

글로벌 ICT 표준 컨퍼런스 2023

Global ICT Standards Conference 2023

(세션7) ICT 표준화 전략 및 활동 지원

개인화된 디지털 휴먼 생성 기술의 표준화

CTO 강지수, 클레온

주최



과학기술정보통신부
Ministry of Science and ICT



특허청
Korean Intellectual
Property Office

주관



국립전파연구원
National Radio Research Agency



IITP

KEA

kista

ETRI

Index

01 Introduction

02 Evolution of digital human technology

03 Digital Human standardization

04 Klleon's standardization

01. Introduction

클레온의 비전과 미션을 소개합니다

Make digital communication sincere

클레온은 디지털 휴먼을 사용해 인간의 소통을 혁신하고자 합니다.
아래와 같은 인간 소통의 시공간적 한계를 기술을 통해 극복하고자 합니다.

1. 시간적 공간적 한계로 인해 나와 대화할 수 있는 사람은 제한적
2. 과거의 위인과 대화하는게 불가능
3. 전세계적으로 인지도 있는 사람과 대화 어려움



01. Introduction

클레온의 서비스를 소개합니다

 Digital Human Content Creation Solution • 2022. 09 Launch Date • 50 Customers Acquired Digital Human Full-body Creation Creating a custom digital human by selecting preferred body type and clothing, and inputting 1 face photo and 30-second voice recording. Message-based Digital Human Video Creation Auto-generation of digital human video upon text input with various features including background, BGM, subtitles, etc. Major Clients 	 Automatic AI Dubbing Solution • 2022. 11 Launch Date • 10 Customers Acquired Voice Generation Dubbing of 5 Languages Multilingual voice generation with the same tone as the original voice in Korean, English, Spanish, Japanese, and Chinese. Creation of Natural Lip Syncs that Match the Voice Various features such as speaker separation, noise removal, subtitles, etc. to generate appropriate lip syncs that match the speaker's voice. Major Clients 	 Digital Human Chatbot Platform • 2023. 06 Launch Date • DAU = 10000 Digital Human Face Creation Automatic generation of a digital human's face upon input of 1 face photo and 30-second voice recording. Sales-oriented digital human video Automatic creation of customized digital human sales content for customers upon input of their name and contact information. Major Clients
--	--	--

01. Introduction

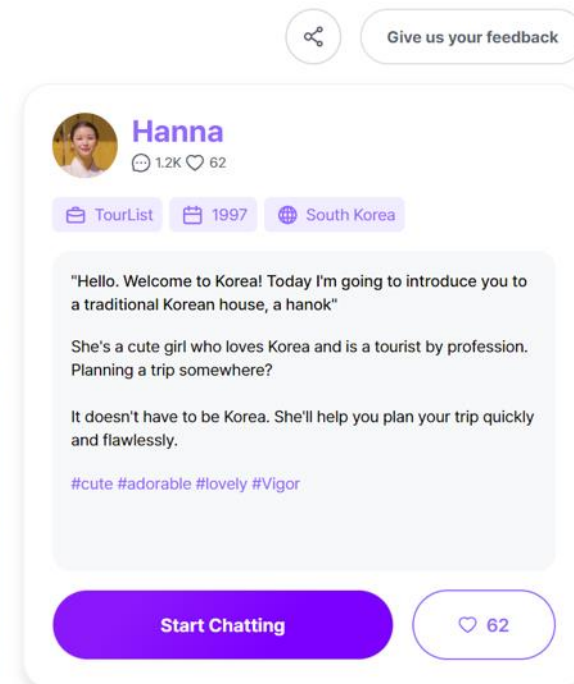
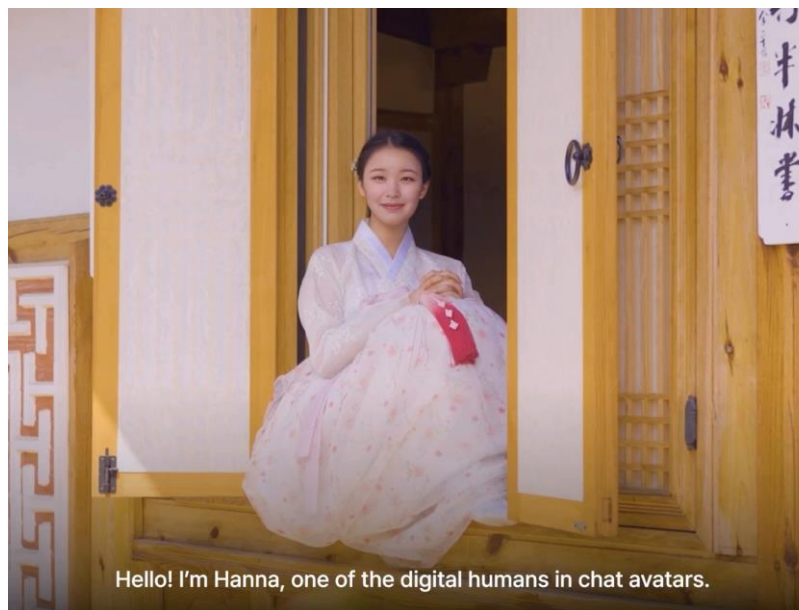
클레온의 서비스를 소개합니다

Chat-Avatar

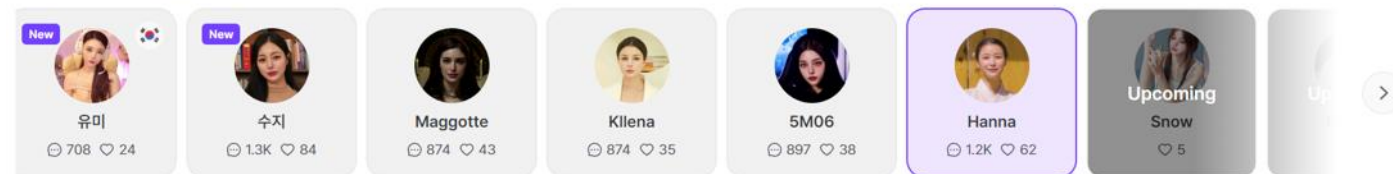
클레온은 디지털 휴먼과 사람이 실시간으로 소통할 수 있는 챗 아바타 서비스를 개발했으며, 모두가 개인의 디지털 휴먼을 제작하고, 소통하는 플랫폼을 개발하고자 합니다.

<https://www.chat-avatar.com/>

Chat Avatar

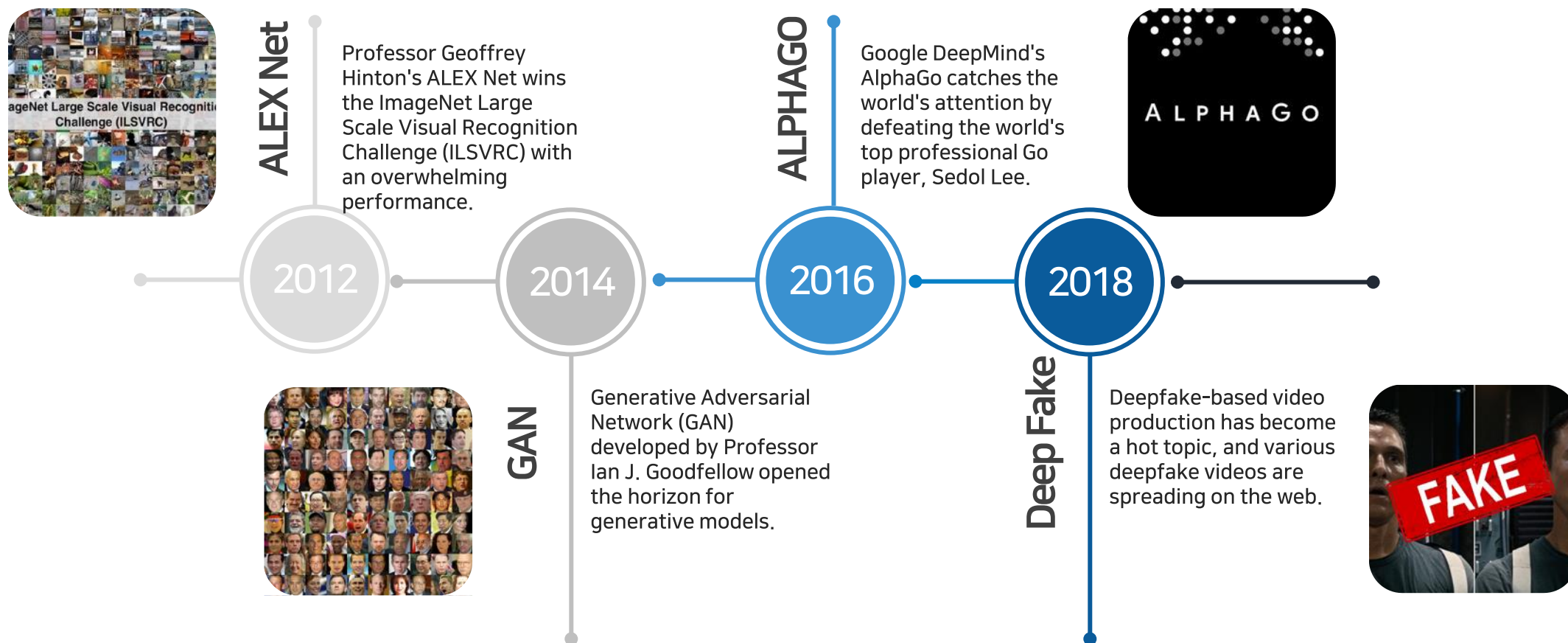


Avatar list 27



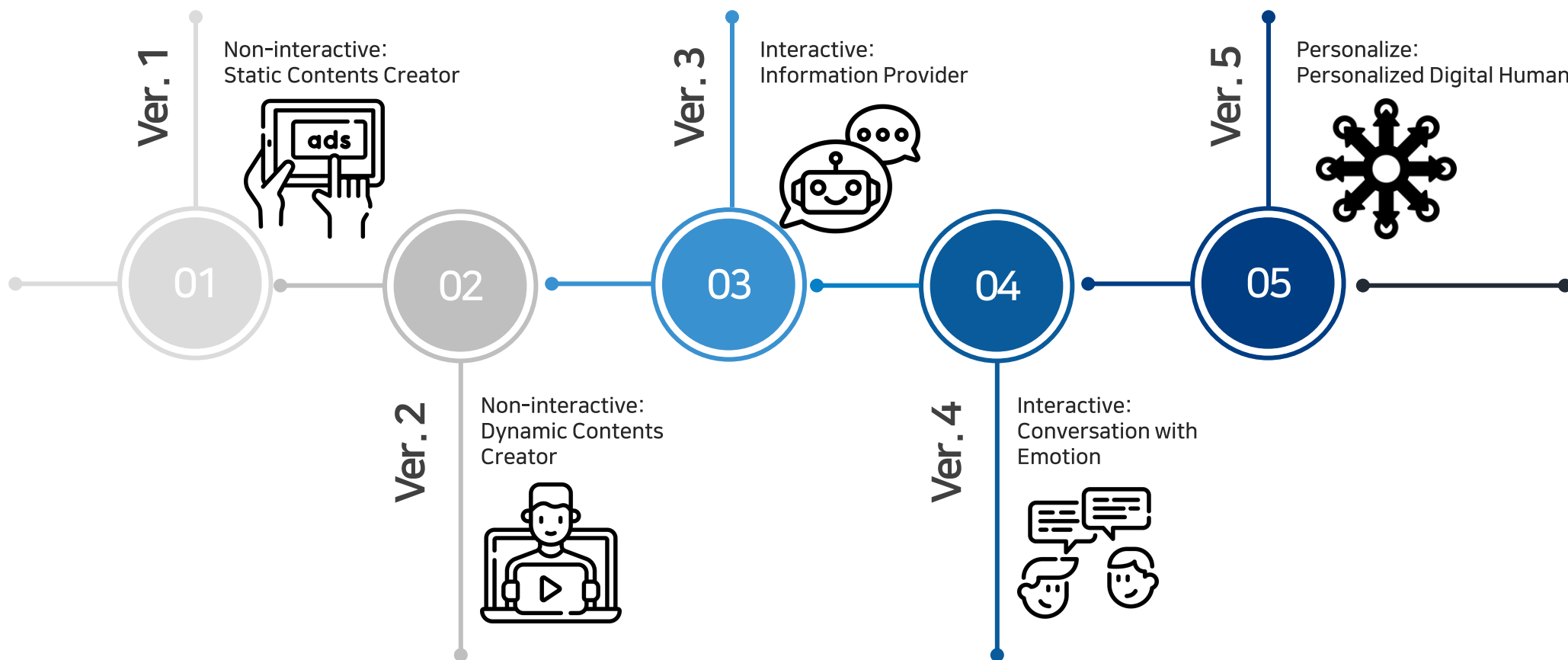
02. Evolution of digital human technology

AI기술이 빠르게 발전하고있습니다



02. Evolution of digital human technology

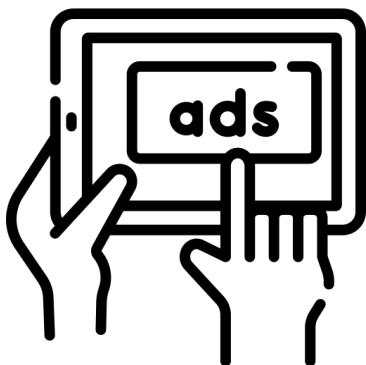
디지털 휴먼 기술이 빠르게 발전하고 있습니다



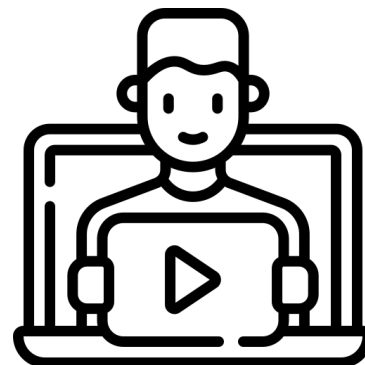
02. Evolution of digital human technology

상호작용이 없는 디지털 휴먼

1. Static



2. Dynamic

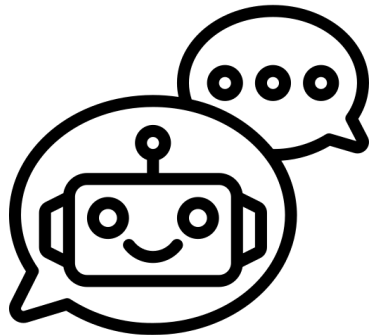


- 상호작용이 없는 디지털 휴먼으로써 디지털 휴먼의 외형만을 구현합니다.
- Static: 사전에 디지털 휴먼을 사용해 콘텐츠를 제작합니다(광고 영상 제작)
- Dynamic: 실시간으로 디지털 휴먼을 사용해 콘텐츠를 제작합니다(버튜버)
- 실시간 생성 기술이 발전하면서 Dynamic 콘텐츠 제작이 가능해졌습니다.

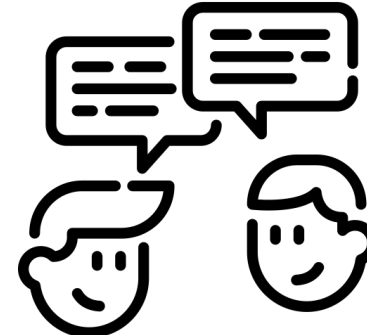
02. Evolution of digital human technology

상호작용이 있는 디지털 휴먼

3. Informative



4. Emotional



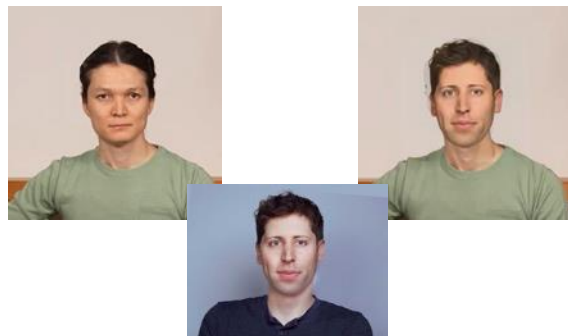
- 상호작용이 있는 디지털 휴먼으로써 디지털 휴먼의 내면도 구현합니다.
- Informative: 디지털 휴먼이 정해져있는 정보를 제공합니다.
- Emotional: 디지털 휴먼이 감정 기반으로 사람과 유사한 대화를 제공합니다.
- Chat-GPT의 발전과 함께 상호작용 가능한 디지털 휴먼이 빠르게 발전하고있습니다.

02. Evolution of digital human technology

개인화 가능한 디지털 휴먼

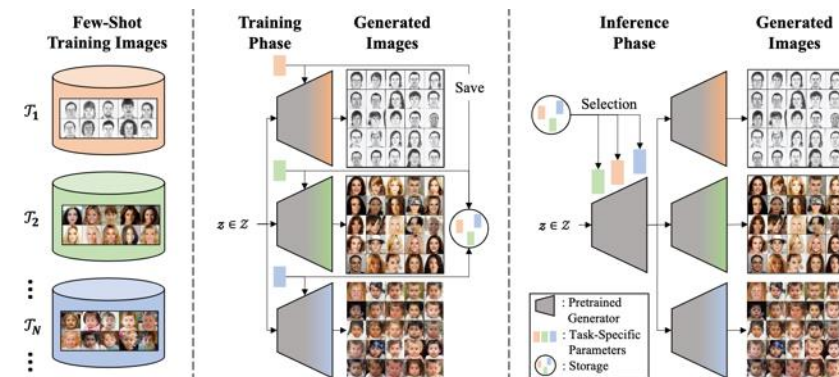
5. Low Cost

Only one face image/30 seconds of voice



6. Evolving

LFS-GAN: Lifelong Few-Shot Image Generation (ICCV 2023)



- 모두가 자신의 디지털 휴먼을 제작 및 활용할 수 있는 보급화 단계입니다.
- Low Cost: 디지털 휴먼 제작 및 활용에 적은 비용이 들어야하며, 적은 데이터만을 활용합니다.
- Evolving: 디지털 휴먼이 사람과 상호작용하며 지속적으로 발전합니다.
- 클레온은 이미 관련 기술을 보유하고있습니다.

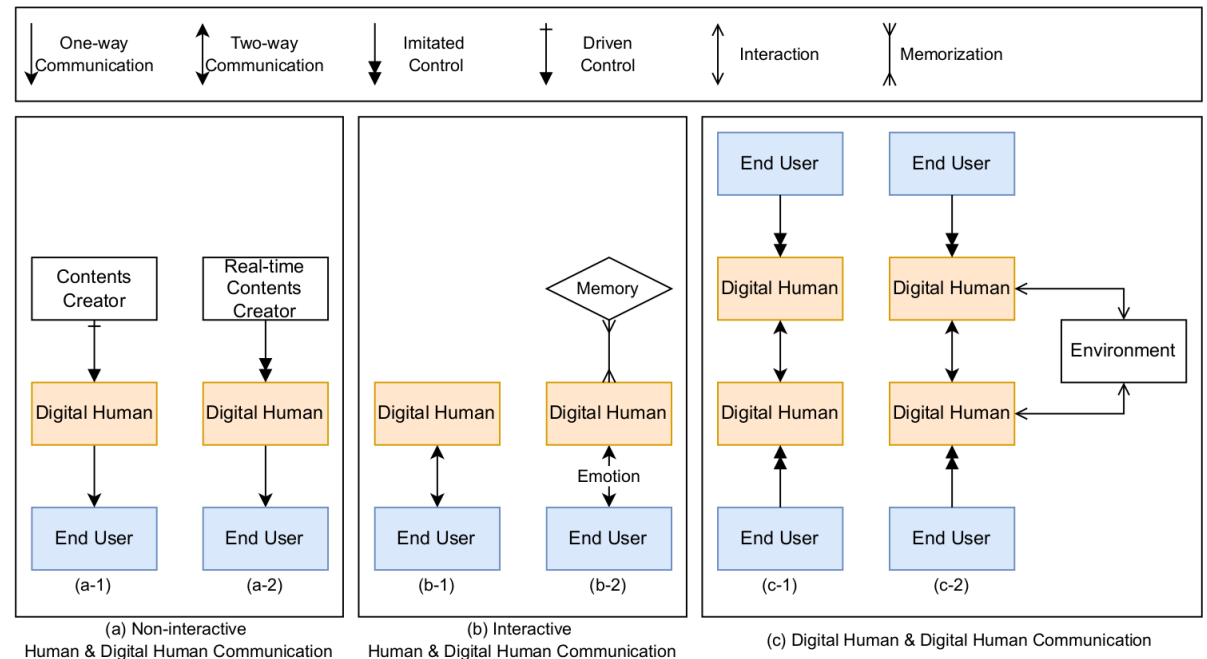
02. Evolution of digital human technology

클레온의 표준 문서를 소개합니다

F.CSDH: Requirements of communication services for digital human

총 3가지 커뮤니케이션 서비스 타입을 정의합니다.

1. Non-interactive Communication Service
2. Interactive Communication Service
3. Digital-Human Digital-Human Communication Service

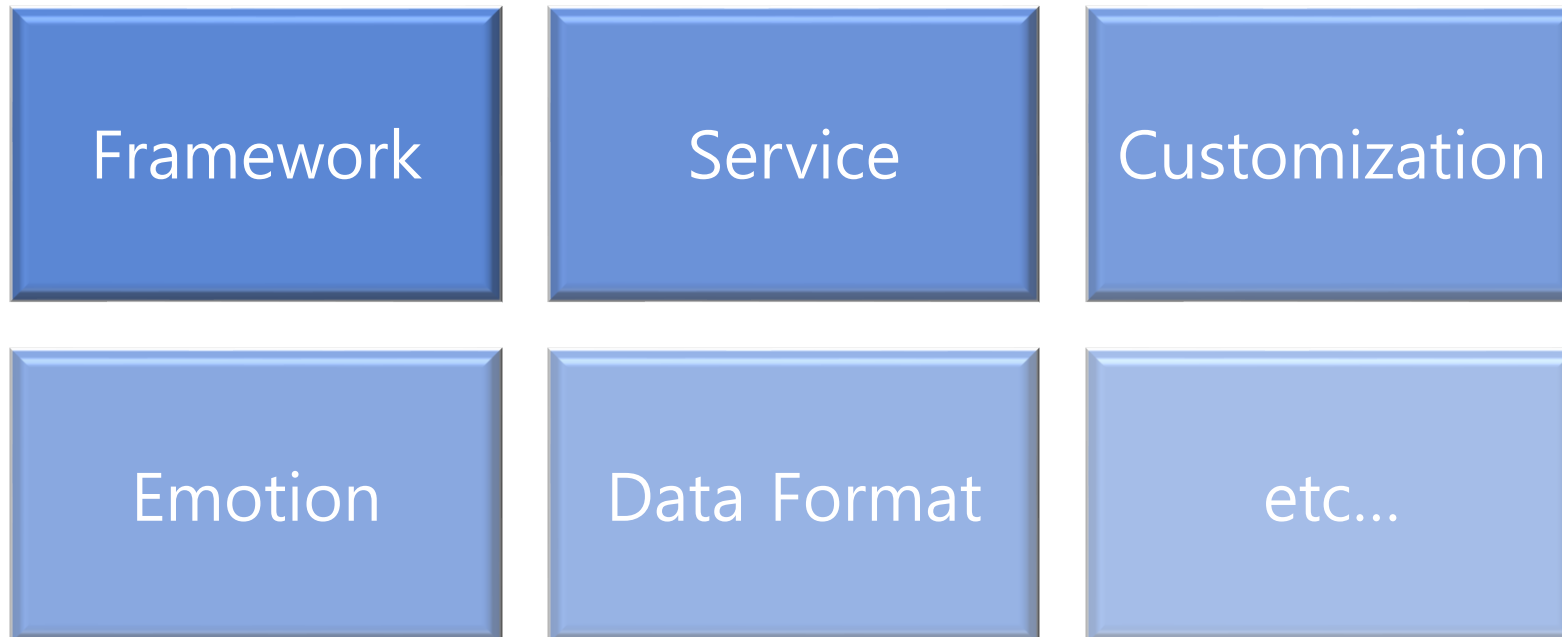


03. Digital Human standardization

디지털 휴먼 표준화 동향

Need for standardization

다양한 주제에서 디지털 휴먼 표준화가 필요합니다.



03. Digital Human standardization

디지털 휴먼 관련한 표준 문서들을 소개합니다

F.748.14
(03/2022)

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

F.748.14
(03/2022)

SERIES F: NON-TELEPHONE TELECOMMUNICATION
SERVICES

Multimedia services

Requirements and evaluation methods of
non-interactive 2D real-person digital human
application systems

F.748.15
(03/2022)

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

F.748.15
(03/2022)

SERIES F: NON-TELEPHONE TELECOMMUNICATION
SERVICES

Multimedia services

Framework and metrics for digital human
application systems

- ITU-T F.748.14는 이미지, 음성, 움직임, 디스플레이 등의 측면에서 비대화형 2차원(2D) 디지털 휴먼 애플리케이션 시스템에 대한 요구 사항 및 평가 방법을 지정합니다.
- ITU-T F.748.15는 디지털 인간 응용 시스템에 대한 프레임워크를 지정하고 이미지, 음성, 애니메이션, 대화형 처리 및 다중 모드 입력/출력의 차원에 대한 해당 주관적 및 객관적 메트릭을 제안합니다.

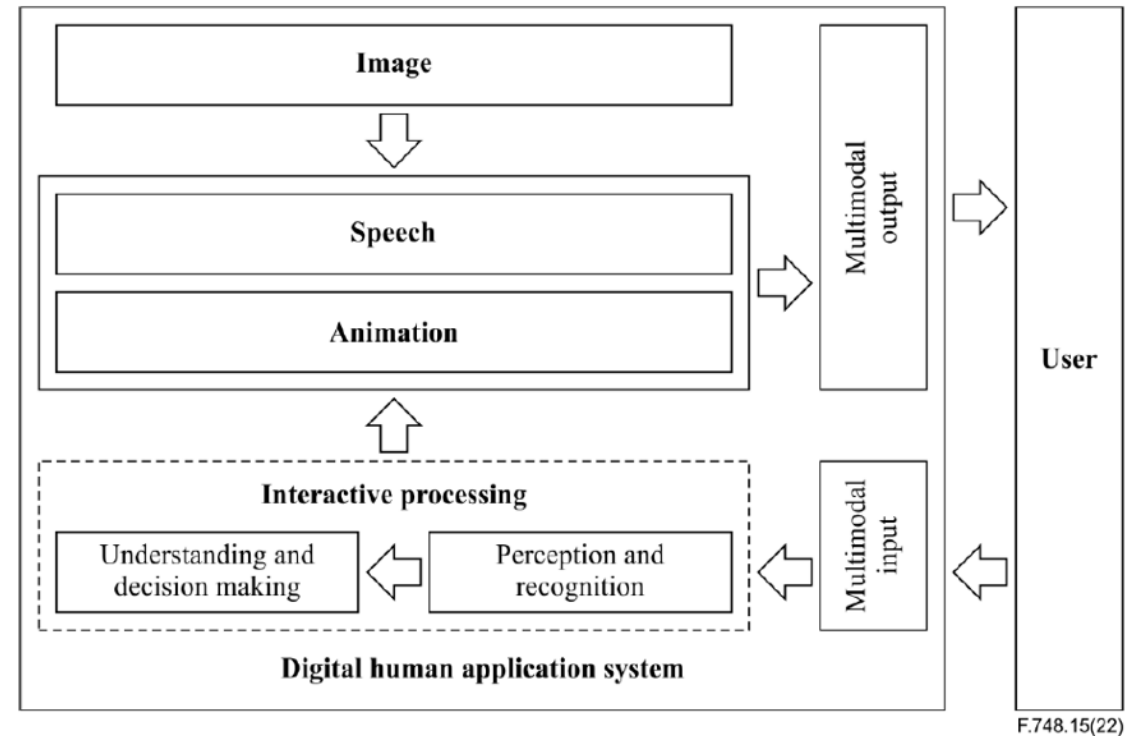
03. Digital Human standardization

디지털 휴먼 관련한 표준 문서들을 소개합니다

F.748.15

지능을 갖고 있는 디지털 휴먼 프레임워크 및 객관적 메트릭을 제안합니다.

1. Multi-modal 데이터를 입력으로 분석 및 대화를 생성
2. Image, Speech, Animation, Interactive Processing, Multi-modal Input, Multi-modal Output Module에 대한 메트릭을 제안



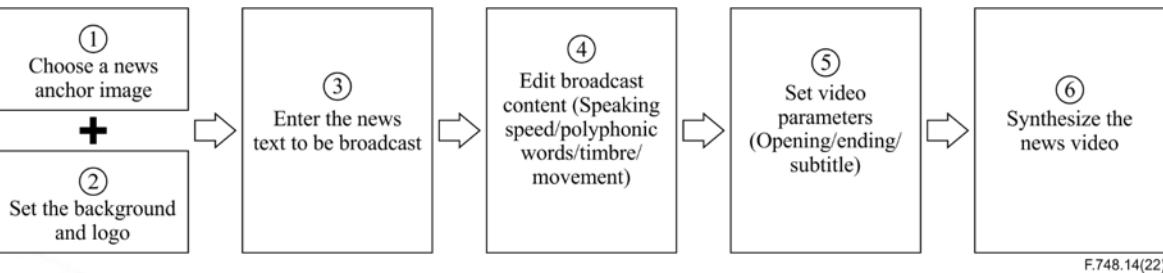
03. Digital Human standardization

디지털 휴먼 관련한 표준 문서들을 소개합니다

F.748.14

비 대화형 2D 디지털 휴먼 애플리케이션 시스템

- 비 대화형 2D 디지털 휴먼을 정의
- F.748.15표준에서 제안한 메트릭을 사용해 성능을 평가하는 방식 제안



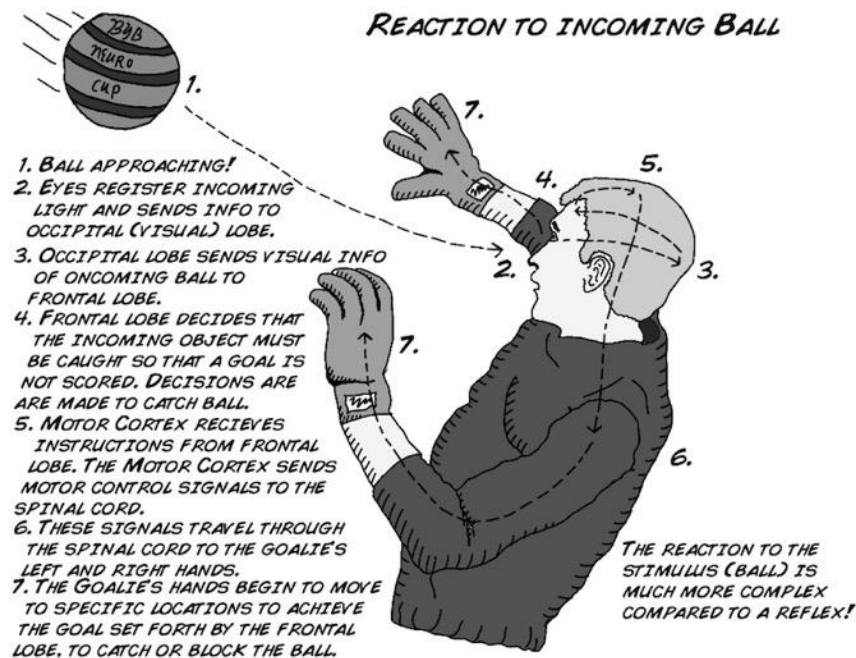
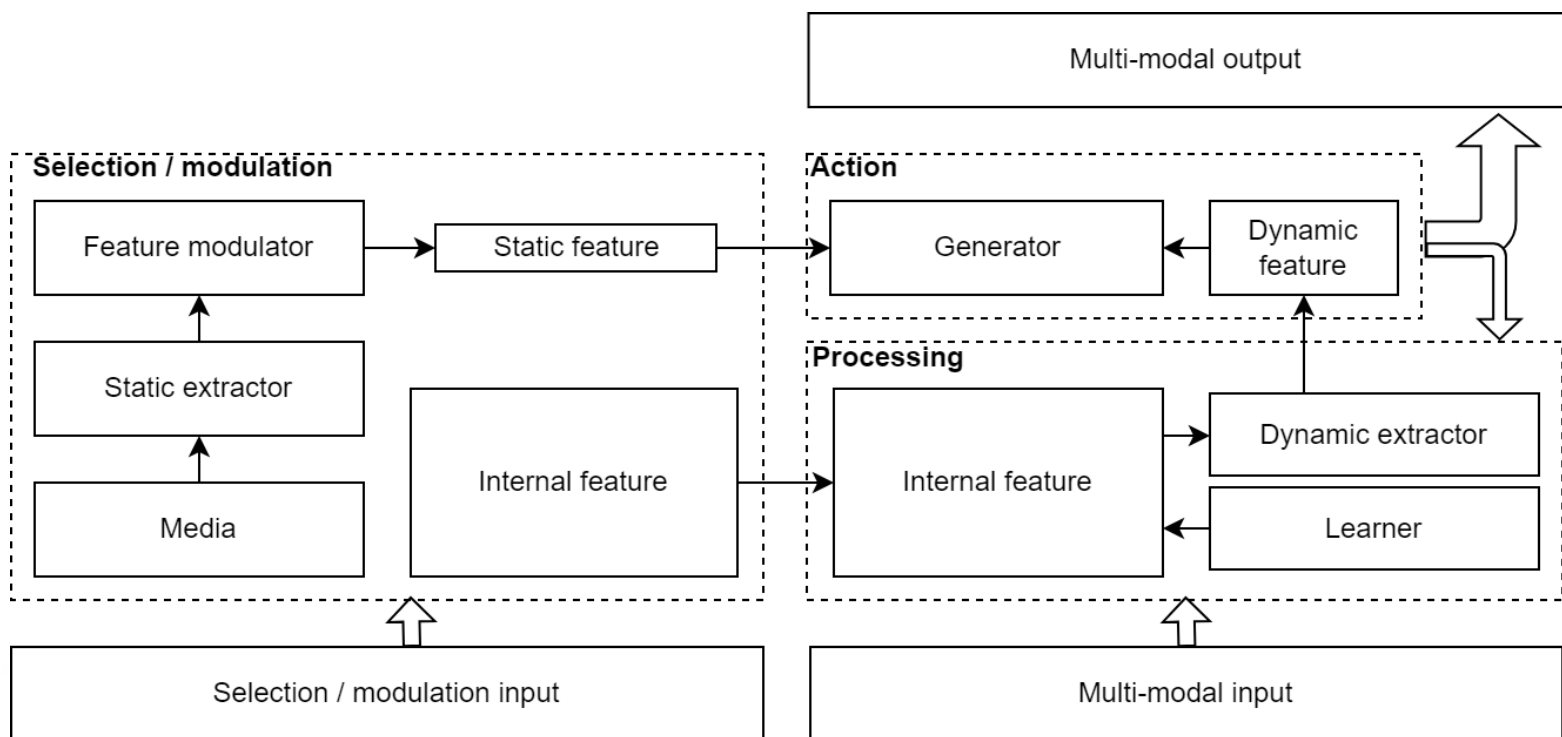
F.748.14(22)

Table 1 – Subjective scoring rules for speech comfort

Evaluation dimension	Description	5	4	3	2	1
Pronunciation and intonation	Is the pronunciation standard as a whole?	Quite standard	Relatively standard	Basically standard	Individually standard	Quite nonstandard
	Is pronunciation clear?	Quite clear	Relatively clear	Basically clear	Not quite clear	Quite unclear
	Is word segmentation and sentence pause proper?	Quite proper	Relatively proper	Basically proper	Not quite proper	Quite improper
	Is tone and intonation natural?	Quite natural	Relatively natural	Basically natural	Not quite natural	Quite unnatural
	Is accent and pronunciation proper?	Quite proper	Relatively proper	Basically proper	Not quite proper	Quite improper
	Is speech speed expression proper?	Quite proper	Relatively proper	Basically proper	Not quite proper	Quite improper
Fluency and coherence	Is speech expression fluent?	Quite proper	Relatively proper	Basically proper	Not quite proper	Quite improper
Emotional fullness	According to the semantics and content of the text, is the emotional expression proper?	Quite proper	Relatively proper	Basically proper	Not quite proper	Quite improper
Anthropomorphic comfort	Is voice as anthropomorphic as a real human?	Completely indistinguishable	Relatively similar and slightly different from real voice	Basically similar	Not quite the same	Quite different
	When listening to the sound, do you feel happy?	Quite happy	Relatively happy	Just so so	Not quite happy	Quite unhappy
	Are you willing to have this voice serve you?	Quite willing	Relatively willing	Just so so	Not quite willing	Quite unwilling

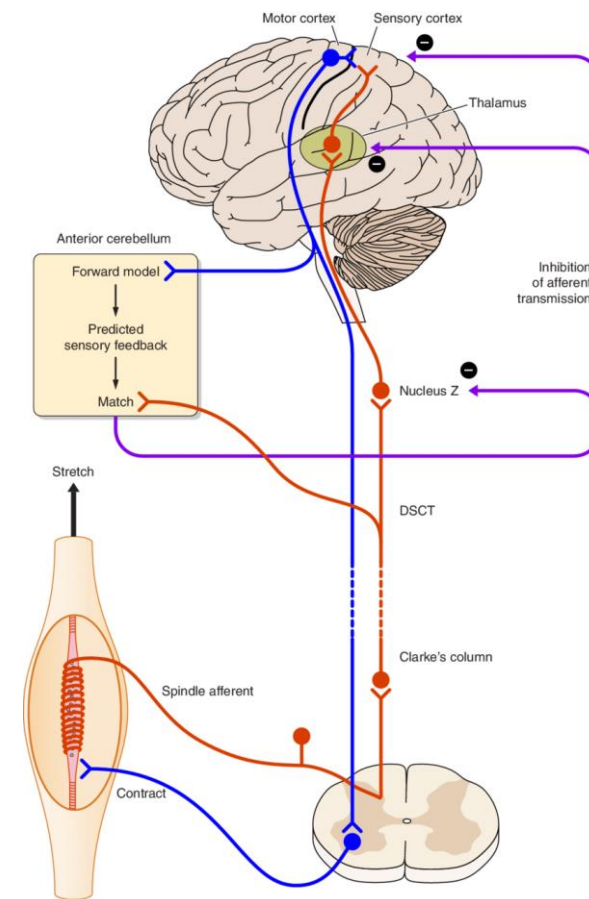
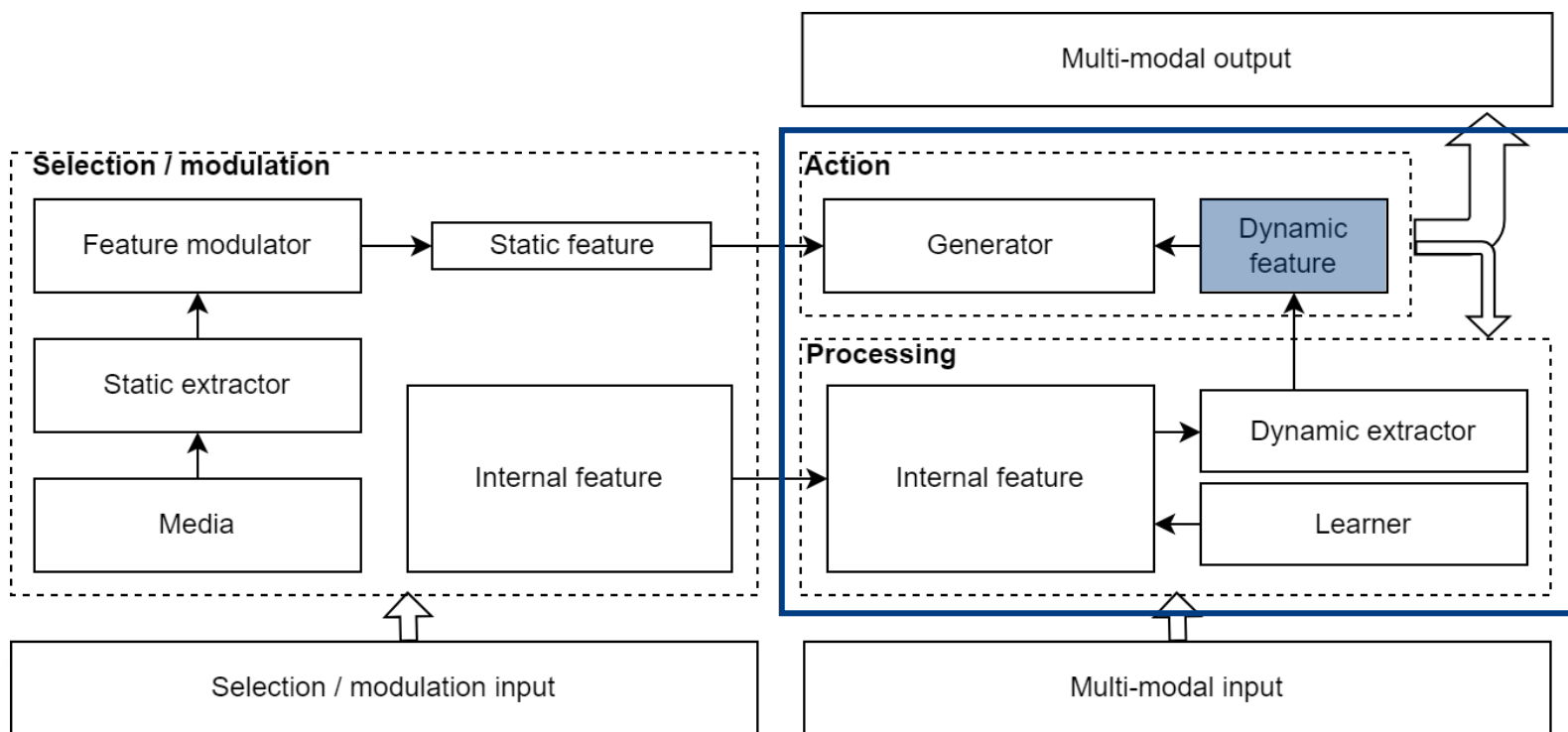
04. Klleon's Standardization

디지털 휴먼 구현에는 수많은 시기술들이 활용됩니다



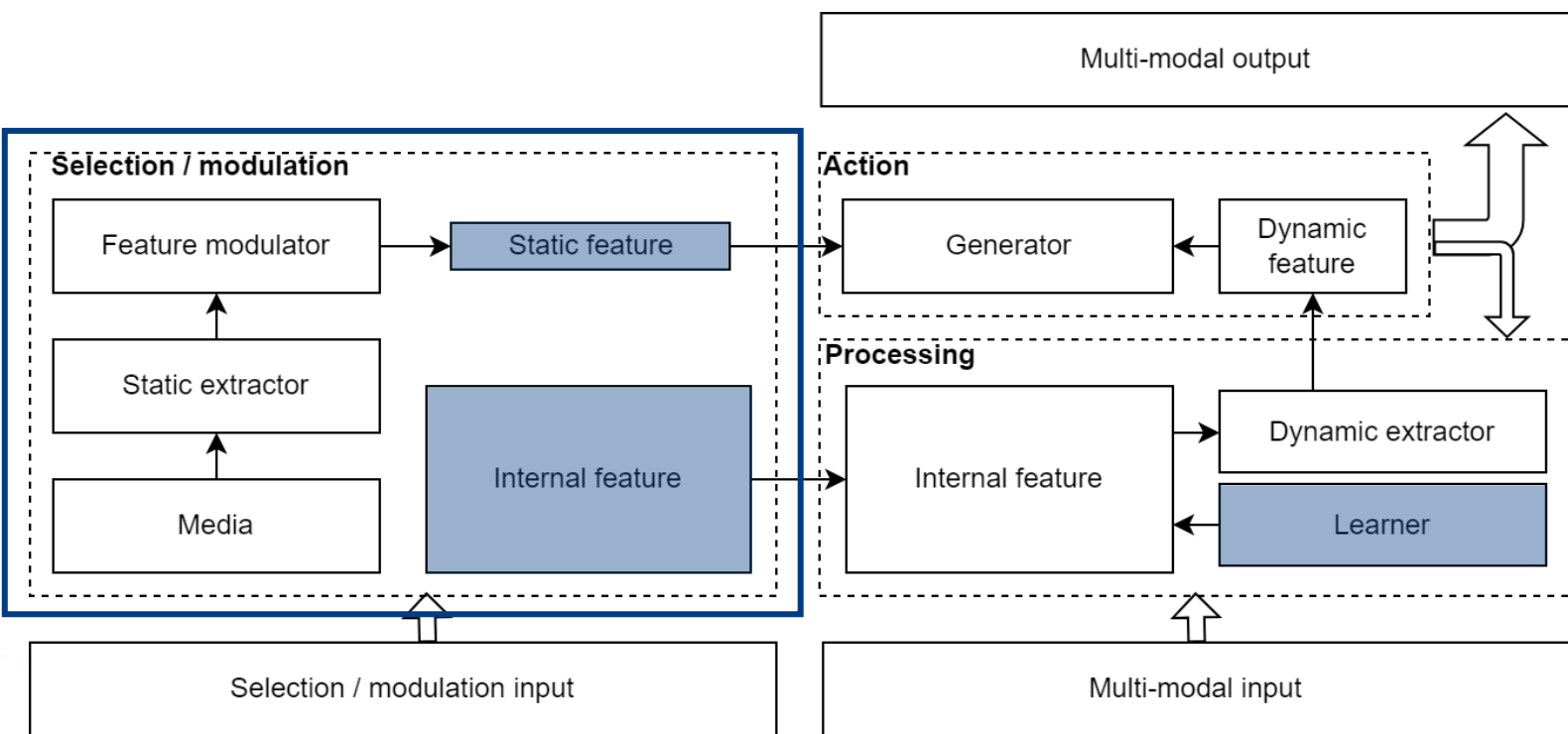
04. Klleon's Standardization

행동 예측과 움직임 구현



04. Klleon's Standardization

개인화된 디지털 휴먼



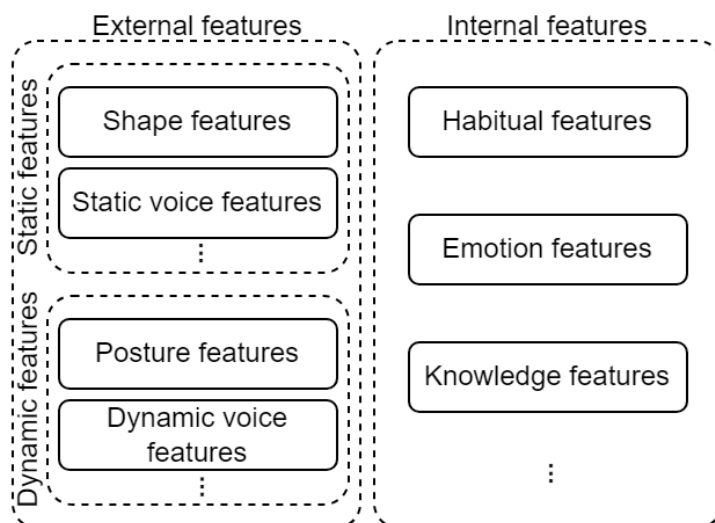
04. Klleon's Standardization

클레온의 표준 문서를 소개합니다

F.FDHC: Framework for digital human customization

원하는 디지털 휴먼을 만들기 위해 디지털 휴먼을 커스터마이징할 수 있는 프레임워크를 제안합니다. 이 과정에서 디지털 휴먼이 가지는 다양한 특성을 정의합니다

1. External, static features
2. External, dynamic features
3. Internal features



INTERNATIONAL TELECOMMUNICATION UNION
**TELECOMMUNICATION
STANDARDIZATION SECTOR**
STUDY PERIOD 2022-2024

SG16-TD71-R1/WP2

STUDY GROUP 16

Original: English

Question(s): 24/16

Geneva, 10-21 July 2023

TD

Source: Editor F.FDHC

Title: F.FDHC "Framework for digital human customization" (New): Output draft (Geneva, 10-21 July 2023)

Contact: Jisu Kang
Klleon
Republic of Korea

Tel: + 82-10-2724-5144
E-mail: jisu.kang@klleon.io

Contact: Seunggeun Baek
Klleon
Republic of Korea

Tel: +82-10-7550-4329
E-mail: seunggeun.baek@klleon.io

Contact: Miran Choi
ETRI
Republic of Korea

Tel: + 82-42-860-6197
E-mail: miranc@etri.re.kr

Abstract: This TD contains the output draft of new Recommendation ITU-T F.FDHC "Framework for digital human customization", based on the contribution C277, RGM-Q24/16-DOC17 (2023-05) and discussion of the Q24/16 experts at the SG16 meeting at Geneva, 10-21 July 2023.

04. Klleon's Standardization

클레온의 표준 문서를 소개합니다

개인화 및 보안

개인화된 디지털 휴먼이 활용되기 시작하면, 사용자의 행동에 대한 데이터를 활용해 디지털 휴먼이 지속적으로 발전할 것이며, 사용자의 데이터에 대한 보안 이슈가 더욱 중요해질 것입니다.

1. 디지털 휴먼의 정보 공개 범위
2. 사용자 정보에 대한 보안



감사합니다.

CTO 강지수, Klleon
Jisu.kang@klleon.io