



CreDB

금융 빅데이터 개방시스템

소개 및 시연

2019.6.3



한국신용정보원
Korea Credit Information Services

1 개요

데이터 개방(data openness)은 디지털 경쟁(digital competition)을 촉진하기 위한 필수요소입니다.



2 한국신용정보원 소개 및 금융데이터의 특성

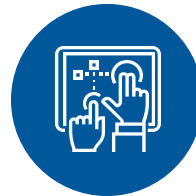
한국신용정보원은 신용정보법에 의한 종합신용정보집중기관으로서 신용정보의 **안전한 집중관리**와 **효과적인 활용**을 위해 2016년 1월 설립되었습니다.



3 금융 빅데이터 개방시스템(CreDB)

신용정보원은 설문조사 및 심층인터뷰(17년), 시범서비스(18년) 등을 통해 금융 빅데이터에 대한 수요를 조사하고, 활용 활성화를 위한 필요사항을 고민하였습니다.

금융 빅데이터 활용 활성화



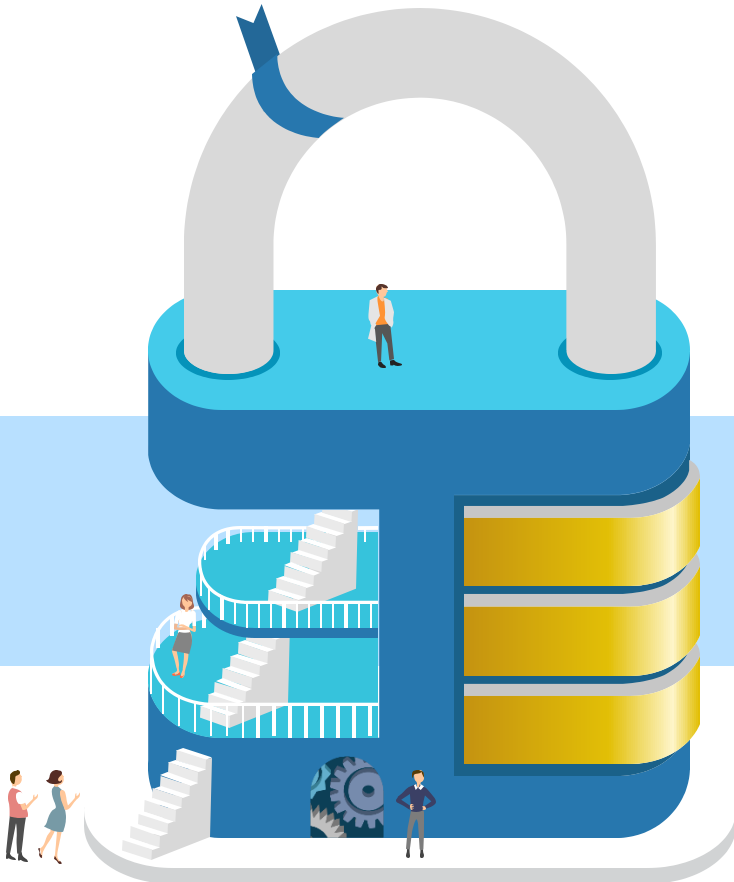
활용의 자유도가 높은
Raw data 형태



다양한 분석 소프트웨어와
고성능 작업 환경



안전한 데이터 활용을 위한
데이터 보안



3 금융 빅데이터 개방시스템(CreDB)

금융 빅데이터 개방시스템은 신용정보를 금융서비스 개발, 금융소비자연구 등의 목적으로 이용하고자 하는 이용자에게 분석 **데이터**와 분석 **환경**을 제공하는 서비스입니다.



4 신용정보 DB

신용정보DB는 활용 목적에 따라 **표본DB**, **교육용DB**와 **맞춤형DB**로 구분됩니다.

신용정보DB

표본DB
('19년 내)

교육용DB
('19년 하반기)

맞춤형DB
('20년 상반기)

**개인신용
표본DB**
('19.6월)

**기업신용
표본DB**
('19년 내)

**보험신용
표본DB**
('19년 내)



4 신용정보 DB

개인신용정보 표본DB는 '19.6.4일 서비스 개시하며, 약 200만명에 대한 차주, 연체, 대출 및 카드개설정보를 제공합니다.

2019년 6월

2019년 하반기

2019년 하반기

2020년 상반기

개인신용정보 표본DB

- 우리나라 전체 신용활동 인구의 5%(약 200만명)을 Random sampling
- 차주, 연체, 대출 및 카드개설정보 등 4개 테이블 25개 속성*으로 구성
* 테이블 및 속성은 단계적으로 확장할 계획 (대출금리, 대출상환방식, 월별카드이용금액등)
- '15.12월 ~ '18.12월(총 37개월) + 매분기 업데이트
- 프라이버시 보호를 위한 비식별조치* 수행
* 데이터 범주화, 총계처리, 데이터 삭제, 저빈도값보정등

차주정보
차주일련번호
생년
성별

대출정보
기준년월
차주일련번호
업권코드
기관일련번호
대출상품코드
대출년월
대출금액

연체정보
기준년월
차주일련번호
업권코드
기관일련번호
연체유형코드
연체사유코드
등록사유코드
연체등록년월
연체등록금액

카드개설정보
기준년월
차주일련번호
기관일련번호
개설사유코드
카드유형코드
개설년월

4 신용정보 DB

개인신용정보 표본DB는 '19.6.4일 서비스개시하며, 약 200만명에 대한 차주, 연체, 대출 및 카드개설 정보를 제공합니다.

2019년 6월

2019년 하반기

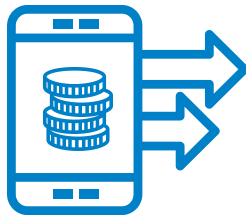
2019년 하반기

2020년 상반기

개인신용정보 표본DB

A 핀테크회사: 개인신용정보 표본DB를 이용한 소액신용대출 상품 개발

소액신용대출
상품 개발



일반DB

1단계 시장탐색

연령·업권·대출금액·기간별
신용대출 이용현황 분석

목표고객군 선정

✓ 30-40대 소액대출 이용자

2단계 서비스준비

목표고객군에 대한 맞춤형
신용평가모형을 개발



3단계 서비스개시



개인신용정보
조회동의 후
조회 가능

개인DB는
비식별조치된 DB이기 때문에
마케팅 활용이 불가능

4 신용정보 DB

2019년 하반기에는 법인과 개인사업자에 대한 **기업신용정보 표본DB**와 생명·장기·손해보험에 대한 **보험신용정보 표본DB**(신용정보원, 보험개발원)를 공개할 계획입니다.

2019년 6월

2019년 하반기

2019년 하반기

2020년 상반기

기업신용정보 표본DB

- 신용정보원에 정보가 등록된 법인·개인사업자의 약 20%를 Random sampling
- 기업개요, 차주, 대출 및 연체정보 등으로 구성

B 창업컨설팅: 소상공인 컨설팅 서비스 제공

Q 창업희망지역음식업 사업자의 평균 부채규모는?

보험신용정보 표본DB

신용정보원

생명·장기보험
중심의 DB 개발

보험개발원

손해보험 중심의
DB 개발

C 보험사: 맞춤형 알뜰 보험상품 개발

Q 내게 부족한보장을 메워주면서도
보험료가 부담스럽지 않은 보험은?

4 신용정보 DB

2019년 하반기에는 법인과개인사업자에 대한 **기업신용정보 표본DB**와 생명·장기·손해보험에 대한 **보험신용정보 표본DB**(신용정보원, 보험개발원)를 공개할 계획입니다.

2019년 6월

2019년 하반기

2019년 하반기

2020년 상반기

기업신용정보 표본DB

B 창업컨설팅: 소상공인 컨설팅 서비스 제공

대출, 얼마를 받아야 하지?



기업DB

지역·업종별
부채·연체 현황 분석



평균부채수준과
한계부채수준을 계산



한식 음식점 사업자들은
평균 □만원 대출을 받아요
△만원 이상으로는 대출을
받지 않는 것이 좋겠습니다.



하마터면 무리하게
대출을 받을 뻔 했네!



4 신용정보 DB

2019년 하반기에는 법인과 개인사업자에 대한 **기업신용정보 표본DB**와 생명·장기·손해보험에 대한 **보험신용정보 표본DB**(신용정보원, 보험개발원)를 공개할 계획입니다.

2019년 6월

2019년 하반기

2019년 하반기

2020년 상반기

보험신용정보 표본DB

C 보험사: 맞춤형 알뜰 보험상품 개발

보험료가 부담스러워
해지했더니 손해가...



보험 계약자



성별·연령별 보험가입·유지
현황을 보험상품별로 분석



가입 보장항목과
조기해약 현황 분석



부족한 보장과
보험료 부담이 없는
실속 있고 알뜰한
보험 가입을 추천합니다.



큰 부담 없이
필요한 보장만 가입



보험 계약자

4 신용정보 DB

신용정보 교육용DB는 개인정보 유출 위험이 없는 가상 데이터로서, 데이터 전문가 양성을 위해 대학 등 교육기관에 제공할 계획입니다.

2019년 6월

2019년 하반기

2019년 하반기

2020년 상반기

신용정보 교육용DB

- 교육용 재현데이터를 대학 등 교육기관에 제공하여 데이터 전문가 양성에 기여
* 재현데이터란? 원본데이터의 통계적 특성을 보유하면서 가상으로 구현한 데이터로서, 개인정보 유출의 위험이 적음
- 표본DB·맞춤형DB와 달리 이용자에게 파일 형태로 제공



실제데이터



재현데이터

대학 신입생 D씨

- Q 통계 소프트웨어 실습을 위한 데이터?
- A 실제 금융산업의 특성과 유사하게 구성된 신용정보 교육용DB를 이용하여 현실 접목이 가능한 지식과 기술을 습득

박사과정 E씨

- Q 신용평가 방법론간의 예측성능 비교?
- A 신용정보 교육용DB를 이용하여 데이터를 탐색하고 소스코드를 작성한 뒤, 개인신용정보 표본DB를 이용하여 실제 연구 수행

4 신용정보 DB

맞춤형DB는 개별 이용자의 분석 목적에 따라 샘플수, 비식별조치 수준, 제공항목 등을 맞춤형으로 가공하여 제공하는 서비스입니다.

2019년 6월

2019년 하반기

2019년 하반기

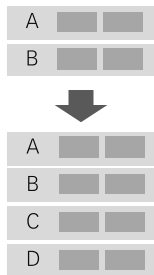
2020년 상반기

맞춤형 DB

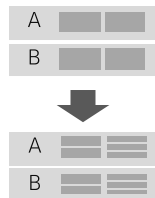
- 맞춤형DB는 개별 이용자의 분석 목적에 따라 정보를 맞춤형으로 추출하여 제공하는 서비스
* 표본DB는 제공 정보량 및 항목이 고정

Customization 방안

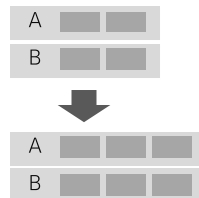
샘플수 확대



비식별조치 수준 변경



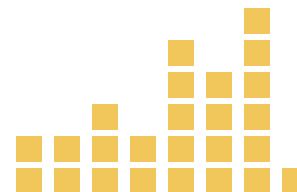
제공항목 확대



F 연구소

- Q 과다한 채무를 가진 20대 초반 청년을 연구?
- A 개인신용정보 표본DB(5% 샘플링)을 기준 20대 초반의 다중채무자는 수백 명에 불과
→ 맞춤형DB를 통해 샘플링 비율을 20%로 확대

5% 샘플링



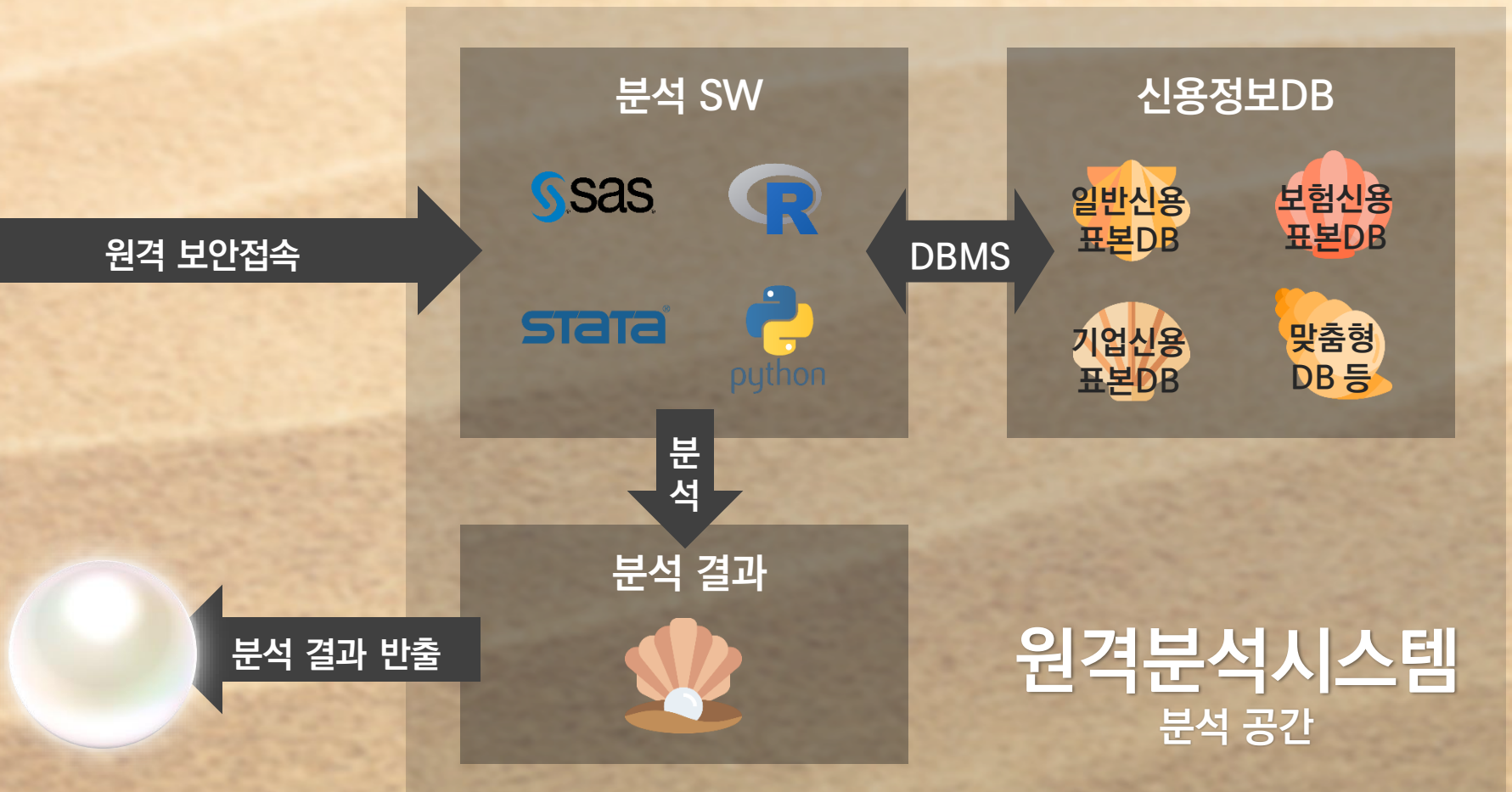
20% 샘플링



5 원격분석시스템

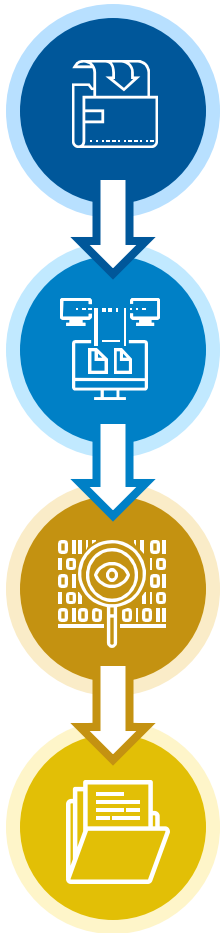
이용자는 원격분석시스템*에서 제공하는 **분석 소프트웨어와 분석 공간** 등을 이용하여 제공DB를 안전하게 분석한 뒤, 분석 결과물을 반출할 수 있습니다.

* 총 40 계정, 윈도우 10(64bit), 메모리 8GB, 저장공간 110GB, SAS·STATA·R·Python·DBMS 이용가능



6 서비스 신청 및 이용 방법

이용 신청은 **서비스 포털(credb.kcredit.or.kr)**을 통해 이루어지며,
제공 대상 및 목적에 대한 심의를 거쳐 신용정보 DB를 제공하게 됩니다.



접수 credb.kcredit.or.kr

연구책임자와 공동연구원은 금융 빅데이터 개방 시스템 포털을 통해 이용신청서를 접수

심의

비식별정보제공심의위원회(외부 전문가 포함 7인으로 구성)를 통해
제공 대상, 목적 및 방법의 적정·타당성을 검토

분석

신용정보원 원격분석시스템을 이용하여 제공 DB를 분석
(SAS, STATA, R, Python 등 통계패키지 제공)

반출

반출신청시 raw data 포함여부 등을 검토한 뒤 반출 승인



7 기대효과

신용정보원의 금융 빅데이터 개방은
빅데이터 시장 조성 및 인공지능 산업 발전을 위한 첫 걸음입니다.

빅데이터 초기 시장 조성

폐쇄적인 금융 인프라,
불합리한 규제 등으로 인해
**금융데이터의
산업적 활용에 제약**



공공인프라로서
신용정보원 데이터를 개방하여
데이터 산업의 성장을 지원

인공지능 산업의 기반 마련

효과적인 딥러닝을 위해서는

- ① 대량의 데이터
- ② 계량화 된 최적화 목표
- ③ 유사한 훈련&판단 대상



딥러닝에 필요한 대량의
기초 데이터 지원을 통해
다양한 영역의 AI 개발을 촉진

8 향후 계획

신용정보원은 금융 빅데이터 개방시스템을 시작으로
다양한 산업의 민간·공공을 아우르는 **데이터생태계 조성**을 위해 노력하겠습니다.



감사합니다