

문화와 기술1 - 메타버스와 메타휴먼 -



발표자: 원종서 Ph.D.

주요경력

- 성균관대학교 글로벌융합학부 컬처애펀테크놀로지융합전공 겸임교수
- KT 융합기술원 AI2XL연구소 AI Metaverse CX기획 팀장/ Ph.D.
- 한국콘텐츠진흥원 평가위원, ICT 대연합 전문멘토 겸 자문위원, KAIT(한국정보통신진흥협회) 심사/자문위원
- 연구분야: AI/메타버스 등 신기술 융합 서비스기획/ 사업화 전략, 디지털플랫폼 생태계, 플랫폼 상호작용
- 연구실적: 고객상담 컨택스트에서 AI콜센터 AI 상담사의 페르소나를 추천해 통화 완결율을 높여주는 방법 및 시스템 ('13. 3/ 특허출원)
3rd Party 기업의 AI플랫폼 참여 방법과 효과, CPND 가치사슬에서 스마트홈 콘텐츠의 선택적 요인 등 논문 다수
- KIA K9/K7 디지털마케팅 담당 (국내최초 공간인식형 AR 매뉴얼 기획/총괄: K9)
- 삼성 멀티캠퍼스 신사업 개발 & 사내벤처 CEO
- SK 네트워크서비스 전략기획 파트장
- i-hub 지식서비스 스타트업 CEO (04~07)

개인 소개

원종서 팀장은 KT메타버스 '지니버스' 서비스기획 및 IMC 마케팅 업무를 맡아 대고객 서비스 및 제휴/협력 업무를 수행하고 있다. 고시공부를 하며 뒤늦게 사회생활을 시작하였는데 이를 사업화 하여 고시생들을 위한 취업컨설팅 및 시장조사 스타트업을 창업해 2000년 초중반까지 활동하였다. 이후 SK, KIA, 삼성, KT 등 주요 대기업을 거치며 기술사업화, 마케팅, 기술영업, 기술연구 등을 경험하였다. 특히 KIA 재직시절 국내최초로 K9에 공간인식형 AR 차량 매뉴얼 개발/기획을 총괄 하였다. 평소 실무 현상을 학문연구와 연계하여 새로운 가치를 창출하는 것을 중요시 하며 관련 특허 및 논문, 기고 등 연구 활동도 함께 하고 있다

Agenda

What is Meta-Human?

Related theory & Theoretical approaches

Meta-Human Tech Trend

Future Landscape

plave

ASTERUM

COMING SOON
02.26 6PM(KST) Release



What is Meta-Human?



Digital
Human

Virtual
Human

Meta
Human

Meta-Human Use Cases

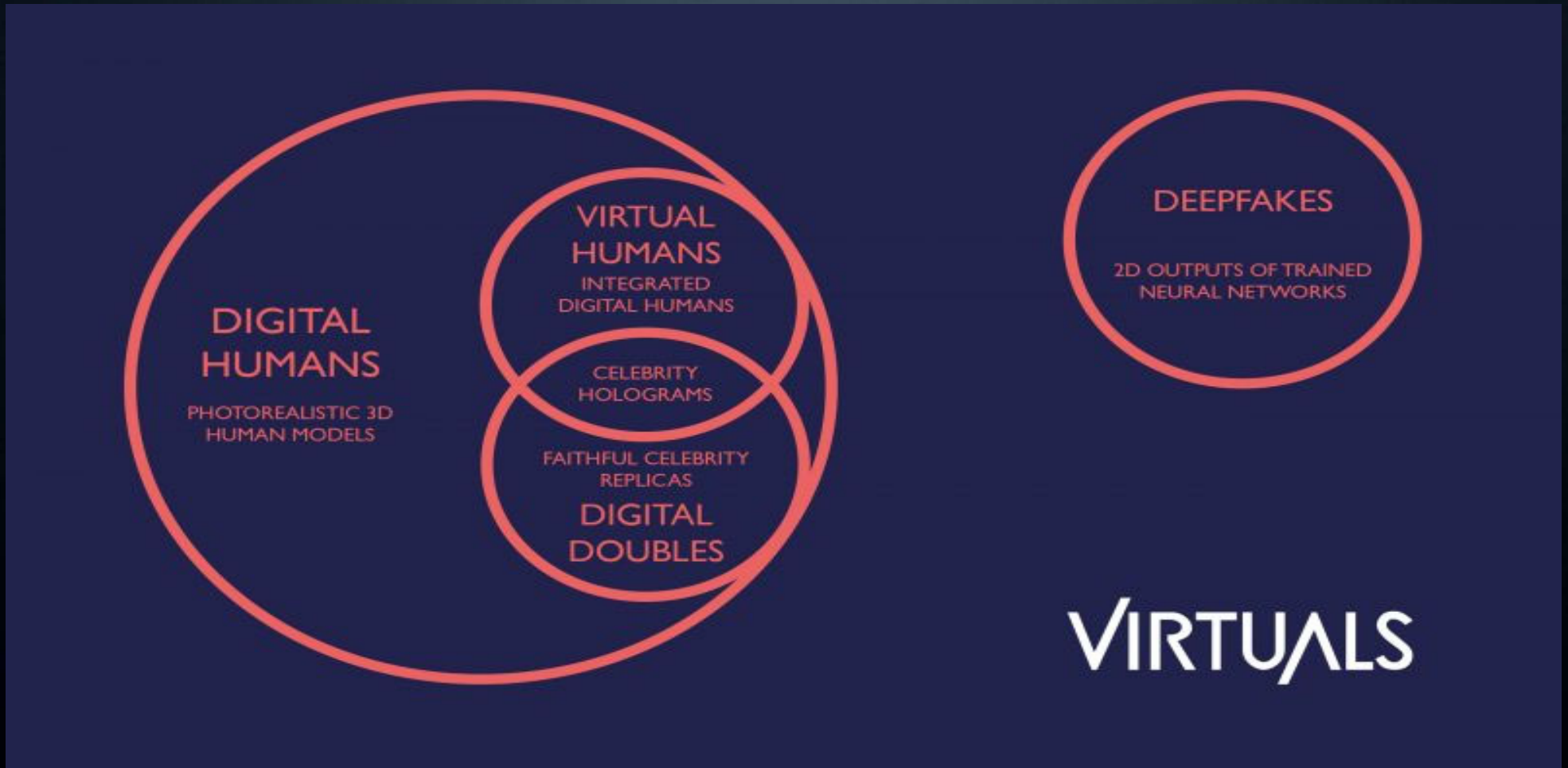
AD-Tech Innovation

- ROZY: Shinhan Life's first Virtual Influencer, over 10 million views in 20 days (2020)
- SUA+ SKT+ OnMind (Start-Up) AD Collaboration (2021)

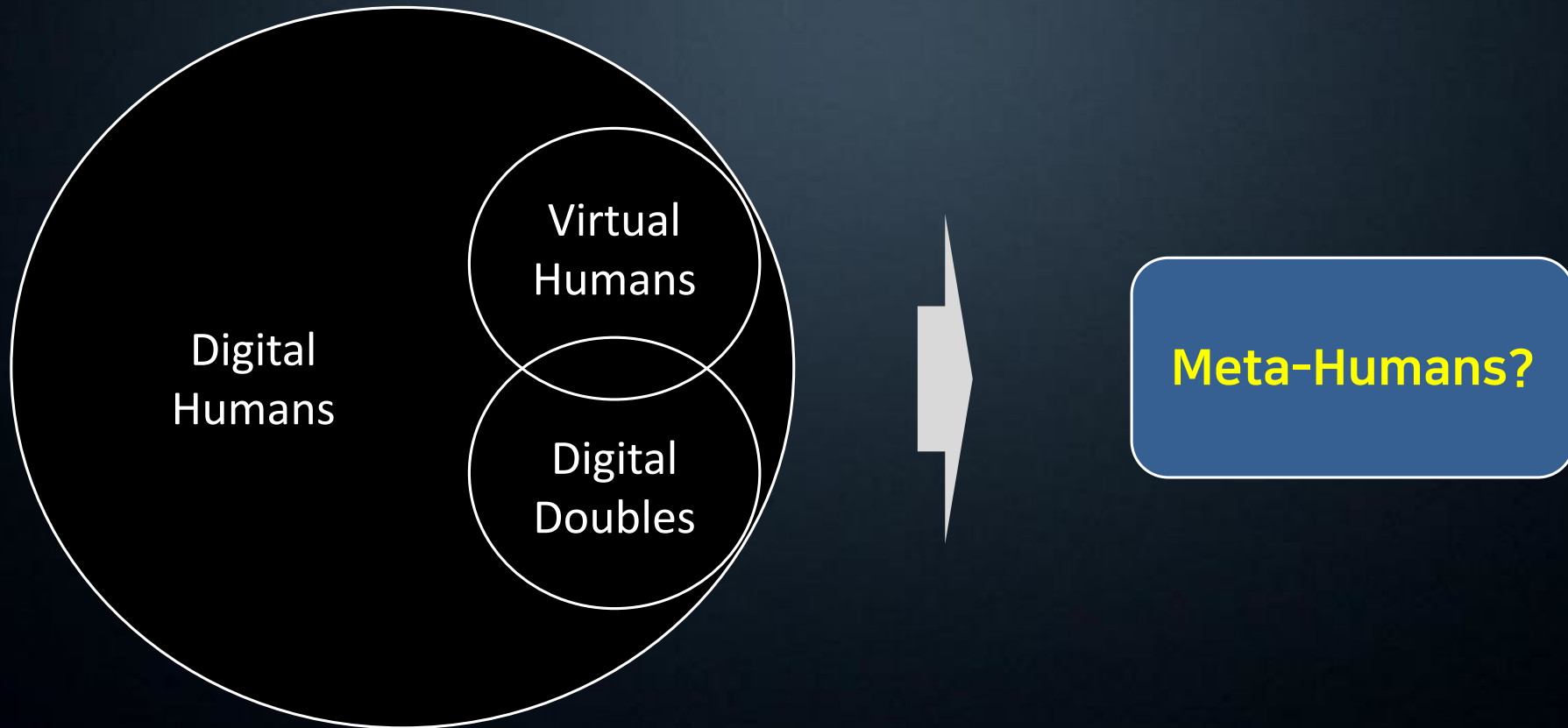


Source: <https://www.xportsnews.com/article/1634202>

What is Meta-Human?



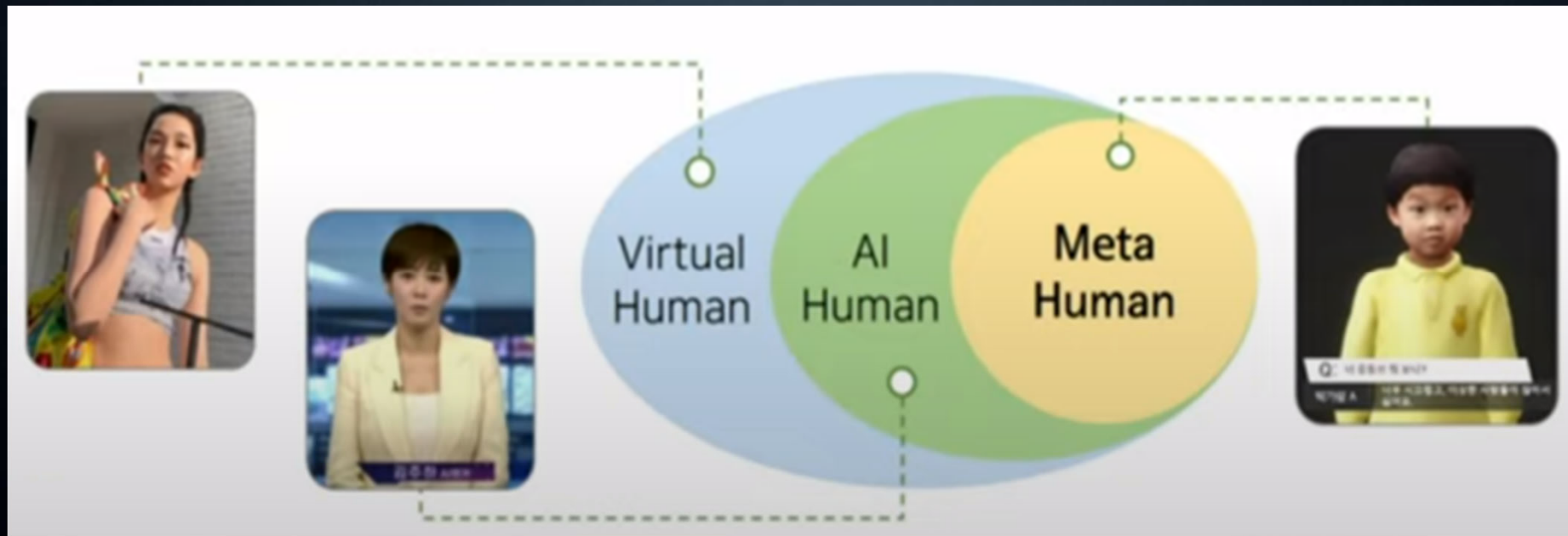
What is Meta-Human?



What is Meta-Human?

Meta-Human

A combination of Meta and Human, Digital humans operating in multiple metaverse, including virtual/augmented worlds, mirror worlds, etc.



Source: Metaverse x AI = Metahuman, etnews korea Vision 2021

Meta-Human Use Cases

Role & Responsibility

Advertising, PR, Fashion Model
Vtuber/ Singer/ Influencer/ Celeb
Reception/ Presenter/ Moderator...

Application & Industry

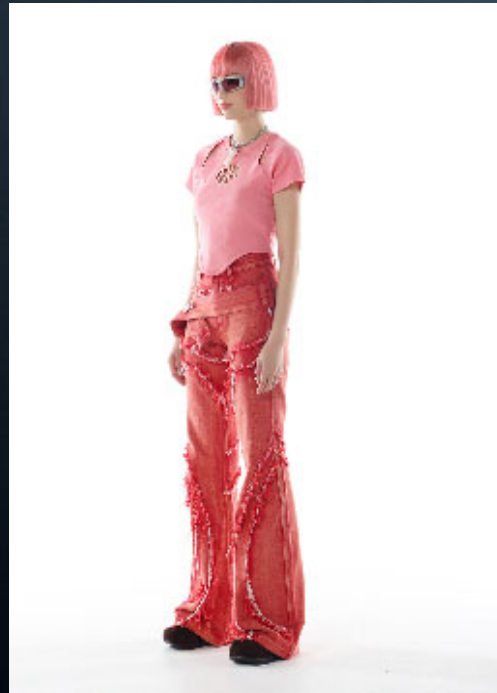
Education
Counseling
Retail
Sales
Customer Relation
Finance

Meta-Human Use Cases

Lil Miquela (2016)



Influencer IMMA(2019)



Model Shudu (2017)



Vtuber Kizuna AI(2016)





Big Tech Perspectives

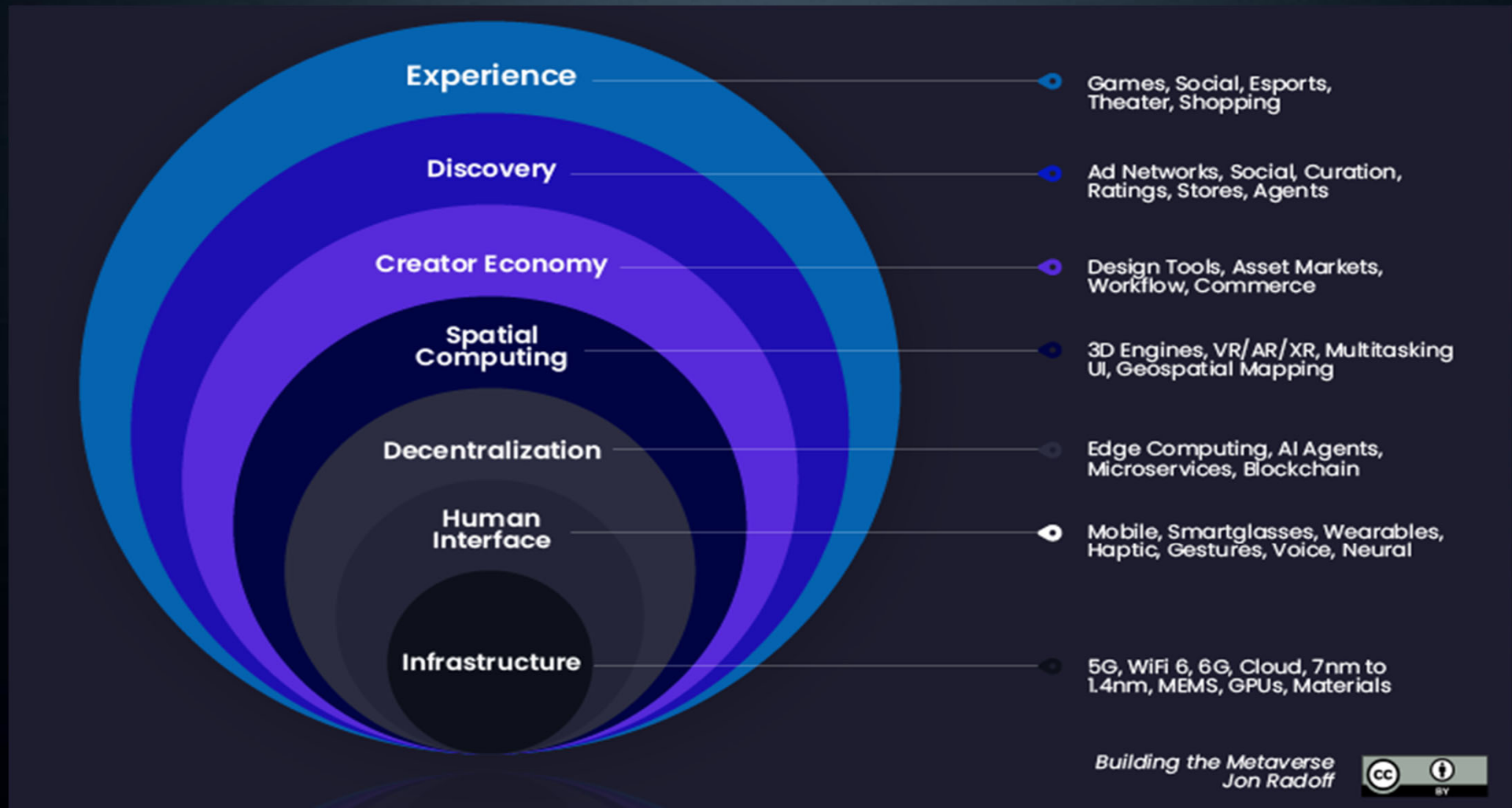
- Samsung Electronics NEON Project ('20.1)
 - Future technology commercialization venture 'STARLabs' AI project
 - Pranav Mistry (MIT Media lab), head of Galaxy Gear development since 2013, presented at CES2020
 - Opening a startup after leaving job (21.10)



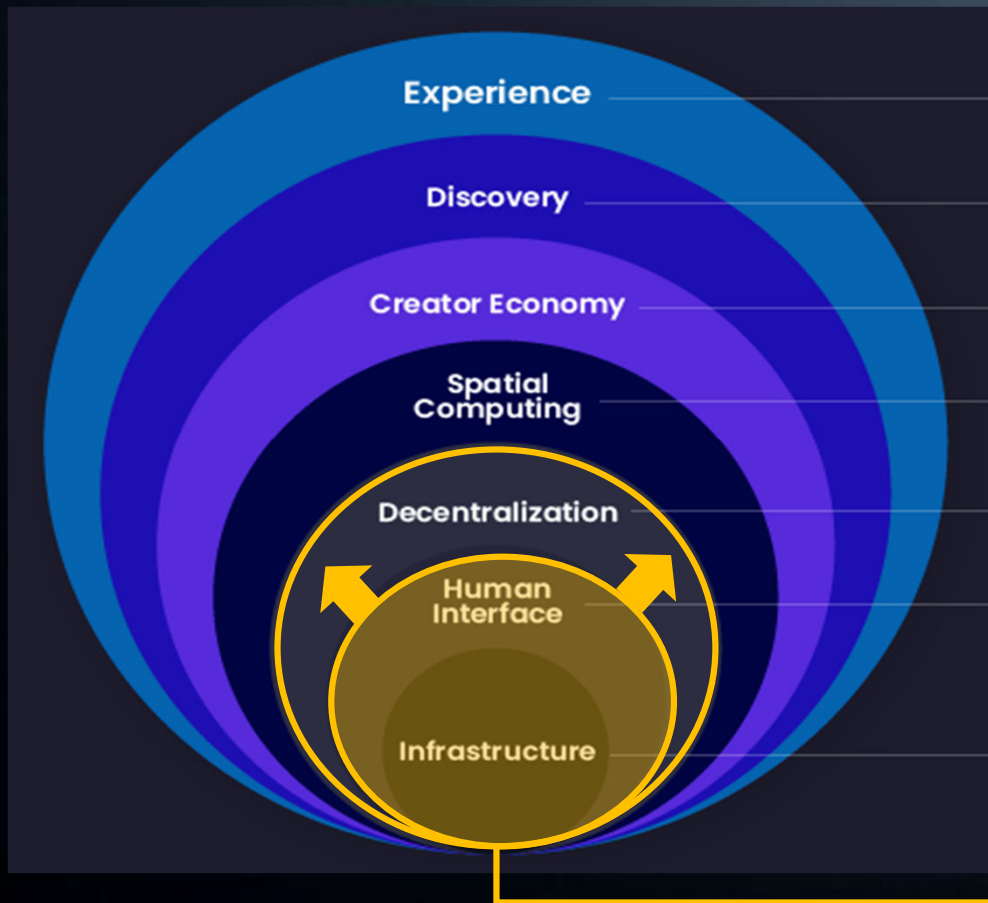




The Seven Layers of the Metaverse



Role and Relationship between Metaverse and Meta-Human

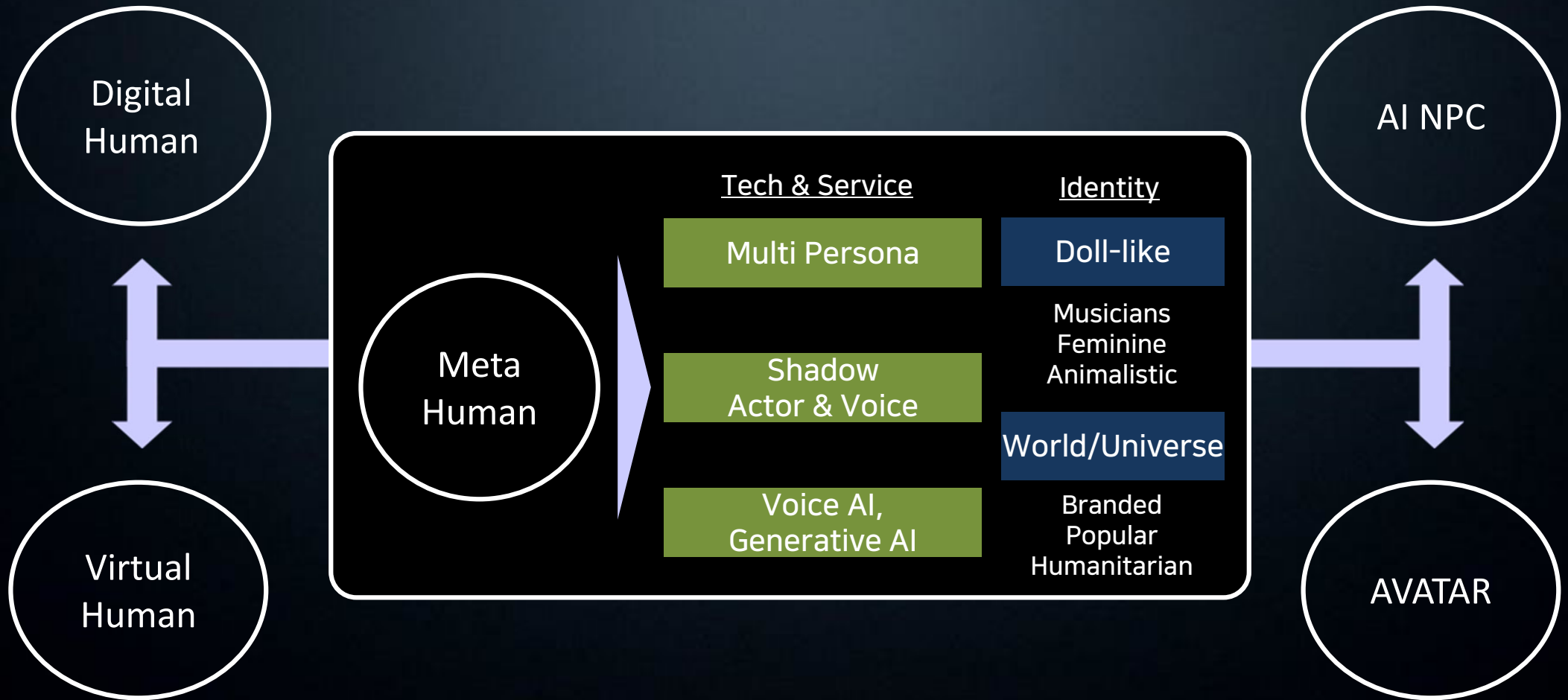


Meta-Human =

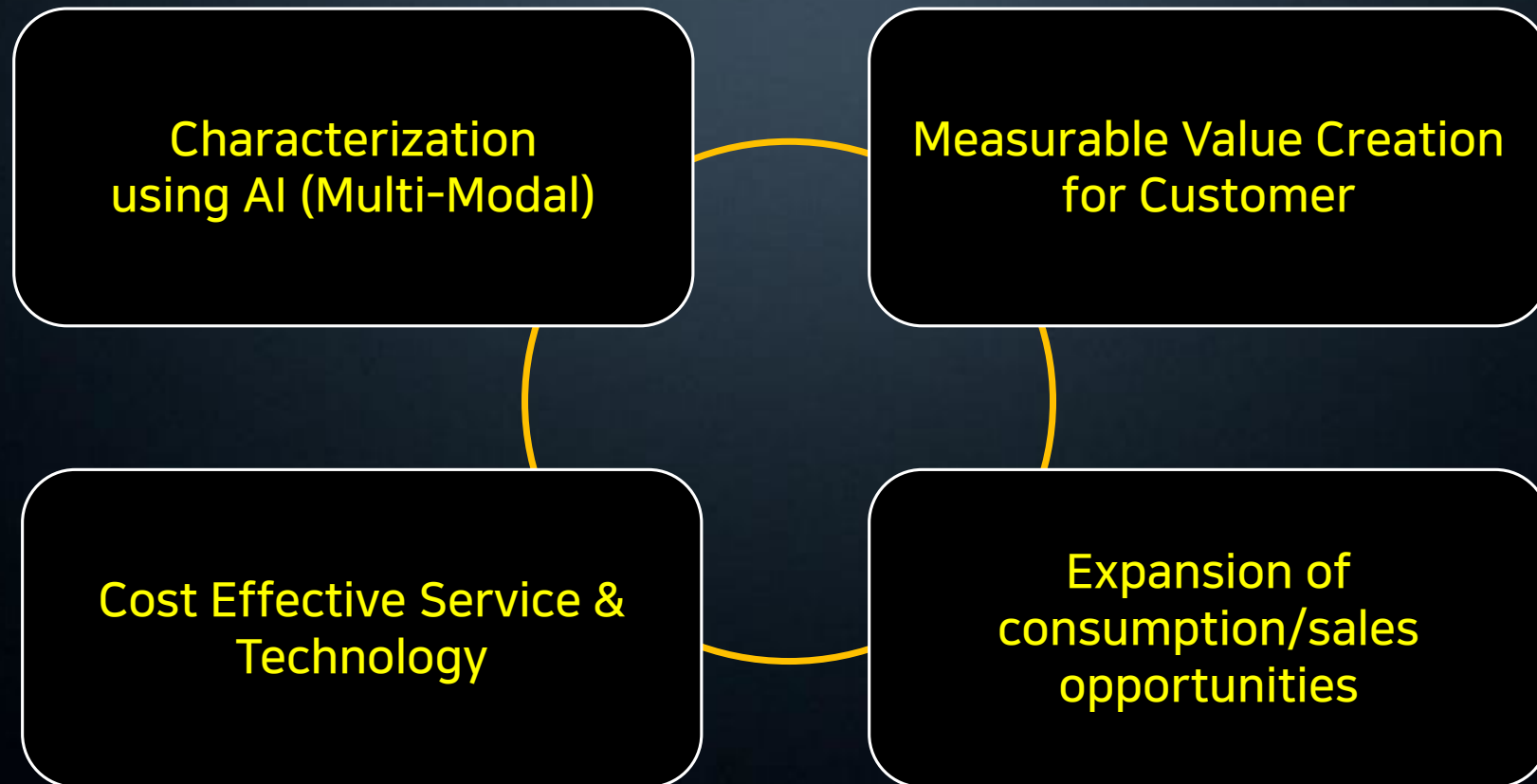
Decentralization &
Human Interface Facilitator

Customer Experience Innovation
CSF in Metaverse

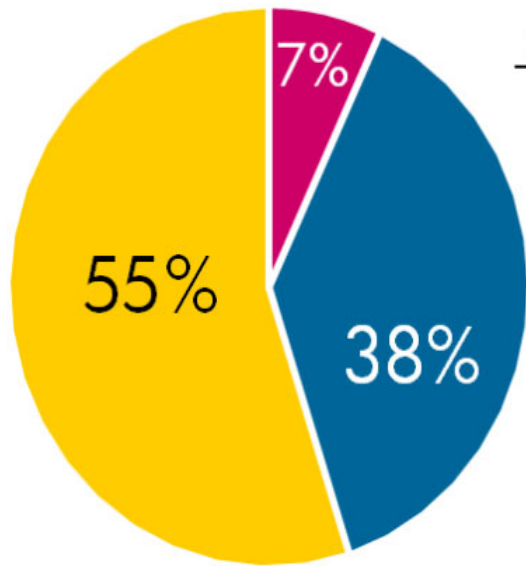
Meta-Humans with various perspectives



What's the success factor?



Related theory & Theoretical approaches



Dr. Albert Mehrabian's 7-38-55% Rule

Elements of Personal Communication

- 7% spoken words
- 38% voice, tone
- 55% body language

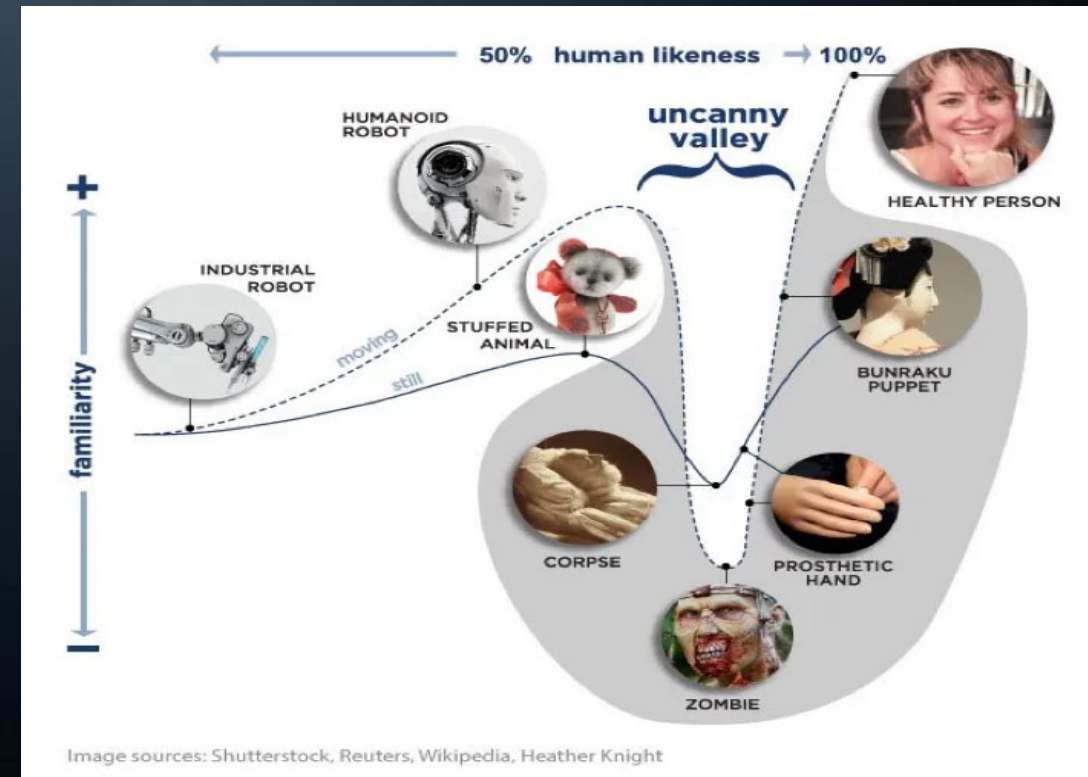
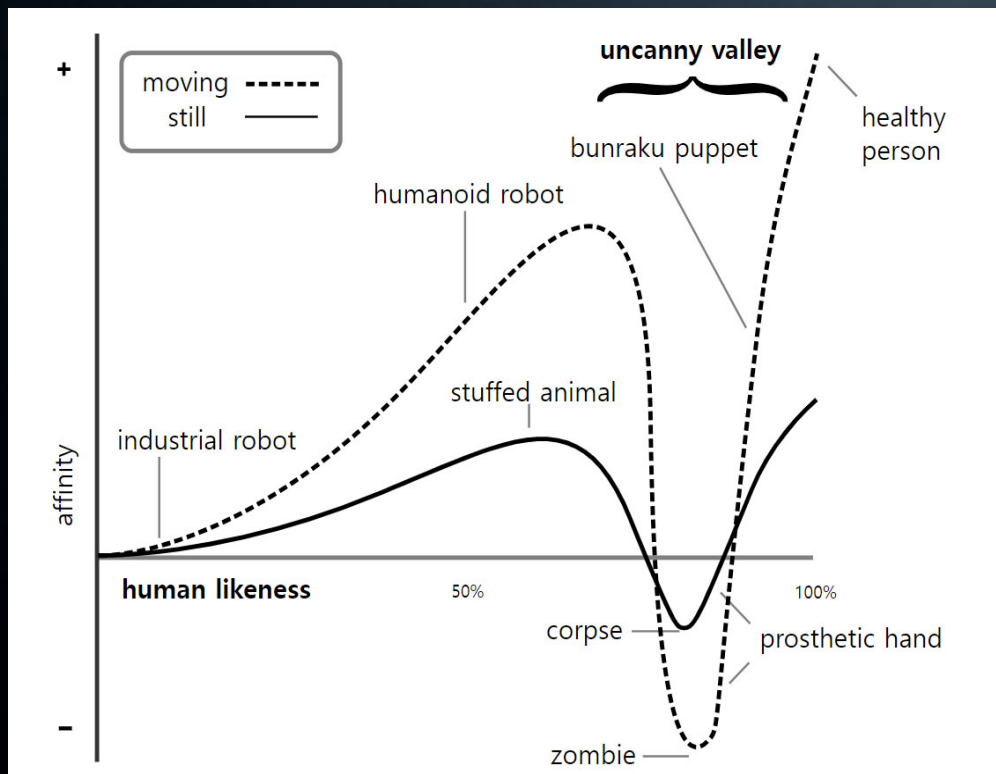
The rule states that 7 percent of meaning is communicated through spoken word, 38 percent through tone of voice, and 55 percent through body language. It was developed by psychology professor Albert Mehrabian at the University of California

- 7%: Feelings and attitudes is in the words that are spoken
- 38%: Feelings and attitudes in the voice
- 55%: Nonverbal behavior, Feeling and attitudes in facial expression, posture, and gestures

Theoretical approaches to Meta-Human

Uncanny Valley

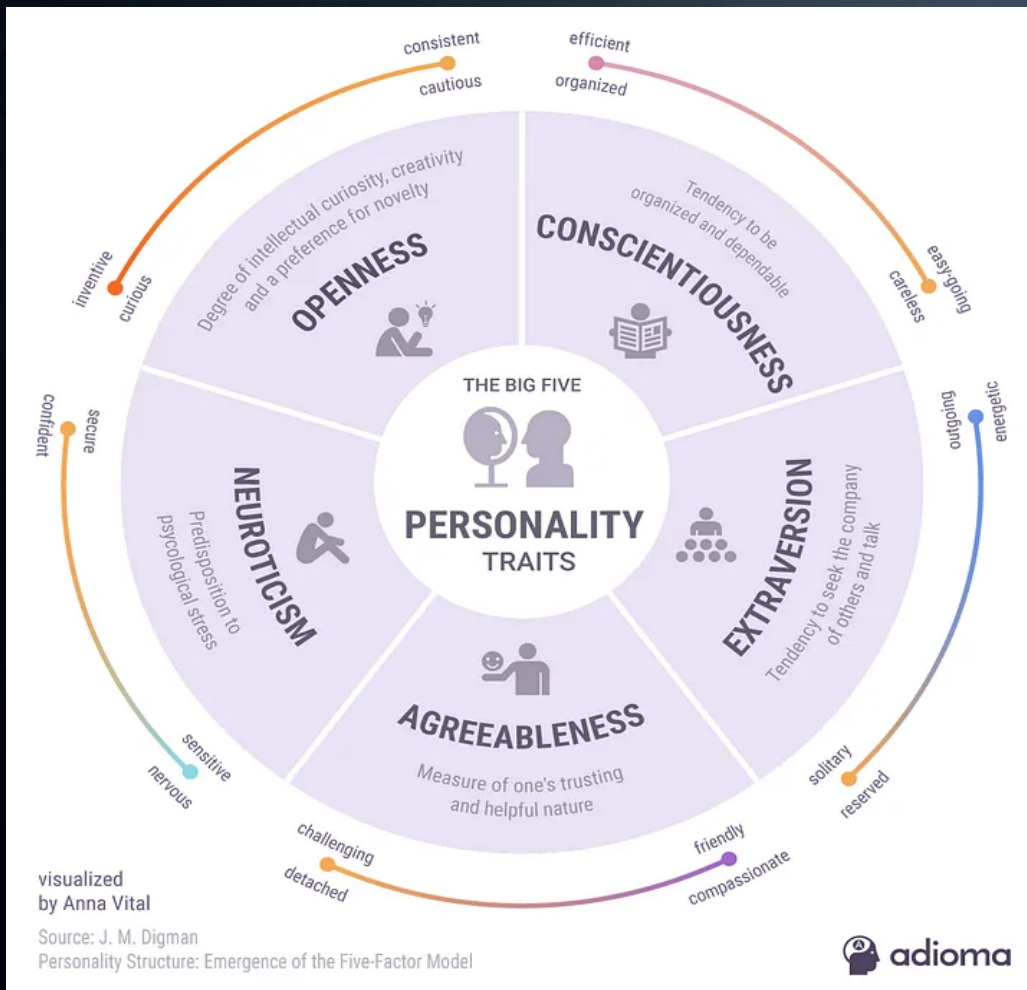
- The uncanny valley is a very precise feeling of weirdness or fear that nearly-humans (but not quite humans) evoke. The uncanny valley is elicited when a robot or AI has some human likeness, but isn't human.
- The uncanny valley is specifically about the relationship between being human-like and likability



Source: <https://virtualrealitypop.com/total-recoil-the-uncanny-valley-is-an-uncanny-cliff-8a35ecfadd3d>

Theoretical approaches to Meta-Human

FFM (Five Factor Model)



The five-factor model of personality (FFM) is a set of five broad trait dimensions or domains, often referred to as the “Big Five”: Extraversion, Agreeableness, Conscientiousness, Neuroticism (sometimes named by its polar opposite, Emotional Stability), and Openness to Experience (sometimes named Intellect).

- 신경성, 외향성, 친화성, 성실성, 개방성 -



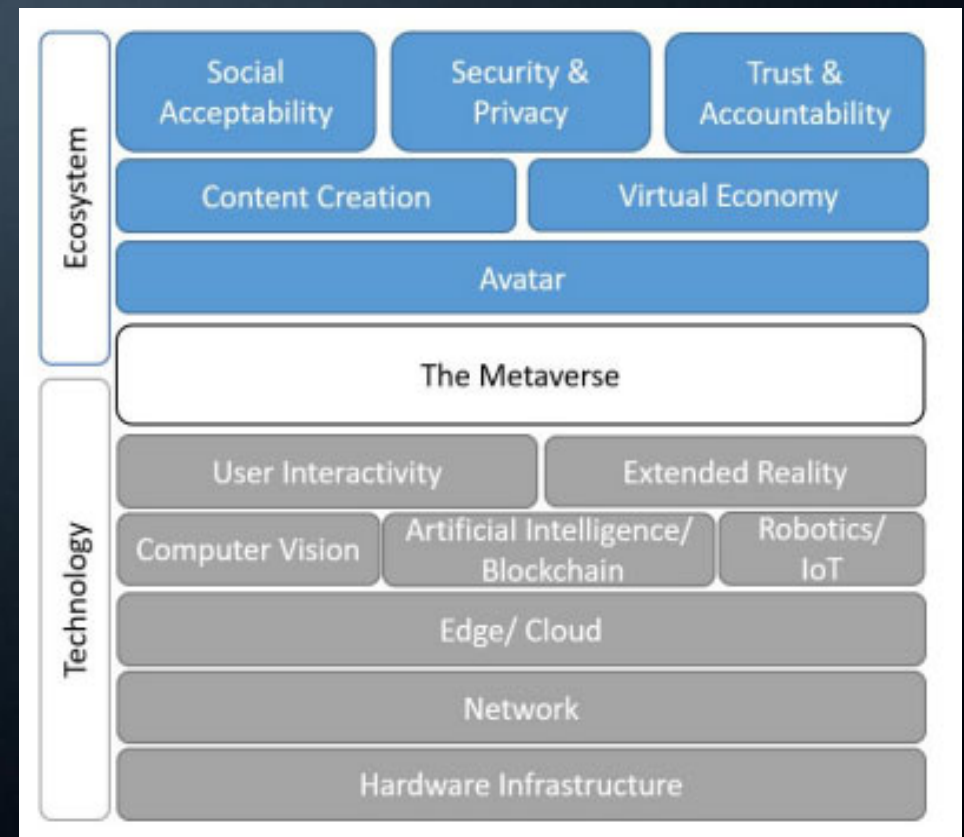
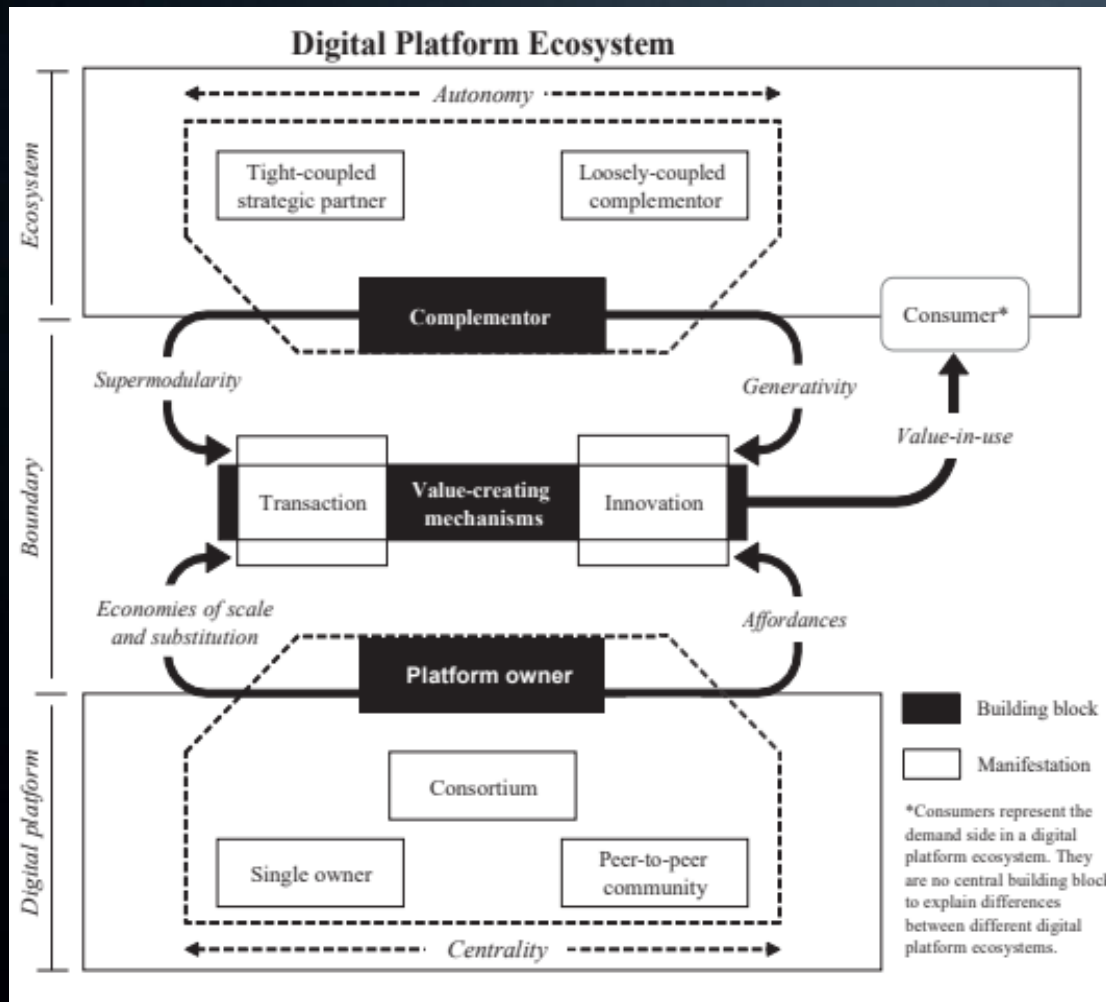
Meta-Human Tech Trend

**Digital Platform
Ecosystem**

**Meta-Human
Ecosystem**



Meta-Human Ecosystem



Meta-Human Ecosystem

Complementor

LOCUS X

GIANTSTEP

Smilegate

afun interactive

NAVER

METAVVERSE
Entertainment



Unity

neptune

onMIND

dobstudio

CG/VFX(Visual Effects)

Game/ Entertainment

Consumer

Transaction

Value Creating
Mechanism

Innovation

Platform Owner

LG
SAMSUNG

NAVER
kakao



Cheil

SK square

kt

NC

KRAFTON



LOTTE Homeshopping

Meta-Human Ecosystem Technology

World/ Universe

2D + AI

3D + AI

Hybrid

Body/
Shape

Motion Retargeting
Synthesis
Reenactment

3D Modeling
Realtime Rendering
Volumetric Capture
Motion Capture

Digital Double

Face

Face Generation/
Manipulation
Face Swap/
Deep Fake
Face Synthesis
Lip Sync

Facial Capture
Facial Rigging
Lipreading (Wav2lip)
Facial Retargeting
Game Engine

Face Reconstruction
& Face Swap

Voice

Voice Generation

ASR/ KWS/ TTS/ STT/ DM/ P-TTS

Synthetic Voice

Voice Styling

Speech Separation

VAD (Voice Activity Detection)

LLM (Large Language Model)

Generative AI (GPT)

NLP/ NLG/ NLU

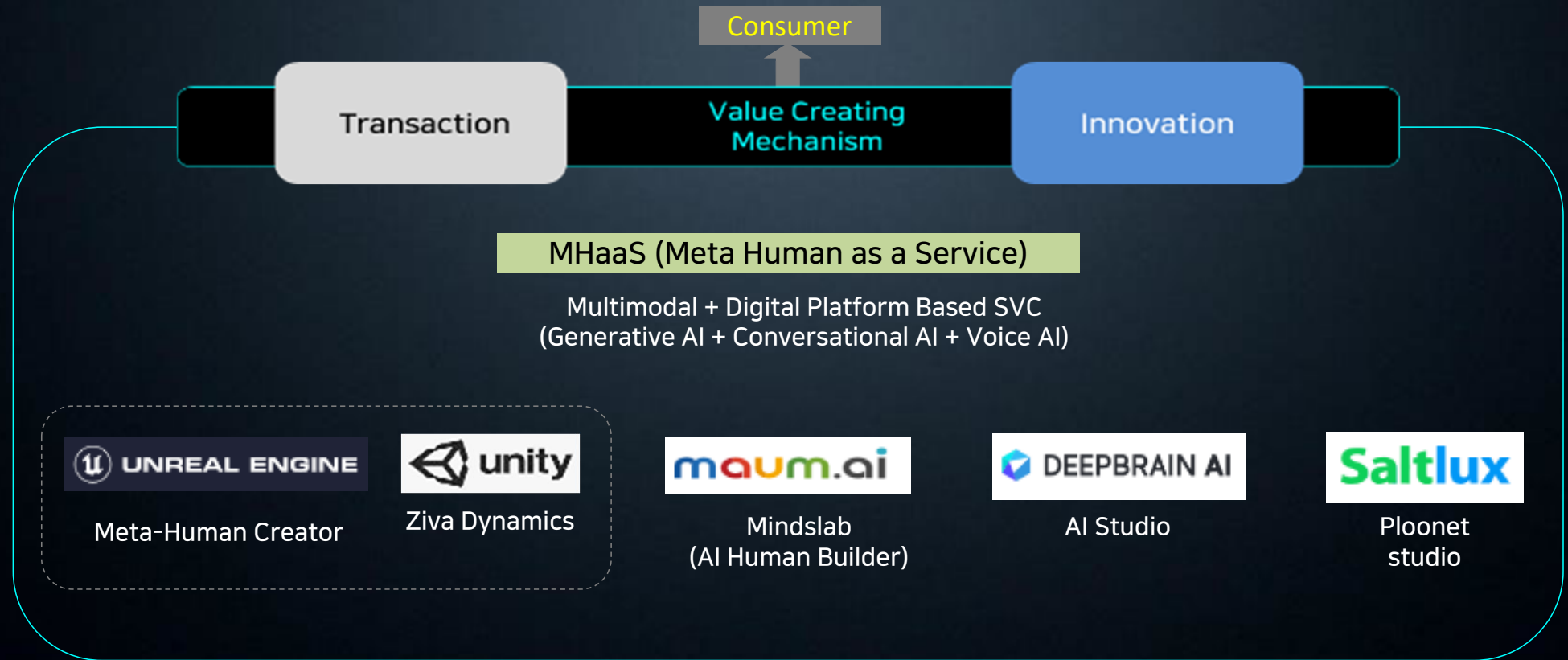
Sentimental Analysis

Hyper Scale & Multi-Modal AI

AI Orchestration

Interaction

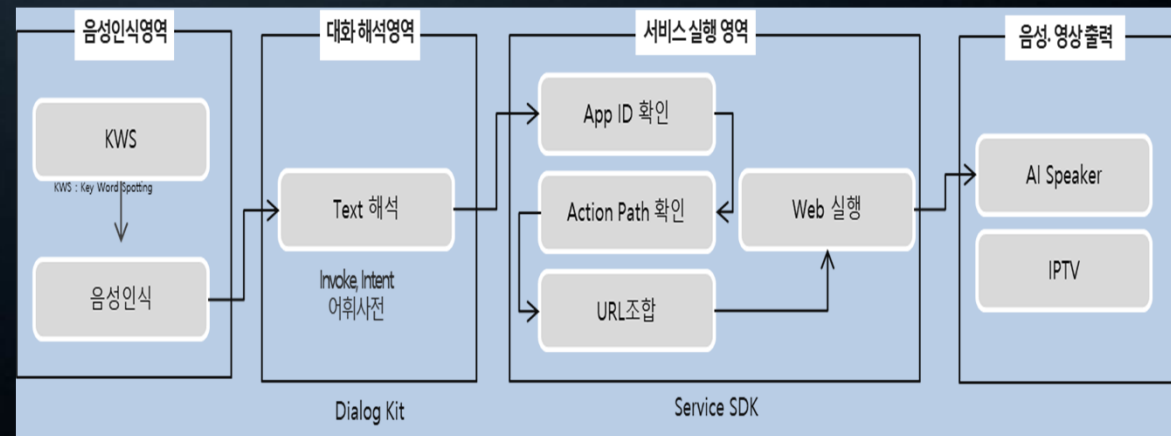
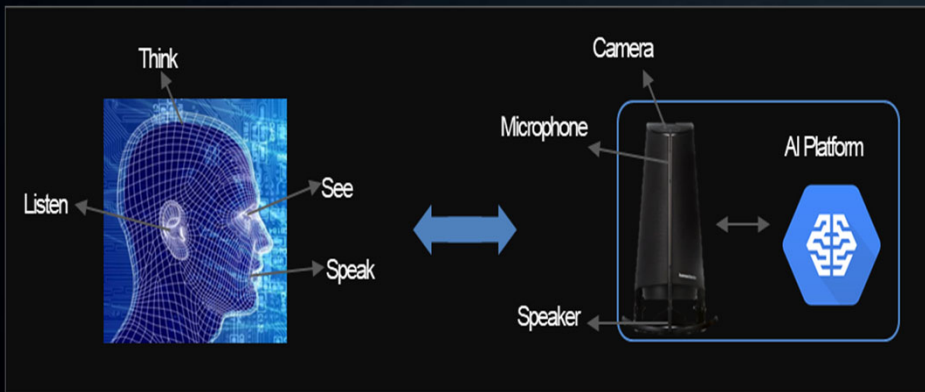
Meta- Human Solution (MHaaS)



Meta-Human Tech Trend

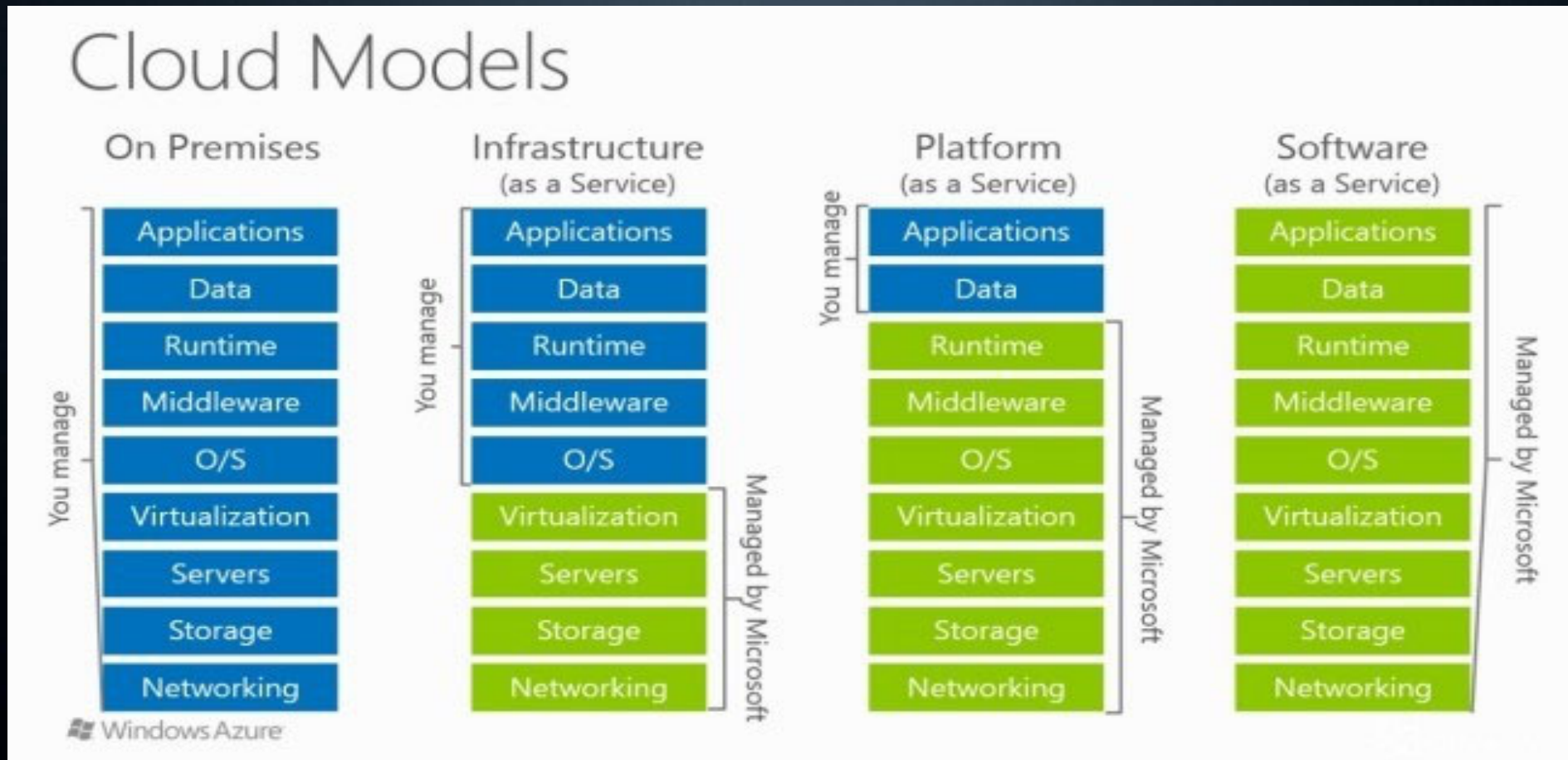
Voice AI : Human-AI communication consists of four main processes. (Wake up – Listen – Think – Speak)

- Wake-up: KWS (Key word Spotting), Hot-word Detection
- Listen: Speech Recognition Stage. ASR (Automatic Speech Recognition) / STT (Speech-to-Text) Tech Adoption
- Think: The process of understanding human language and the steps to generate language that AI can understand. Dialog Manager (DM)/ NLU(Natural Language Understanding)/ NLG(Natural Language Generation)
- Speak: Speech Synthesis Stage / Human voice commands are heard, understood, and judged within the platform and re-presented as human speech : TTS(Text-to-Speech) → P-TTS (Personalized TTS) & Conversation Styling Tech



Meta- Human Tech Trend

- Meta Human, a new way of AI → MHaaS (Meta-Human as a Service)



Meta-Human Tech Trend

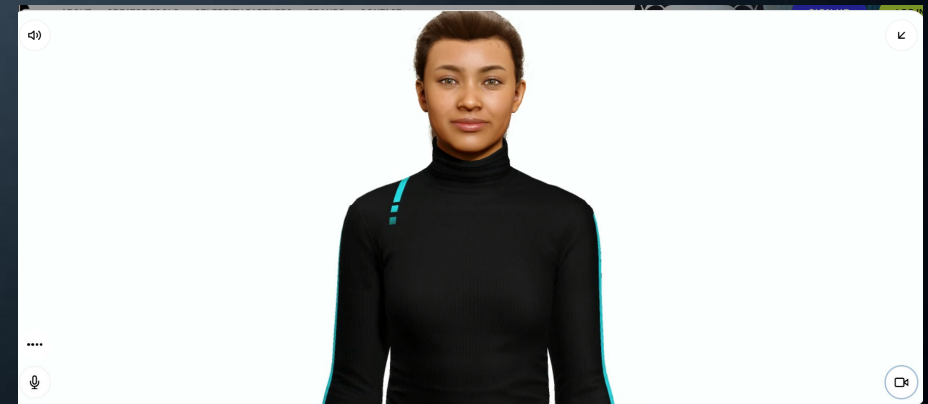
MHaaS

Meta-Human as a Service

<https://www.soulmachines.com/>



The screenshot shows the homepage of Soul Machines. At the top left is the 'sm' logo. Navigation links include 'ABOUT', 'CREATOR TOOLS', 'CELEBRITY PARTNERS', 'BRANDS', and 'CONTACT'. On the right, there are 'SIGN UP' and 'LOG IN' buttons. The main headline reads 'Create dynamic Digital People™ in minutes'. Below this, a sub-headline states: 'Soul Machines™ provides easy-to-use creator tools, so anyone can create astonishingly life-like, GPT-enabled Digital People'. A central image features a digital woman's face with a play button and the text 'TRY DIGITAL DNA STUDIO [FREE TRIAL]'. A 'TRY DIGITAL DNA STUDIO' button is located below the text. At the bottom, a chat window is open with the digital woman saying 'Hi! Let's talk!' and a 'See what Digital People can do for you' headline.



Talk to GPT-4 Nova

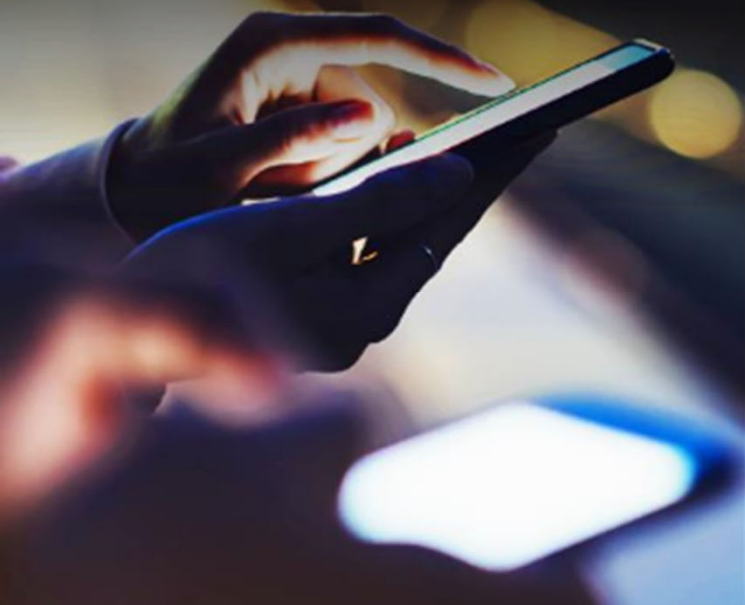
Want to see the future...or at least a beta version? Get a sneak peek at the very latest in self-animating, generative AI.

TALK TO GPT4 NOVA

인터페이스 혁신

터치

Smart phone



지능 혁신

음성/영상

GiGA Genie



Human-like
Intelligence

Everything



AI Brain

사람 같이 자연스러운

Natural Interaction



-영화 her 中



사람의 능력을 높여주는

Empowered Thinking



-영화 아이언맨 中



나를 이해하고
친구처럼 공감



언제나 나를
파악하고 챙겨주는



사람 같이 자연스러운

Natural Interaction

사람의 능력을 높여주는

Empowered Thinking



반복업무를
스스로 처리

복잡한 일은
지능적으로 분석

위험하고
힘든 일 대신 수행

Natural Interaction

감성/언어지능



영상/행동 인식 및 표현

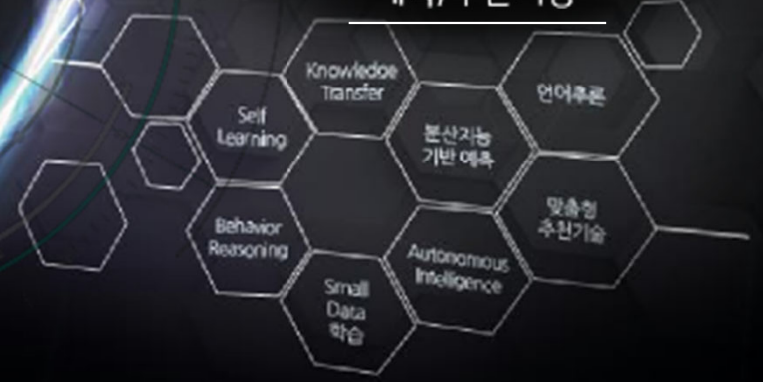


Empowered Thinking

분석/판단지능



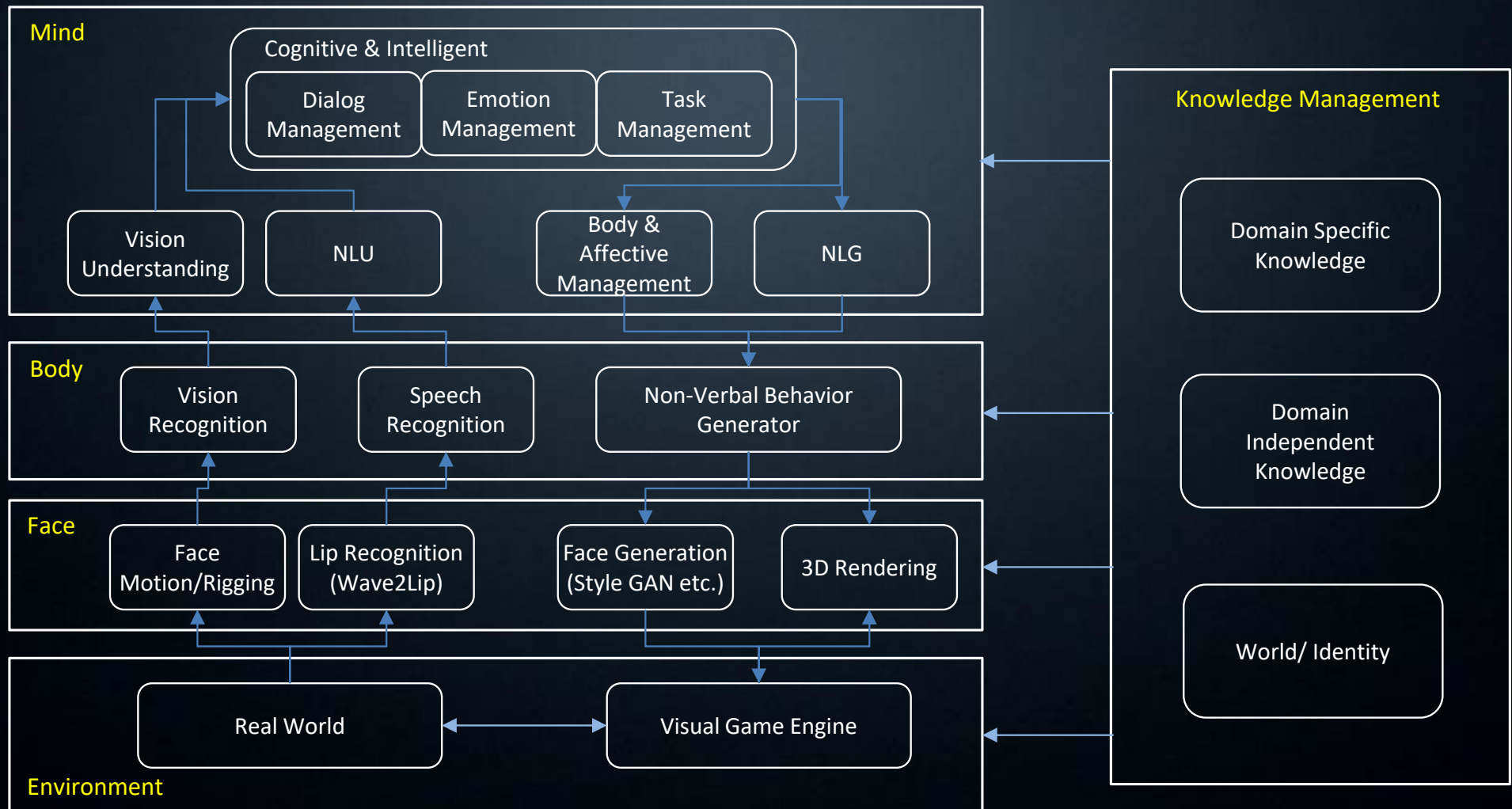
예측/추론지능



AI Brain

Future Landscape

Meta- Human Architecture Design



AI Tech *vs* Service *vs* Business

AI Tech R&D

AI Service Plan

AI Business

AI Tech vs Service vs Business

AI Business Model

AI Tech R&D

AI UX/CX/HCI

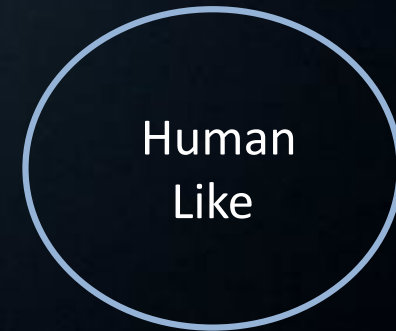
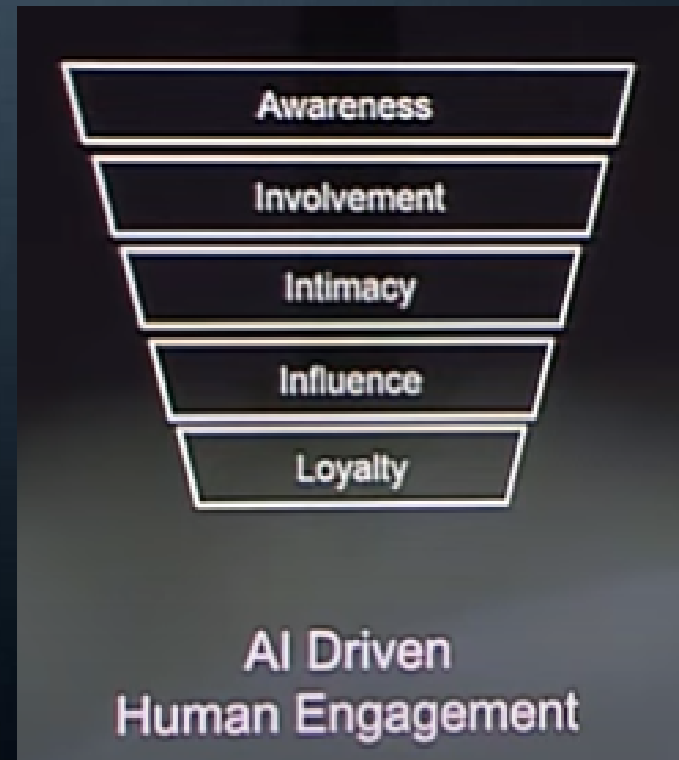
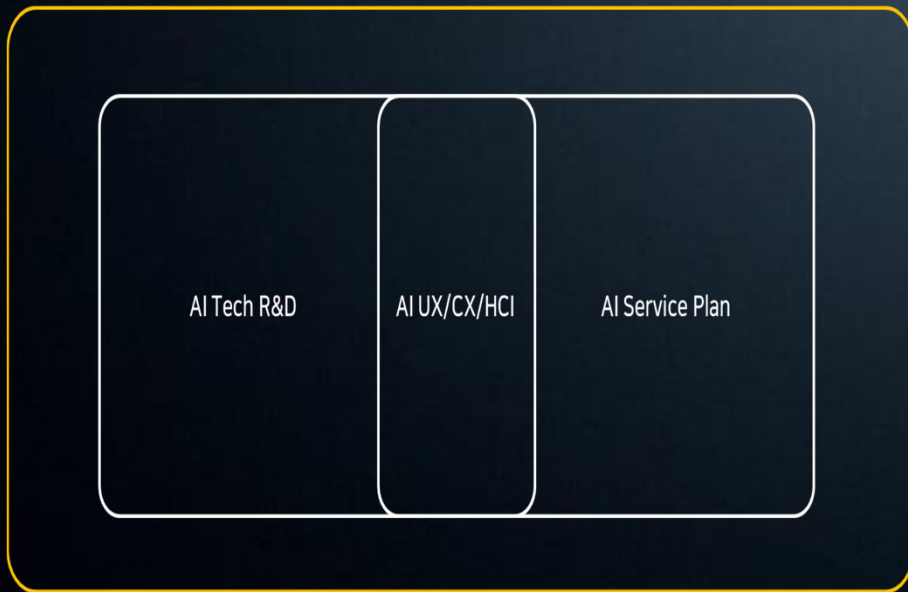
AI Service Plan

(Customer Experience/
Human-Computer Interaction)

Academic Insight & Approach

AI Contents Core

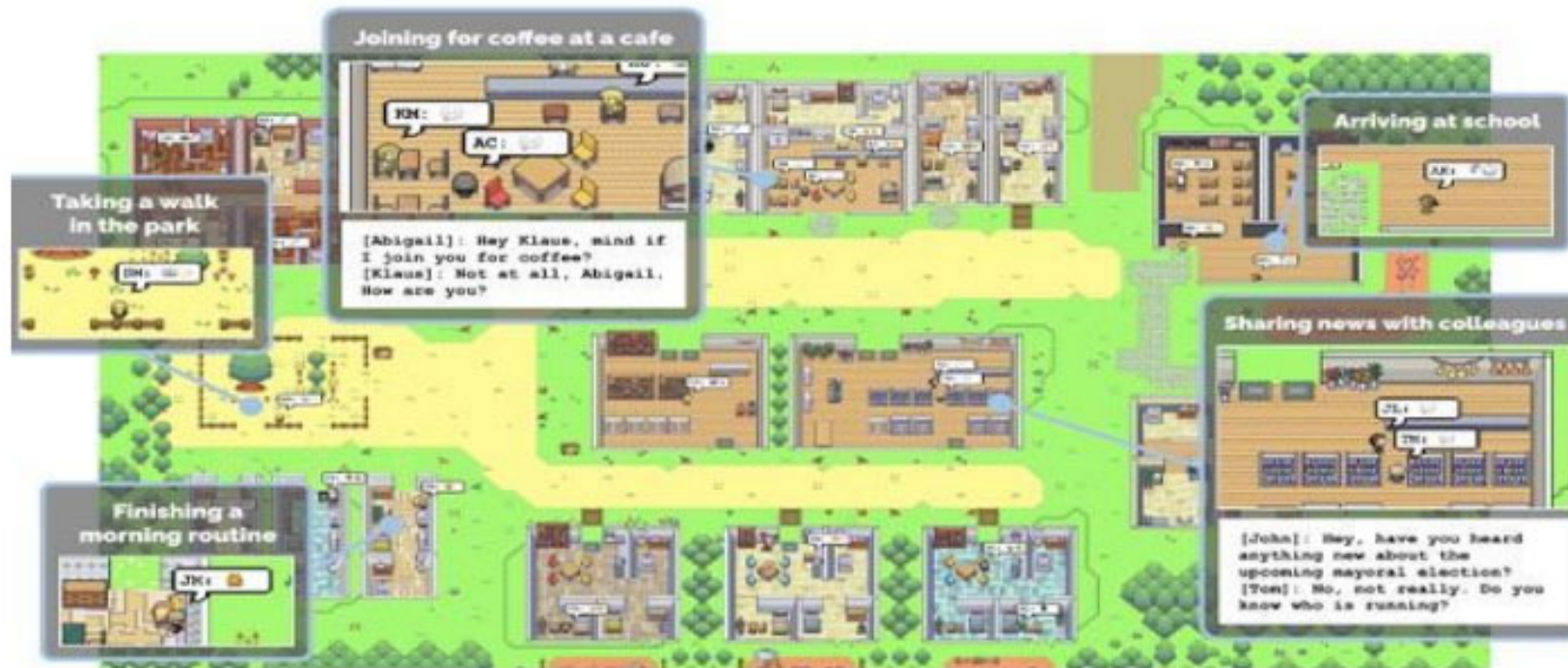
AI Business Model



Source: Korea Investment Festival KIF 2022

Thank You

Meta- Human or AI NPC (Metaverse & Generative AI)



게임 내 생성 모델을 통해 NPC들이 인간의 행동을 모방, 상호작용하는 것을 확인

- AI 기반 NPC들(총 25명)이 적극적으로 상호작용(예시 : 발렌타인 데이 파티를 계획하며 각자의 역할 수행)

[Source] "Generative Agents: Interactive Simulacra of Human Behavior (Stanford University, Google)

AI Contents Core : Tech Side

Motion Capture

Motion capture is the process of capturing an actor's physical movements and then recreating them from a digital character model.

→ 배우 움직임 모션캡처 비실시간 기술을 실시간 라이브 환경구현, Retargeting 문제 해결





Voice Interaction
AI Speaker
Or
Mobile Phone



Complex
Interaction
Vision AI + Sound
AI + Virtual Agent

멀티모달 상호작용이 바꿀 미래는?



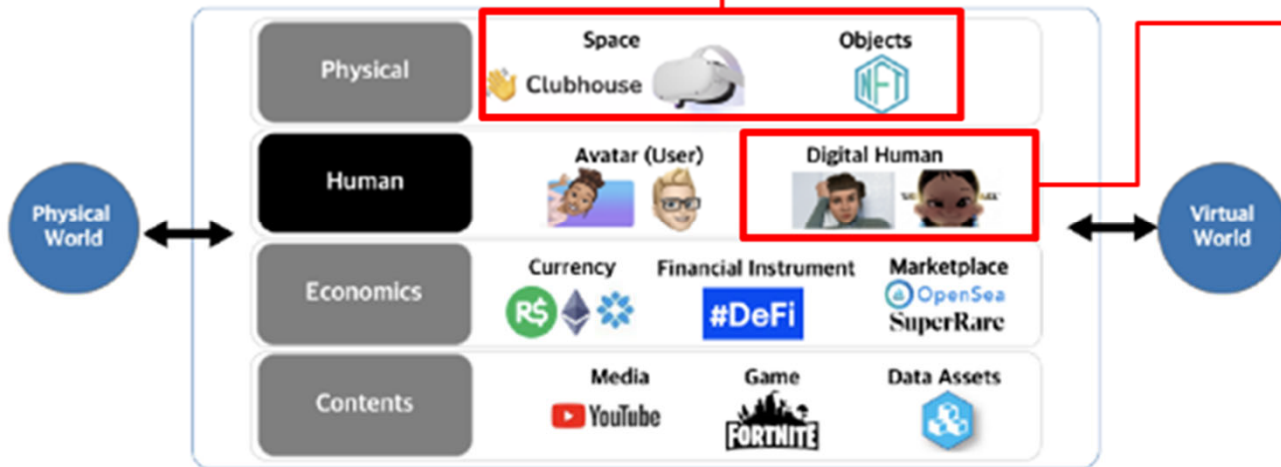
터치인터랙션 → 아바타/AI 음성, 대화, 시각정보인터랙션

Natural Multi-Turn 대화
- QA, 초거대 언어모델 AI 적용



음성인식, 음성합성
- STT, TTS 기반 사람과 AI 소통

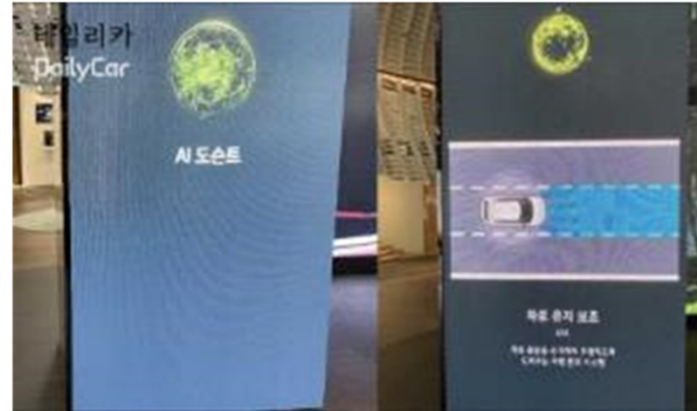
객체인식 (Vision AI)
- 얼굴 탐지, 식별, 제스처, 표정
인식 사람과의 감성케어



멀티모달 상호작용

AIBM : Value-Creating Mechanism

- ✓ [예시] 시비즈니스 측면의 멀티모달 상호작용 활용
- 멀티모달 기반 고객경험 혁신



무엇을 도와드릴까요?



스마트폰의 QR코드와 마이크를 활용해 음성으로 해당 기능을 물어 보면 AI 도슨트는 음성과 디지털 사이니지로 관련 물음에 답변



디지털휴먼이 멀티모달 상호작용하며 AI도슨트 제공

출처 : 데일리카 (<http://www.dailycar.co.kr>)

AI Virtual Human 제작 과정

로지 제작

로지팀 = 로지 → 이들의 이야기, 생각이 로지로 투영

<팀/조직 구성>

- 제작/관리: 총괄 디렉터
- 캐릭터 연출: 미디어 아트, 크리에이티브 디렉터
- 모델링 매니저: 표정, 헤어스타일링
- 합성 매니저: 2D 아트, 영상, 애니메이션 합성
- 라이팅 매니저: 라이팅 렌더링, 3D R&D
- PD : 제작, 스케줄 관리
- 홍보/마케팅 (싸이더스스튜디오): 제휴/협력



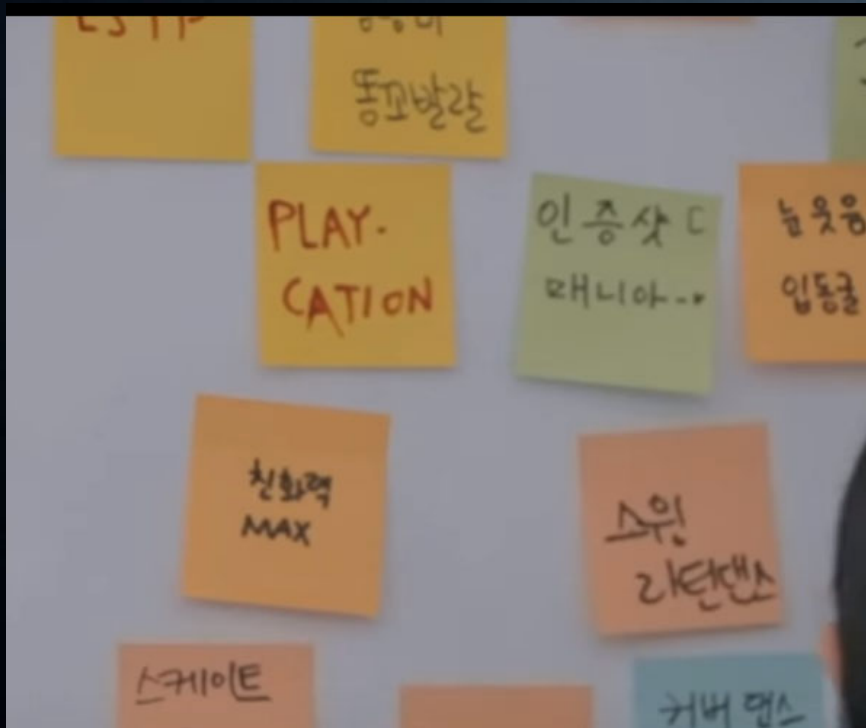
AI Virtual Human 제작 과정



마야 기반의 3D로 제작, 바디 모델 (Shadow Actor) 에 로지의 얼굴을 합성한 매치 무브 기반 Full 3D 방식, 디지털 더블 기술 적용

매치무브(MatchMove): 영화제작 과정 중 3D VFX(Visual FX, Visual effects, 시각 효과) 파이프라인(pipeline)에서 가장 중요한 작업 단계중 하나로서, 영화 씬이나 비디오로부터 이 장면을 찍은 카메라의 위치(궤적)를 복원, 추적해 내거나 공간을 재구성하는 작업
장면을 촬영한 세트장 환경을 그대로 옮겨 컴퓨터에 재구성

AI Virtual Human 제작 과정



로지의 인기비결은?

- 구체적인 프로필
- 명확한 세계관
- 인스타 업로드 메시지 말투 고민
- 아무도 컨펌 하지 않는다



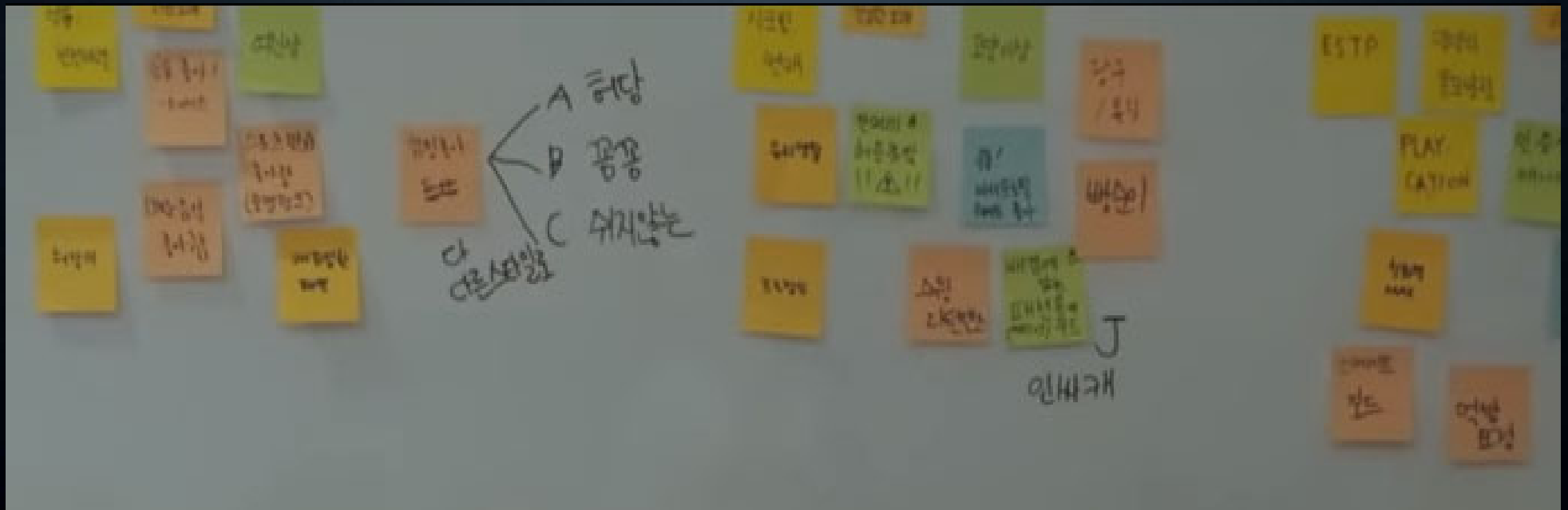
세계 최초와 국내최초의 만남

AI Virtual Human 제작 과정

로지 제작 과정

1) 기획

- MZ세대의 관심 키워드 브레인 스토밍 (라이프 스타일 전반)



AI Virtual Human 제작 과정

2) 촬영

- 동작, 의상착용 등 대역 모델 촬영



3) 모델링 작업

- 3D 모델링- 룩 디벨롭먼트- 합성

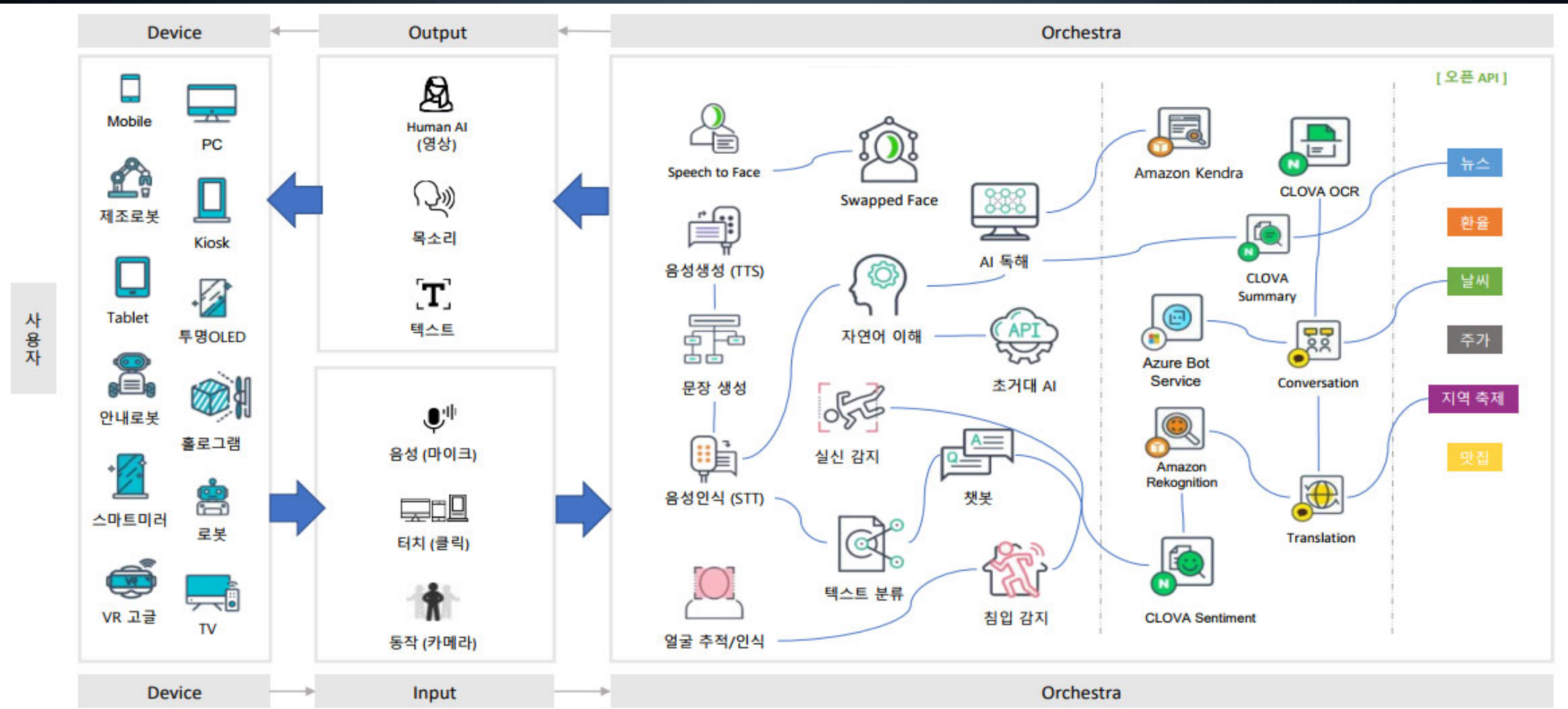
- 3D 모델링: 3차원 가상공간에서 입체적인 그림을 그려 실제처럼 보이도록 사실감을 구현

*룩 디벨롭먼트: 3D 화면에 구현된 대상의 조명, 질감, 색감 등 시각적 요소를 조정 (360카메라 촬영 후 부족한 빛, 질감 등을 조정)



- 2D 합성: 로지 얼굴과 대역모델을 합성 후 컬러, 명암 등 조정 (합성 매니저)

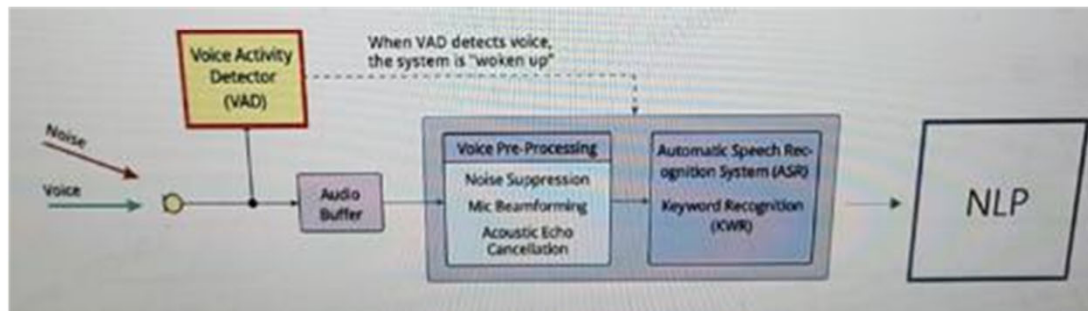
Virtual Human Service Orchestration



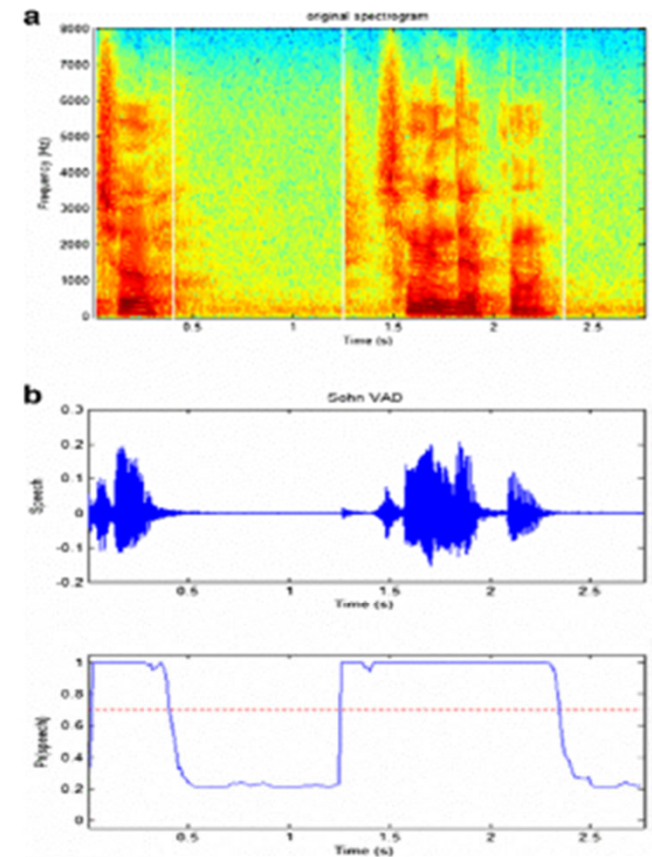
사용자

AI Business Model Canvas

- ✓ 잡음환경에서데이터수집, 음성인식, 음성코딩등 전처리단계기술에 대한내용 보완
 - ✓ VAD (Voice Activity Detection) 적용제안
 - 노이즈가있는 환경에서 음성구간을 검출하는 기술
 - 각프레임에서 음성유무를 Binary Classification으로 판단
 - 실제음성구간만 검출 전처리단계이후 음성인식/화자인식후처리단계전달
- 실시간모든음성이 아닌 실제음성만 검출하여 연산량증가에 따른 과부하, 지연방지

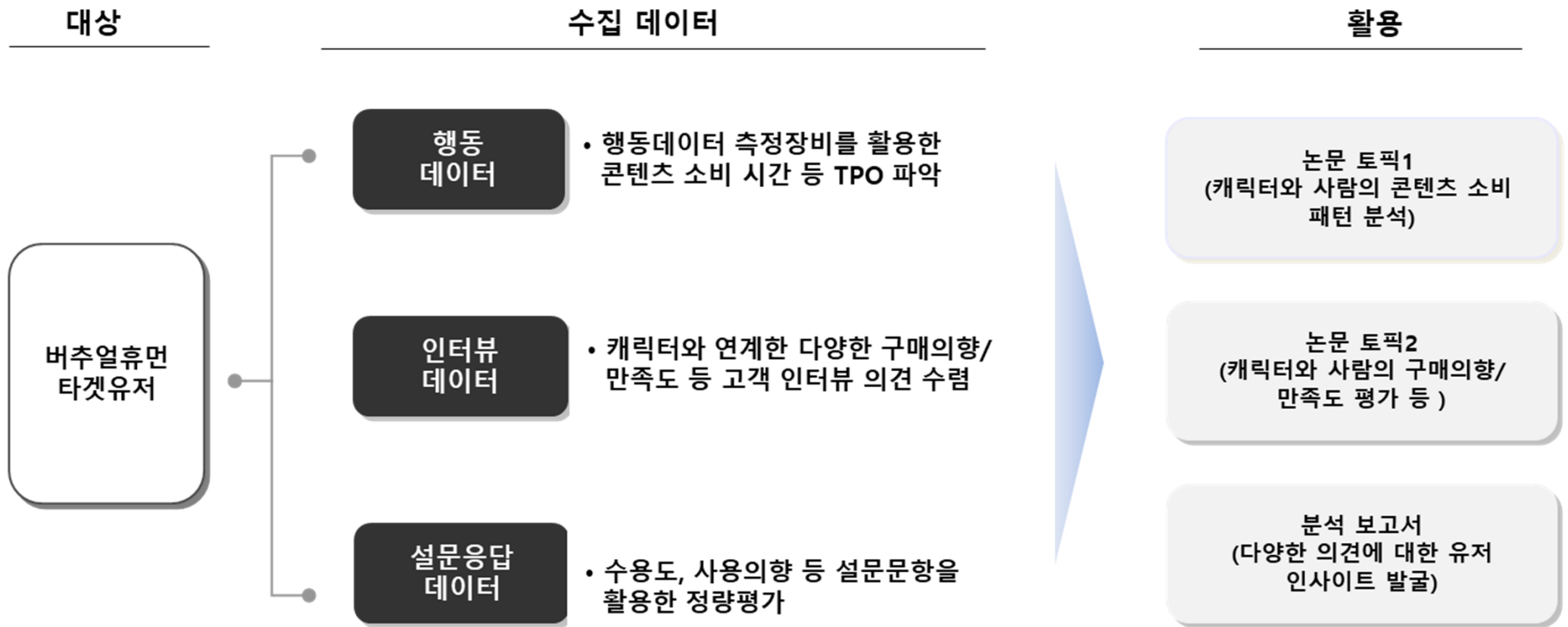


AIBM : Value-Creating Mechanism



수집 데이터 및 활용

타겟유저의 실제 행동과 인터뷰, 설문 응답 데이터를 분석하여 논문 주제 및 인사이트 도출



연구 방법

본 연구는 FGI, Survey, Usability Test를 Hybrid 방식으로 통합 수행할 계획임

수행방법

- 본 연구 수행 시 특정 장소를 섭외하여 Hybrid 방식으로 FGI+ Survey + Usability Test 시행하고자 함 (단, 해외는 Survey)

FGI

타겟 유저를 5~10명 그룹으로 구성하여
사전 논의 및 궁금한 가설/질문에 대한
검증 및 니즈 전반에 대한 평가
→ FGI Guideline 협의 필요



국내

Survey

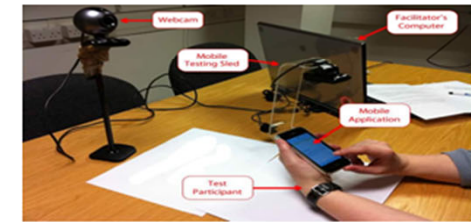
버추얼 휴먼 관련 국내외 논문에서
설문문항을 발췌하여 논문 연구/결과로
활용하며, 평소 궁금한
가설/질문에 연계하여 시행



국내, 해외

Usability Test

실제 타겟유저가 사람과 버추얼휴먼의
콘텐츠 관점의 소비가 다른지를
HW or SW Tool을 활용하여 측정하여
실증연구 활용



국내

연구 실행 계획

② Usability Test (사용성 테스트)

[실험 1]아이트래킹

[연구 목표]

연구 모형 ①
인간 vs. 버추얼 캐릭터 주목효과 비교

연구 모형 ②
버추얼 캐릭터 외형간 주목효과 비교

[실험 설계]

1) Calibration

2) 인간, 버추얼 캐릭터가 섞인 썸네일 제시 (5초) | 아이트래커로 시선데이터 측정

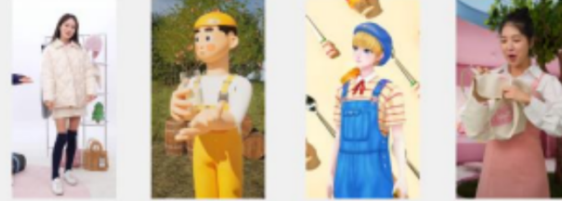
3) 다른 종류의 인간, 버추얼 캐릭터가 섞인 썸네일 제시 (5초) | 아이트래커로 시선데이터 측정

4) 버추얼 캐릭터만 있는 썸네일 제시 (5초) | 아이트래커로 시선데이터 측정

5) 다른 종류의 버추얼 캐릭터만 있는 썸네일 제시 (5초) | 아이트래커로 시선데이터 측정

6) 시선데이터 측정 종료

(예시)



(예시)

