

# Technology for Team Vodia

**i** 본 문서는 Team Vodia에서 사용하는 기술 스택을 정의한 문서입니다.

먼저, 사용한 기술 스택을 정리하면 아래와 같습니다.

Sector	Technology Stack
Design	Figma
Android	Kotlin, Jetpack Compose(Library)
Backend	Python, FastAPI(Frame work), PostgreSQL(DB) Restful API(API Format)
AI	Python, Transformer(Library), Pytorch(Library)
Server	Linux(Ubuntu)

본 어플리케이션은 Kotlin의 Jetpack Compose 라이브러리를 활용하여 구현되었습니다.  
백엔드는 Fast API로 구현되었으며, DB는 PostgreSQL를 사용합니다. 또한 모바일과 서버의 통신은 Https 프로토콜을 따르며, Restful API의 형태로 API를 주고 받습니다.  
상품 이미지를 텍스트로 변환, 댓글 정보 요약을 위해 2가지 종류의 AI 모델(VLM, LLM)을 사용하며, AI와 백엔드 API는 리눅스 서버환경에서 가동됩니다.

## #Main Technology

Team Vodia에서 사용한 주요기술을 정리하면 아래와 같습니다.

- 1. 화면 낭독기 친화적 데이터 포메팅
- 2. AI, VLM(Vision Language Model)를 활용한 Image to Text
- 3. AI, LLM(Large Language Model)를 활용한 상품리뷰 요약
- 4. 화면 줌 인, 줌 아웃
- 5. 단축키를 통한 음성검색 활성화
- 6. 음성 검색 기능

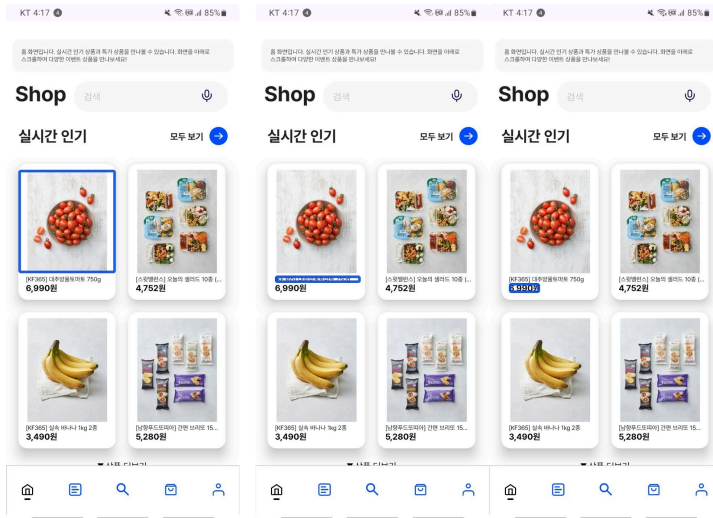
### 1. 화면 낭독기 친화적 데이터 포메팅

#### 1.1 Modifier.semantics

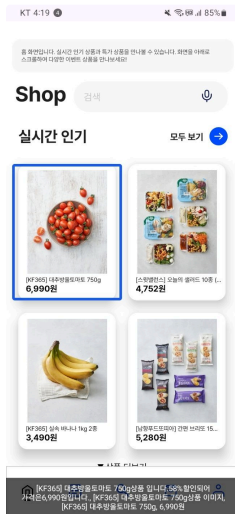
화면 낭독기는 기본적으로 한 번에 하나의 아이템을 읽고 음성으로 변환합니다.  
ex) 상품 제목 → 상품 가격 → 상품 이미지 → 상품 제목 → 상품 가격 → 상품 이미지 → ...  
이렇게 모든 정보를 하나씩 읽는 것 보다는, 하나의 상품을 하나의 그룹으로 처리하여 한 번에 해당 상품의 정보를 읽는 것이 효과적일 것입니다.  
ex) [상품 제목 + 상품 가격 + 상품 이미지] → [상품 제목 + 상품 가격 + 상품 이미지] → ...  
이를 위해 Team Vodia는 **Jetpack Compose**의 **Modifier.semantics**를 활용하였고, 데이터를 한 번에 읽게하여 보다 편리한 쇼핑이 되도록 기능을 구현하였습니다.

```
Box(
  modifier = Modifier
    // ScreenReader를 위해 텍스트를 한 묶음으로 처리
    .semantics(mergeDescendants = true){
      text = AnnotatedString(text: name + "상품 입니다." + discount + "할인되어 가격은" + price + "입니다.")
    }
    .width(180.dp)
    .padding(
      start = 10.dp,
      end = 10.dp,
    )
)
}
```

데이터를 묶어서 스크린 리더기에 제공



before semantics



After semantics

## 1.2 contentDescription [🔗](#)

화면 낭독기가 접근 하지 못하는 데이터 형식(상품 이미지, 버튼..)의 경우 대체 텍스트(contentDescription)을 제공하여 해당 정보에 접근 할 수 있도록 구현 하였습니다.

위 기능은 Html의 'alt' 속성을 설정하는 것과 유사합니다.

```

Icon(
    painter = painterResource(R.drawable.arrow_button),
    contentDescription = "모두 보기 버튼",
    tint = Color.Unspecified
)

```

대체 텍스트 설정

## 2. VLM을 활용한 상품정보 텍스트화(Image to Text) [🔗](#)



VARCO Model

**VARCO-VISION-14B** 모델(VLM)을 활용하여 상품이미지를 텍스트로 변환하여 제공합니다.

( [🤖 NCSOFT/VARCO-VISION-14B · Hugging Face](#) )

해당 모델을 가동하기 위해 별도의 리눅스 서버를 구축하였고, 모델을 실험한 결과 상품 이미지를 요약하는데 약 11초 정도가 소요됩니다.

아래는 모델이 생성한 요약문 결과입니다.

상품 : 고려은단 비타민C ( [🛒 고려은단 비타민C 1000 + 쇼핑백, 180정, 1개](#) )

위 링크에 존재하는 상품 이미지 정보를 텍스트로 변환한 결과입니다.

```

Result of 8////////////////////////////////////
고려은단 비타민 C 1000mg 제품입니다. 이 제품은 1일 2회, 1회 1정씩 복용하는 것이 좋습니다. 주요 성분으로는 비타민 C가 1000mg 들어 있으며, 비타민 C의 흡수율을 높이기 위해 비타민 E와 마그네슘도 포함되어 있습니다. 또한, 천연 감귤 추출물과 비타민 C를 함께 함유하여 면역력 강화에 도움을 줍니다. 포장에는 60정이 들어 있으며, 100% 식품의약품안전처 기준을 충족합니다.

Result of Time : 11.323305606842041
////////////////////////////////////

```

모델 결과(상품 정보 텍스트화)

## 3. LLM을 활용한 리뷰 정보 요약 [🔗](#)

**Llama-VARCO-8B-Instruct** 모델(LLM)을 활용하여 상품의 주요 리뷰 3개를 요약합니다.

( [🤖 NCSOFT/Llama-VARCO-8B-Instruct · Hugging Face](#) )

해당 모델을 가동하기 위해 별도의 리눅스 서버를 구축하였습니다.

아래는 모델의 요약문을 저희 어플리케이션에 적용한 사진입니다.

#### 리뷰 정보

방울토마토를 활용한 다양한 요리 레시피에 대한 리뷰입니다. 피자과 타코에서 방울토마토의 신선함이 돋보이며, 에어프라이어 치킨 가라아게와 함께 즐기는 방법도 소개되었습니다. 새송이 쉬림프 파스타는 버터와 마늘, 굴소스로 감칠맛을 낸다는데, 대추방울토마토로 감바스를 만든 후에는 바게트 곁들여 먹는 방법도 언급되어 있습니다. 방울토마토의 감칠맛과 신선함이 각 요리에 잘 어울린다는 평가가 지적되었고, 집에서 만든 음식의 건강함과 만족도도 높다는 점을 강조하고 있습니다.

6,990 원

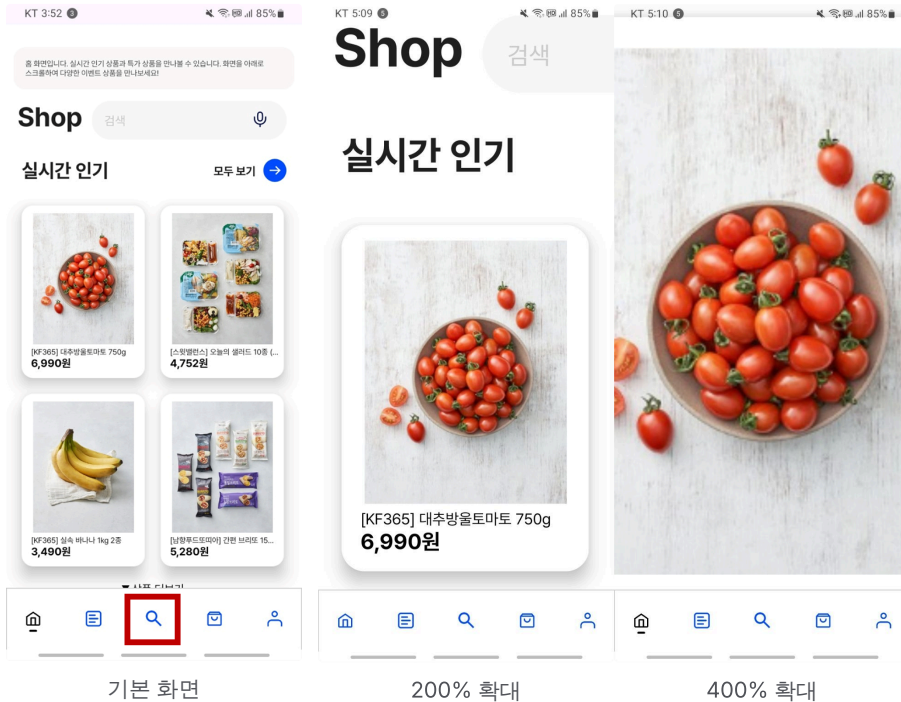
장바구니

결제하기

모델 결과(리뷰 정보 요약)

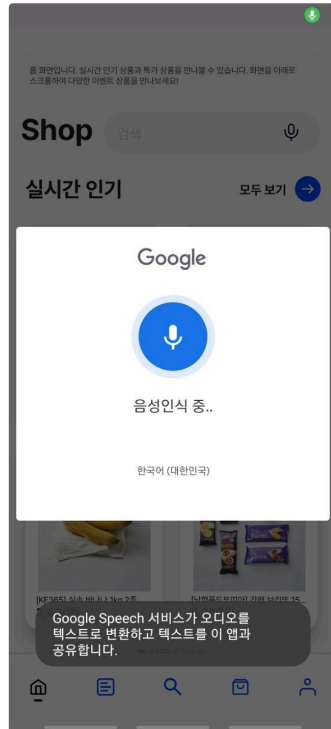
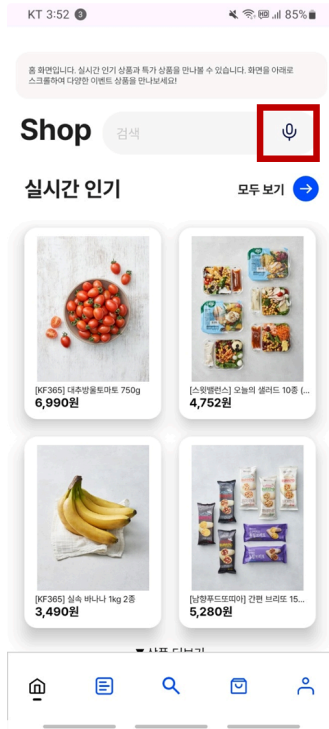
## 4. 화면 줌 인, 줌 아웃 기능

jetpack compose의 NavHost Composable를 활용하여 화면 줌 인, 줌 아웃이 가능하게 기능을 구현하였습니다. 최대 400% 까지 화면이 확대가 되며, 하단 네비게이션 바의 돋보기 버튼을 클릭하거나, 화면을 연속으로 두 번 터치하면 화면이 확대됩니다.



## 5. 음성 검색 기능

음성 검색 버튼을 누르면 음성으로 원하는 상품을 검색할 수 있는 서비스를 제공합니다. 해당 기능은 구글 음성 검색 API를 활용하였습니다.



음성 검색 화면

## 6. 단축키를 통한 음성검색 활성화 [🔗](#)

시각 장애인 분들이 쇼핑앱을 이용하실 때 검색 입력란을 못 찾아 불편함을 겪는 경우가 많습니다.

Vodia는 단축키를 통해 음성검색이 활성화 되는 기능을 제공합니다.