

Big Data 기술 개요 및 미래 발전 방향 및 대표 서비스 사례 Top 5

네트워크 기초 주제 7.

INDEX 목차

1. Big Data 개요
2. Big Data 특징
3. 미래 발전 방향
4. 대표 서비스 사례 Top5

Big data?

- 기존 데이터 처리 방법으로 처리하기 힘든 **거대한 양의 데이터**
- 여러 데이터가 모여 **의미 있는** 단위로 묶인 데이터
- 거대한 양의 데이터를 **처리**하는 방법



사진 1. BIG DATA 표현 사진,
<https://www.pharmstoday.com/news/articleView.html?idxno=321147>

전통적인 data와 Big data의 차이점

	전통적인 데이터	빅 데이터
크기	Giga Byte or Tera Byte	Peta Byte 이상
구성 방법	Static Schema	Dynamic Schema
관리 아키텍처	중앙 집중식 아키텍처	분산식 아키텍처
데이터 소스	ERP ¹⁾ , CRM ²⁾ , 온라인 트랜잭션 ³⁾	SNS, 장치 Sensor, 시청각 데이터

표 1. 전통적인 데이터와 빅 데이터의 비교

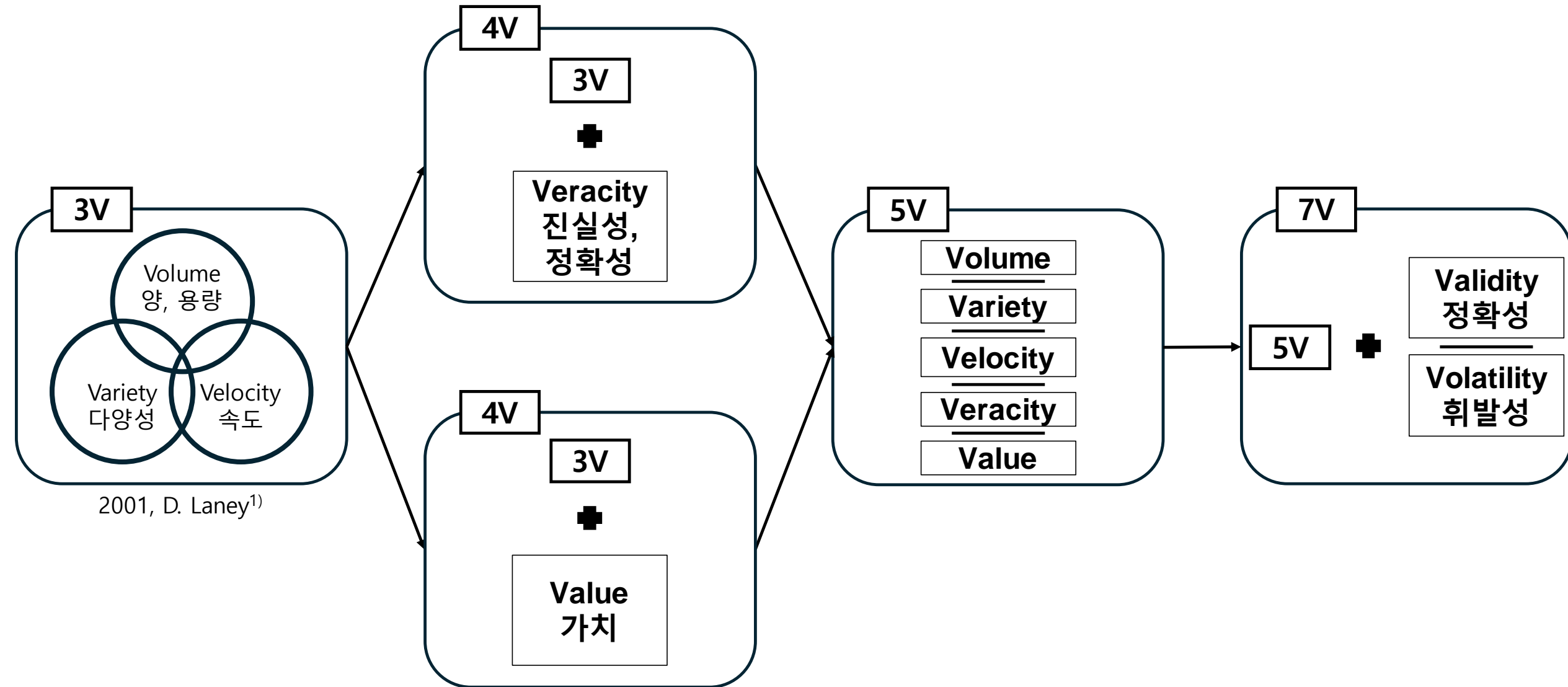
¹⁾ERP : 재무, 인사관리, 제조, 공급망 등 비즈니스의 여러 부문의 자동화 소프트웨어

²⁾CRM : 고객 관계 관리, 연락처, 담당자와의 상호작용 내역, 구매 이력 등의 고객 데이터

³⁾온라인 트랜잭션 : 온라인 banking, 쇼핑, 주문 입력 등 동시에 발생하는 다수의 데이터를 즉각적으로 처리

전통적인 data 장점	Big data 장점
<ul style="list-style-type: none">- 보안이 쉬움 개인적인 데이터, 기밀 데이터에 적합 크기가 작아 하나의 서버에서 운용 가능- 조작 및 해석이 더 쉬움 단순하고 관계형, 비전문가도 활용 가능	<ul style="list-style-type: none">- 심층분석 가능 시장 동향, 소비자 행동에 대한 심층 분석- 경쟁력 우위 빅데이터로부터 실시간으로 많은 정보 국제 사회에서 생성되는 엄청난 데이터로부터 의미 있는 인사이트 도출

표 2. 전통적인 데이터와 빅 데이터 각각 가지고 있는 장점



¹⁾ D. Laney:시장조사업체 애널리스트


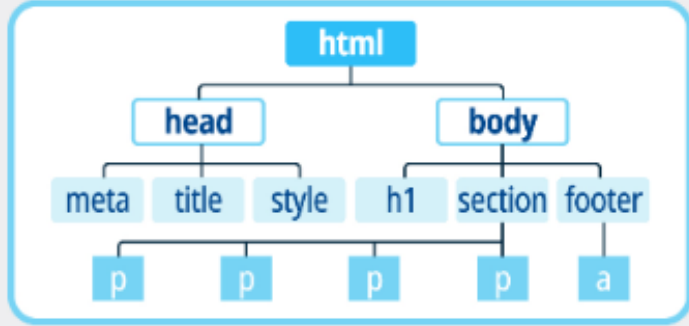

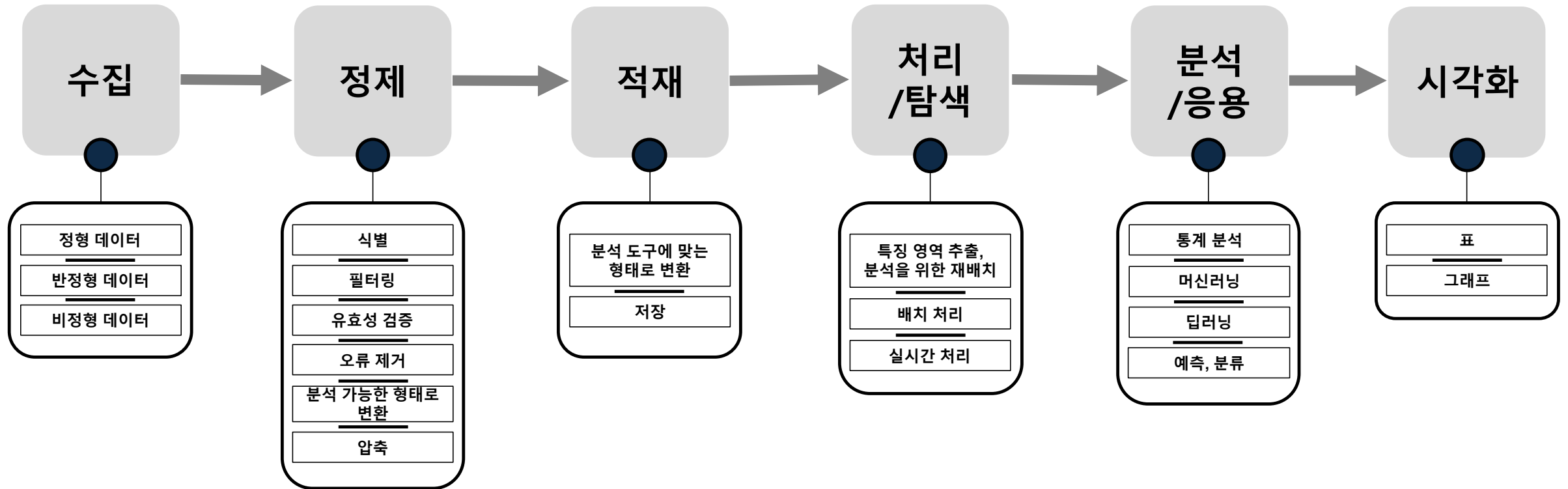
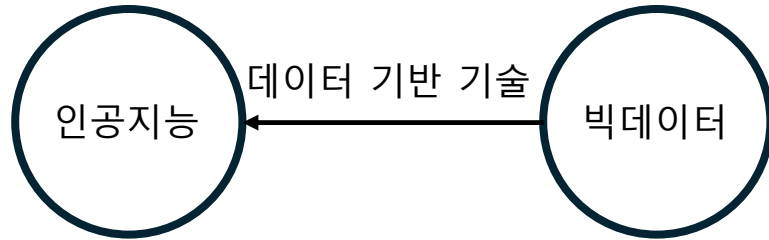
정형 데이터 Structured Data	반정형 데이터 Semi-Structured Data	비정형 데이터 Unstructured Data
<ul style="list-style-type: none">미리 정해 놓은 형식과 구조에 따라 저장된 데이터명확한 데이터 구조	<ul style="list-style-type: none">데이터의 형식과 구조가 변경될 수 있는 데이터데이터의 구조 정보를 데이터와 함께 제공	<ul style="list-style-type: none">정의된 구조가 없는 데이터데이터 자체만으로 내용에 대한 질의 처리가 불가
CSV	html, XML, JSON	오디오, 이미지, 영상
		

표 3. 여러 형태의 데이터, 정보통신용어사전

기존의 데이터 처리 방법으로는 비정형 데이터에 대한 처리가 힘들



<빅데이터 분석과 활용에 대한 수요 급증>



4차 산업혁명 기반 기술



사진 2. 세계 빅데이터 시장 전망
, VMR(Verified Market Research)

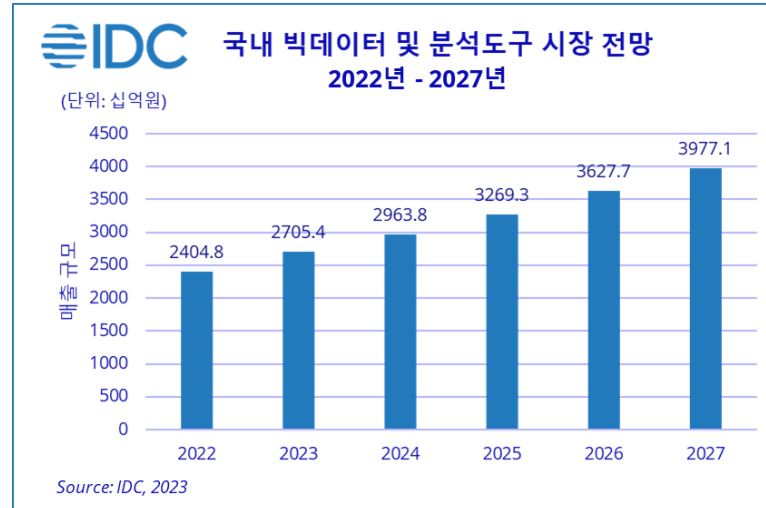


사진 3. 국내 빅데이터 시장 전망, IDC



사진 4. 빅데이터 보안 시장 전망
, AMR(Alied Market Research)

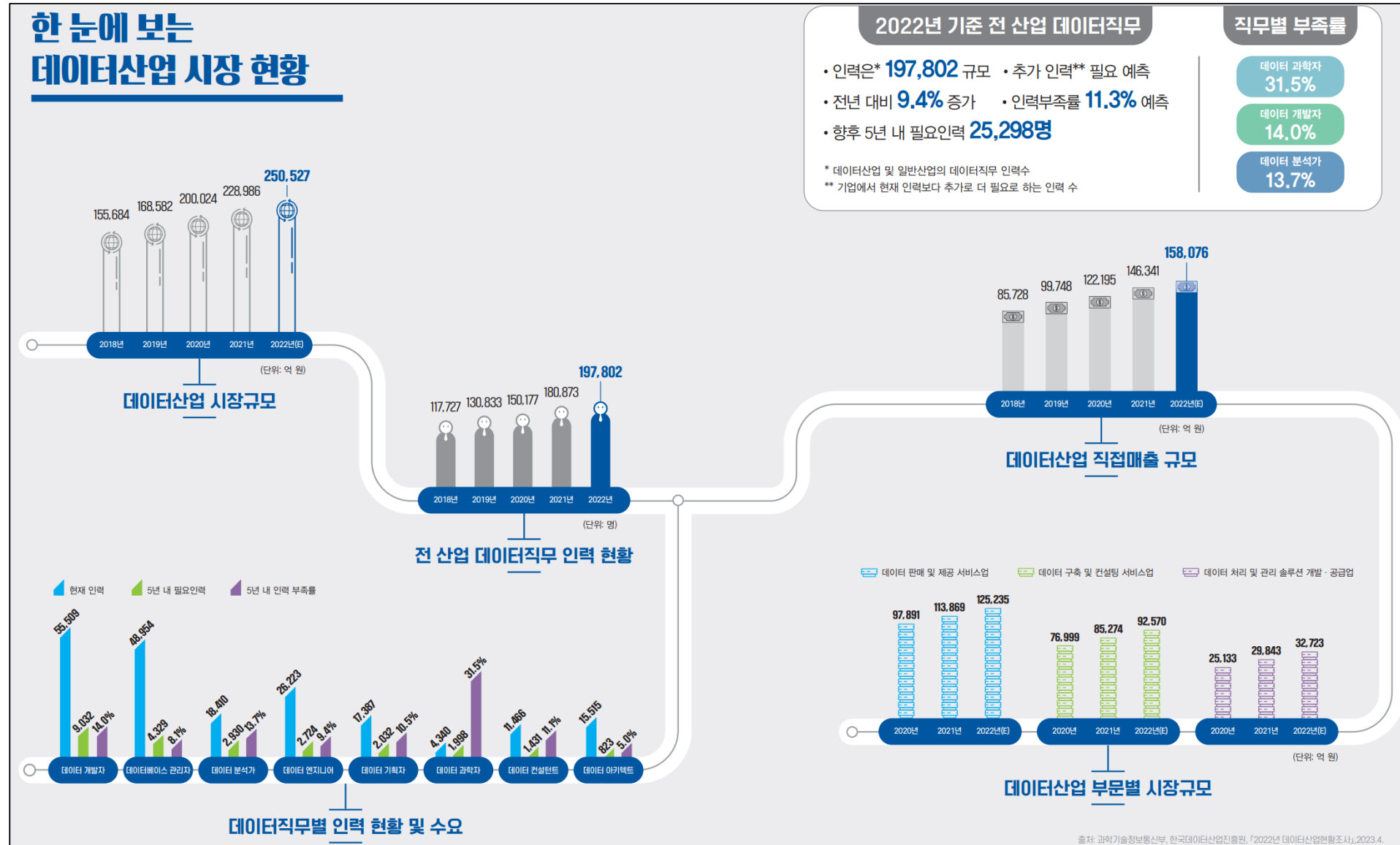
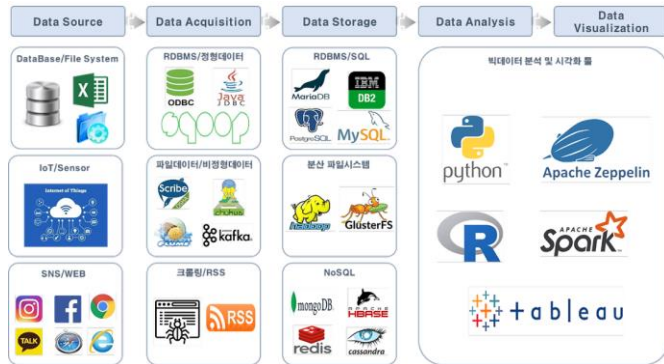


사진 6. 국내 빅데이터 산업 법제화 현황, 한국데이터산업진흥원

빅데이터



인공지능

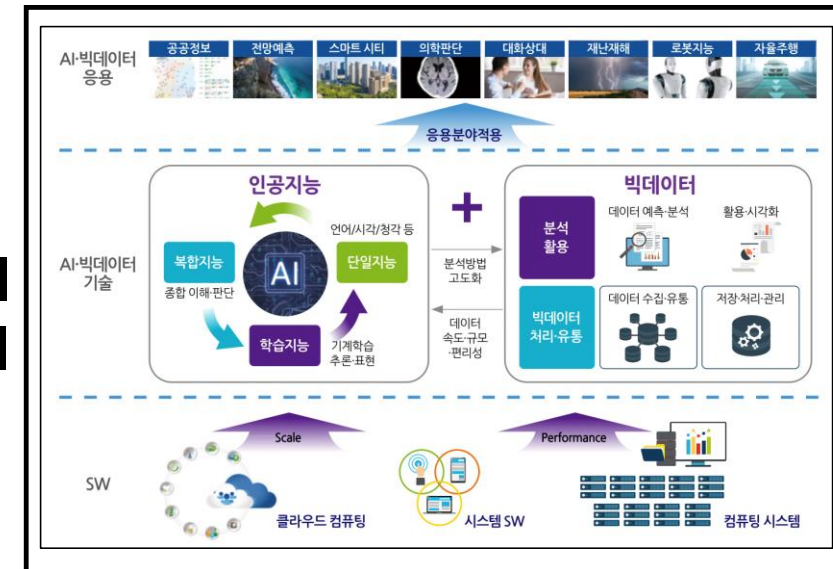
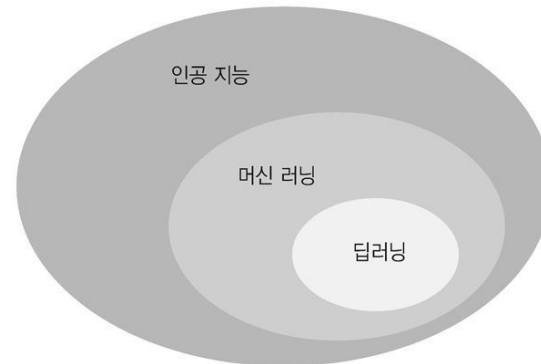


사진 7. 빅데이터 처리 프로세스,
<https://spacefordata.tistory.com/12?category=1113292>

사진 8. 인공지능, 머신러닝, 딥러닝 관계,
<https://tensorflow.blog/%EC%BC%80%EB%9D%BC%EC%8A%A4-%EB%94%A5%EB%9F%AC%EB%8B%9D/1-%EB%94%A5%EB%9F%AC%EB%8B%9D%EC%9D%B4%EB%9E%80-%EB%AC%B4%EC%97%87%EC%9D%B8%EA%B0%80/>

사진 9. 인공지능과 빅데이터 기술 동향, 한국정보통신기술협회

		연도별 시장규모							증감률	연평균 증감률
		2016년	2017년	2018년	2019년	2020년	2021년	2022년 (E)	2020-2021	2020-2022(E)
데이터 처리 및 관리 솔루션 개발·공급업	데이터 수집·연계 솔루션 개발·공급업	1,345	1,393	1,622	1,871	2,122	2,499	2,715	17.8%	13.1%
	데이터베이스 관리 시스템 솔루션 개발·공급업	6,148	6,121	6,775	7,510	8,979	11,021	11,787	22.7%	14.6%
	데이터 분석 솔루션 개발·공급업	1,249	1,325	1,782	2,014	2,586	2,932	3,247	13.4%	12.1%
	데이터 관리 솔루션 개발·공급업	5,367	4,628	4,972	5,203	6,022	7,137	7,963	18.5%	15.0%
	데이터 보안 솔루션 개발·공급업		1,213	1,517	1,975	2,558	2,894	3,015	13.1%	8.6%
	빅데이터 통합 플랫폼 솔루션 개발·공급업	1,611	1,776	1,949	2,231	2,866	3,359	3,995	17.2%	18.1%
소계		15,720	16,457	18,617	20,805	25,133	29,843	32,723	18.7%	14.1%
데이터 구축 및 컨설팅 서비스업	데이터 구축/가공 서비스업	54,571	57,207	58,993	62,223	72,805	80,403	87,366	10.4%	9.5%
	데이터 관련 컨설팅 서비스업	1,279	1,687	2,297	3,189	4,194	4,871	5,204	16.1%	11.4%
	소계	55,850	58,894	61,290	65,412	76,999	85,274	92,570	10.7%	9.6%
데이터 판매 및 제공 서비스업	데이터 판매/중개 서비스업	6,123	6,608	8,198	11,332	16,054	20,861	22,194	29.9%	17.6%
	정보제공 서비스업	59,854	61,570	67,580	71,033	81,838	93,008	103,040	13.6%	12.2%
	소계	65,977	68,179	75,778	82,364	97,891	113,869	125,235	16.3%	13.1%
데이터산업 전체		137,547	143,530	155,684	168,582	200,024	228,986	250,527	14.5%	11.9%

사진 10. 빅데이터 시장 규모, 한국데이터산업진흥원

4. 대표 서비스 사례 Top5

1. Big Data 조회 서비스
2. 서울시 올빼미 버스
3. ZARA Fast Fashion
4. 쿠팡
5. 빅데이터 기반 사기 예방 서비스

검색어 인사이트

다음 통합검색의 검색어 트렌드를 확인해보세요.
다양한 옵션으로 조회할 수 있습니다.

검색어를 입력한 다음 + 버튼을 눌러주세요.

[서비스 가이드](#)

기간 검색어 조회는 2018년 부터 제공됩니다.

전체
1개월
3개월
✓ 1년

2023.05.11. - 2024.05.11.

기준
일간

기기 PC 선택 시 지역 정보는 제공하지 않습니다.

✓ 전체
모바일
PC

성별

✓ 전체
여성
남성

연령

✓ 전체
10대
20대
30대
40대
50대
60대이상

지역

전국

초기화
검색어 조회하기

사진 11. 카카오 데이터 트렌드 홈페이지

NAVER DataLab.

[데이터랩 홈](#)
[검색어트렌드](#)
[쇼핑인사이트](#)
[지역통계](#)
[댓글통계](#)

쇼핑인사이트

쇼핑 분야별 클릭 추이와 분야별 검색어 현황을 확인할 수 있습니다.

분야별 인기 검색어
인기분야

화장품/미용

일간

2024.05.08.(수)	2024.05.09.(목)	2024.05.10.(금)	2024.05.11.(토)
1 디올립글로우	1 와이플러스여성청결제	1 와이플러스여성청결제	1 와이플러스여성청결제
2 퍼퓸샴푸	2 마데카르뮴	2 루켄도르샴푸바	2 루켄도르샴푸바
3 마스크팩	3 디올립글로우	3 피토메스남자화장품세트	3 피토메스남자화장품세트
4 설화수	4 피토메스남자화장품세트	4 마데카르뮴	4 클리오팩트에나멜에디션
5 우먼여성청결제	5 샴푸	5 샴푸	5 마데카르뮴
6 설화수자음2종	6 트리트먼트	6 바디워시	6 바디워시
7 ahc아이크림	7 바디워시	7 트리트먼트	7 샴푸바
8 샴푸	8 헤라블랙쿠션	8 디올립글로우	8 디올립글로우
9 마데카르뮴	9 매스티나광채쿠션	9 마스크팩	9 샴푸
10 피토메스남자화장품세트	10 설화수	10 샴푸바	10 트리트먼트

[검색어 통계 자세히 보기 >](#)

사진 12. 네이버 데이터랩 홈페이지

카카오데이터트렌드와 네이버 데이터랩		
	카카오데이터트렌드	네이버 데이터랩
개시	2020년 7월	2016년 4월
정보 기준	다음 검색어	네이버 검색어
제공 정보	기간, 성별, 연령, 지역 등	
특징	조회 내용을 파일로 다운로드할 수 있어 연구 등에 활용 용이	일반 분야 외 쇼핑 특화 검색 정보가 있어 사업자에 유용

〈자료: 카카오·네이버〉

사진 13. 카카오 데이터 트렌드와 네이버 데이터 랩 비교,
<https://www.kmib.co.kr/article/view.asp?arcid=0014855973>

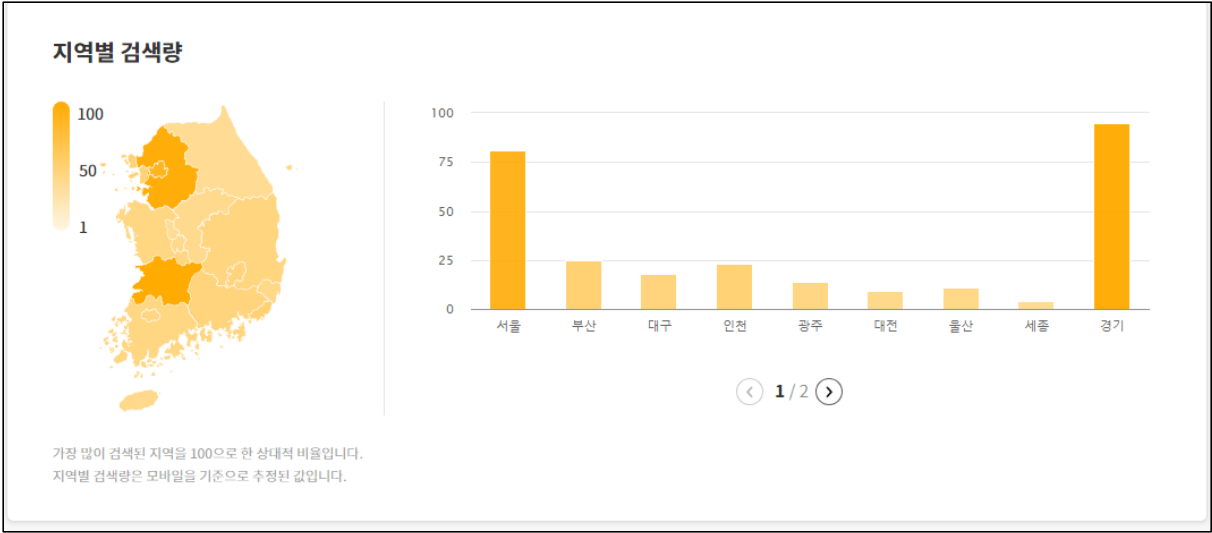
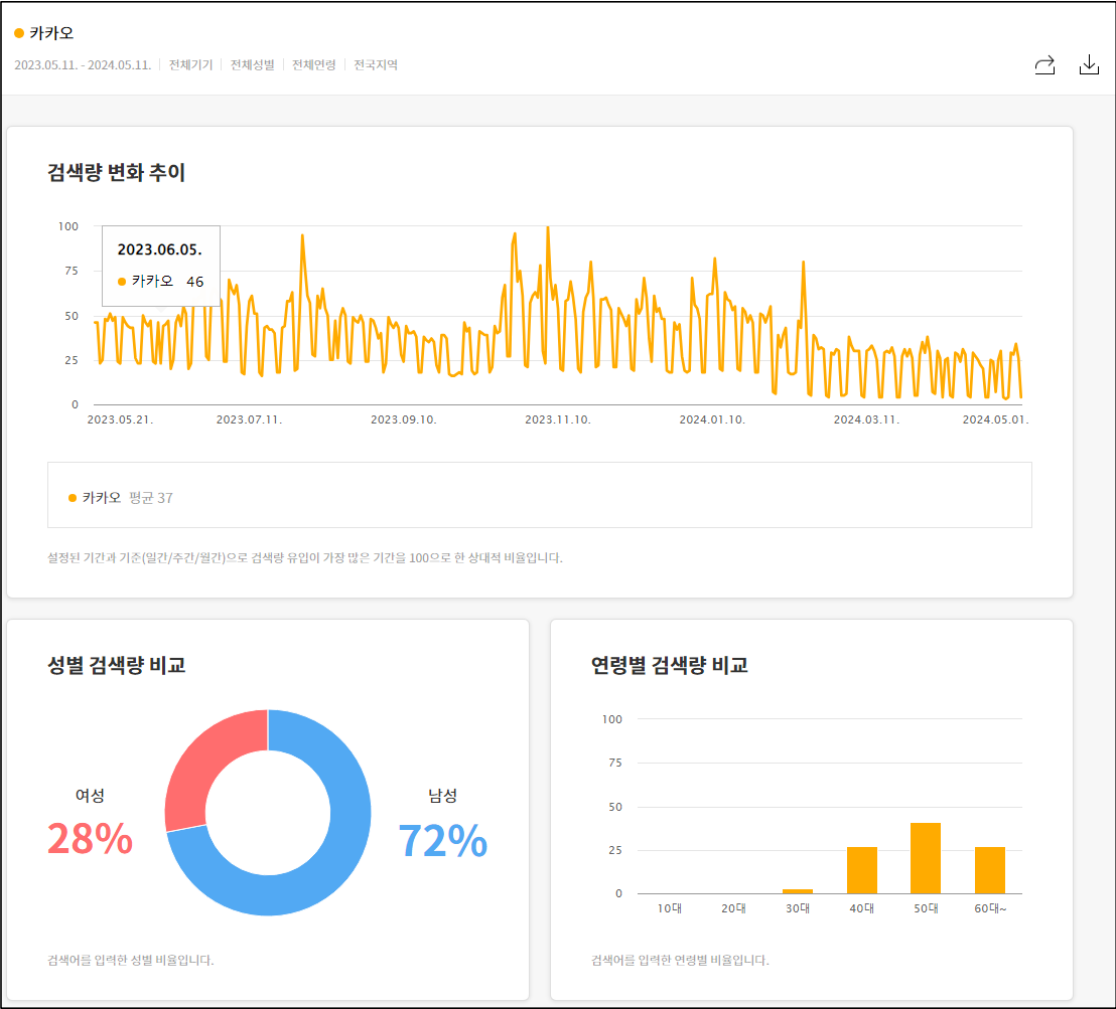


사진 14. 카카오 데이터 트렌드 '카카오' 검색 결과

1	기간	2023-05-11 ~ 2024-05-11	
2	기기	전체	
3	성별	전체	
4	연령	전체	
5	지역	전체	
6			
7	주제	통합검색, 검색어 변화량 추이	
8	일	카카오	
9	2023-05-11	46	
10	2023-05-12	46	
11	2023-05-13	23	
12	2023-05-14	25	
13	2023-05-15	48	
14	2023-05-16	47	
15	2023-05-17	51	
16	2023-05-18	47	
17	2023-05-19	49	
18	2023-05-20	24	
19	2023-05-21	23	
20	2023-05-22	49	
21	2023-05-23	46	
22	2023-05-24	44	
23	2023-05-25	43	
24	2023-05-26	43	
25	2023-05-27	26	
26	2023-05-28	23	
27	2023-05-29	23	
28	2023-05-30	50	
29	2023-05-31	46	
30	2023-06-01	44	
31	2023-06-02	47	
32	2023-06-03	24	

주제	통합검색, 성별 검색량 비교
성별	카카오
여성	28
남성	72
주제	통합검색, 연령별 검색량 비교
연령	카카오
10대	0
20대	0
30대	3
40대	27
50대	41
60대 이상	27
주제	통합검색, 지역별 검색량
지역	카카오
강원	4
경기	95
경남	18
경북	17
광주	14
대구	18
대전	9
부산	25
서울	81
세종	4
울산	11
인천	23
전남	15
전북	100
제주	9
충남	15
충북	7

사진 15. 카카오 데이터 트렌드 엑셀 파일 변환

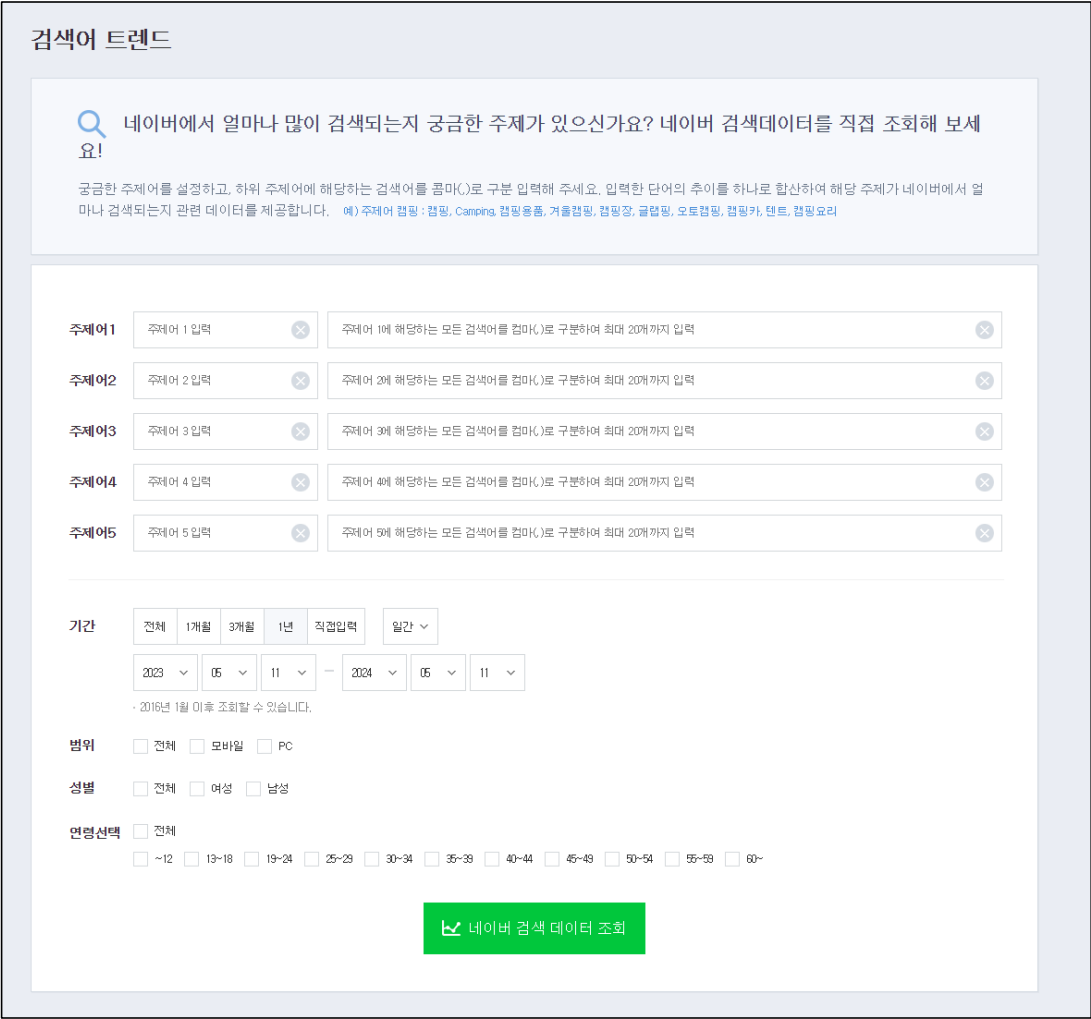


사진 16. 네이버 데이터랩 검색어 트렌드

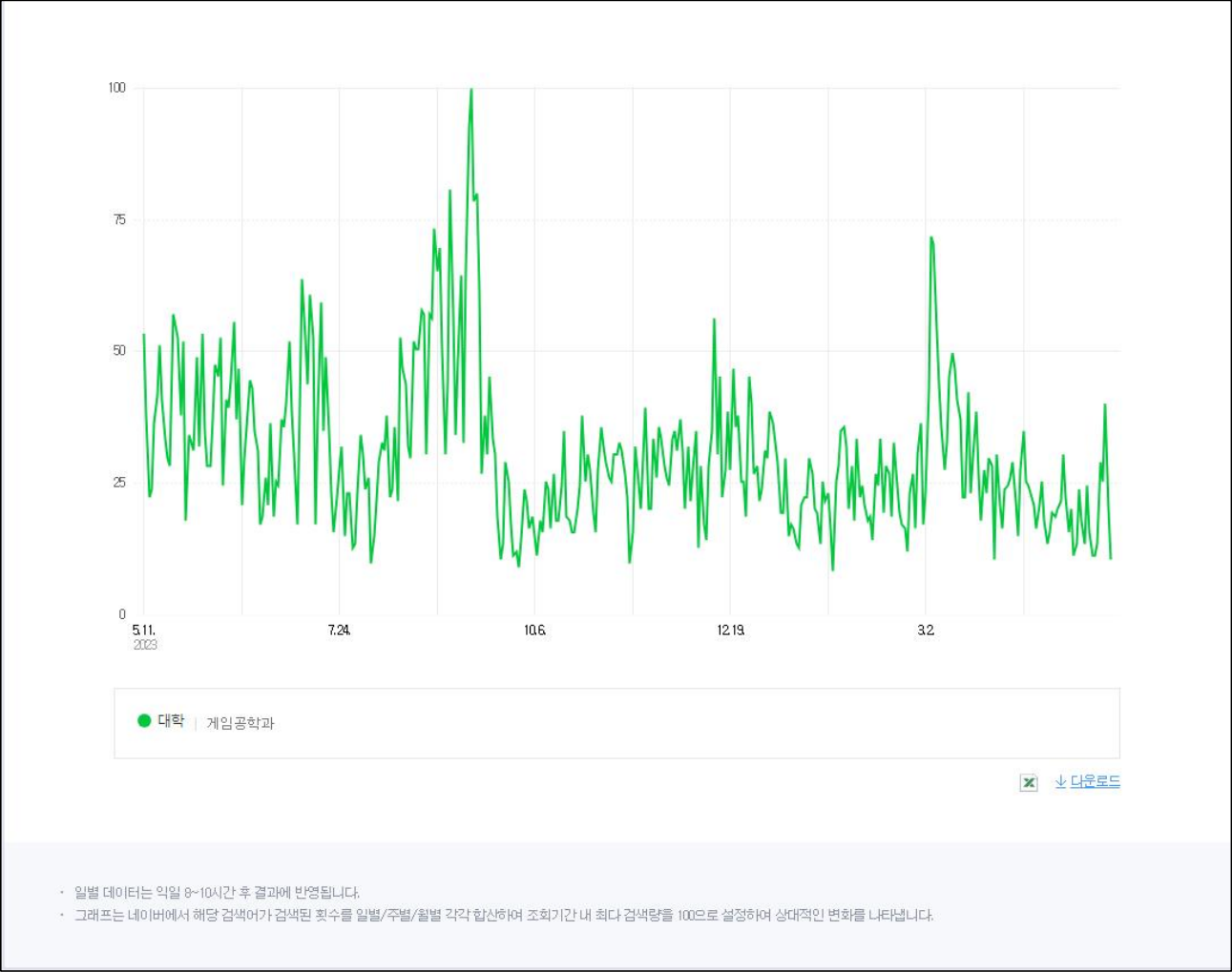


사진 17. 대학-게임공학과 입력 결과

✓ 지역별 검색 관심도 업종별 검색 관심도

☐ 전체선택 초기화

- 업종은 하나만 선택가능하며, 대분류 업종 또는 중분류 업종의 관심도 변화를 확인하고 싶은 경우는 하위 분류를 선택하지 않고 조회합니다.
- 지역은 복수선택이 가능하나, 상위 지역이 다르거나 지역 단위가 다른 지역들은 같이 선택할 수 없습니다.
(ex. 속초시/교동, 춘천시/신북읍 지역은 함께 선택 불가)
- 하위 지역 선택에서 '전체'를 선택하는 경우 해당 지역 상위 지역의 모든 하위 지역들의 개별적인 관심도를 모두 조회하며, 하위 지역을 선택하지 않은 경우는 선택된 지역 단위까지만 조회합니다.

STEP1 업종선택

업종 대분류	중분류	소분류
기업,기관,일반장소	선택없음	
음식점	주점	
패션,뷰티	카페,디저트	
취미,오락	치킨	
1차산업	중식	
2차산업	음식점 기타	
기타서비스업	분식	
레저	일식	
생활	한식	

STEP2 지역선택

광역시/도	시/군/구	읍/면/동
<input checked="" type="checkbox"/> 서울특별시	<input type="checkbox"/> 전체	
<input type="checkbox"/> 부산광역시	<input type="checkbox"/> 종로구	
<input type="checkbox"/> 대구광역시	<input type="checkbox"/> 중구	
<input type="checkbox"/> 인천광역시	<input type="checkbox"/> 용산구	
<input type="checkbox"/> 광주광역시	<input type="checkbox"/> 성동구	
<input type="checkbox"/> 대전광역시	<input type="checkbox"/> 광진구	
<input type="checkbox"/> 울산광역시	<input type="checkbox"/> 동대문구	
<input type="checkbox"/> 세종특별자치시	<input type="checkbox"/> 중랑구	
<input type="checkbox"/> 경기도	<input type="checkbox"/> 성북구	

조회할 내용은 다음과 같습니다.

음식점 업종에 대한 서울특별시 지역의 검색 관심도

사진 18. 네이버 데이터랩 검색 관심도 서비스

음식점 업종에 대한 서울특별시 > 종로구 하위 지역들의 검색 관심도

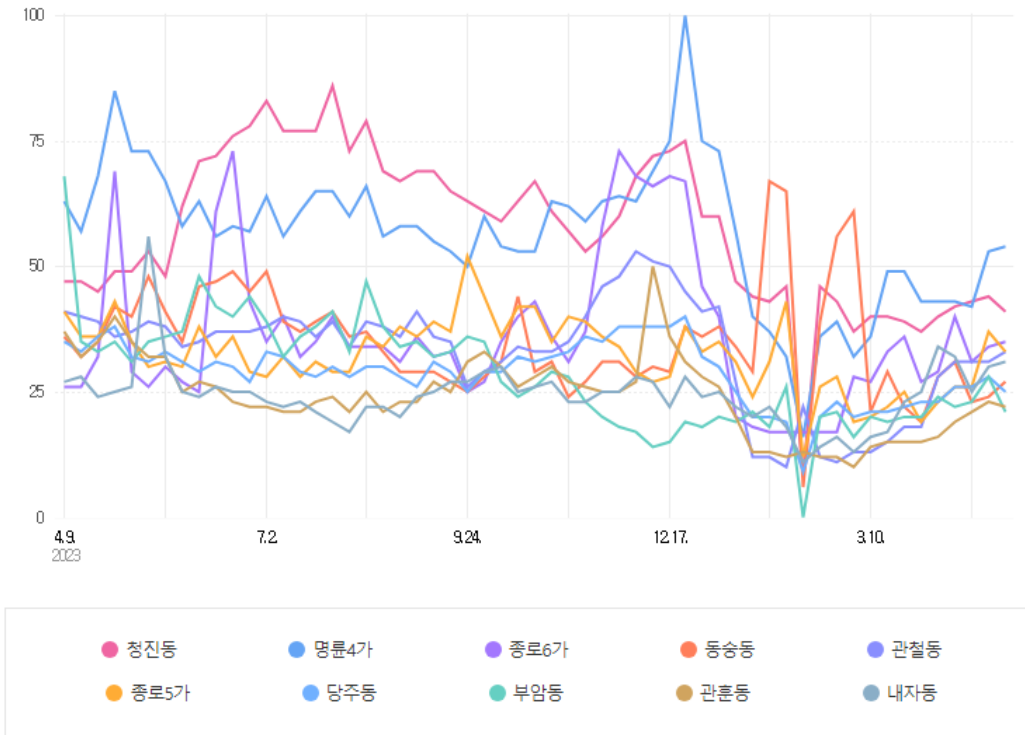


사진 19. 음식점-종로구 업종별 검색 관심도 검색 결과

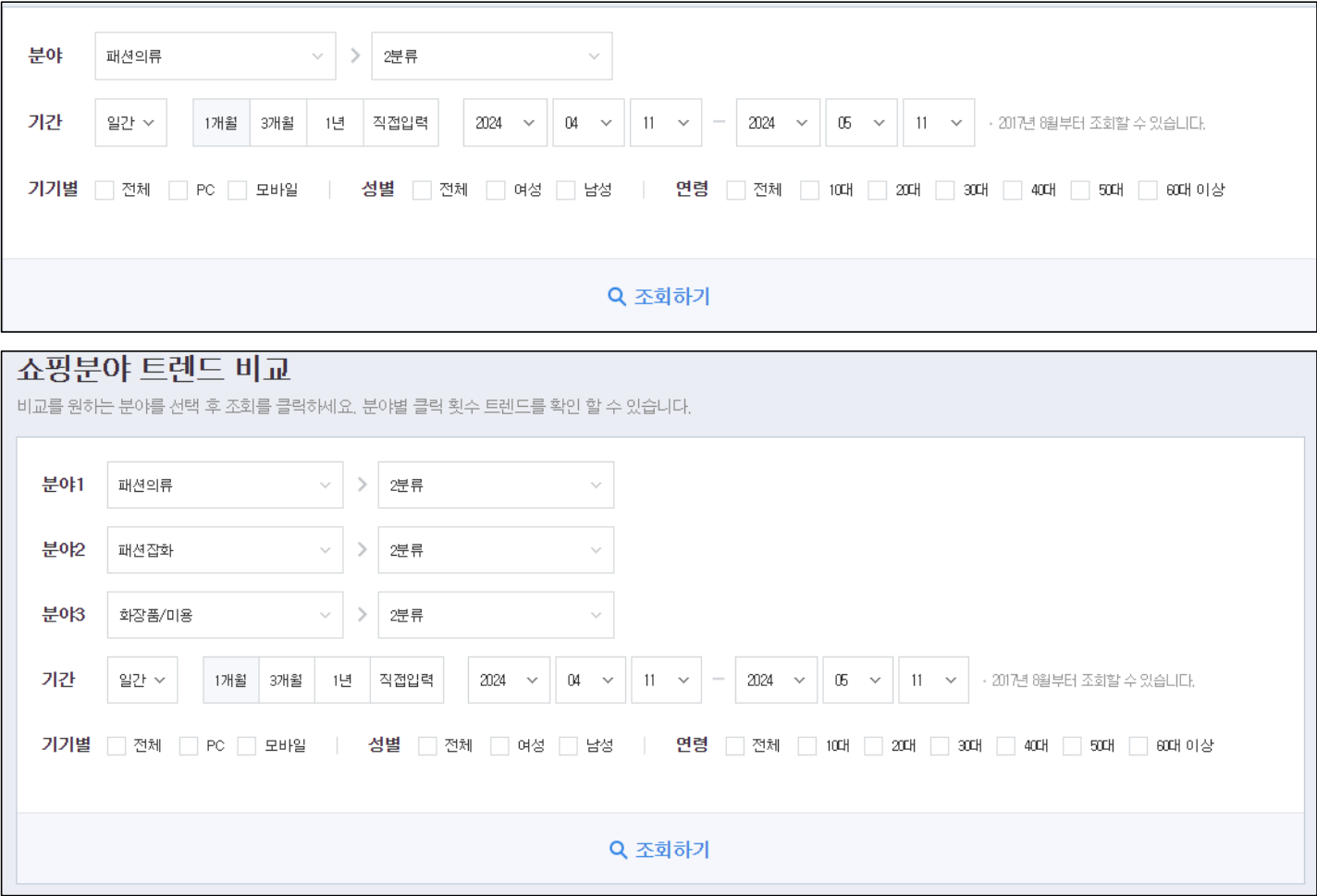




사진 20. 네이버 데이터랩 분야 통계 서비스



4. 대표 서비스 사례 Top5

1. Big Data 조회 서비스
2. 서울시 올빼미 버스
3. ZARA Fast Fashion
4. 쿠팡
5. 빅데이터 기반 사기 예방 서비스

구분	KT 유동인구 데이터(1차 검증)	택시승·하차 데이터(2차 검증)
데이터	 KT고객의 통화 기지국 위치 및 청구지 주소	 스마트카드사(KSCC)의 택시 iDTG* 정보
OD구분 (O:출발지 D:도착지)	<ul style="list-style-type: none"> • O : 통화 시 기지국 위치 • D : 거주지 주소 	<ul style="list-style-type: none"> • O : 승차 위치정보 • D : 하차 위치정보
기간	<ul style="list-style-type: none"> • 1개월치(3/1 ~ 3/31) • 약 1억 건/1일, 약 30억 건 	<ul style="list-style-type: none"> • 7일치(3/18 ~ 3/24) • 1일 약 70~80만 건
요일 기준	• 주중 · 토요일 · 일요일 : 00시 ~ 05시	

* 통합형 디지털운행기록계(iDTG, Integrated Digital Tachograph) : 기존 택시미터기에 비해 GPS 위치정보, 주행거리, 요금 관련 사항 등의 정보 수집이 가능하다

사진 22. 활용 데이터

[유동인구 데이터 처리 흐름도]



[택시 iDTG 데이터 처리 흐름도]

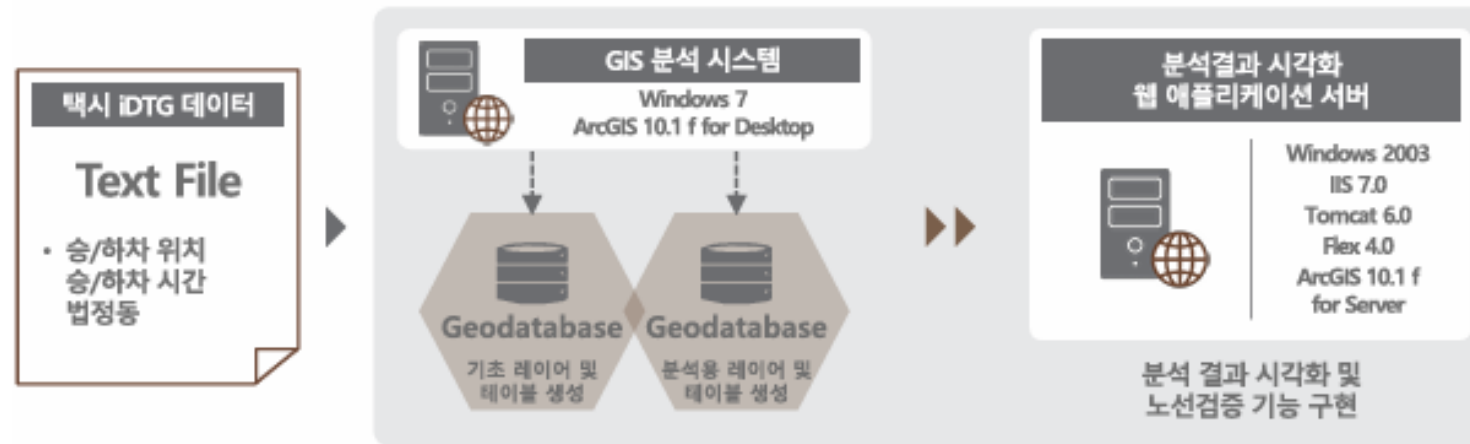


사진 23. 활용 데이터 처리 흐름도

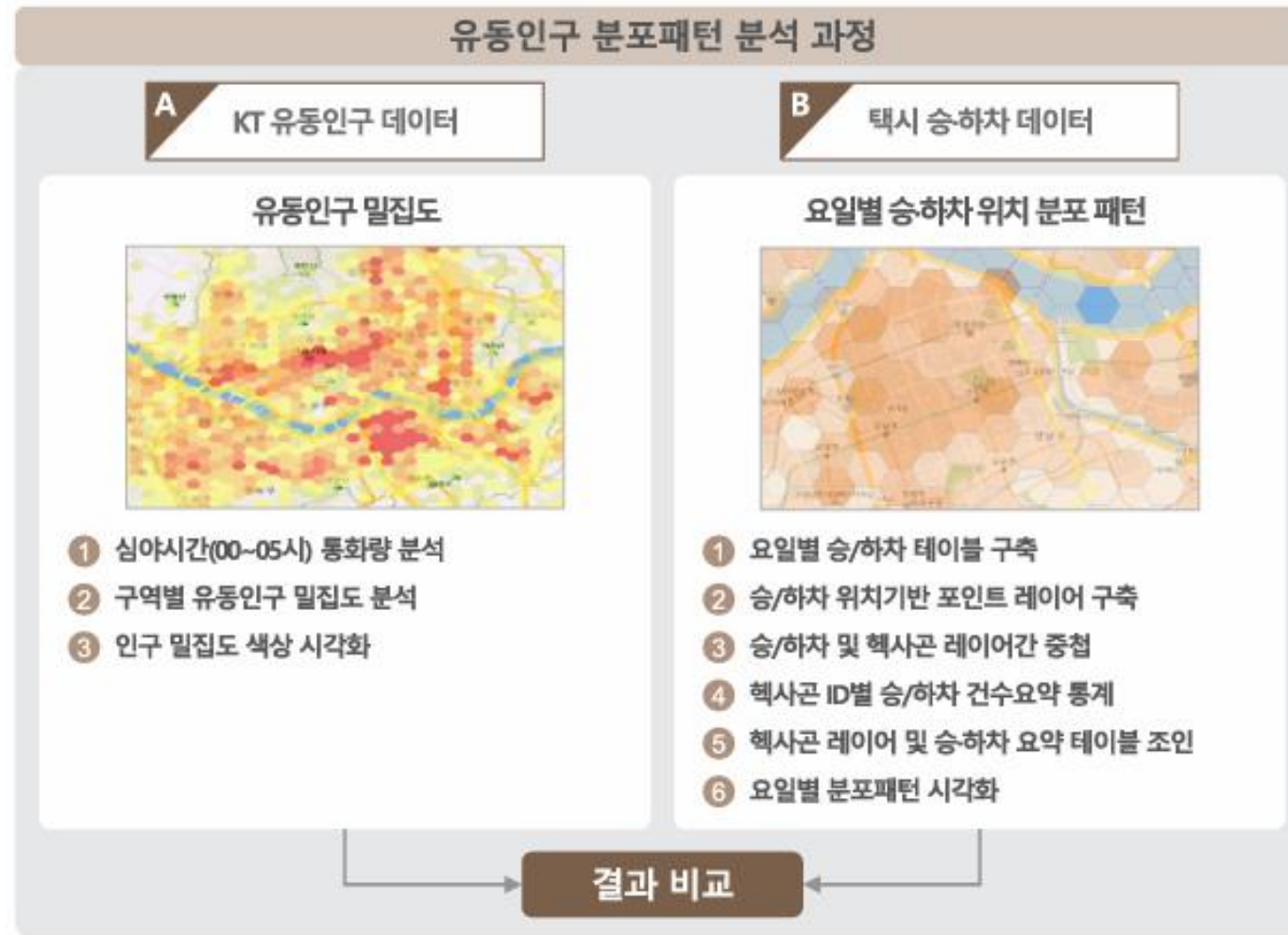


사진 24. 활용 데이터 기반 유동인구 분포패턴 분석

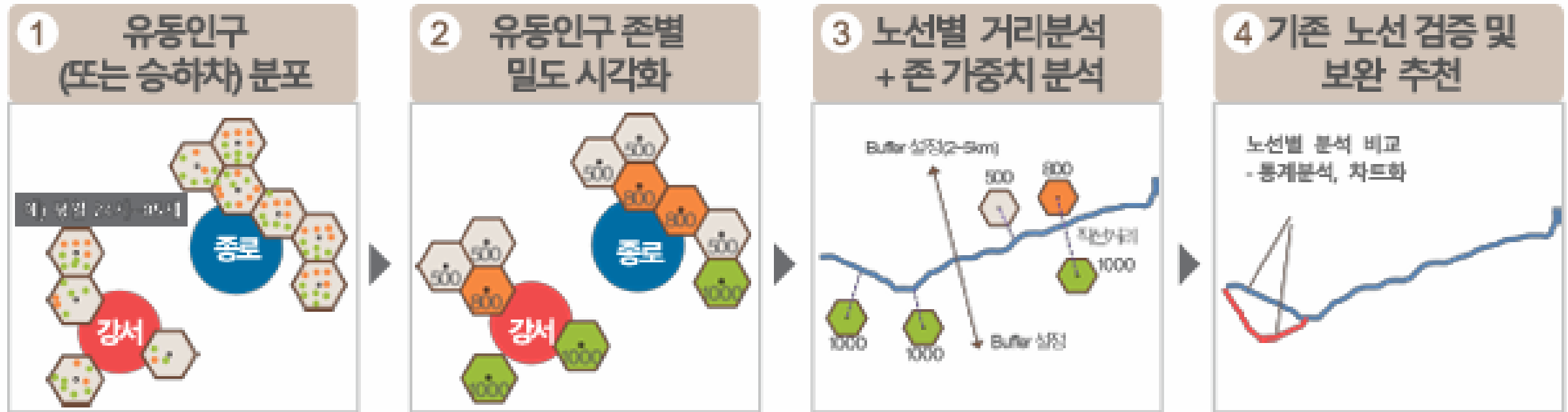


사진 25. 노선 수요 분석 과정

보완 필요 지역



심야버스 노선 수립 지원 시스템 화면 예시



사진 26. 심야 버스 노선 수립 지원 시스템 적용

심야버스 확대노선 최종(안)



사진 27. 심야 버스 확대 노선 최종 결정

4. 대표 서비스 사례 Top5

1. Big Data 조회 서비스
2. 서울시 올빼미 버스
- 3. ZARA Fast Fashion**
4. 쿠팡
5. 빅데이터 기반 사기 예방 서비스



사진 28. 벨기에 브뤼셀 ZARA 매장

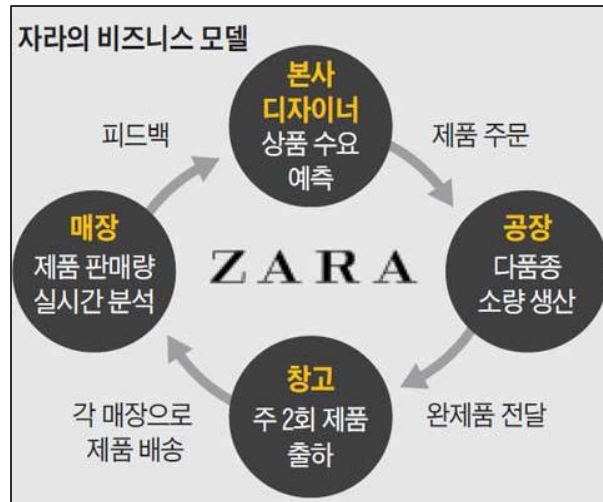


사진 29. ZARA의 비즈니스 모델

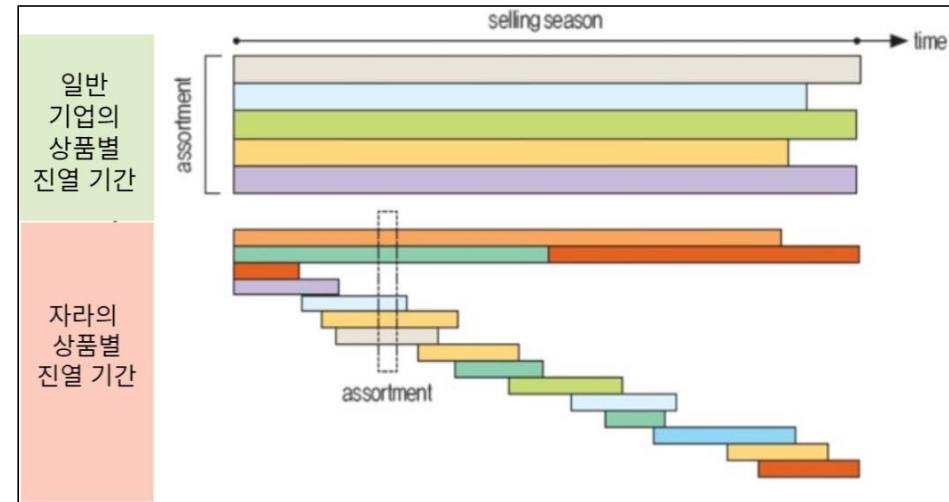


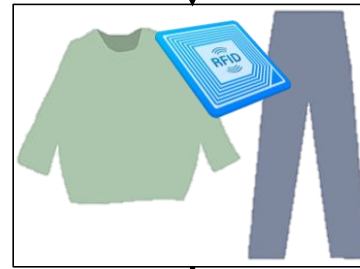
사진 30. ZARA와 일반 패션 기업 시즌 상품 진열 기간 비교,
<https://betterberry.co.kr/entry/%EC%9E%90%EB%9D%BCZARA%EC%9D%98-%EB%8D%B0%EC%9D%B4%ED%84%B0-%EA%B2%BD%EC%98%81-%EB%94%94%EC%A7%80%ED%84%B8-%ED%98%81%EC%8B%A0%EC%9D%98-%EB%B2%A0%EC%8A%A4%ED%8A%B8-%EC%82%AC%EB%A1%80>



사진 31. RFID 태그

무선 주파수 식별 기술, RFID(Radio-Frequency Identification)

ZARA는 RFID 활용하여 매장 운영에 필요한 데이터 수집 및 분석



의류에 RFID 삽입하여 매장에서 발생하는 모든 데이터 수집



자라는 POS 단말기, 전자상거래 판매, 고객 조사, PDA 기기 및 RFID(rad io-frequency identification, 무선인식시스템) 태그에서 데이터를 캡처한다. 매장 직원은 버튼, 지퍼, 컬러, 컷 등 PDA에서 고객의 선호도를 포착하도록 교육을 받고 매일 밤 피드백을 입력하고 업로드하며 지역 매니저는 그들 지역의 선호와 취향을 분석한다.



RFID 태그는 재고를 효율적으로 관리할 수 있는 능력을 줌과 동시에 어떤 옷이 피팅룸으로 들어가고, 나오며 실제 판매로 이어지는지에 대한 데이터를 수집할 수 있게 해 준다.



자라는 인스타그램, 설문 조사, 온라인 소셜 미디어에서 데이터를 수집하여 각각의 마켓 및 모든 마켓에서 고객의 패션 센서빌리티에 대한 중앙 분석 시스템을 구축한다.



모든 데이터는 스페인 Arteixo 데이터 센터에 수집된다. 분석가들은 하루 24시간 일하고, 새로운 디자인을 개발하기 위해 데이터를 처리하며, 일주일에 두 번씩 신제품을 매장으로 내보낸다.

초기 물량은 적게 가져가고, 반응에 따라 신속하게 추가 생산을 하기 때문에 소진되지 않는 재고는 거의 없으며, 할인율 또한 경쟁사에 비해 현저히 낮다. 자라는 데이터를 통해 고객을 속속들이 파악하고 있으며, 만약 분석 및 예측에 오류가 있다 할지라도 빠르게 대응할 수 있는 시스템을 갖추었다.

사진 32. ZARA 빅데이터 접근 방식, FASHION POST

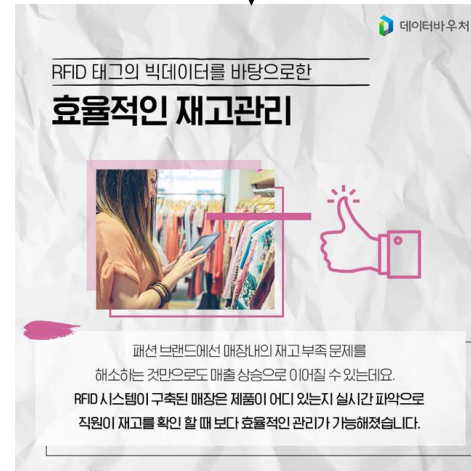


사진 33. RFID 도입을 통한 얻은 이점, 데이터바우처

4. 대표 서비스 사례 Top5

1. Big Data 조회 서비스
2. 서울시 올빼미 버스
3. ZARA Fast Fashion
- 4. 쿠팡**
5. 빅데이터 기반 사기 예방 서비스

로켓배달 서비스

사진 34. coupang newsroom, <https://news.coupang.com/archives/29293/>

쿠팡 검색 시스템

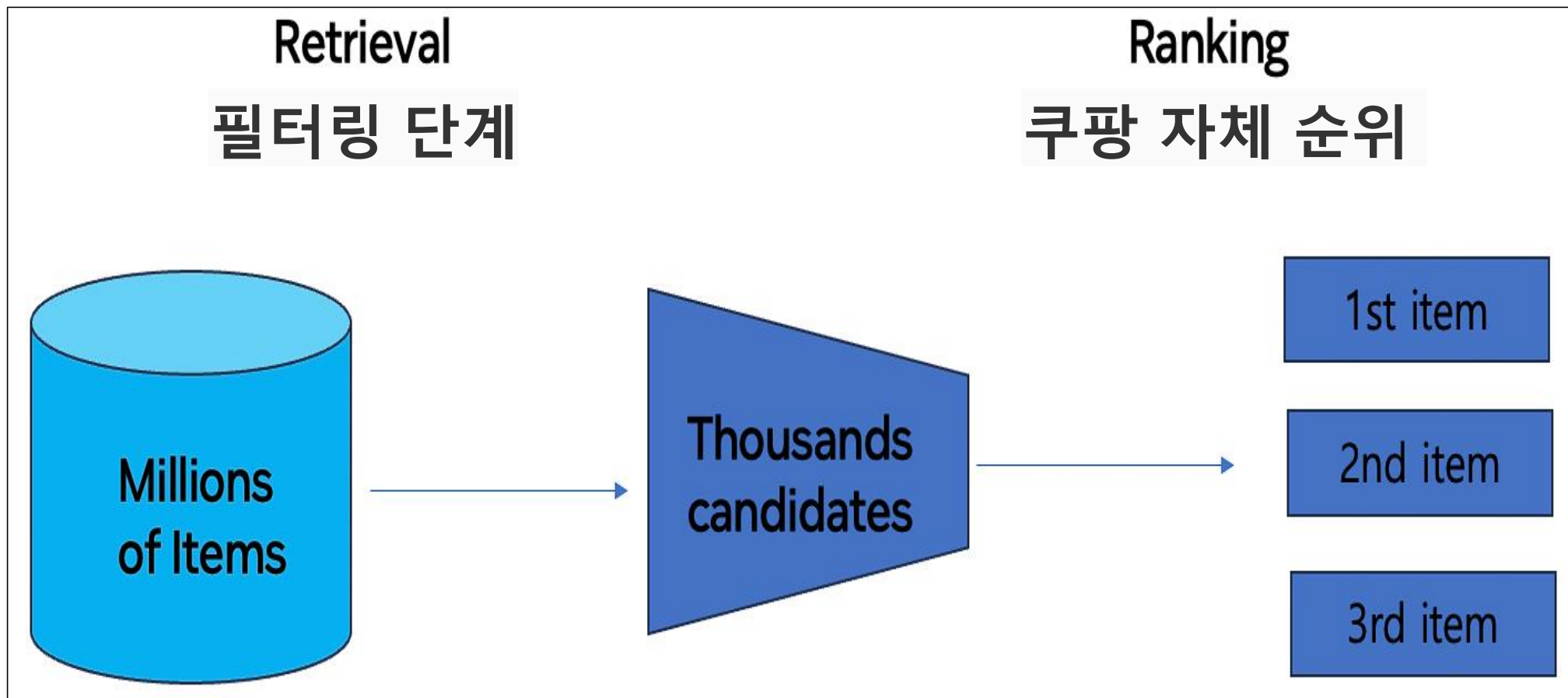


사진 35. coupang newsroom, <https://news.coupang.com/archives/29293/>

4. 대표 서비스 사례 Top5

1. Big Data 조회 서비스
2. 서울시 올빼미 버스
3. ZARA Fast Fashion
4. 쿠팡
5. Big Data 기반 사기 예방 서비스



PayPal의 공동 창업자 피터 필

사기 예방 데이터베이스 분석 소프트웨어를 바탕으로
빅데이터 분석 서비스 회사 창업

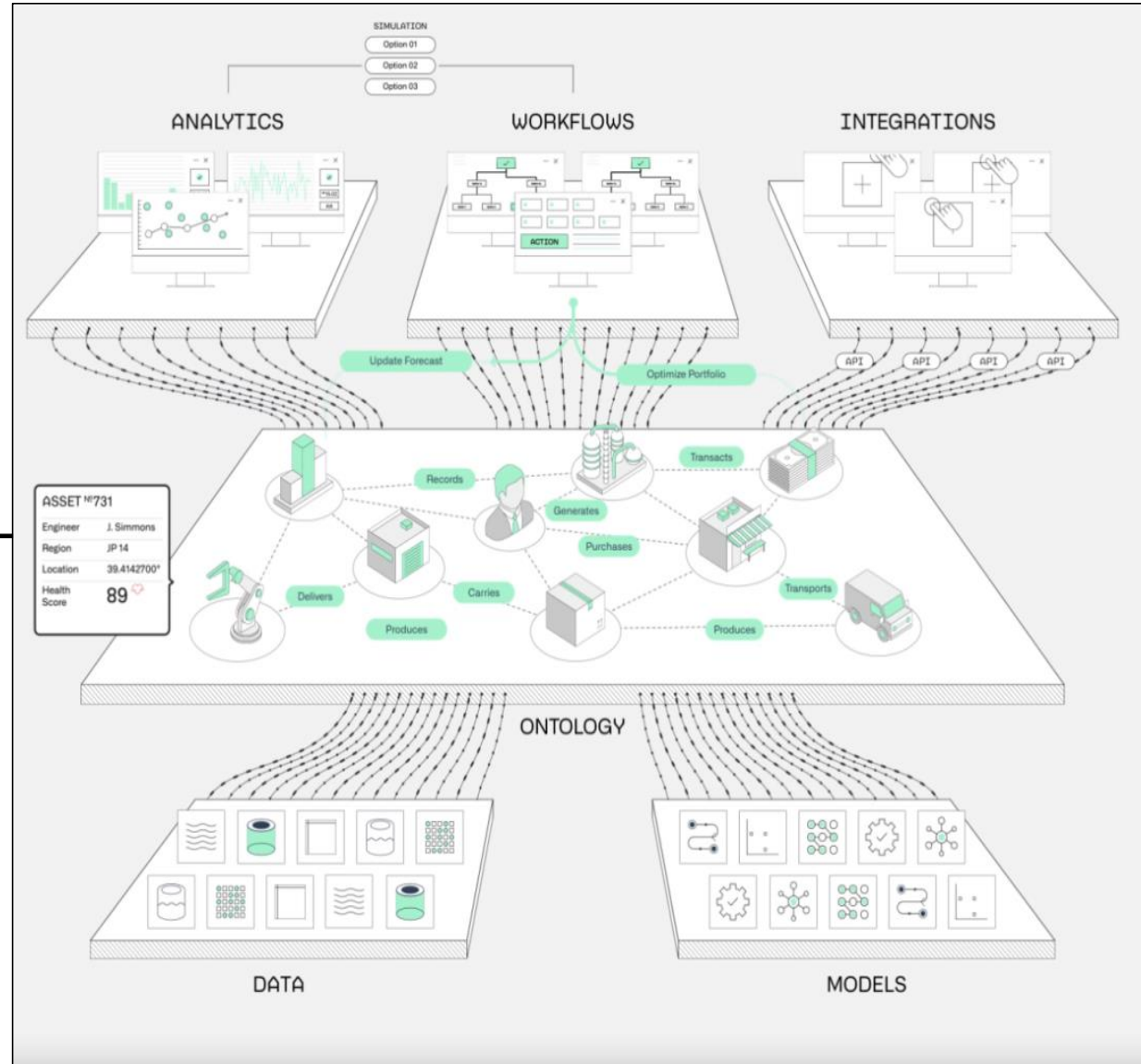


사진 36. PayPal의 빅데이터 기반 사기 예방 서비스,
<https://www.palantir.com/platforms/gotham/>

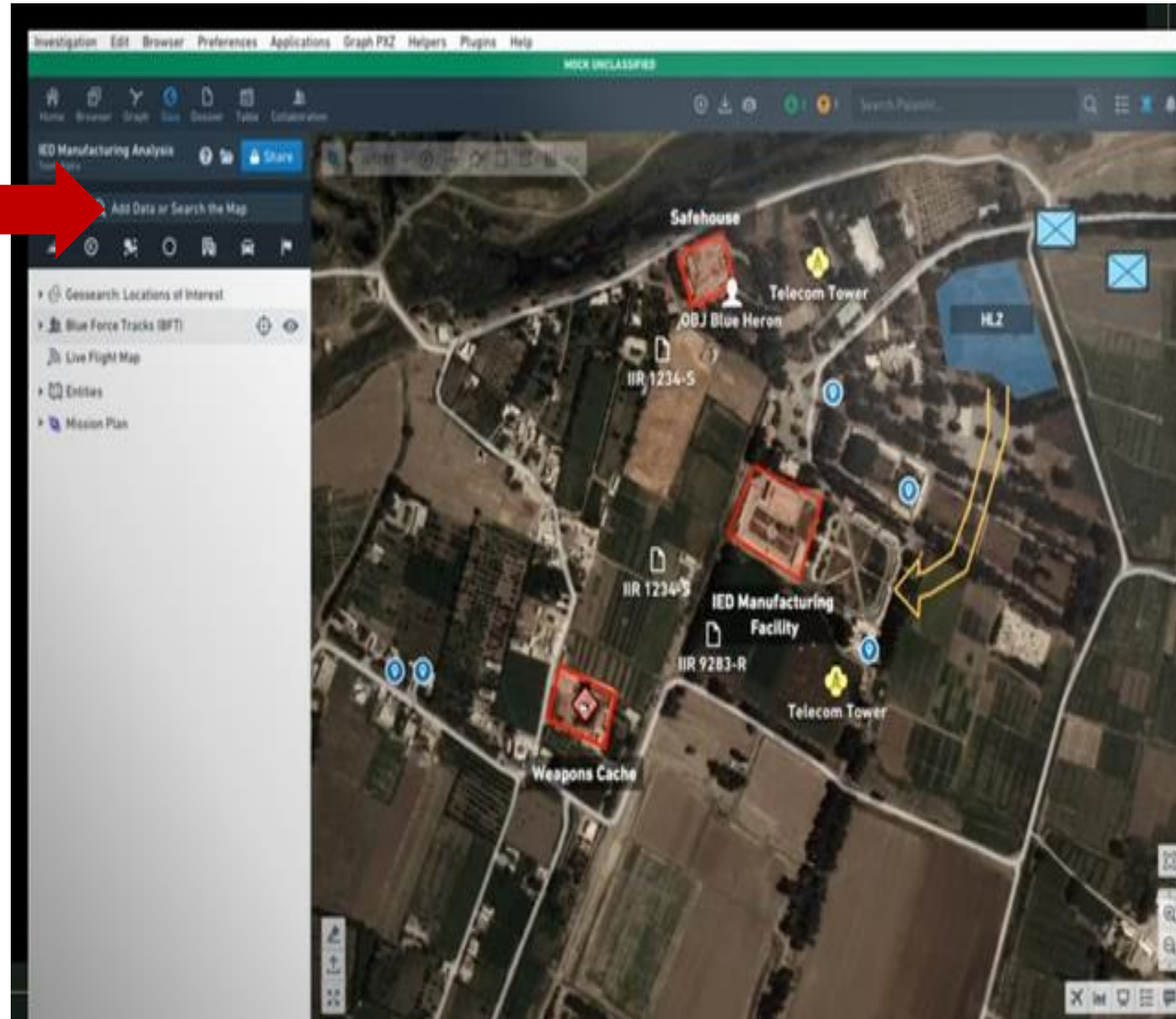


사진 37. Palantir 고담 서비스

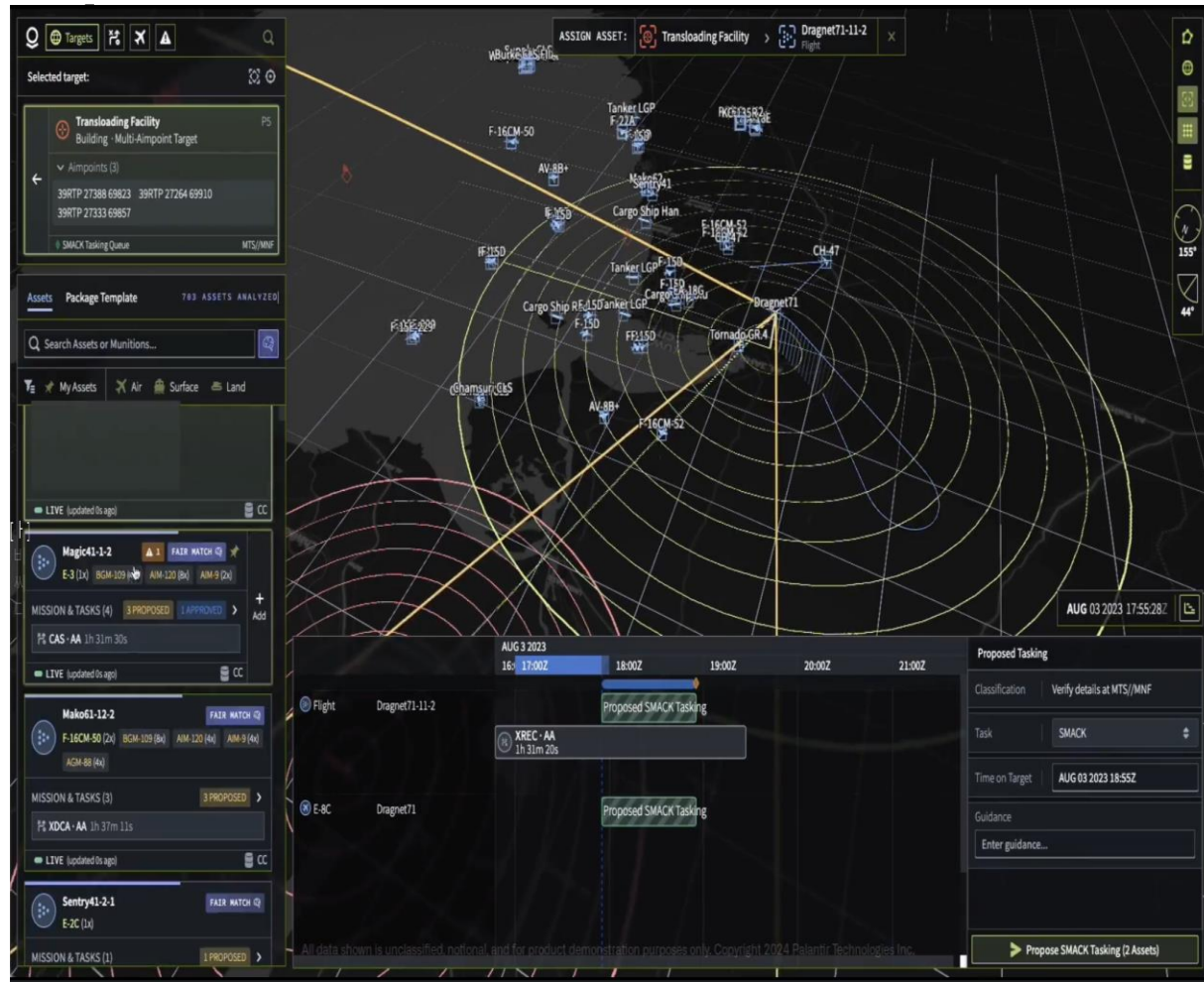
2011년 넵툰 스피어 작전



사진 39. 연합뉴스

사진 38. 실제 Palantir 고담 서비스 속 모습, <https://www.palantir.com/platforms/gotham/> 36

러시아 vs 우크라이나 전



이제 전쟁도 AI·빅데이터 시대



Big Data 개요

- 빅데이터 정의
<https://www.sap.com/korea/products/technology-platform/what-is-big-data.html>
https://sjeec.or.kr/board/2023/06/15/1686811719_2053930f9494787fc80e.pdf
<https://modulabs.co.kr/blog/big-data/>

Big Data 특징

- 빅데이터 3V, 4V, 5V, 7V
<https://wikidocs.net/22652>
<https://velog.io/@garam/DE%EB%B9%85%EB%8D%B0%EC%9D%B4%ED%84%B0%EC%9D%98-%ED%8A%B9%EC%A7%953V-5V-7V>
- 데이터 종류
<https://computer-science-student.tistory.com/471>
- 빅데이터 프로세스
<https://aws.amazon.com/ko/what-is/data-analytics/>
<https://wikidocs.net/65239>
<https://ikkison.tistory.com/69>

미래 발전 방향

- 빅데이터 보안
<https://www.gttkorea.com/news/articleView.html?idxno=9735>
- 빅데이터 글로벌 동향
<https://www.gttkorea.com/news/articleView.html?idxno=6415>
- 빅데이터 비즈니스 트렌드
<https://yozm.wishket.com/magazine/detail/2276/>
<https://www.gttkorea.com/news/articleView.html?idxno=8424>
- 이외 자료 참조 출처
 - 한국데이터산업진흥원
<https://www.kdata.or.kr/kr/whitePaper/view.do>
 - 소프트웨어정책연구소
https://spri.kr/posts/view/19814?code=data_all&study_type=industry_trend

대표사례 TOP 5

- **Big Data 조회 서비스**
 - 카카오 데이터 트렌드와 네이버 데이터 랩 비교
<https://www.kmib.co.kr/article/view.asp?arcid=0014855973>
- **서울시 올빼미 버스**
 - 서울시 빅데이터 분석 사례집
- **ZARA Fast Fashion**
 - https://www.chosun.com/site/data/html_dir/2019/06/06/2019060600954.html
 - https://weeklybiz.chosun.com/site/data/html_dir/2015/08/28/2015082801936.html
 - <https://betterberry.co.kr/entry/%EC%9E%90%EB%9D%BCZARA%EC%9D%98-%EB%8D%B0%EC%9D%B4%ED%84%B0-%EA%B2%BD%EC%98%81-%EB%94%94%EC%A7%80%ED%84%B8-%ED%98%81%EC%8B%A0%EC%9D%98-%EB%B2%A0%EC%8A%A4%ED%8A%B8-%EC%82%AC%EB%A1%80>
 - 월드패션리포트
https://fpost.co.kr/board/bbs/board.php?wr_id=17&bo_table=fsp1
 - 데이터바우처
<https://blog.naver.com/datavoucher/222080175765>
- **쿠팡**
 - coupang newsroom
<https://news.coupang.com/>
- **Big Data 기반 사기 예방 서비스**
 - 디지털 콘텐츠 학회 논문지 Journal of Digital Contents Society Vol. 25, No. 2, pp. 383-393, Feb. 2024
 - Palantir 홈페이지
<https://www.palantir.com/platforms/gotham/>
 - Palantir 관련 기타 자료
<https://aimrich.co.kr/56/?idx=16642238&bmode=view>
<https://time.com/6691662/ai-ukraine-war-palantir/>
<https://medium.com/paypal-tech/how-paypal-uses-real-time-graph-database-and-graph-analysis-to-fight-fraud-96a2b918619a>
https://www.chosun.com/international/international_general/2024/02/17/72FIOBMNGNFLBALDMYZFPEANQU/