



# 정보독점의 해소와 개인정보보호간 상충문제의 해법과 정책과제

강경훈 (동국대 경영), 임요한 (서울대 통계)  
전성인 (홍익대 경제), 천정희 (서울대 수학)

정보집중 문제와 데이터 공유

# 데이터 경제의 도래와 데이터 집중 또는 독점 문제

- ❖ 경제활동에서 데이터의 중요성이 급속하게 커지고 있으며 이와 관련하여 데이터 집중 또는 독점에 대한 우려도 증가
- ❖ 특히 데이터가 중요 투입요소가 되는 data-driven industry (market)에서 데이터 집중 문제에 대한 논의가 활발
  - data-driven industry의 등장은 (1) 데이터 양의 증가, (2) 데이터의 저장 비용 하락, (3) 데이터 분석 능력 향상 등의 요인으로 설명
  - data-driven market에서 사업자 경쟁력의 원천은 데이터 양 및 데이터 분석 능력임이 널리 인정되고 있으며 특히 플랫폼 이용자에 대한 데이터를 매우 낮은 한계비용으로 수집, 축적할 수 있는 플랫폼 사업자와 관련된 논의가 많음
  - 스타트업의 cold start problem

# 데이터 집중의 경쟁 제한성

## ❖ 지배적 플랫폼의 시장지배적 지위 강화 및 경쟁 제한 가능성

- 빅데이터에 접근하는 것이 특정 시장에서의 경쟁력 확보에 중요 요소인 경우
- Competition can't work if just a few companies control a vital resource that you need to be able to compete – and if they refuse to share it with you. Right now, it looks as though data is becoming one of those vital resources. (Margrethe Vestager, Feb. 2019)
- Natural tendency of data-driven markets towards monopolization (J. Prüfer, 2020)
- 데이터를 필수설비(essential facility)로 간주해야 한다는 주장: 대부분의 나라에서 필수설비 요건을 강하게 규정하고 있어 빅데이터가 필수설비로 인정되는 경우는 아직까지 없음

## ❖ 대규모 데이터가 진입장벽이나 시장지배력의 원천이 아니라는 주장도 다수

- 데이터는 도처에 있으며 데이터 수집 및 유통 비용도 낮고 데이터 가치의 짧은 유효기간 등

# 데이터 집중의 경쟁 제한성 해소 방안

## ❖ 데이터 공유(pooling or sharing)에 대한 다양한 제안

- One thing we may need to do, to open up competition, is to require companies to give rivals access to their data. (Margrethe Vestager, Apr. 2019)
- Mandatory sharing of user information in data-driven markets (J. Prüfer, 2020)
- 독일 사민당(SPD)은 data sharing obligation을 포함한 Daten-für-alle-Gesetz (Data-for-all law, 2019)안을 발표
- Mayer-Schönberger and Ramge (2018): “[E]very company above a certain size [...] that systematically collects and analyzes data would have to let other companies in the same market access a subset of its data. The larger a firm’s market share, the more of its data others would be allowed to see.”
  - The Economist(2018.1.12)도 유사한 방안을 제시

# 데이터 공유제도의 주요 원칙

## ❖ 데이터 공유 명령제(mandated data sharing 또는 data sharing mandates)의 원칙

- 개인정보 보호 원칙 및 빅데이터 수집, 결합, 분석 등의 유인이 유지되어야 함
  - Margrethe Vestager (Apr. 2019): Very often, the data you need to compete is data about particular individuals. So any access to data would need to be in line with the data protection rules. And collecting data also takes effort and time. So if we insist that companies share it with others, without proper compensation, we could discourage others from putting in those efforts in the future.

## ❖ GDPR의 데이터 이동권 및 PSD2의 고객정보 공유

- 개인정보 보호 원칙을 유지하기 위해 동의에 기반하여 데이터를 공유하도록 한 제도로 이해

개인정보 활용과 데이터 공유

# 개인정보의 활용 동의 방식의 문제점

## ❖ 기존의 개인정보 활용 방식은 “동의” 기반 방식(consent-based approach)

- 정보 수집자의 정보 수집 및 제3자에 대한 정보 제공에 대한 개인정보주체의 동의
- 제3자에 대한 정보의 “제공”이 과연 경제적 대가를 주고 받는 “데이터 매매”까지 포괄할 수 있는 지에 대해서는 잠재적 논란 가능성
- 고전적인 동의 모형에서 동의자인 개인정보주체는 명시적인 경제적 대가를 수령하지 않고, “제공”의 형태로 발생하는 데이터 매매의 모든 이익은 정보 수집자가 향유

## ❖ 개인정보 활용 방식에 대한 재검토는 거대 정보수집자의 데이터 독점 해소에도 중요한 함의를 지님

- 개인정보주체의 권리 강화는 거대 정보수집자의 데이터 독점권을 통제할 수 있기 때문



# 개인정보의 이동권(data portability) 또는 전송 요구권

## ❖ 개인정보의 이동권 도입은 개인정보 활용에 대한 새로운 기회를 제공할 가능성

- 정보 수집자가 수집한 개인정보에 대하여 개인정보주체의 자료 전송 요구권을 인정
- GDPR Article 20(Right to data portability), 신용정보법 제33조의2(전송 요구권)
- 자료 이동 요구시 기존 데이터에 대한 삭제 요구권은 재량적으로 행사

## ❖ 개인정보 이동권은 정보사업자들 간의 기존 시장 외에 개인정보주체와 정보사업자 간의 **새로운 시장을 출현시킴**

- 기존 데이터 거래 시장: 정보수집자/정보사업자와 또 다른 정보수집자/정보사업자 간 시장
- 추가적 데이터 거래 시장: **개인정보주체와 정보수집자/정보사업자 간 시장**
- 새로운 시장의 추가 출현은 **기존 데이터 거래 시장의 경쟁 양태에 영향**을 미칠 가능성 큼

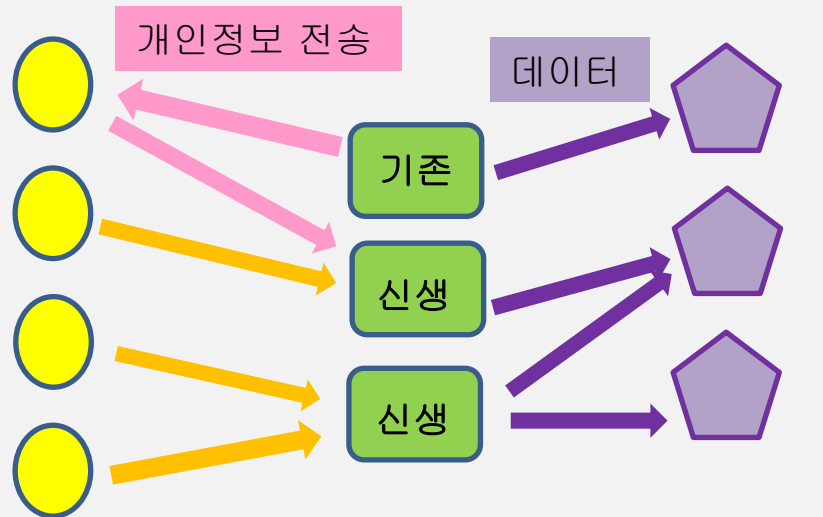
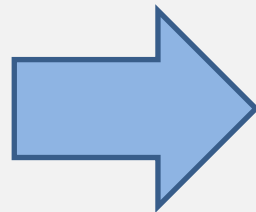
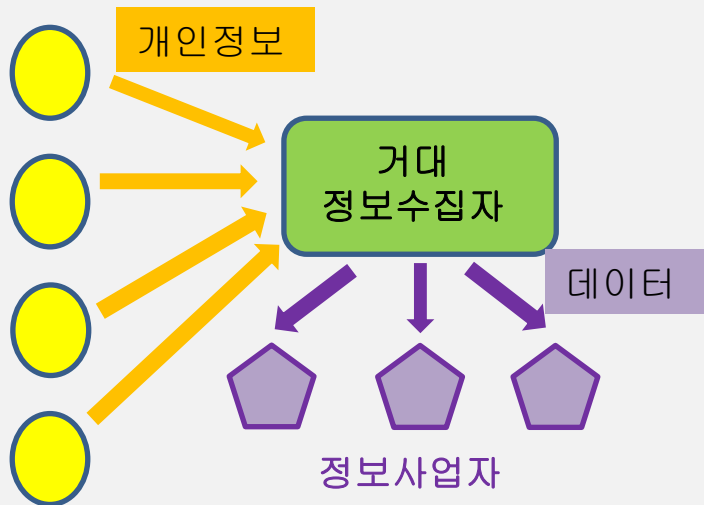
# 개인정보의 이동권 도입이 데이터 독점에 미치는 영향

- ❖ 개인정보의 이동권 도입은 거대 정보수집자의 데이터 독점 현상을 완화 가능
- ❖ 제1측면: 거대 정보수집자의 데이터 수집 및 보유 능력에 대한 견제
  - 개인정보주체는 자신에게 유리한 경제적 대가가 지불될 경우 기존 거대 정보수입자로부터 자신의 데이터를 다른 정보수집자로 이동(필요시 삭제 요구까지)시킴으로써 기존 거대 정보수집자의 데이터 보유 능력을 견제할 수 있음
- ❖ 제2측면: 신생 정보수집자의 데이터 수집 능력 강화
  - 신생 정보수집자는 과거에는 거래기회의 제공과 동의를 통해 스스로 데이터를 구축하거나, 기존의 거대 정보수집자로부터 데이터를 매입하는 방법으로만 데이터 구축이 가능했음
  - 개인정보의 이동권 도입 이후에는 직접 개인정보주체로부터 그 정보주체의 기존 데이터를 전송받아 데이터 베이스를 구축할 수 있는 추가 통로가 열리게 됨

# 개인정보의 이동권과 데이터 거래 시장의 경쟁 촉진

❖ 개인정보의 이동권 도입은 그 자체로 데이터 거래 시장의 경쟁을 촉진

- 공권력에 의한 강제적인 "데이터 공유 명령제" 도입 필요성은 그에 비례하여 감소



개인정보주체

개인정보주체

정보수집자

정보사업자

차분정보보호(Differential Privacy)  
동형암호(Homomorphic Encryption)

# 차분정보보호 Differential Privacy

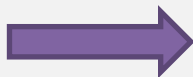
원 자료나 통계량에 적절한 잡음(noise)를 추가하여

임의의 개인의 데이터베이스에 참여 여부에 따른 결과값의 차이를

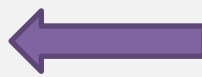
정하여진 수준 이하로 제어하는 방법.

# 개념형상화

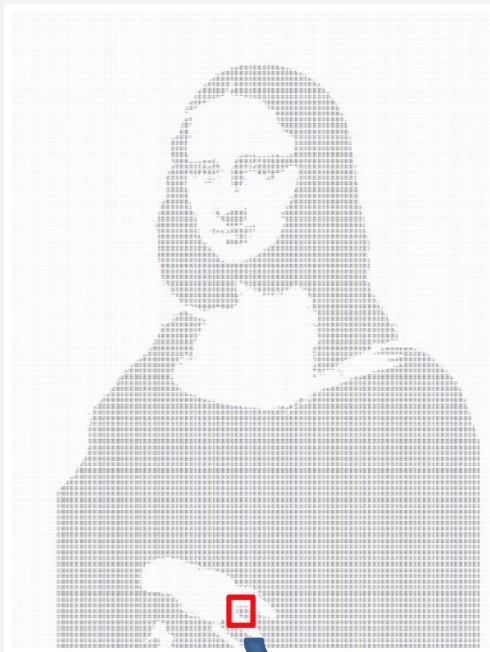
원자료



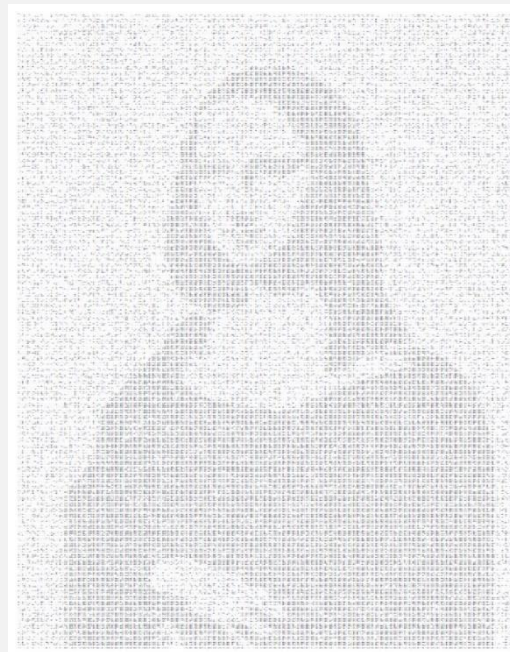
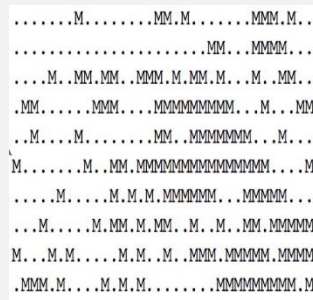
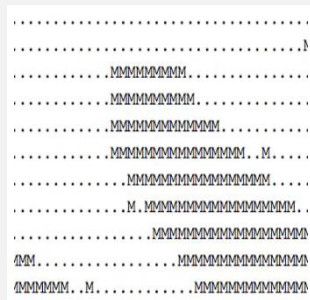
모나리자



국소차등정보 자료

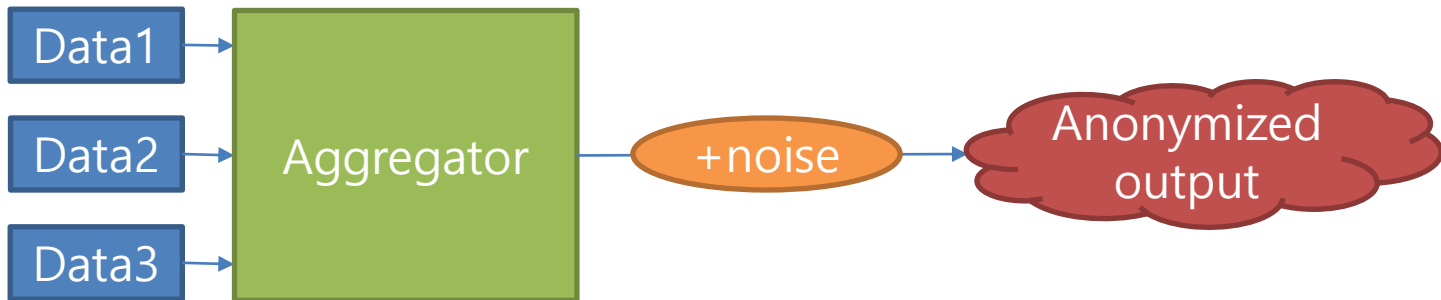


+잡음

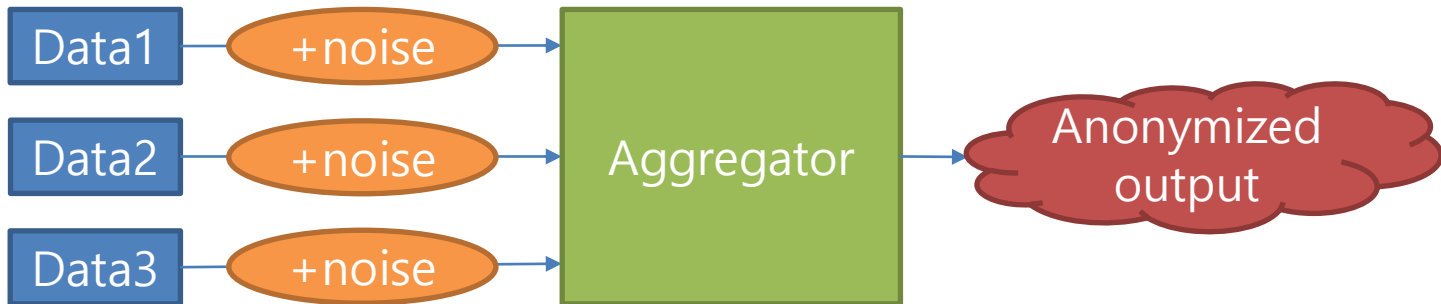


# 잡음 추가 대상에 따른 구분

Global DP



Local DP : TTP 불필요



# 차분정보보호의 사례

- 김철수가 이용하는 통신회사는 네비게이션을 무료로 사용하는 대신 **차량의 운행기록 데이터**(운행 경로와 속도 등)를 회사의 **서버에 저장**하고 **분석**하고자 함.
  - 장점: 회사가 운행기록을 분석하여 더 좋은 네비게이션 서비스를 제공할 수 있음.
  - 단점: 통신회사와 보험회사가 고객 정보를 공유, 특히 보험회사는 운전자의 운전 행태를 분석하여 보험료를 차등적으로 부가할 수 있음.
- 예를들어, 김철수의 운전 기록과 행태를 분석한 결과, 보험회사는 김철수가 주로 이용하는 경로들이 다른 경로보다 사고율이 높은 사실을 알아내고 김철수의 사고확률이 10%가 아닌 30%라고 결론을 내림. 이에 따라 보험료를 현재 10만원에서 30만원으로 인상할 수 있음.
- 만일 epsilon-차분정보보호가 적용된 자료가 제공된 경우 향상된 네비게이션 서비스를 제공 받음과 동시에 김철수의 사고확률 추정치의 상한이  $(1 + \text{epsilon}) * 10\%$ 로 제한되어 보험료의 최대 증가분은  $\text{epsilon} * 10\text{만원}$ 이 됨.

\* 2018년 하반기 통계개발원 연구보고서(차등정보보호에 대한 연구, 박민정,이용희, 권성훈 저)에서 발췌 수정함



# 동형암호 Homomorphic Encryption

- 동형암호 (Homomorphic Encryption; HE)  
: 암호화된 데이터를 복호화 없이도 연산할 수 있는 암호 기술
- 10 Emerging Technologies (MIT Technical Review 2011)
- “완벽한 하인”



# 동형암호화한 개인정보는 익명정보로 볼수 있는가?

❖ 개인과 연결된 정보가 주어졌을 때..

❖ 기존 비식별화

- 개인의 고유식별자를 가명화 (혹은 제거)
- 개인을 식별할 수 있는 **데이터 제거**

❖ 동형암호화

- 개인의 고유식별자를 가명화
- 개인을 식별할 수 있는 **데이터 암호화**

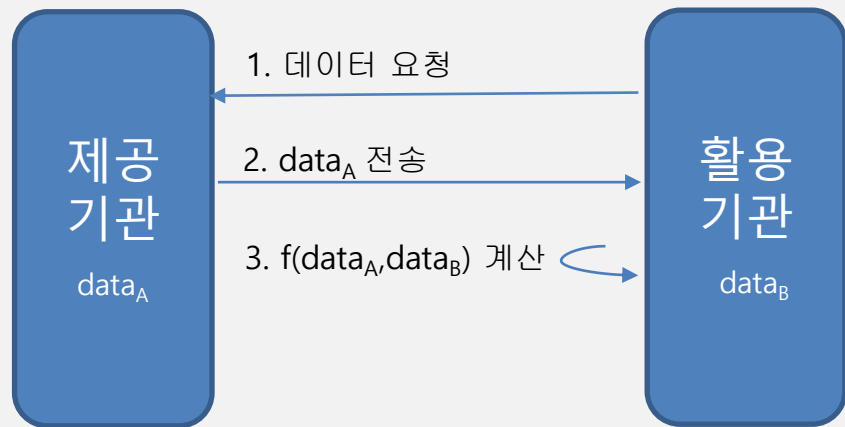
(복호화키는 **데이터소유주** 혹은 **신뢰기관**에서 안전하게 보관)

- 차분정보보호 추가

HE-DP기반 프라이버시 보존  
데이터공유방안

# 데이터 의무 공유 방안 및 이의 문제점

- ❖ 포괄적 동의: 공유되는 데이터의 종류나 활용범위를 확인하기 어려움
- ❖ 철회의 어려움: 데이터를 이동한 후 철회시 이를 확인하기 어려움
- ❖ 재이동 방지 어려움: 데이터를 타 기관에 공유하거나 재판매시 적발이 어려움
- ❖ 보안성 약화: 데이터의 집중화로 데이터 해킹시 피해가 증가

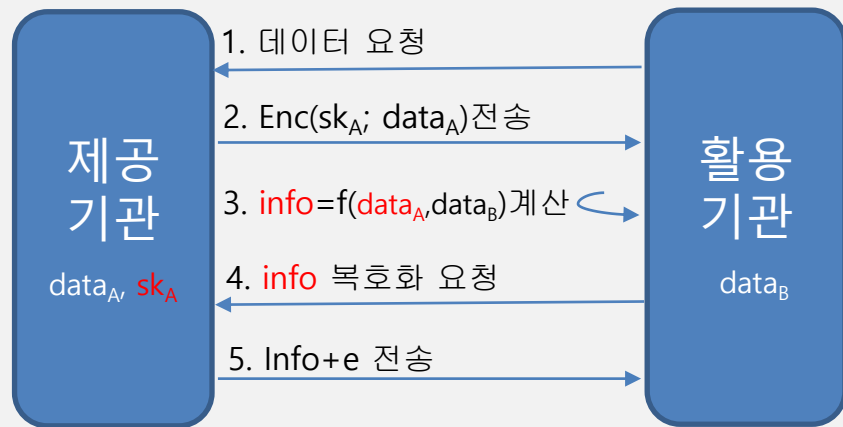


데이터 공유 흐름도

# 동형암호기반 공유

- ❖ 목적: 기관 B는 자신의 데이터  $data_B$ 와 기관 A의 데이터  $data_A$ 를 활용하여  $info=f(data_A, data_B)$ 를 얻고 싶다.
- ❖ 기관 A는 동형암호 비밀키  $sk_A$ 로 자신의 데이터  $data_A$ 를 암호화하여 B에 제공
- ❖ 기관 B는  $info=f(data_A, data_B)$  계산
- ❖ 기관 A는  $info$ 를 복호화하고 DP적용하여 기관 B에게 전송

(붉은 색은 암호화된 데이터)



동형암호기반 데이터공유

# 동형암호기반 공유 (클라우드 버전, Encrypted Twin)

(붉은 색은 암호화된 데이터)



# 제안 방법의 특징

## ❖ 단점

- Info를 얻기 위한 기관 B의 계산 시간 증가
- Data Scientist가 다양한 기법을 적용하기 불편함

## ❖ 장점

- 보안성: 기관 A의 데이터가 남용되거나 재공유 되지 않음
- 데이터수집 기관의 인센티브 (데이터 사용기관 대비)
  - 복호화 과정에서 과금
  - 사용통계 확보

실행방안: 거래소, 전문기관, 법규 등



# 동형암호화 및 DP 처리된 데이터 전문 거래소

## ❖ 동형암호화 및 DP 처리된 데이터가 거래되는 거래소를 설립 운영

- 빅데이터가 경쟁 제한의 수단이 못되거나 이용되지 않는다면 데이터 거래나 결합은 프라이버시 침해 문제만 없으면 기업들이나 소비자 모두의 후생을 증가시킴
- 동형암호화 및 DP 처리된 데이터는 익명성이 매우 높아 프라이버시 침해 문제를 회피하여 데이터 거래와 결합을 활성화시킬 가능성
- 거래소 활성화와 기업들에 대한 데이터 원산지 표시 의무화 등을 결합하여 데이터 지하시장을 제한하는 한편 장기적으로 개인정보 유출이나 해킹의 유인을 축소
- 한편 앞에서 살펴본 데이터 공유 명령제가 도입되더라도 애초의 데이터 보유 기업에 대해 데이터 제공에 대한 대가를 지불하도록 해야 하는데 거래소에서 평소 거래되는 데이터의 가격을 기초로 공유에 따른 대가를 산정

# 데이터 공유 처리 전문기관

- ❖ 만일 데이터 공유 명령제가 필요한 경우 해당 기업들 간 자율적인 공유 방식을 취하거나 기업들 사이에서 데이터 공유 업무를 전담하는 기관을 둘 수도 있음
  - 정책당국에서 강제적 데이터 공유의 필요 여부, 공유 데이터 종류 및 공유 대상 기업의 범위 등에 대해 판단이 내려지면 데이터 보유 기업으로부터 (동형암호 처리된) 데이터를 받아 공유 범위 내의 기업들에게 제공하고 추후 각 기업으로부터 복호화 요청을 받아 이를 수행
  - 데이터 공유가 1 대 多 또는 多 대 多로 이루어지기 때문에 전담 기관(agent, dam or pool...)을 두는 것이 비용 중복을 줄일 수 있으며 공유업무 처리 전문기관 없이 기업들끼리 직접 공유하도록 하는 경우 기술 표준화의 문제나 의도적 지연 등의 문제도 우려될 수 있음
  - 어느 나라나 은행 등 대출기관들은 신용정보(credit information)을 공유하고 있는데 Credit Bureau 또는 Public Credit Registry가 중간에서 공유 작업을 수행하는 것과 유사

# 관련 법규 해석 및 개정 이슈

## ❖ 데이터 집종의 경쟁 제한성은 누가, 어떤 절차로 판단?

- 공정위가 데이터 필요 기업으로부터 문제 제기를 받아서? 심사 기간 동안의 데이터 가치 하락 문제, 중요 데이터의 사전적 공유 명령제가 필요?

## ❖ 금융데이터의 공유 명령은 누가?

- 공정위 or 금융위?
- 금융데이터와 비금융데이터의 공유가 필요한 경우는?
- 개인정보보호위원회의 역할은?

## ❖ 동형암호화 및 DP 처리된 데이터는 익명정보인가 가명정보인가?

- 개인정보보호법 및 신용정보법 해석 문제

감사합니다.