

## Lecture - 1

### Medium vs modality :

Medium = Is a technical means to store transfer and convey information. Physical channel = air/lights etc.

Modality = is a human sensory channel. For Ex : Visual Channel, Auditory channel, Olfactorial Channel ( smell ) , Haptic ( touch ) , Gustatory ( taste ) , Vestibular ( Balance ) .

Modality : Is all about the capability of the system to mimic such human sensory channels.

Sense and organs are defined for humans.

Modality and sensors are defined for Systems.

Nose, tongue and balance are less used.

### Multimedia vs. Multimodality :

Multimedia - Multiple media. It is a system with more than 1 way to transfer information from system to user.

Multimodality - A system that can stimulate more than one human sensory channel.

Ex :

1) TV - has a screen ( has visual modality ) and loudspeaker ( has auditory modality ) - both multimedia and multimodality. So TV is a multimedia.

NOTE : We cannot call a system both as multimedia and multimodality.

2) Movie hall, smartphone, PC = can be both ( but not possible in real world to be both )

### Alternate Definition :

Multimodal system = A system that provides more than one user input mode to interact with it.

Ex :

1) Smartphone is a multimodal system - since it has a touchscreen ( for touch input ) and microphone ( for speech input ).

2) Robot = can be a multimodal system , can be operated with gesture, speech etc.

### 2nd modified definition :

A multimedia system is an interactive system that provides information via several output ( output can be sound , graphics etc. ) channels.

A multimodal system process two or more combined user input modes ( such as speech, touch, gesture,

head, body ) and the output gives that many no. of modes which means that : If Speech + touch at input side, then output side also contains speech + touch.

Sharon Oviatt = 2012 gave this definition.

" for a truly multimodal system , information we provide at input must be related and coordinated to give the O/P "

" Multimedia system focussed on output of system "

3rd definition : ( Given by researcher Benoid )

A multimodal system represent and manipulate information from different human communication channels. These systems automatically extracts meaning from multimodal raw data and gives desired information .

Any system that uses multiple human sensory channel as an input, we call it multimodal.

TV = multimedia

PC = multimodal

Smartphone = multimodal ( since input has more than one human sensory channels )

"Put that there" - voice + gesture ( 1979 at MIT )

Domains of multimodal system :

- 1) Active I/P mode
- 2) Passive I/P mode
- 3) Virtual Environment