

Glue + Athena를 이용한 데이터 분석

저렴한 비용으로
데이터 전후처리와 데이터 분석

김선우

2025. 7. 15



Introduce

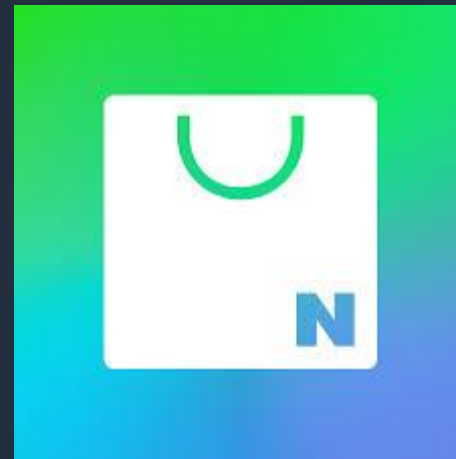
- Ex) Backend Developer
- 현재 Devops Engineer
- Game, Personal Mobility, Travel
- AWS Community Builder
- AWSKRUG 자격증 소모임 운영진
- AWS 8 Certified



문득 궁금해지는 순간

- 배달의 민족은 어떻게 내가 좋아하는 음식점을 알고 있을까?
- 키보드는 어떻게 내가 필요할 때 1분 안의 거리에 배치해둘까?
- 네이버는 어떻게 내가 찾는 상품의 최저가를 알고 있을까?

데이터로 비즈니스를 해결한 수많은 회사들



내 목소리 마저 마케팅에 활용되는 세상

"ACTIVE LISTENING" SOFTWARE

PARIS

FRANCE 24

TO
ATIONS
ND IOT

COMBINING
VOICE DATA WITH
ONLINE BEHAVIORAL DATA

HYPER-TARGETED
ADVERTISING

SCIENCE

**US firm Cox Media accused of
listening to private conversations**

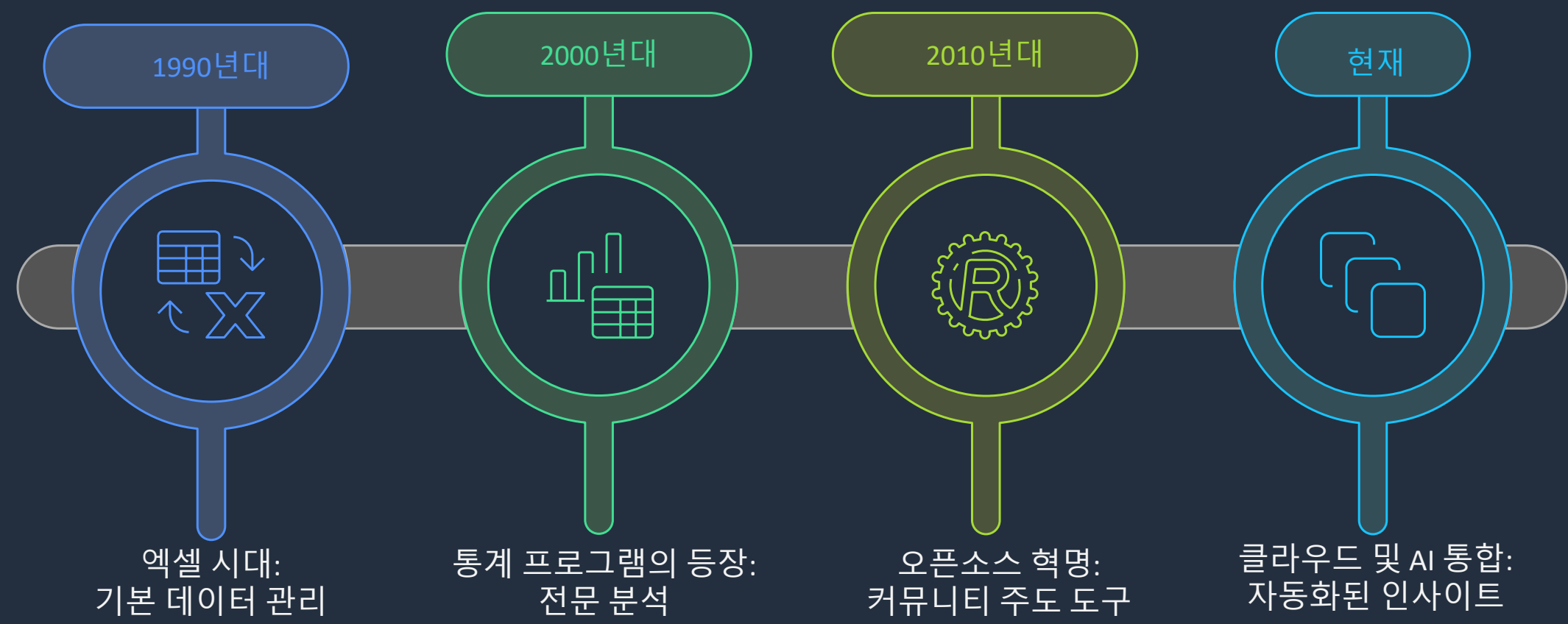
```
graph LR; A[TO ACTIONS AND IOT] --> B[COMBINING VOICE DATA WITH ONLINE BEHAVIORAL DATA]; B --> C[HYPER-TARGETED ADVERTISING];
```

내 목소리 마저 마케팅에 활용되는 세상

- 친구와 여행 얘기 → 항공권 광고 등장 ✈️
- 카페에서 화장품 얘기 → 뷰티 광고 폭격 💄
- 운동 얘기 → 헬스장/운동용품 광고 💪
- 반려동물 얘기 → 펫샵 광고 🐕
- 토익 공부 이야기 → 토익 학원 광고

위치, 검색기록, 친구 관계, 앱 사용 패턴 등 수백 가지 데이터를 AI가 조합한 결과!

데이터 통계의 흐름



지금은 No-Code/Low-Code 시대



요즘 많이 사용되는 도구들



Tableau



Amazon QuickSight

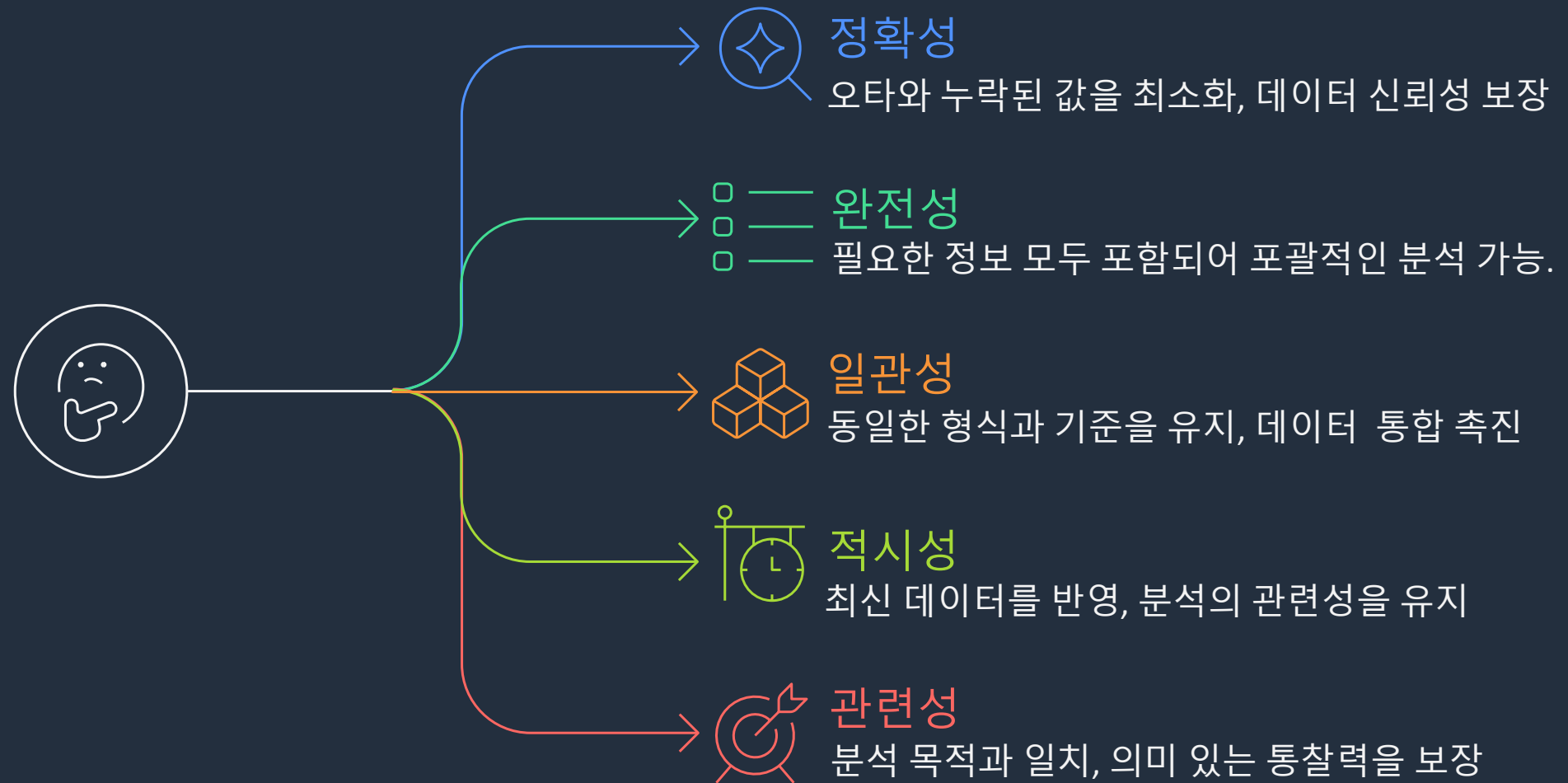


MS Power BI



Google Looker Studio

데이터셋의 중요성



데이터 구하기 좋은 곳

- Kaggle (World Largest Data Scientist Community)
- Hugging Face (AI/ML 모델과 Dataset Hub)
- Google Dataset Search (구글 데이터셋 검색엔진)
- AWS Open Data (아마존 제공 공개 데이터)
- 정부 공공데이터 포털 (정부 제공 데이터 포털)

이제 실습을 해볼까요?

<https://m.site.naver.com/1MFot>