, 제24조제3항, 제24조제2	· 주체자의 사생활을 침해할 수 있는 식별정보 (ex. 의료정보, 정신적 성향 등) · 주체자의 신분 확인을 위한 일반 식별정보 (ex. 이름, 주민등록번호, 주소 등)
공공부문		· 전자정부법 제42조	· 정당한 사용자임을 인증하는 식별정보 (ex. 인증서 일련번호, 유효기간 등)
		· 주민등록법 10조	· 신분 확인정보와 가족구성원 정보를 통해 확인될 수 있는 식별정보 (ex. 성명, 성별, 세대주와의 관계 등)
		· 공공기관의 정보공개에 관한 법률 제18조 · 공공기록물 관리에 관한 법률 제37조	· 주체자의 신분 확인을 위한 일반 식별정보 (ex. 이름, 주민등록번호, 연락처 등)
		· 민원사무처리에 관한 법률 제26조 · 국가정보화 기본법 제39조	· 본인·대리인 확인을 위한 식별정보 (ex. 주민등록번호, 대리인 신분증 등)
민간	정보통신	· 정보통신망 이용촉진 및 정보보호 등에 관한 법률 · 전자서명법 제24조	· 회원제 관리를 위한 사용자 식별 정보 (ex. 이름, ID, PWD 등) · 정당한 사용자임을 인증하는 식별정보 (ex. I-PIN인증, 단말정보, 휴대폰정보 등)
		· 전자금융거래법 제25조	· 휴대폰 결제 서비스 수행을 위한 식별정보 (ex. 결제수단별 개인정보, 카드번호, 비밀번호 등)
		· 전기통신사업법 제83조	· 주체자의 신분 정보 및 통신상의 사용자 정보에 대한 식별정보 (ex. 이름, ID, 주민등록번호 등)
		· 위치정보보호법 · 통신비밀보호법	· 업무 수행 및 처리를 위한 통신상의 식별정보 (ex. 접속 IP정보, GPS 정보 등)
		· 청소년보호법 제29조,	· 제한된 연령 확인에 대한 식별정보

III. 비식별화와 법규제

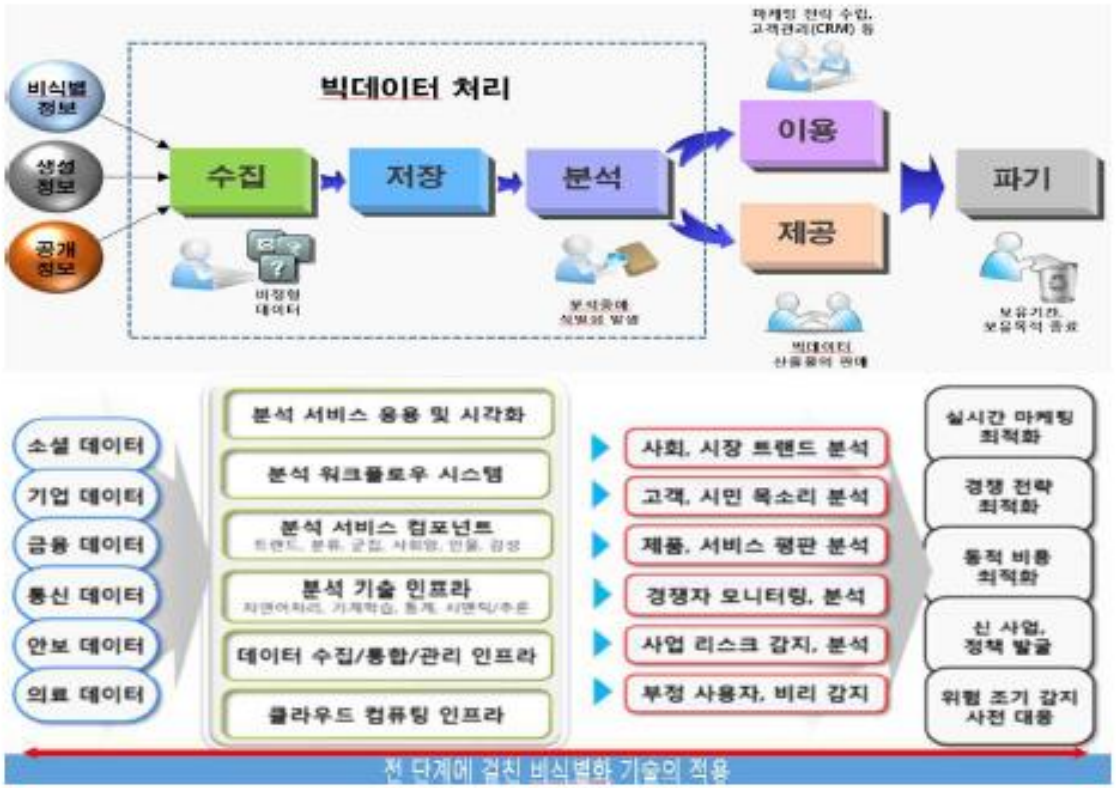
구분		근거	개인 식별 정보 항목
		제16조	(ex. 법정 생년월일, 법정 대리인 정보 등)
	상거래	· 전자문서 및 전자거래기본법 제12조 · 정보통신망 이용촉진 및 정보보호 등에 관한 법률 제23조, 제24조 · 전자상거래 등에서의 소비자보호에 관한 법률 제12조	· 전자문서 서비스를 위한 식별정보 (ex, 공인전자주소, 송신자, 수신자 등) · 통신의 안전한 조치를 위해 확인할 수 있는 식별정보 (ex. 비밀번호, 계좌번호, 주민등록번호 등) · 거래 기록 및 배송을 확인하기 위한 식별정보 (ex. 배송 주소지, 수령인 연락처 등)
		· 전자서명법 제24조	· 정당한 사용자임을 인증하는 식별정보 (ex. 가입자 이름, 전자서명검증정보, 인증서 일련번호)
	금융·신용	· 신용정보의 이용 및 보호에 관한 법률 제33조 · 금융실명거래 및 비밀 보장에 관한 법률 제4조	· 신용정보 및 거래능력을 판단할 수 있는 식별정보 (ex. 재산, 소득, 대출 보증 등) · 금융기관의 거래내역을 판단할 수 있는 정보 (ex. 주민등록번호, 계좌번호, 거래실적 자료 등) · 이용자 및 거래내용의 정확성을 확인하기 위한 식별정보 (ex.전자금융업자에 등록된 이용자번호, 이용자의 생체정보, 등)
		· 전자금융거래법 제26조 · 전자금융 감독규정 제5조의 3	
		· 특정 금융거래 정보의 보고 및 이용 등에 관한 법률 제5조의 3	· 자금이체를 수행을 위한 식별정보 (ex. 송금인 성명, 계좌번호, 수취인의 정보)
	보건·의료	· 의료법 제21조 · 응급의료에 관한 법률 제22조의 2조 · 산업안전보건법	· 정확한 환자의 진료를 위해 확인가능 한 식별정보 (ex. 주민등록번호, 의료기록, 가족력 등) · 신체의 질병정보를 통해 인지될 수 있는 식별정보 (ex. 감염병명, 혈액정보, 조직정보 등)
		· 후천성면역결핍증예방법 · 감염법의 예방 및 관리에 관한 법률 제74조	
		· 장애인 차별금지 및 권리규제 등에 관한 법률 제22조	· 신체 장애정보를 통해 확인 가능한 식별정보 (ex. 주민등록번호, 신체장애, 장애등급 등)
		· 국민건강보험법 제5조	· 가족구성원의 정보를 통해 확인할 수 있는 식별정보 (ex.가족구성원의 이름, 출생지, 소득 등)

● 무차별적 개인 정보 보호

비식별화 기술 활용 안내서 (미래부, 정보화진흥원 공동 발간)

o 빅데이터는 '수집 → 저장 → 분석 → 이용·제공 → 파기' 단계를 거쳐 활용

< 빅데이터 활용 단계 정의 >



● 무차별적 개인정보 보호

비식별화 기술 활용 안내서 (미래부, 정보화진흥원 공동 발간)

1. 빅데이터 분석 및 활용은 법률상 허용되는 목적 및 범위 내에서만 가능
2. 개인정보가 포함된 정보는 비식별화 조치 필요
3. 생성된 개인정보는 목적 달성 후 파기 또는 비식별화 해야 함
4. 전송중인 이메일 문자메시지 등 통신 내용은 조합, 분석 또는 처리 불가
5. 공개 개인정보도 수집-이용 및 제공을 위하여 정보주체의 동의 필요

자급자족 형태 이외는 데이터 산업이 허용되지 않음!

● 식별정보 vs. 인증정보

비식별화 기술 활용 안내서 (미래부, 정보화진흥원 공동 발간)

1. 빅데이터 분석 및 활용은 법률상 허용되는 목적 및 범위 내에서만 가능

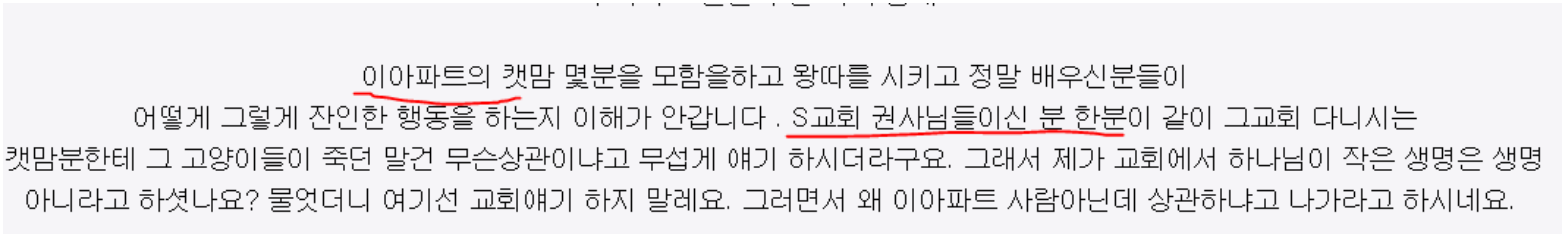
2. 개인정보가 포함된 정보는 비식별화 조치 필요

   - **비식별화 조치된 정보가 조합, 분석 처리 과정에서 재식별화되면 안됨.**

3. 생성된 개인정보는 목적 달성 후 파기 또는 비식별화 해야 함

4. 전송중인 이메일 문자메시지 등 통신 내용은 조합, 분석 또는 처리 불가

5. 공개 개인정보도 수집-이용 및 제공을 위하여 정보주체의 동의 필요


이아파트의 캣맘 몇분을 모함을하고 왕따를 시키고 정말 배우신분들이
어떻게 그렇게 잔인한 행동을 하는지 이해가 안갑니다 . S교회 권사님들이신 분 한분이 같이 그교회 다니시는
캣맘분한테 그 고양이들이 죽던 말건 무슨상관이냐고 무섭게 얘기 하시더라구요. 그래서 제가 교회에서 하나님이 작은 생명은 생명
아니라고 하셧나요? 물엇더니 여기선 교회얘기 하지 말래요. 그러면서 왜 이아파트 사람아닌데 상관하냐고 나가라고 하시네요.

어느 정도로 비식별화해야 재식별화가 안되나요?

● 제안된 비식별화 기술

k-익명성

나. 정의

⊙ 주어진 데이터 집합에서 준식별자 속성값들이 동일한 레코드가 적어도 k개 존재 해야 함

다. 의미

⊙ 데이터 집합의 일부를 수정하여, 모든 레코드가 자기 자신과 동일한(구별되지 않는) k-1개 이상의 레코드를 가짐

⊙ 예를 들어, 〈표 1〉의 의료 데이터가 익명화 된 〈표 3〉에서 1~4, 5~8, 9~12 레코드는 서로 구별되지 않음[17]

⊙ 예를 들어, 〈표 1〉의 의료 데이터가 익명화 된 〈표 3〉에서 1~4, 5~8, 9~12 레코드는 서로 구별되지 않음[17]

	준식별자			민감한 정보
	지역 코드	연령	성별	질병
1	130**	< 30	*	전립선염
2	130**	< 30	*	전립선염
3	130**	< 30	*	고혈압
4	130**	< 30	*	고혈압
5	1485*	> 40	*	위암
6	1485*	> 40	*	전립선염
7	1485*	> 40	*	고혈압
8	1485*	> 40	*	고혈압
9	130**	3*	*	위암
10	130**	3*	*	위암
11	130**	3*	*	위암
12	130**	3*	*	위암

⊙ 따라서, 익명화된 데이터 집합에서는 공격자가 정확히 어떤 레코드가 공격 대상인지 알아낼 수 없도록 하여 "프라이버시 보호"

예) 김민준 → 레코드 1~4 → 전립선염 또는 고혈압

출처: NIA. "개인정보비식별화 안내서"

〈k-익명성의 취약점〉

⊙ 데이터가 k-익명화 되었더라도 민감한 정보가 충분히 다양하지 않으면 프라이버시 문제 발생 가능

⊙ 취약점 1. 동질성 공격 (Homogeneity attack)

- 데이터 집합에서 동일한 민감한 정보를 이용하여 공격 대상의 민감한 정보를 알아내는 공격
- 〈표 3〉에서, 레코드 9~12의 민감한 정보는 모두 '위암'이므로, k-익명성 모델이 적용되었음에도 불구하고 민감한 정보가 직접적으로 노출됨

⊙ 취약점 2. 배경지식에 의한 공격 (Background knowledge attack)

- 주어진 데이터 이외의 공격자의 배경 지식을 통해 공격 대상의 민감한 정보를 알아내는 공격
- 〈표 2〉와 〈표 3〉에서, 공격자가 '이지민'의 질병을 알아내려고 할 때, 준식별자 조합(13068, 29, 여)에 따라 '이지민'은 1~4 레코드 중 하나이며, 질병은 전립선염 또는 고혈압임을 알 수 있음
- 이 때, '여자는 전립선염에 걸릴 수 없다'라는 배경 지식에 의해 공격 대상 '이지민'의 질병은 고혈압으로 쉽게 추정 가능함

⊙ k-익명성의 취약점의 원인

- 다양성의 부족 (lack of diversity)
  - 익명화할 때 민감한 정보의 다양성을 고려하지 않음
  - 동일한 민감한 정보를 가진 (다양하지 않은) 레코드가 익명화되어 하나의 '동질 집합'으로 구성될 경우, 동질성 공격에 무방비
- 강한 배경지식 (strong background knowledge)
  - k-익명성은 '여자는 전립선염에 걸리지 않는다', 또는 '남자는 자궁암에 걸리지 않는다'와 같은 공격자의 배경지식을 고려하지 않아 이를 이용한 공격에 취약함

ℓ-Diversity

정의

– 주어진 데이터 집합에서 함께 익명화되는 레코드들은 (동질 집합에서) 적어도 1개의 서로 다른 민감한 정보를 가져야 함

다. 의미

⦿ 익명화 과정에서, 충분히 다양한(1개 이상) 서로 다른 민감한 정보를 갖도록 동질 집합을 구성

⦿ 민감한 정보가 충분한 다양성을 가지므로, 다양성의 부족으로 인한 공격에 방어가 가능하고, 배경지식으로 인한 공격에도 일정 수준의 방어력을 가짐

⦿ 예를 들어, 〈표 4〉에서 모든 동질 집합은 3-다양성을 통해 익명화 되어 3개 이상의 서로 다른 민감한 정보를 가짐

– 〈표 3〉과 같이 동일한 질병으로만 구성된 동질 집합이 존재하지 않음

– 공격자가 질병에 대한 배경지식(예: 여자는 전립선염에 걸리지 않음)이 있더라도, 어느 정도의 방어력을 가지게 됨 (예: 여성 이지민이 속한 동질 집합 2, 3, 11, 12에서 전립선염을 제외하더라도 고혈압, 위암 중 어느 질병이 이지민의 것인지 여전히 알 수 없음)

ℓ-Diversity

	준식별자			민감한 정보
	지역 코드	연령	성별	질병
1	1305*	≤ 40	*	전립선염
4	1305*	≤ 40	*	고혈압
9	1305*	≤ 40	*	위암
10	1305*	≤ 40	*	위암
5	1485*	> 40	*	위암
6	1485*	> 40	*	전립선염
7	1485*	> 40	*	고혈압
8	1485*	> 40	*	고혈압
2	1306*	≤ 40	*	전립선염
3	1306*	≤ 40	*	고혈압
11	1306*	≤ 40	*	위암
12	1306*	≤ 40	*	위암

표 4 l=3

마. 정보 유용성 (data utility)

◉ 정보 손실 (information loss)

– 데이터를 익명화하게 되면, 프라이버시를 보호하는 대가로 일정량의 정보가 필연적으로 손실됨

– 〈표 3〉, 〈표 4〉의 '*' 나 '〉' 등과 같이 일부 정보를 지우거나, 원본 값을 구간 값 또는 더 상위 개념의 값으로 일반화(generalization)하는 과정에서 원본 데이터의 정보가 일부 손실됨

– 예: 연령 '23'을 구간 값 (20 ~ 25)으로 익명화, 성별 '여성'을 '*' (남성/여성을 구분 없이 모두 지칭함)로 익명화

◉ 프라이버시 보호–정보 손실 간의 관계

– k-익명성과 l-다양성 모델에서 k, l값은 곧 프라이버시 보호 수준을 의미

– k, l값이 증가할수록 프라이버시 보호 수준은 증가하지만, 이에 따라 많은 정보가 손실되므로 정보 유용성은 감소하는 경향을 보임

t-Closeness

나. 정의

◉ 동질 집합에서 민감한 정보의 분포와, 전체 데이터 집합에서 민감한 정보의 분포가 이하의 차이를 보여야 하는 것

다. 의미

◉ 민감한 정보의 분포를 고려

- 각 동질 집합에서 '민감한 정보의 분포'가 전체 데이터 집합의 그것과 비교하여 너무 특이하지 않도록 함
- 〈표 5〉에서, 전체적인 급여 값의 분포는 30 ~ 110
- 이 때, 레코드 1, 2, 30이 속한 동질 집합에서 급여의 분포는 30 ~ 50으로, 이는 전체 급여 값의 분포(30 ~ 110)와 비교할 때 극히 일부 → 공격자는 근사적인 급여 값을 알 수 있음
- *t*-근접성 모델은 이러한 동질 집합과 전체 데이터 집합 사이 분포의 과도한 차이를 *l*-다양성 모델의 취약점으로 규정함

t-Closeness

◉ '민감한 정보의 분포'를 조정하여 프라이버시를 보호
- 민감한 정보가 특정 값으로 쏠리거나, 유사한 값들이 뭉치는 경우를 방지
- 〈표 6〉에서 *t*-근접성 모델에 따라 레코드 1,3,8이 하나의 동질 집합으로 익명화됨
- 이 때, 레코드 1, 3, 8의 급여의 분포는 (30 ~ 90)으로 전체적인 급여의 분포(30 ~ 110)와 큰 차이가 나지 않음
- 또한, 레코드 1, 3, 8의 질병 분포는 (위궤양, 만성위염, 폐렴)으로 병명이 서로 다르면서, 질병이 '위'와 관련된 것 이외에 '폐'와 관계된 것이라서, 특정 부위의 질병임을 유추하기 어려움
- 따라서 〈표 5〉의 경우와 비교하여, 공격자가 공격 대상의 민감한 정보를 추측하기가 더욱 어려워짐

◉ 민감한 정보의 의미(semantics)까지 파악하는 프라이버시 모델
- 민감한 정보의 의미를 고려하여 값의 분포를 계산함
- 연속 속성(continuous attribute)의 경우 숫자 값을 통해 의미가 유사한 정도를 파악 (예: 급여)
- 범주 속성(categorical attribute)의 경우 분류 트리(taxonomy tree, 〈그림 1〉 참고)를 이용해 의미가 유사한 정도를 파악 (예: 질병)

- **k-Anonymity, l-Diversity, t-Closedness의 문제점**

  - **General k-Anonymity is NP-Hard**

    - Practical algorithms exist using approximation!

● 비식별화된 정보가 재식별화가 가능한지를 둘러싼 논쟁

비식별화가 불가능하다고 보는 논문 및 보고서

1. **백악관 보고서, "Big Data: Seizing Opportunities, Preserving Values," Executive Office of the President, May 2014**

   "When data is initially linked to an individual or device, some privacy-protective technology seeks to remove this linkage, or 'de-identify' personally identifiable information—but equally effective techniques exist to pull the pieces back together through 're-identification.'"

2. Public Knowledge et al., "Petition for Declaratory Ruling Stating that the Sale of Non-Aggregate Call Records by Telecommunications Providers without Customers' Consent Violates Section 222 of the Communications Act," December 11, 2013

3. Timothy Lee, "There's No Such Thing as an Anonymized Dataset," TechDirt, November 30, 2007

● 비식별화된 정보가 재식별화가 가능한지를 둘러싼 논쟁

비식별화가 불가능하다고 보는 논문 및 보고서

4. **Yves-Alexandre de Montjoye, César A. Hidalgo, Michel Verleysen & Vincent D. Blondel, "Unique in the Crowd: The Privacy Bounds of Human Mobility," Scientific Reports 3.**

   The study found that when an individual's location is specified hourly within the reception area of a mobile phone antenna, knowing as few as four random spatio-temporal points was enough to uniquely identify 95 per cent of the mobility traces in the dataset of one and a half million individuals.

   It is admittedly very difficult to de-identify mobility traces, while maintaining a sufficient level of data quality necessary for most secondary purposes, due to their high degree of uniqueness.

5. **EU Data Protection Working Party (Article 29), "Opinion 05/2014 on Anonymisation Techniques," April 10, 2014.**

   The removal of direct identifiers alone is generally insufficient to properly de-identify datasets.

● 비식별화된 정보가 재식별화가 가능한지를 둘러싼 논쟁

비식별화가 불가능하다고 보는 논문 및 보고서

6. **Latanya Sweeny, Uniqueness of Simple Demographics in the U.S. Population, Carnegie Mellon University, Laboratory for International Data Privacy, Pittsburgh, PA, 2000.**

   It used 1990 U.S. census data to show that 87 per cent of the U.S. population could be uniquely identified through the combination of gender, date of birth, and ZIP code.

7. Phillippe Golle, Revisiting the Uniqueness of Simple Demographics in the US Population, Palo Alto Research Center, 2006

   - Only 63 per cent of the U.S. population is uniquely identifiable given those data categories.
   - If an individual's date of birth is replaced with only the month and year of birth, the percentage of those uniquely identifiable drops to 4.2 per cent.
   - If one further replaces the ZIP code with an individual's county, then the percentage of the population capable of being uniquely identified drops dramatically to 0.2 per cent.1

● 비식별화된 정보가 재식별화가 가능한지를 둘러싼 논쟁

비식별화가 불가능하다고 보는 논문 및 보고서

8. Nate Anderson, "'Anonymized Data Really Isn't—and Here's Why Not," Ars Technica, September 8, 2009.
9. Caroline Perry, "You're Not So Anonymous," Harvard Gazette, October 18, 2011.
10. Arvind Narayanan and Vitaly Shmatikov, "Robust De-anonymization of Large Sparse Datasets," Proceedings of the 2008 IEEE Symposium on Security and Privacy, 2008.

● 비식별화된 정보가 재식별화가 가능한지를 둘러싼 논쟁

비식별화가 가능하다고 보는 논문 및 보고서

1. Jane Yakowitz, “Tragedy of the Data Commons,” Harvard Journal of Law and Technology 25 (2011): 1–40.

2. Felix T. Wu, “Defining Privacy and Utility in Data Sets,” University of Colorado Law Review 84 (2013): 1117–1175.

3. Khaled El Emam, Guide to the De-Identification of Personal Health Information (Boca Raton, FL: CRC Press, 2013).

4. Daniel C. Barth-Jones, “The ‘Re-identification’ of Governor William Weld’s Medical Information: A Critical Re-Examination of Health Data Identification Risks and Privacy Protections, Then and Now.” June 4, 2012.

5. Ann Cavoukian and Khaled El Emam, “Dispelling the Myths Surrounding De-identification: Anonymization Remains a Strong Tool for Protecting Privacy,” June 2011.

III. 비식별화와 법규제

● 비식별화된 정보가 재식별화가 가능한지를 둘러싼 논쟁

비식별화가 가능하다고 보는 논문 및 보고서

6. Ann Cavoukian and Daniel Castro, “Big Data and Innovation, Setting the Record Straight: De-identification *Does* Work,” Jun 2014.

“While nothing is perfect, the risk of re-identification of individuals from properly de-identified data is significantly lower than indicated by commentators on the primary literature.”

● 비식별화 기술이 처리하는 문제

1. They protect an individual's records from being uniquely identified in the dataset.
2. They prevent an individual's records from being linked to other datasets.
   - If a set of attributes uniquely identifies an individual within a de-identified dataset and those same attributes are found in a personally identifying dataset, then that individual may be re-identified by linking the two datasets together.
3. They make it difficult to infer sensitive information about an individual from the de-identified dataset.
   - If groups of individuals are identified in a dataset and all the individuals in a certain group have a certain property, then if an individual is known to belong to that group, one could easily find out the value of his/her group property.

Source: EU Data Protection Working Party (Article 29), "Opinion 05/2014 on Anonymisation Techniques," p. 11–12

● 데이터 유통 금지에 따른 가장 큰 피해자는?

1. 개인 정보 부재에 따른 중금리 대상자 (서민)
   - 은행 안전 대출 선호로 인해 우리나라는 중금리시장이 형성되지 못한 채 제2금융권, 대부업체, 불법사금융으로 내몰리고 있음.
   - 중금리 대출시장의 부재로 신용도로 5~6등급 (중신용자)에 해당하는 1180만명이 연 15%~34.9 (50%이상이 30%이상임)의 약탈적 금리에 시달리는 중.
   - 금리 양극화를 해소하기 위해서는 데이터가 유통되어야 함.
2. 제2금융권 및 대부업체
   - 대출자에 대한 정보가 전무함으로 인해 높은 대출 부실률에 시달리는 중.

12.6%(2016.2)에 달하는 청년실업자들이 가장 큰 피해자!

● 데이터 생산자는 누구?

1. 기간 산업-인프라 제공 서비스사
- 전기통신, 전화, 수도, 철도, 고속도로, 인터넷

2. 편의제공 서비스 제공사
- 은행-신용카드-보험 등 금융사 및 금융관련 서비스 제공사
- 식품점, 마트

3. 온라인 플랫폼
- 페이스북, 트위터, 카카오톡, 유튜브, 위키피디아 등 소셜미디어
- 네이버, 다음 등 포털 서비스
- 아마존, 11번가, G Market 등 e-Commerce 서비스
- 구글 등 검색제공사

데이터 생산자로 보면 빅데이터(=벌크데이터)는 상품화할 이유 없음!

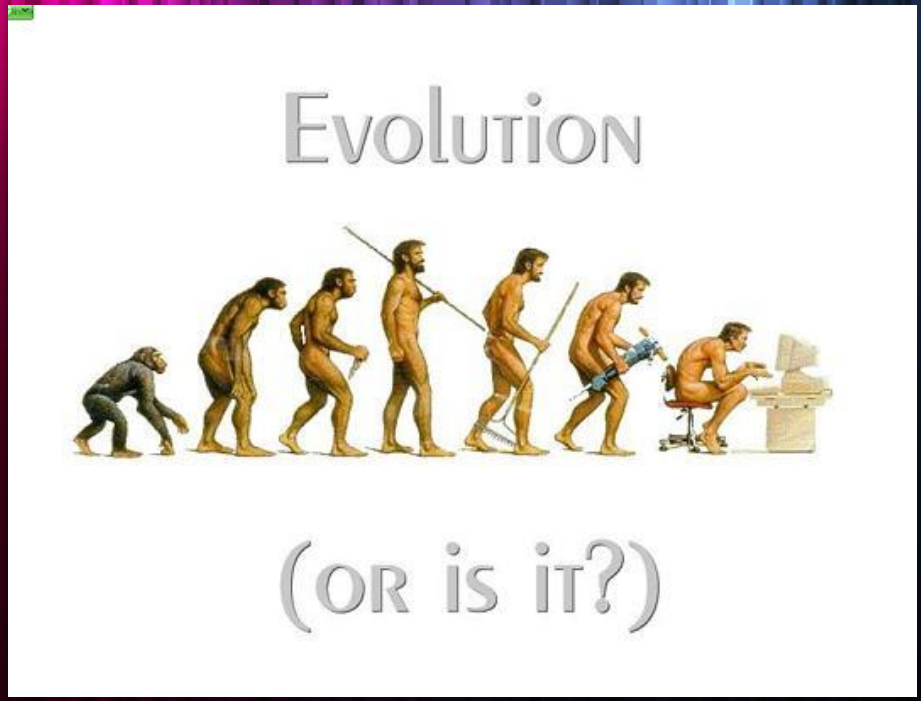
● 개인정보자기결정권: 판매결정도 자기가 할 수 있어야 함

개인정보의 공개와 이용에 관하여 스스로 결정할 수 있는 권리

- 헌법재판소는 개인정보자기결정권을 헌법상 기본권으로 인정.
- 개인정보자기결정권의 보호대상이 되는 개인정보 : 성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보 등.

개인정보자기결정권은 개인정보 판매 유통권이 포함되어야 함!

II. 진화의 관점에서 본 데이터 산업



- 위험하면 막아야 할까요?

II. 진화의 관점에서 본 데이터 산업

- 산업의 진화

산업사회 시대

가내수공업 시대

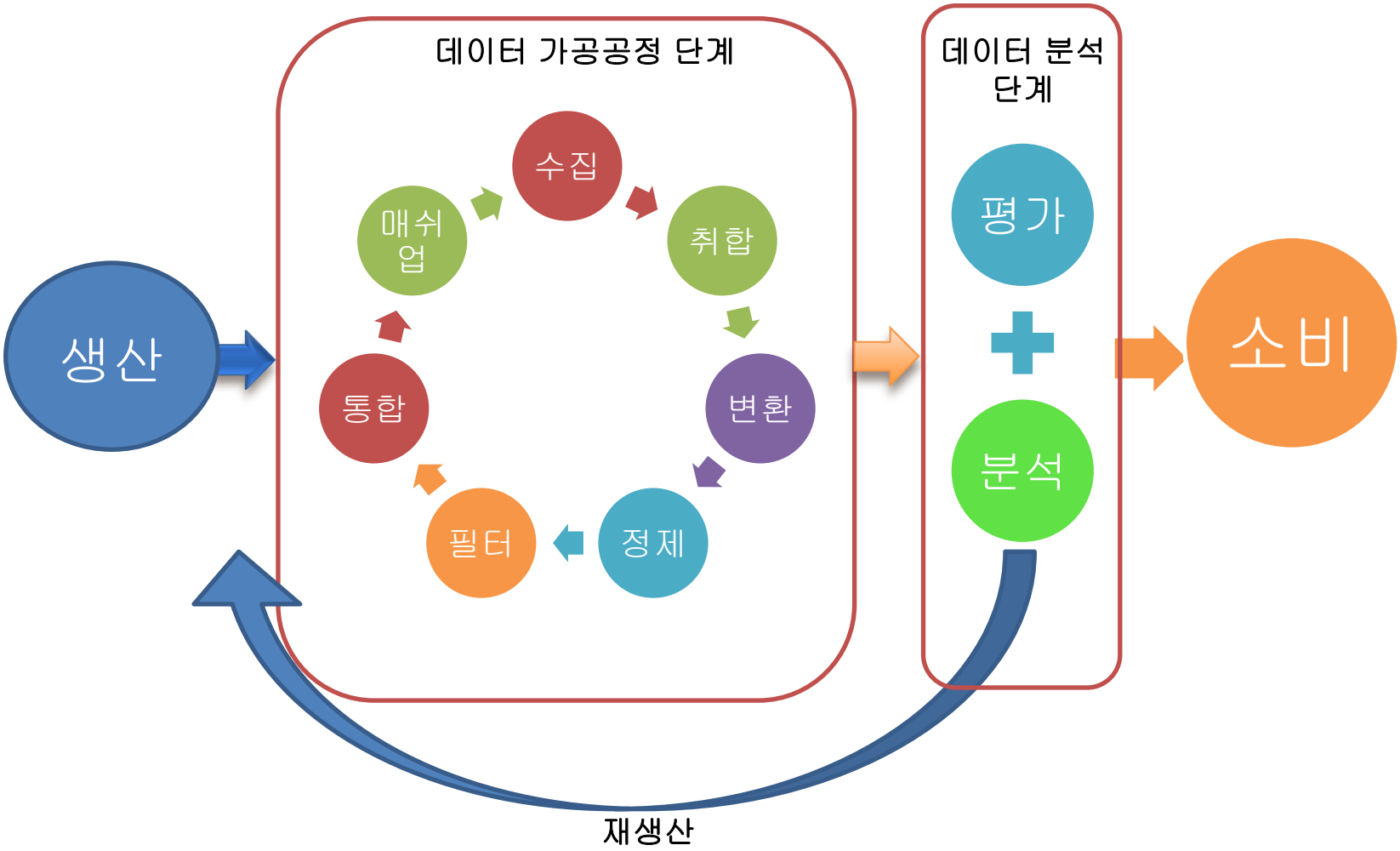
자급자족 시대

수렵채취시대

진화로 본 한국 데이터 산업은?

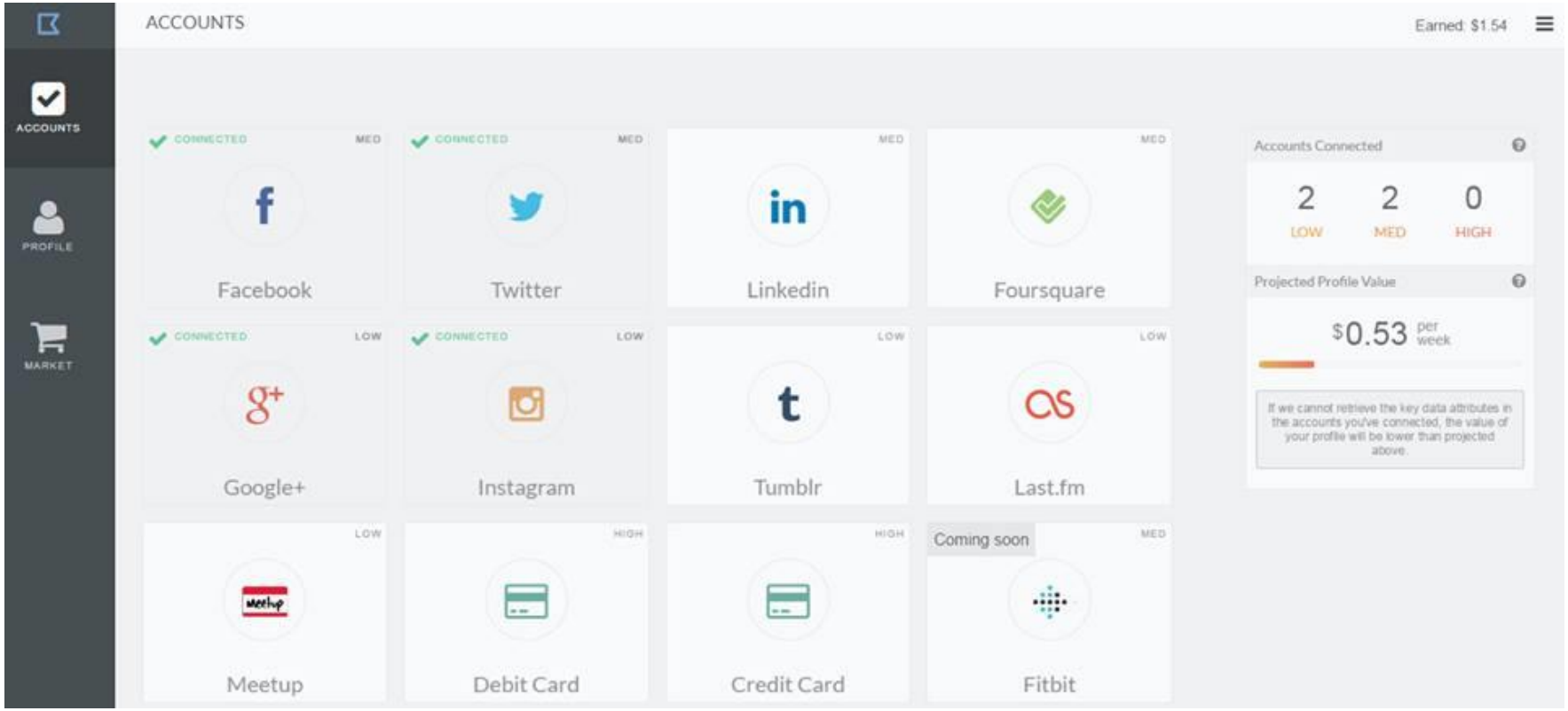
I. 진화의 관점에서 본 데이터 산업

● 이상적인 데이터 생태계



II. 진화의 관점에서 본 데이터 산업

- 외국의 경우: DataCoup



개인 정보 일주일에 0.53센트!

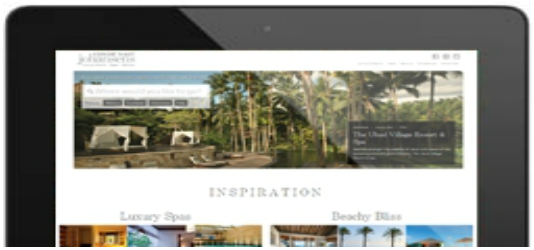
II. 진화의 관점에서 본 데이터 산업

- 외국의 경우: Dutch 대학생 Shawn Buckles

WIRED.CO.UK

'Data soul' of Shawn Buckles sells for £288

TECHNOLOGY / 15 APRIL 14 / by OLIVIA SOLON



Dutch student Shawn Buckles has auctioned all his personal data to the highest bidder and earned a grand total of €350 (£288).

In March, Buckles set up a website with an online bidding system in order to make a comment about privacy and the value of personal data. He put all of his personal data

Shutterstock

출처 : http://bit.ly/1m5djw3

2014년 한 대학생이 경매를 통해 개인정보 판매로 벌어들인 돈은 350유로!

III. 결론

● 데이터산업 활성화

- **과도한 정보보호는 데이터산업 발전을 저해**
  - 정보보호에 대한 정확한 인식이 필요
  - 현행 개인정보보호법은 위헌 소지 있음
- **데이터 산업은 유통 생태계부터 시작**
  - 데이터를 유통시킬 수 있는 법제도 정비필요
  - 완벽한 비식별화를 요구하는 것은 무리라는 인식이 필요
- **데이터 유통을 위한 거래소에 투자할 필요있음**

- **참고논문**
  - 이영환, 전희주, 윤정연. 데이터 산업에서 창업 활성화를 위한 데이터 거래소 제안 : 금융거래소형 데이터거래소를 중심으로. 창업학회지. 2015년 6월.
- **신문칼럼**
  - 개인정보보호법은 무효다. TechM (2016/1/6).
  - "'비식별화'라는 코끼리" 전자신문. (2015/11/9)
  - "사하라 사막에 인터넷을 공급하려면" 서울경제. (2015/10)
  - ."인터넷 전문은행, 빅데이터 분석이 특화전략 핵심" 머니투데이. (2015/8)

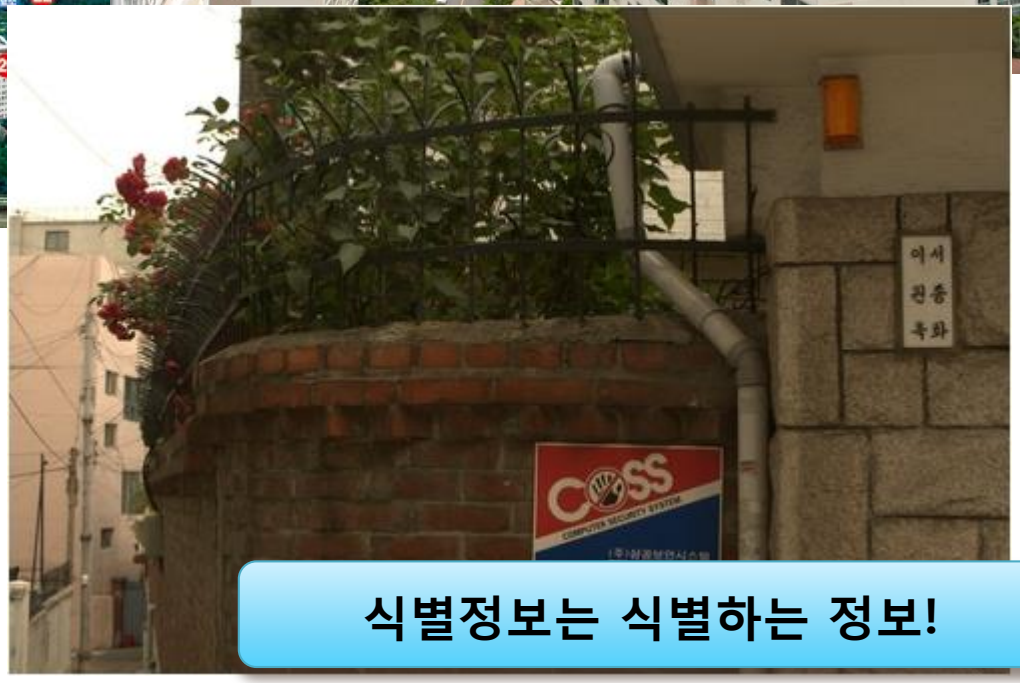
Q&A

부록

얼굴이 식별 정보 인가요, 인증 정보인가요?

식별정보와 인증정보

식별정보의 예



식별정보는 식별하는 정보!

식별정보와 인증정보

식별정보의 예

식별정보와 인증정보

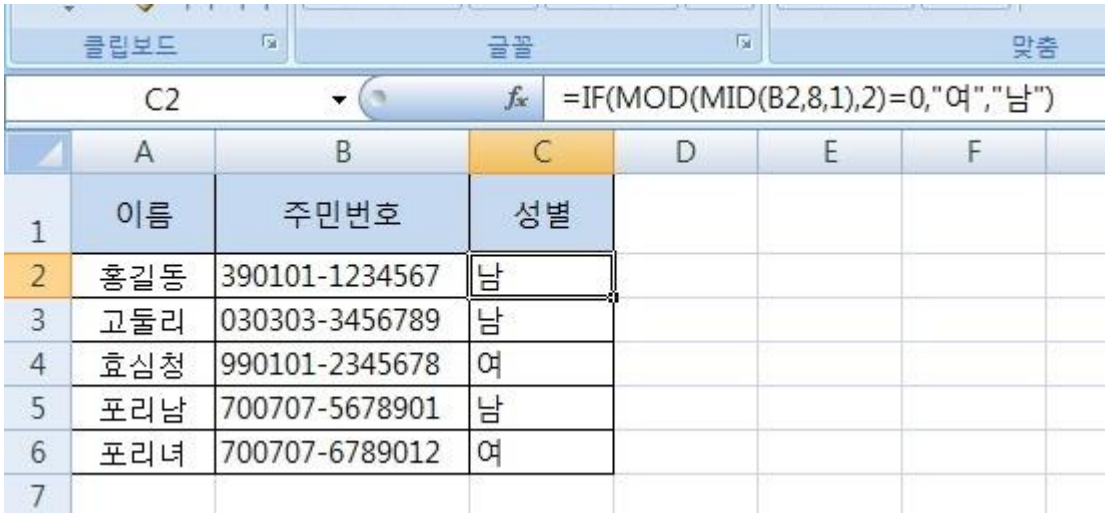
● 식별정보 vs. 인증정보

식별정보의 예

식별정보를 감추면?

식별정보와 인증정보

식별정보의 예



	A	B	C	D	E	F
1	이름	주민번호	성별			
2	홍길동	390101-1234567	남			
3	고둘리	030303-3456789	남			
4	효심청	990101-2345678	여			
5	포리남	700707-5678901	남			
6	포리녀	700707-6789012	여			
7						

주민번호는 식별정보!

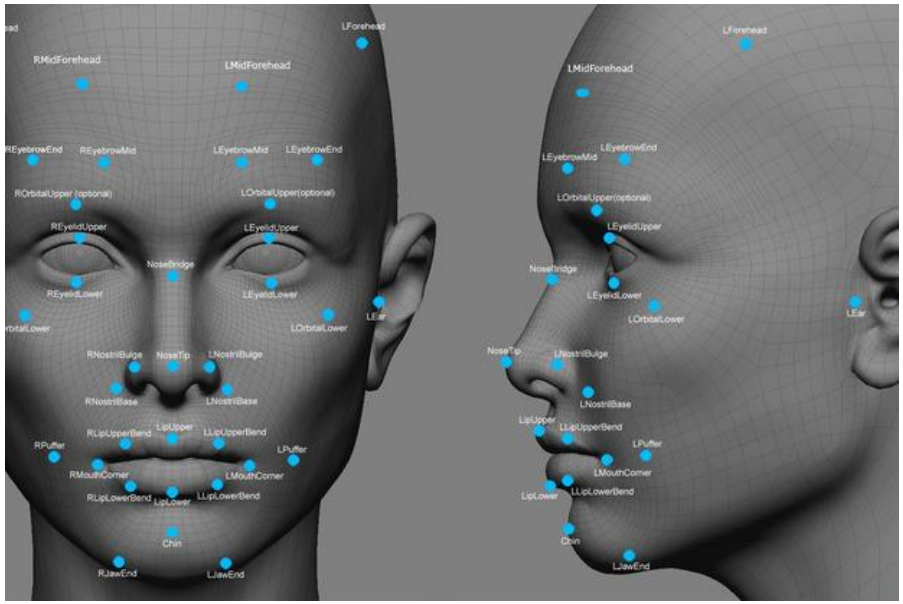
식별정보와 인증정보

인증정보의 예

Q 지금까지 알려진 인증 기술이 안전한가요?

식별정보와 인증정보

식별정보와 인증정보의 혼동 존재



안면 인식이 인증이 되나?