

SCUOLA DI SCIENZE  
Corso di Laurea in Informatica

**ViVi**  
**Progettazione e realizzazione di un voice  
powered retail assistant nella realtà virtuale**

**Relatore:**  
**Chiar.mo Prof.**  
**Marco Roccetti**

**Presentata da:**  
**Jari Tarabelli**

**Co-Relatore:**  
**Prof.**  
**Gustavo Marfia**

**III Sessione**  
**Anno Accademico 2018/19**

*Ad Agnese e a chiunque  
mi abbia sopportato...*

## Introduzione

La realtà virtuale, negli ultimi anni, ha riscosso un crescente successo dovuto principalmente ad un miglioramento implementativo della tecnologia, che ne ha permesso la diminuzione dei costi, e all'aumentare della disponibilità di applicazioni in qualsiasi campo, dall'intrattenimento all'istruzione. Per primi i nuovi pionieri della realtà virtuale da desktop come Oculus e Valve, in seguito anche iOS ed Android con i relativi ARKit e ARCore più incentrati verso la realtà aumentata tramite mobile. La strada che tutti stanno cercando di percorrere è chiara, e non è difficile immaginare come la naturale evoluzione di questa tecnologia porti sempre più verso l'e-commerce virtuale.

Nello specifico, le applicazioni di shopping immersive sono ancora nello stato sperimentale, e i grandi nomi del settore stanno cercando soluzioni innovative per prepararsi all'ingresso della VR nel mercato mainstream. L'avvento della VR potrebbe costituire una rivoluzione per lo shopping, come lo è stato l'e-commerce 10 anni fa.<sup>[1]</sup>

Durante lo sviluppo di un'esperienza retail (sia essa fisica o virtuale) ci si trova di fronte a una serie di sfide e scelte strutturali dal punto di vista UX e della retail experience per far sì che l'utente si trovi piacevolmente a suo agio. Tali scelte impattano significativamente sul prodotto finale, e con esso l'immagine del brand a cui viene associato. Le limitazioni che si trovano nello sviluppo di un'interfaccia per la VR iniziano ad essere motivo di omologazione della forma, fattore che gioca pesantemente a sfavore di quei marchi che vogliono avere un look and feel unico e caratteristico.

Un'altra tecnologia che ha raggiunto una discreta utenza e che ha guadagnato una fetta molto ampia di mercato, è lo Smart Assistant. L'impiego di un assistente virtuale, con cui si interagisce attraverso la voce, che riconosce l'utente, esegue comandi e raccoglie informazioni sulla base delle preferenze dell'utente, è entrato appieno nel quotidiano: molte realtà hanno deciso di integrarlo tra i loro servizi per dare più valore al proprio brand<sup>[2]</sup> e ovviamente alla user experience.

Lo scopo di questa tesi è di cercare di migliorare diversi aspetti dell'esperienza (tra cui: semplicità d'uso e di reperimento delle informazioni), applicando un approccio che mira ad unire le due tecnologie, sfruttando le potenzialità date dall'esperienza d'uso dell'assistente (quali semplicità d'uso e gestione/reperimento delle informazioni) e quelle visive della VR (espressività visiva, immersione ed engagement). Adottando questa soluzione si può lasciare più libertà creativa in molti aspetti del retailing ed al contempo alleggerendo il contesto visivo, delegando funzioni che normalmente sarebbero a carico dell'interfaccia direttamente all'assistente. Ricorrendo all'uso dell'assistente l'utente verrà agevolato nell'esperienza, come già abituato nei negozi fisici, dove il personale è formato per essere l'addetto a chiarimenti, assistenza e soprattutto suggerimenti.

Inoltre, adottando questo approccio si cercherà, per quanto possibile, di liberare le mani dall'incombente uso dei controllers.

Si prende ispirazione dal precedente lavoro di tesi di Marco Valerio Manzini,<sup>[3]</sup> e dall'articolo di Sonia Di Vaio.<sup>[4]</sup> Negli elaborati risulta che la gamification nell'ambito retail sia apprezzata: in particolare c'è un certo entusiasmo nel cercare di capire come l'esperienza di acquisto all'interno della realtà virtuale possa essere migliorabile, rendendola più divertente e quindi più piacevole per l'utente.

Per semplicità, in fase di progettazione, si prende in oggetto un generico negozio in quanto rappresenta il campione più variegato, in seguito per poter dare una valutazione puntuale in fase sperimentale si specializzerà nel settore moda, che risulta essere uno dei più interessati dalla riuscita di questo sistema.

A differenza dell'elaborato da cui prendiamo spunto,<sup>[3]</sup> l'applicazione si focalizza più sull'acquisto che sull'iterazione tra utente e articolo. Ci aspettiamo che l'esperienza generale sia percepita positivamente grazie all'approccio adottato, raccoglieremo e analizzeremo questi dati in fase sperimentale.

Questo elaborato non ambisce a dimostrare un risultato rigoroso, ma si prefigge di aprire la strada a nuovi sviluppi e ricerche.

La struttura della tesi prevede nelle prime due parti una disquisizione su ciascuna tecnologia. Per entrambi i capitoli si definirà lo stato dell'arte, una panoramica delle alternative presenti in commercio, dei settori in cui tali tecnologie operano, ed inoltre verrà fornita una rapida spiegazione dei principi teorici.

Nella terza parte della tesi vengono descritti i punti salienti dello sviluppo e progettazione generale dell'applicazione. Nell'ultima parte descriveremo le metodologie di test ed analizzeremo i risultati ottenuti tramite sperimentazione su un piccolo campione di studenti del Master in Design and Technology for Fashion Communication dell'Università di Bologna.<sup>[5]</sup>

# Indice

<b>I</b>	<b>Virtual Reality</b>	<b>3</b>
<b>1</b>	<b>Stato dell'arte</b>	<b>5</b>
1.1	Visore . . . . .	6
1.2	Sensori di Tracking . . . . .	8
1.2.1	Lighthouse Tracking . . . . .	9
1.2.2	Oculus Insight . . . . .	11
1.3	Controllers . . . . .	13
<b>2</b>	<b>Panoramica dei dispositivi</b>	<b>14</b>
2.1	Oculus Rift S . . . . .	16
2.2	Oculus Quest . . . . .	18
2.3	HCT Vive . . . . .	20
2.4	HTC Cosmos . . . . .	22
2.5	Valve Index . . . . .	24
2.6	Sony PSVR . . . . .	26
<b>3</b>	<b>Campi d'uso</b>	<b>27</b>
3.1	Intrattenimento . . . . .	27
3.2	Formazione e Simulazione . . . . .	28
3.2.1	Gamification e Serious Games . . . . .	29
3.3	Retail . . . . .	29
<b>4</b>	<b>Concetti teorici</b>	<b>31</b>
4.1	Immersione . . . . .	31
4.2	Presenza e illusione . . . . .	32
4.3	Percezione dell'utente . . . . .	33
4.3.1	Percezione dello spazio . . . . .	33
4.3.2	Percezione del tempo . . . . .	34
4.3.3	Percezione del movimento . . . . .	34
4.4	Effetti nocivi della VR . . . . .	34
4.4.1	Motion Sickness . . . . .	35
4.4.2	Effetti fisici e igiene . . . . .	36
4.5	Linee guida alla progettazione in VR . . . . .	37

<b>II</b>	<b>Virtual Assistant</b>	<b>39</b>
<b>5</b>	<b>Stato dell'arte e campi d'uso</b>	<b>41</b>
5.1	Voice Commands vs Voice Assistant . . . . .	44
5.2	VR + AR . . . . .	45
<b>6</b>	<b>Panoramica delle alternative</b>	<b>46</b>
6.1	Alexa . . . . .	47
6.2	Google Assistant . . . . .	48
6.3	Siri . . . . .	49
6.4	Mycroft AI . . . . .	50
6.5	Snips . . . . .	51
<b>7</b>	<b>Concetti teorici</b>	<b>52</b>
7.1	Anatomia di un comando vocale . . . . .	52
7.2	Processo di riconoscimento vocale . . . . .	54
7.2.1	NLU . . . . .	55
7.3	Percezione dell'utente . . . . .	55
7.3.1	Fattori positivi . . . . .	56
7.3.2	Privacy e Sicurezza . . . . .	56
<b>III</b>	<b>Realizzazione di ViVi</b>	<b>59</b>
<b>8</b>	<b>Analisi delle problematiche</b>	<b>61</b>
8.1	Sfide strutturali . . . . .	61
8.2	Carenza di informazioni . . . . .	62
<b>9</b>	<b>Approccio e soluzioni</b>	<b>64</b>
9.1	Panoramica . . . . .	65
9.2	L'assistente . . . . .	66
9.2.1	Interazione vocale . . . . .	67
9.3	Business Logic . . . . .	68
9.4	Limitazioni . . . . .	70
<b>10</b>	<b>Tecnologie impiegate</b>	<b>72</b>
10.1	Visore . . . . .	72
10.2	Voice Assistant . . . . .	72
10.3	Setup in laboratorio (VarLab) . . . . .	73
10.4	Ambiente di sviluppo 3D . . . . .	74
10.4.1	Assets e plugins . . . . .	74
10.4.2	Suoni e musica . . . . .	74
10.5	Ambiente di sviluppo Backend . . . . .	75
<b>IV</b>	<b>Risultati</b>	<b>77</b>
<b>11</b>	<b>Somministrazione del test</b>	<b>79</b>
11.1	Svolgimento . . . . .	79

11.2	Valutazioni e metriche . . . . .	80
11.3	Contesto moda . . . . .	83
<b>12</b>	<b>Risultati ottenuti</b>	<b>85</b>
12.1	Valutazioni qualitative . . . . .	85
12.2	Domande aperte . . . . .	88
<b>13</b>	<b>Future migliorie</b>	<b>90</b>
<b>14</b>	<b>Conclusione</b>	<b>93</b>

# Elenco delle figure

1.1	Schermata dalla procedura di setup iniziale di SteamVR . . . . .	9
1.2	Schematizzazione del tracking di Oculus Insight . . . . .	11
5.1	Infografica riassuntiva delle principali caratteristiche riguardanti le varie tipologie di assistente . . . . .	44
6.1	Il modello di interazione di Alexa . . . . .	47
7.1	Un esempio di comando vocale . . . . .	52
7.2	Un esempio di modello per voice recognition . . . . .	55
9.1	Schermata inattiva che ritrae ViVi . . . . .	66
12.1	Valutazione dell'interfaccia . . . . .	86
12.2	Domande sull'usabilità . . . . .	87
12.3	Apprezzamento generale . . . . .	88
12.4	Valutazioni specifiche delle esperienze . . . . .	89



# Elenco delle tabelle

11.1	Domande di valutazione dell'interfaccia . . . . .	81
11.2	Domande sull'apprezzamento generale del retail VR . . . . .	81
11.3	Domande sulle valutazioni specifiche delle esperienze . . . . .	81
11.4	Domande sull'usabilità . . . . .	82



Parte I  
Virtual Reality



# Capitolo 1

## Stato dell'arte

La realtà virtuale, per sua stessa definizione, simula la realtà effettiva permettendo di navigare in ambientazioni fotorealistiche in tempo reale ed interagendo con gli oggetti presenti in esse. Attualmente il termine è applicato a qualsiasi tipo di simulazione virtuale creata attraverso l'uso di un computer e un visore, anche se, a livello teorico, con il termine si intende un generico sistema che coinvolga e permetta l'utilizzo di tutti i sensi. In questa trattazione si fa riferimento esclusivamente al livello tecnologico presente nell'attuale mercato consumer degli ultimi anni, in quanto le sue forme alternative e precedenti, seppur interessanti, sono poco significative per quello che si andrà a dimostrare. Nello specifico dal 2010 si è assistito ad un rilancio del concetto di realtà virtuale immersiva come un ambiente costruito intorno all'utente con cui poter interagire tramite l'uso di specifiche periferiche, descritte nelle sezioni seguenti.

Spesso si citerà Oculus, a cui va una menzione particolare, in quanto è stata l'azienda (allora start-up, e successivamente acquisita da Facebook) che ha riaperto l'interesse verso questo settore, facendo da pioniere anche dal punto di vista della ricerca, con la sua prima implementazione del Rift.

La realtà virtuale si basa principalmente sui concetti di immersione e presenza.

**Immersione:** il termine è inteso nel senso metaforico di "immergersi" mentalmente all'interno della rappresentazione (finzione o simulazione). Può essere anche definito come uno stato in cui la consapevolezza dell'io fisico del visitatore è portata a pensare di essere all'interno di un ambiente artificiale. Descrive anche la sensazione di parziale o completa suggestione nel rispondere e reagire a stimoli artificiali in un ambiente virtuale. Aumentando la suggestione, cresce il grado di presenza percepito. L'immersione può essere divisa in quattro categorie: sensoriale-motoria, cognitiva, emozionale e spaziale.<sup>[6]</sup> Quest'ultima riguarda le applicazioni in cui l'utente prova la sensazione di trovarsi "lì" e il cui mondo simulato dà la percezione di essere realistico.

**Presenza:** termine derivante da "telepresenza", è un fenomeno che consente agli utenti di interagire e sentirsi connessi col mondo al di fuori della loro fisicità grazie all'uso della tecnologia. Definisce inoltre la sensazione soggettiva di sentirsi all'interno di una scena rappresentata solitamente nella realtà virtuale.<sup>[7]</sup>

E naturale pensare che la presenza sia alimentata principalmente dalla tecnologia, tuttavia, i fattori psicologici coinvolti nell'ottenere la percezione di presenza sono molti e di importanza rilevante: la qualità dell'illusione è principalmente data dalla percezione soggettiva dell'utente, che viene semplicemente manipolata dalla tecnologia.<sup>[8]</sup>

Di seguito viene descritto ciò che compone l'insieme di tecnologie che rende possibile il raggiungimento di tale obiettivo.

## 1.1 Visore

Il Visore o (Head Mounted Display - HMD) è un casco o dei semplici occhiali comprendenti due schermi, posti vicini agli occhi, che mirano ad annullare il mondo reale dalla visuale dell'utente. Il visore può inoltre contenere dei sistemi per la rilevazione dei movimenti, in modo che girando la testa da un lato, ad esempio, si ottenga la stessa azione anche nell'ambiente virtuale. Se non vi sono dei controllers associati al dispositivo, allora questo è l'unico tipo di comando concesso e l'esperienza sarà concepita al solo scopo di far immergere lo spettatore in maniera passiva.

Il visore è la minima componente necessaria a raggiungere l'obiettivo di simulare la presenza all'interno di un mondo virtuale. Secondo Oculus, è necessario che tale sensazione venga percepita a livello viscerale tramite una bassa latenza e movimenti tracciati in maniera precisa.<sup>[8]</sup> In seguito, durante una conferenza indetta dal team di ricerca sulla VR di Valve (nota azienda che opera nel settore videoludico), Michael Abrash ha ampliato questi requisiti indicando come necessari:<sup>[9]</sup>

- un ampio campo visivo (ovvero il grado di visibilità permessa all'utente, a cui riferiremo come FOV - Field Of View)
- una risoluzione adeguata (superiore a 1080p)
- un tempo di risposta dei pixel rapido (minore di 3ms GTG - Gray To Gray, ovvero una metrica che indica il tempo necessario ad un pixel per passare da grigio a bianco a grigio)
- un refresh-rate molto elevato (almeno 60Hz, ovvero lo schermo deve riuscire a visualizzare una nuova immagine 60 volte in un secondo)
- un display con tutti i pixel illuminati simultaneamente
- lenti speciali con apposita calibrazione (distanza interpupillare, gradazioni, profondità, etc...)
- un sistema di tracciamento con precisione al millimetro o migliore per la traslazione, al quarto di grado per la rotazione e con volume di 1.5 metri o più
- latenza di rilevazione dei movimenti più bassa possibile (20 ms come ottimale)

I primi headset sono stati prodotti negli anni 90 per poi quasi svanire. Dal 2010 la loro popolarità è cresciuta grazie allo sviluppo di diversi sistemi dedicati alla fascia enthusiast (ovvero quella fascia di clientela attenta alle innovazioni e disposta a spendere pur di avere l'ultimo ritrovato di un prodotto giudicato interessante) e resi accessibili dall'aumento delle capacità di calcolo. Fino a poco tempo fa i dispositivi di realtà virtuale in commercio non erano molti, ma nel corso di pochi anni è stata ampliata la scelta grazie all'ingresso nel mercato di nuove aziende già forti nel settore tecnologico: HTC, con Vive<sup>[10]</sup> (2016), Vive Pro<sup>[11]</sup> (2018) e più recentemente Vive Cosmos<sup>[12]</sup> (2019), Steam col suo Index<sup>[13]</sup> (2019) e Facebook/Oculus con Rift S<sup>[14]</sup> e Quest<sup>[15]</sup> (2019). Al di fuori dei sopracitati big players iniziano a spuntare anche le prime alternative che si rivolgono ad un pubblico più di nicchia come Pimax 8K/5K VR<sup>[16]</sup> (2018) o Sony Playstation VR<sup>[17]</sup> (2016).

La tecnologia attuale si avvale di un computer per l'elaborazione dell'ambiente virtuale da ricreare e un visore VR ad esso collegato per visualizzare tale ambiente. Tale soluzione permette ad oggi di ottenere il massimo delle performance e del realismo, in quanto un pc è in grado di compiere molti calcoli complessi in frazioni di secondo. Ciò che l'utente vede attraverso il visore non è altro che un render in real time del mondo virtuale. La VR Desktop ha il vantaggio di garantire performance elevate, ma ha una serie di svantaggi che possono renderla impraticabile, tra cui la necessità di un PC abbastanza potente e molto spesso ingombrante. Per questo motivo esistono in commercio appositi PC denominati "VR Ready", che rispondono a tutti i requisiti tecnici necessari a far funzionare il visore.<sup>[18]</sup> La maggior parte di questi device hanno come obiettivo la portabilità: parliamo appunto di laptop o unità PC a forma di zainetto (con batterie), da mettere in spalla e collegare al visore. Questo sforzo collettivo dell'industria mette in luce l'importanza della portabilità tra le problematiche che si incontrano nella fruizione dell'esperienza VR.

Dall'altro lato dello spettro, un supporto VR la cui componente video viene elaborata e visualizzata dallo smartphone (Google Cardboard,<sup>[19]</sup> Samsung Gear VR, etc...) costituisce un'alternativa piuttosto economica per introdurre l'utente alla realtà virtuale. In tali soluzioni due apposite lenti dividono lo schermo a metà per simulare un FOV e avere l'illusione di indossare un vero headset. Tale supporto ha lo scopo di sovrapporre diversi modelli di smartphone, inoltre è compatibile con la realtà aumentata, cosa che non sempre è vera per un HMD.

A questo punto è importante comprendere la differenza tra i vari gradi di libertà concessi dai diverse tipologie di visore. I cosiddetti DoF, o Degrees of Freedom, si riferiscono al numero di assi su cui un oggetto solido può spostarsi all'interno di uno spazio tridimensionale. In totale sono sei: tre riferiti al movimento rotatorio sugli assi x, y, e z (ovvero quelli a cui di solito ci riferiamo come inclinazione, rotolio e imbardata) e tre riferiti al movimento traslatorio lungo suddetti assi (quindi muoversi avanti/indietro, lateralmente e su/giù).<sup>[20]</sup>



Con 3DoF possiamo girare la testa e vedere la scena a 360 gradi,



mentre nella 6DoF possiamo anche spostarci all'interno di essa

I visori per smartphone rientrano nella categoria 3DoF e quindi risultano estremamente limitanti nell'esperienza d'immersione. Questo evidente limite tecnico è dato dalla mancanza della necessaria sensoristica, troppo costosa sia in termini di performance che di praticità del sistema. Per questi motivi è preferibile utilizzare un HMD quando l'interazione con l'ambiente è necessaria.

## 1.2 Sensori di Tracking

Si tratta di sistemi che hanno il compito di tracciare i movimenti nello spazio, essi disegnano un'area all'interno della quale ci si può spostare e captano dei segnali dal visore e dai controllers per determinarne la posizione. La mole di dati e la complessità delle operazioni effettuate su di essi rendono quest'aspetto il più complesso da gestire quando si tratta di implementare un sistema di realtà virtuale.

Come vedremo, nei nuovi modelli attualmente in produzione si sta spostando la responsabilità del tracking direttamente al visore, consentendo più libertà e semplicità d'uso. Addirittura alcuni produttori iniziano a predisporre la possibilità di tracciare particolari movimenti come ad esempio il movimento delle mani o degli occhi.



Per semplicità di seguito verranno descritte le due principali categorie disponibili in commercio, anche in previsione del capitolo riguardante i modelli in circolazione, in quanto scopriremo che i sistemi di tracciamento in uso dai vari visori sono gli stessi (con qualche ininfluente differenza di implementazione).

## 1.2.1 Lighthouse Tracking

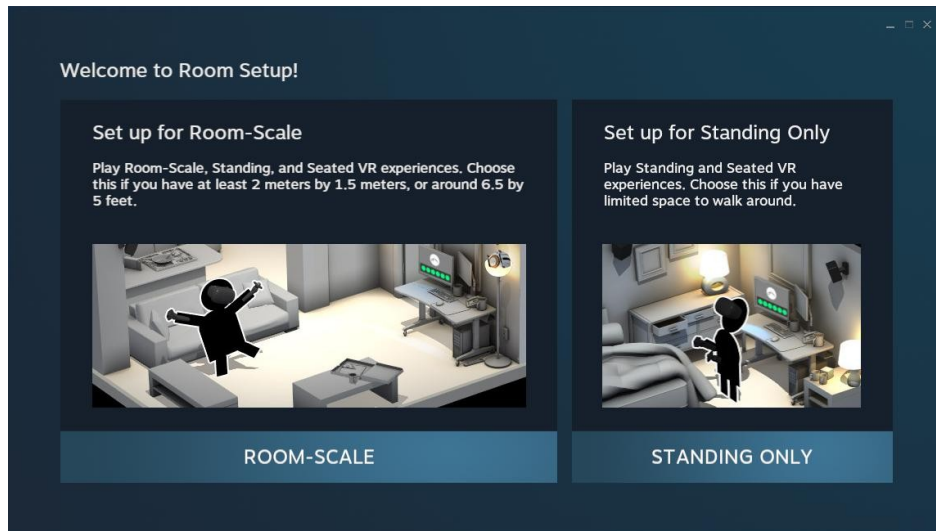


Figura 1.1: Schermata dalla procedura di setup iniziale di SteamVR

Questo sistema di tracciamento (più formalmente Steam VR Tracking<sup>[21]</sup>) è stato annunciato da Valve per il visore sviluppato in collaborazione con HTC nel 2015. Differentemente da quanto proposto già da diversi anni da Oculus, grazie a questo innovativo sistema era possibile alzarsi e spostarsi all'interno di una stanza fisica, piuttosto che rimanere seduti o potersi solo ruotare interamente col corpo.

Da qui una prima importante distinzione viene evidenziata tramite l'uso dei termini Room Scaled VR e Seated VR (o Standing Only VR), dove la prima indica un'esperienza sviluppata attorno alla libertà di movimento possibile all'interno di un'area delimitata dai sensori, mentre l'altra indica che l'esperienza si svolgerà stabilmente in un solo punto fisico senza possibilità di muoversi, in ogni caso entrambe lasciano la possibilità di interagire con l'ambiente "raggiungibile".

Le applicazioni sviluppate attorno al primo paradigma hanno il vantaggio di risultare altamente immersive e di garantire all'utente la massima flessibilità nell'ambiente. Essa è particolarmente indicata per contesti altamente dinamici o dove viene richiesta la possibilità di valutare diversi punti di vista (come nel caso dei videogames o dell'architettura). Nonostante questo risulti preferibile se non ottimo, abbiamo dei compromessi in termini di costo e ovviamente di spazio. Difatti prendendo in esempio l'implementazione più recente (SteamVR Base Station 2.0) possiamo arrivare a coprire 10m<sup>2</sup> di superficie calpestabile, ma va presupposta la disponibilità, non sempre ovvia, di quell'area. Il requisito minimo per usufruire della Room Scale VR è di 2m per 1.5m di spazio libero da ingombri, che sembrano pochi

ma in realtà potrebbero essere difficili da ottenere in quanto è richiesto che i sensori vengano posizionati in alto uno di fronte all'altro e piegati verso il basso di circa 30°, questo significherebbe forare i muri o predisporre delle aste telescopiche e dover attraversare svariati metri di cavo per la stanza. Tutto questo è spesso impossibile da attuare per la maggior parte della clientela, rimanendo in ogni caso molto poco pratico.

In maniera del tutto contrastuale la Seated VR invece cerca di evitare la maggior parte di questi inconvenienti facendo rimanere seduto (o in piedi) l'utente in una posizione stazionaria richiedendo molto meno spazio e un minor numero di sensoristica. Permettere di posizionare i sensori sopra una scrivania e non lasciare che l'utente inciampi in alcun cavo (sia esso del visore o dei sensori) renderebbe l'esperienza più gradevole secondo l'ormai vecchia visione di Oculus, in contrasto Vive, seppur rendendolo possibile, sconsiglia questo approccio, affermando che la vera espressività della VR si ottenga in Room Scaled, in quanto la possibilità di muoversi liberamente dà un grande senso di immersione e presenza. La Seated VR rimane comunque consigliata se si dovesse applicare ad un uso più statico come una simulazione di guida o di un tavolo operatorio.

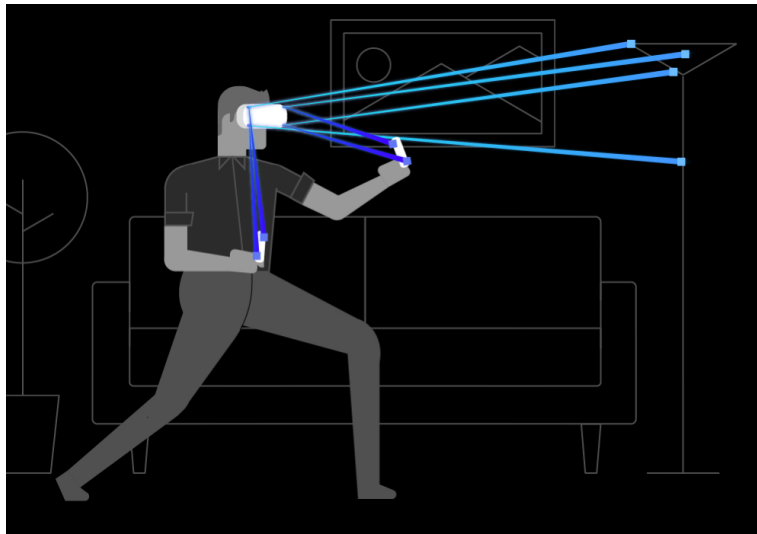
Appresa questa distinzione, torniamo alla composizione del sistema.

I componenti più importanti sono sicuramente le base stations, sensori rettangolari che si occupano di disegnare l'area di tracking partendo dai vertici della sua diagonale e di funzionare da punto di riferimento per le periferiche. Questi sensori emettendo un flusso continuo di raggi infrarossi che vengono captati dai ricettori nei dispositivi da tracciare, così che possano carpire la loro posizione in relazione agli emettitori. Il sistema necessita almeno di due o tre base stations per funzionare in quanto esse sono suscettibili all'occlusione dei raggi, quindi se la visuale tra l'utente e il sensore è ostacolata il dispositivo non verrà tracciato correttamente, per questo si consiglia di posizionare i sensori in posizioni elevate. Aumentare il numero di basi consente di ottenere un tracking più preciso e un'area maggiore.

Entrando nel tecnico, ogni base contiene un emittente infrarossi (chiamata Sync Blinker) che emette un segnale di sincronizzazione 60 volte ogni secondo e due laser rotanti che alternatamente sparano raggi per la stanza. I ricettori dei dispositivi, di forma circolare, sono coperti di materiale fotosensibile che rilevano gli specifici segnali luminosi. Quando riconoscono il segnale di sincronizzazione, i ricettori calcolano il delay (ovvero il ritardo) con cui il ricettore viene colpito dal laser, a questo punto Lighthouse calcola la posizione in relazione alla base che ha lanciato il segnale, approssimandone l'eventuale errore cercando di predire il movimento successiva (per cercare di ridurre il drifting, ovvero un movimento continuo non previsto). Molto importanti in questo processo sono inoltre le IMU (Inertial Measuring Unit), componenti elettroniche all'interno dei dispositivi che inglobano accelerometro, giroscopio e magnetometro, e che come nei moderni smartphone, tracciano velocità, orientamento e forza gravitazionale (nel caso di Lighthouse sono estremamente performanti, ad una frequenza di 1000Hz circa).<sup>[22]</sup> Questo procedimento applicato a tutte le basi del sistema consente di tracciare la posizione in 6DoF dell'utente.

Un ulteriore fondamentale vantaggio di SteamVR è la possibilità di poter implementare il tutto con un costo ampiamente ridotto, in quanto OpenVR, il software su cui è basato, è open source. Per cui qualsiasi qualsiasi produttore o sviluppatore che voglia creare il proprio prodotto per la realtà virtuale può usare a proprio piacimento le funzionalità messe a disposizione dall'SDK (o Software Development Toolkit, una suite di codice che consente la creazione di altro software attraverso l'uso di interfacce programmabili). La potenzialità di Lighthouse non si limita solo al campo degli HDM e relative app, ma grazie alla sua generalità, può essere usato come sistema di tracciamento in altri ambiti, come ad esempio la robotica, come vedremo in seguito.

## 1.2.2 Oculus Insight



**Figura 1.2:** Schematizzazione del tracking di Oculus Insight

Oculus Insight è il nuovo metodo di inside-out tracking della posizione sviluppato da Facebook AI per i nuovi visori Rift S e Quest (di cui parleremo nel Cap.3).

Prima di iniziare a parlare del funzionamento specifico della tecnologia vanno fatte le dovute premesse, in quanto è comune vedere confrontate Lighthouse e inside-out tracking come se non fossero in alcun modo collegate. Questo è frutto di una generalizzazione incorretta data nome del metodo di tracciamento: inside-out tracking è un termine che indica un metodo di tracciamento posizionale in cui la telecamera è attaccata al dispositivo da tracciare, mentre la sua controparte, l'outside-in tracking è invece composto da un set di telecamere statiche nella scena. Lighthouse si basa sul cosiddetto marker-based inside-out tracking, mentre gli HDM come quelli di Oculus usano la markerless inside-out. Nel primo caso si è già discusso dell'uso dei ricettori (da cui marker-based), per il secondo caso è ovvia la loro assenza, ma ciò non cambia che chi ha la responsabilità di acquisizione della posizione sia il dispositivo, per cui è errato dissociare Lighthouse dalla sua categoria.

Ad oggi inside-out tracking viene usato come termine di marketing per evidenziare un'esperienza d'uso svincolata dalla necessità di ingombranti base station. Fatte queste premesse, per chiarezza col resto del medium, ci atterremo comunque all'uso di quest'ultimo, che è il più comune.

Tornando nello specifico, Insight fa uso di sensori ad ampio raggio per tracciare la posizione dell'utente in tempo reale attraverso appositi algoritmi di computer vision, una tecnica per estrarre conoscenza da immagini o video, in questo caso dell'ambiente circostante e sfruttando un'intelligenza artificiale (AI). Questo sistema privo di sensori esterni consente un maggiore senso di immersione, presenza e mobilità, oltre che permettere una maggiore scalabilità in termini di superficie tracciabile.<sup>[23]</sup>

La principale sfida di questi sistemi è cercare di essere il più efficienti possibili sia in termini di prestazioni che di spazio, dato che tutto sarebbe localizzato nell'HMD. Ovviamente il tracking deve anche essere estremamente preciso. Per raggiungere tale obiettivo, Insight fa uso di SLAM (Simultaneous Localization And Mapping) che ogni millisecondo esegue una mappatura precisa dell'ambiente calcolata sulla base dei dati ottenuti dalle ottiche del visore, che proiettano svariati raggi infrarossi nella stanza. Questo algoritmo, oltre che a tracciare la sua posizione nello spazio, riesce a localizzare gli oggetti, fornirne una ricostruzione, identificare aspetti della geometria della stanza (come la presenza di spigoli o la distanza dal pavimento) nonché correggere il drift dei controllers, i quali emettono a loro volta raggi infrarossi recepiti dalle ottiche per aumentare la precisione dell'algoritmo. Attraverso lo studio dei dati ottenuti nel corso dell'esperienza, Oculus Insight è capace di imparare ogni più piccolo dettaglio dell'ambiente e rendere la mappatura più precisa di volta in volta.<sup>[24]</sup>

Nell'ultimo periodo la tendenza da parte dei produttori è quella di adottare implementazioni di inside-out tracking, in quanto risulta essere più fruibile e pratica per l'utente finale. Rendere accessibile la realtà virtuale all'utente finale senza avere troppe pretese sull'ambiente in cui viene utilizzata migliora ampiamente l'esperienza finale e amplifica il bacino d'utenza, semplificando anche il setup iniziale. Un altro vantaggio di questa implementazione riguarda la sua scalabilità, ovvero la possibilità di ampliare le capacità della tecnologia: al crescere dello spazio disponibile sarà possibile tracciare un'area sempre più grande senza incontrare ulteriori complicazioni, a differenza delle altre tecnologie che richiedono l'aggiunta di ulteriori sensori.

Purtroppo al momento non tutti possono adottare questa soluzione, anche se sembrerebbe essere ottima, perché è molto costosa in termini di risorse e sviluppo. Difatti non a caso è stato preso in studio un prodotto di Facebook che di sicuro non ha di questi problemi. Insight, a differenza di SteamVR, non è ne open, ne riutilizzabile da terzi, è unicamente disponibile un SDK per rendere possibile agli sviluppatori l'integrazione lato applicativo. A questo proposito viene reso noto che questa non è l'unica implementazione di markerless inside-out tracking, ma è comunque esemplare in quanto le basi del funzionamento sono simili.

Inoltre allo stato attuale dell'arte molti utenti nutrono dei forti dubbi nella precisione del tracking, in particolar modo se le applicazioni di cui si fanno uso richiedono di posizionare spesso i controller negli angoli ciechi delle ottiche del visore (quindi ai lati e dietro la testa), in quanto a quel punto entrano in gioco gli algoritmi predittivi e il risultato potrebbe essere impreciso.

In conclusione, chi tra gli appassionati può permettersi di usare base station lo

preferisce, mentre l'inside-out tracking viene relegato più per un pubblico casual o meno disposta a scendere a compromessi con tali tecnologie.

## 1.3 Controllers

I controllers sono gli ultimi componenti necessari a rendere davvero completo un sistema di realtà virtuale, grazie all'aggiunta della possibilità di interazione con la scena. Viene posta come ultima in quanto la loro presenza non è obbligatoria, tuttavia la loro inclusione costituisce la possibilità di rendere l'immersione nell'esperienza davvero sofisticata. Questo avviene in gradi diversi, in base alla loro implementazione e al loro utilizzo. Esistono due principali categorie di controllers:

**Joystick:** i controlli vengono impartiti tramite l'uso di un controller tradizionale (di solito usato per i videogames), consente input rapidi e precisi, prima dello sviluppo dei controllers posizionali era molto in voga, ma non dava alcun valore aggiunto all'immersione, bensì la peggioravano introducendo un'inconsistenza tra il movimento della telecamera e quello dell'utente

**Motion Controls:** anche detti controlli di movimento, sono degli strumenti capaci di inviare la loro posizione all'unità centrale tramite un sistema di tracciamento. Nel caso della realtà virtuale abbiamo 6DoF per tutti i controllers posizionali usati dai kit VR in commercio. In questa categoria rientrano anche tutte le tecnologie che non usano un tradizionale controller fisico da mettere in mano all'utente, ma interi sistemi esterni capaci di tracciare la posizione dell'utente nello spazio, ad esempio tute che registrano il movimento di tutto il corpo, sensori che rilevano il movimento delle mani o delle dita o, ancora, intere stanze che tracciano uno o più persone...

Seppur possa sembrare che i controlli tramite tracciamento esterno siano la migliore scelta in assoluto, al momento i più utilizzati sono controllers posizionali specifici per piattaforma, che integrano controlli analogici, di movimento, touch e feedback aptici.<sup>[25]</sup> Difatti è evidente che, se l'applicazione richiede azioni più complesse delle semplici interazioni di base (tutti quei movimenti che sfruttano le mani come nella realtà, quali prendere oggetti, indicare, sfiorare, etc.), è più semplice e immersivo usare dei controllers, ad esempio, basta pensare ad un videogame che fa uso di diversi tipi di abilità fantascientifiche, esprimere queste azioni innaturali è molto più complesso da gestire sia per lo sviluppatore che per l'utente, per cui è preferibile l'uso di pulsanti e touchpads. L'uso di controlli naturali è preferito in situazioni realistiche che hanno a che fare con la simulazione.

## Capitolo 2

# Panoramica dei dispositivi

In questo capitolo si introdurranno le principali alternative in commercio nel campo della VR, ma prima è bene evidenziare quali siano le metriche di paragone che di solito vengono prese in considerazione dell'acquirente nell'acquisto di un headset. Nella lista sono presenti esclusivamente i visori delle aziende più importanti, in quanto quelli sviluppati da terzi sono variazioni che si appoggiano su Steam VR, non troppo differenti dai visori HTC e quindi non necessitano di una descrizione approfondita (qualcosa è stato introdotto nel primo capitolo).

**Schermo:** le caratteristiche tecniche degli schermi del visore (combinati alle lenti) quali risoluzione in pixel, diagonale (ovvero la dimensione degli schermi) in pollici, ppi (pixel per inches) ovvero la densità dei Pixel: un valore troppo basso peggiora l'immersione in quanto i pixel sono "troppo grandi" in relazione alla vicinanza agli occhi (500 ppi sono sufficienti), FOV in gradi, refresh rate in Hertz e altro.

Più sono alti i valori, migliore è quella specifica.

**Sistema di tracking:** come abbiamo già discusso, la principale differenza è tra inside-out tracking e SteamVR tracking, siccome quest'ultimo è un protocollo standard, mentre l'inside-out tracking esiste in implementazioni diverse a seconda del visore, daremo più informazioni riguardo la tecnica utilizzata.

**Audio:** i visori possono avere soluzioni audio integrate o jack per la connessione di cuffie con microfono.

**Peso ed Ergonomia:** un fattore molto importante per la comodità dell'HMD è il peso e come esso viene distribuito alla testa. Inoltre ogni headset da la possibilità di effettuare diverse regolazioni così da garantire un'esperienza comoda per tutti gli utenti. Queste regolazioni possono riguardare la dimensione della testa, la distanza intrapupillare (ovvero tra una pupilla e l'altra), la distanza focale, etc...

**Ecosistema:** ovvero la presenza di software, hardware o servizi esclusivi collegati all'apparato. Siano essi inclusi di base o acquistabili separatamente. In questa sezione verrà descritta brevemente la piattaforma ma non i sistemi operativi compatibili (in quanto abbiamo già fatto un'assunzione a riguardo).

**Requisiti minimi/consigliati:** nel caso in cui il visore necessiti di un PC per elaborazione verranno indicati i requisiti minimi e consigliati per il corretto funzionamento del visore.

Generalmente a requisiti minori corrisponde un costo minore e una migliore resa con hardware più performante. In questa sezione non verranno indicati gli eventuali prerequisiti di montaggio e spazio per Lighthouse se presente. Inoltre, siccome l'unico sistema operativo compatibile che permette di usare agevolmente la VR è Windows 10 si ignorerà questo dato, a scampo di particolari eccezioni. Se l'HMD necessita di connettersi ad un'unità esterna per processare l'immagine ci sono diverse configurazioni possibili sulla base delle scelte del produttore (date dalla quantità dei dati da trasportare e dal protocollo usato).

**Controlli:** in questa sezione verranno descritti i controllers inclusi (se presenti) e le loro specifiche. Nel caso dei visori è scontata la possibilità di muoversi con 6DoF. Potrebbero comunque essere presenti controlli derivanti dal visore (per esempio hand tracking).

**Altre caratteristiche:** ogni visore per cercare di differenziarsi dalle offerte dei competitor includono alcune funzionalità o particolarità per rendersi più interessanti per l'utente. Se presenti verranno evidenziate.

**Prezzo:** il costo di un sistema completo per la VR. Ovviamente minore è preferibile.

Nel costo complessivo non è inclusa un'eventuale configurazione del PC (ottimale o meno), in quanto i prezzi sono molto volatili ed è a discrezione dell'utente quanto investire nel sistema o a che livello debba essere l'esperienza finale. Inoltre il PC che viene usato per la VR non deve per forza avere come unico scopo quello di far funzionare il visore ma può essere comunque usato come un normalissimo computer, per cui si potrebbe già essere in possesso di una macchina capace e quindi talvolta risulta di poco interesse per l'utente tenere quel costo in considerazione. È invece considerevole quando quel costo è del tutto evitabile.

Per completezza assumiamo che una minima configurazione funzionante si aggiri attorno ai 700 Euro.

## 2.1 Oculus Rift S



Top di gamma di Oculus, successore del Rift originale. A differenza del suo predecessore questo visore fa uso di inside-out tracking, piuttosto che dei sensori proprietari per Oculus. È sviluppato in collaborazione con Lenovo, che ha già avuto esperienze nel settore VR con i suoi Mirage ed Explorer. Il setup è piuttosto semplice, si collega al PC, si segue la procedura di installazione e si connettono gli Oculus Touch. Questo dispositivo si concentra principalmente sulla facilità d'uso.<sup>[14]</sup>

**Schermo:** un pannello LCD con risoluzione  $2560 \times 1440$  e frequenza 80 Hz. Fa uso di un paio di lenti di "nuova generazione" con FOV di  $115^\circ$ , sviluppate appositamente per ridurre il cosiddetto screen door effect ed annullare l'effetto "raggiere di luce".

**Sistema di tracking:** a differenza del Rift originale, il Rift S sfrutta Oculus Insight, ovvero il loro sistema inside-out tracking. L'area massima di tracking è di  $3 \times 2.5\text{m}$  mentre la minima è di circa  $1 \times 1\text{m}$  per quelle applicazioni che fanno uso dei controllers. Questo risultato viene ottenuto tramite cinque obiettivi ottici posizionati nel visore: tre frontali, di cui uno superiore, e due laterali. Questi obiettivi fanno uso di una funzionalità chiamata Passthrough+ che processa il segnale video degli obiettivi e li proietta direttamente al display, così che l'utente possa muoversi consapevolmente all'interno della stanza nel mondo reale senza dover togliere l'headset. Passthrough+ si occupa anche di attuare diversi altri protocolli tra cui l'"Asynchronous SpaceWarp" che rende l'esperienza più immersiva alterando l'immagine, inoltre ha la possibilità di abilitare l'hand tracking (ovvero tracciare i movimenti delle mani incluse le dita).

**Audio:** il visore ha due speakers che puntano verso il basso sopra le orecchie. Il visore dispone inoltre di un jack nel caso in cui l'utente voglia utilizzare un paio di cuffie. Ha il microfono integrato anch'esso sostituibile dal microfono integrato con le cuffie collegate al jack.



**Peso ed Ergonomia:** il visore pesa 500g, il supporto che regge il dispositivo sulla testa è chiamato halo headband (letteralmente una fascia per la testa ad aureola), regolabile in larghezza con un apposita ruota posta sul retro del visore. Il visore ha la possibilità di regolare la distanza intrapupillare solo via software, dato che ha un unico schermo invece che due.

**Ecosistema:** ha a disposizione una libreria di videogame esclusivi e non compatibili con altri sistemi VR. Nonostante ciò supporta Steam.

**Requisiti minimi/consigliati:** la configurazione minima richiede almeno una scheda grafica NVIDIA GTX 960/AMD Radeon R9 290, una CPU Intel Intel i3-6100/ AMD Ryzen 3 1200, 8 GB di RAM. L'headset va connesso al PC tramite un cavo DisplayPort 1.2 e un USB 3.0.

**Controlli:** l'headset ha a disposizione i controllers Oculus Touch di seconda generazione. Essi sono molto simili alla loro versione precedente, la principale differenza risiede nella presenza di un anello sulla parte superiore che ha il compito di rendere tracciabile il movimento (tramite gli obiettivi del visore). I controller sono alimentati con una pila AA, che consente di usarli per 30h e una volta finita la carica è possibile cambiare le pile al volo e riprendere.

**Altre caratteristiche:** questo visore fa uso di un tool chiamato Guardian che visualizza una barriera quando si raggiungono i limiti dello spazio nel mondo reale.

**Prezzo:** il prezzo di lancio del visore è 450 Euro.

## 2.2 Oculus Quest



Questo visore è l'alternativa unthetered (senza fili) a Rift S. Destinato ad un pubblico con un approccio molto più "casual", Quest dà la possibilità di usufruire della VR praticamente ovunque senza l'incombenza di un'unità di calcolo esterna, in quanto tutto ciò che serve sono il visore e i controllers, tutti i calcoli necessari vengono svolti all'interno del visore tramite una CPU da smartphone di fascia alta.<sup>[15]</sup> Attualmente è l'unico nel mercato con queste caratteristiche. È interessante notare come Oculus, dopo l'acquisizione da parte di Facebook, abbia spostato interamente l'effort sul creare l'esperienza d'uso più semplice e comoda possibile, agevolando l'ingresso nel mercato a nuovi clienti non particolarmente affezionati (cosiddetti entry level), piuttosto che al suo precedente mercato di enthusiast che avevano caratterizzato il successo iniziale della compagnia. Attualmente Oculus offre le soluzioni più economiche sul mercato.

**Schermo:** il display ha un pannello OLED con risoluzione  $1440 \times 1600$  per occhio con un refresh rate di 72 Hz. Le lenti sono le stesse di Rift S.

**Sistema di tracking:** anche questo dispositivo fa uso di Oculus Insight. Secondo Facebook la dimensione massima potenziale della play area dal visore è di  $350\text{m}^2$  (la cosiddetta Arena Scale VR, ovvero la possibilità di immergersi in un ambiente virtuale particolarmente ampio insieme ad altri utenti che condividono lo stesso spazio nella realtà). In ogni caso le dimensioni del Guardian sono bloccate a  $3 \times 2.5\text{m}$  attualmente.

**Audio:** ha una coppia di speaker con audio posizionale build-in e la possibilità di collegare un jack.

**Peso ed Ergonomia:** il visore pesa 571 grammi ed è tenuto sulla testa con una fascia elastica che avvolge tutta la testa. È possibile regolare la distanza intrapupillare tramite uno slider sul fondo del dispositivo. È incluso uno spessore in foam per aumentare la distanza dalle lenti per chi indossa occhiali.

**Ecosistema:** ha a disposizione una libreria di titoli limitata e non compatibili con altri sistemi. Essendo un sistema non eccessivamente potente le applicazioni vanno rielaborate e sistemate dagli sviluppatori in maniera tale da ottenere un risultato fluido. Da questo ne deriva che Quest non supporta nativamente tutte le funzionalità Steam in quanto le applicazioni vanno caricate direttamente nella memoria interna del dispositivo. Le applicazioni sono acquistabili tramite l'Oculus Store, nessun cross-buy (acquisto incrociato) con altre piattaforme o supporti.

**Requisiti minimi/consigliati:** l'unica maniera per utilizzare tali programmi è acquistare a parte un Oculus Link, un cavo che connette tramite USB C il visore ad un PC compatibile lasciando ad esso la responsabilità del rendering. Il visore ha un processore Snapdragon 835 che si occupa di tutte le operazioni.

**Controlli:** utilizza gli stessi controller Oculus Touch di Rift S.

**Altre caratteristiche:** il visore è capace di adattarsi automaticamente a qualsiasi ambiente ripetendo il setup di volta in volta tramite un app per smartphone. La durata della batteria si aggira intorno alle 2/3 ore, dipendentemente dal tipo di applicazione usata.

**Prezzo:** il prezzo di lancio di questo dispositivo è di 550 Euro per la versione 64 Gb, 100 in più per la versione 128 Gb, in questo caso è importante tenere in considerazione che il visore non richiede un'unità esterna, evitando un'eventuale spesa riguardante quella macchina.

Come postilla, è interessante far notare come la semplice esistenza di questo dispositivo avvalori la tesi proposta. Il mercato della VR sta effettivamente concentrando i propri sforzi nel rendere accessibile questa tecnologia alle masse. La speranza che gli HMD diventino di uso comune non è più fantasia, ma una concreta possibilità che apre alla necessità di sviluppo di applicazioni per gli usi quotidiani, tra cui l'esperienza qui proposta.

## 2.3 HCT Vive



Concorrente del Rift originale prodotto da HTC in collaborazione con Valve.<sup>[10]</sup> Questo dispositivo sancisce la nascita di Valve VR e del relativo sistema di tracking a costellazione. Viste le limitazioni del suo rivale, questo visore è rimasto per diversi anni imbattuto come uno dei visori più venduti e tuttora gode di un discreto seguito nonostante le evoluzioni del settore. Il suo successo è dovuto in parte alle caratteristiche hardware ottime per l'epoca e soprattutto all'ecosistema strettamente legato al più grande retailer virtuale di videogiochi, ovvero Steam. A questo si aggiunge una libreria software (SteamVR) che ha reso lo sviluppo di applicazioni molto più rapido, lasciando la possibilità di sviluppare applicazioni VR anche a piccoli studi e permettendo così al catalogo di titoli di ingrandirsi a dismisura. Nonostante il suo rilascio nel 2016 il visore non ha visto molti cambiamenti, se non per una versione Pro che non verrà discussa in quanto risulta in un upgrade neanche troppo evidente per ambiti professionali dello stesso Vive.<sup>[11]</sup>

**Schermo:** due pannelli OLED 3.4" da  $1080 \times 1200$  (per una risoluzione totale di  $2160 \times 1200$ ) con un refresh rate di 90Hz e FOV di  $110^\circ$  .

**Sistema di tracking:** fa uso della prima versione di Steam VR Tracking, il sistema a costellazione con stazioni IR discusso nella sezione precedente.

**Audio:** jack audio da 3.5mm per cuffie e microfono integrato.

**Peso ed Ergonomia:** il visore pesa 470g e sfrutta il supporto di una fascia elastica regolabile in larghezza e altezza che avvolge la testa. È possibile regolare la profondità del visore, la distanza intrapupillare e aggiungere uno spessore per gli occhiali. I cavi, molto lunghi e pesanti, possono essere sostituiti da un adattatore wireless da posizionare nella parte superiore del supporto (venduto separatamente).

**Ecosistema:** il visore fa pieno uso di tutte le funzionalità e della libreria software di Steam. Attraverso l'overlay di Steam VR è possibile avviare anche applicazioni esterne allo store, grazie anche al supporto ai framework standard di Windows e Steam. Inoltre è possibile utilizzare la piattaforma proprietaria di HTC (chiamata VivePort) che consente di abbonarsi ad una libreria sempre crescente di titoli on demand (ed esclusivi).

**Requisiti minimi/consigliati:** Il visore richiede l'uso di un PC con una GPU Nvidia GeForce GTX 970 o AMD Radeon R9 290 o migliori, come CPU: Intel i5-4590, AMD FX 8350 sono il minimo necessario, 4GB di RAM e due possibilità di collegamento tra cui HDMI 1.4, DisplayPort 1.2 in combinazione ad una USB 3. Inoltre avere un monitor collegato al PC che si occupa del rendering per ottenere un corretto funzionamento del visore; nel monitor potrà essere visualizzato un mirroring del contenuto del visore.

**Controlli:** i Vive Controllers sono due unità con input multipli tra cui un track pad, due grip e un grilletto. Hanno una batteria a ioni di litio della durata di 6h. Negli anelli del controller sono presenti 24 sensori a infrarossi che aiutano a tracciare la posizione dei controller al millimetro.

**Altre caratteristiche:** il visore è equipaggiato con una videocamera frontale che consente la visione dello spazio circostante (mostrato in aloni blu) attraverso il sistema di sicurezza chiamato Chaperone che si occupa di visualizzare il mondo esterno quando esso è troppo vicino al visore.

**Prezzo:** al lancio il prezzo del kit era di 800 Euro, nel 2018 è sceso a 600.

## 2.4 HTC Cosmos



L'HTC Cosmos è il nuovo visore di punta della HTC, si focalizza principalmente sulla semplicità d'uso e sull'estetica, cercando di scartare il vecchio monolite nero e sostituirlo con un'oggetto di design, cercando di attrarre la clientela "business", comunque strizzando l'occhio agli utenti meno esperti.<sup>[12]</sup> Questo cambiamento radicale è stato possibile soprattutto tramite l'implementazione di un sistema di inside-out tracking ottico, che ha liberato la parte frontale dai sensori per lo Steam VR Tracking, e che soprattutto ha rimosso l'incombenza delle stesse base station. Nonostante il suo prezzo elevato, Cosmos si pone come un buon investimento destinato a durare nel tempo, grazie soprattutto al sistema di upgrade modulari.

**Schermo:** due pannelli RGB LCD da 3.4", 1440 × 1700 per occhio (2880 × 1700 combinati), 90Hz di refresh rate e 110° di FOV.

**Sistema di tracking:** principalmente fa uso della sua implementazione di inside-out tracking, che traccia i controller e la posizione attraverso il visore. Siccome il sistema sfrutta delle fasce luminose per orientare i controllers, ci riferiamo a questo metodo col nome di "tracciamento ottico", infatti questo metodo risulta spesso inefficace in condizioni di scarsa luminosità, a differenza dei sistemi ad infrarossi. In alternativa si può rendere compatibile con Steam VR Tracking V2, una versione migliorata del Lighthouse discussa nel capitolo precedente, attraverso l'uso di un modulo apposito.

**Audio:** il visore è equipaggiato di un paio di cuffie molto simili a quelle accessorie per Vive e un microfono integrato.

**Peso ed Ergonomia:** 645g comprese le cuffie e le straps, il visore di per se pesa 330g. Il visore si regge su una helo headband regolabile con supporto superiore, ha la possibilità di regolare profondità, distanza intrapupillare e consente inoltre di alzare il visore di 90° senza dover sfilare il supporto. La ghiera frontale oltre che la funzione estetica offre un sistema di ventilazione per evitare che la porzione del volto racchiuso nel visore sudi eccessivamente.

**Ecosistema:** oltre all'implementazione di Steam VR di cui godeva il suo predecessore, al visore è stata integrata un'applicazione social esclusiva chiamata Origin Environment, uno spazio che si configura come una casa virtuale in cui rilassarsi, accedere alle app e invitare altri utenti.

**Requisiti minimi/consigliati:** necessita un PC con Intel Core i5-4590 o AMD FX 8350, un'NVIDIA GeForce GTX 1060 or AMD Radeon RX 480 o migliore, 8 GB RAM o più. È collegato tramite un DisplayPort 1.2 e un USB 3.0

**Controlli:** il visore fa uso di una coppia di controllers alimentati da una pila AA ciascuno. La fascia luminosa sovrastante fa sì che il visore riesca a tracciare precisamente la posizione delle mani nella realtà virtuale. Oltre ai vari pulsanti sono presenti un bottone per mostrare la realtà circostante in maniera immediata e un trackpad.

**Altre caratteristiche:** il visore è stato concepito in maniera tale da consentire un'esperienza modulare: in futuro sarà possibile aggiungere varie mod al visore (specie sul frontale) per ampliarne le capacità, ad esempio le prime due annunciate sono una mod di compatibilità per i sensori Lighthouse V2 (per ottenere un tracciamento più preciso) e una mod per trasmettere il segnale wireless.

**Prezzo:** il prezzo del visore per il mercato italiano è di 829 Euro.

## 2.5 Valve Index



Il primo PC VR HDM prodotto interamente da Valve,<sup>[13]</sup> la società che si è occupata della libreria Steam VR. La compagnia è nota per rivolgersi principalmente ai clienti più esigenti in materia di VR ed in generale di videogiochi, in quanto essa è anche titolare della piattaforma Steam, il maggior rivenditore digitale. Index è quindi un prodotto di fascia alta che punta molto sulla sua potenza e sulla resa dell'esperienza in contesti molto frenetici dove la maggior fluidità e i migliori tempi di risposta sono essenziali e da ottenere senza alcun compromesso.

**Schermo:** due pannelli RGB LCD da  $1440 \times 1600$  con refresh rate di 120Hz ed una modalità sperimentale da 144Hz e  $130^\circ$  di FOV grazie alle lenti migliorate.

**Sistema di tracking:** il visore usa Steam VR Tracking V2, una versione migliorata del sistema di Valve, descritto nel capitolo precedente.

**Audio:** il visore monta un paio di cuffie off-ears (distanti dalle orecchie) e un microfono integrato nel visore.

**Peso ed Ergonomia:** il visore pesa ,particolare enfasi è stata posta sui materiali traspiranti a contatto con la pelle, scelti appositamente per garantire il miglior comfort e pulizia per le lunghe sessioni d'uso. Il visore è supportato da una fascia regolabile stile halo headband con due fasce regolabili invece di una.

**Ecosistema:** ovviamente il visore di Valve fa uso di tutte le integrazioni di Steam precedentemente citate, in particolar modo Lighthouse 2.0

**Requisiti minimi/consigliati:** necessita un PC con almeno una CPU dual core, un'NVIDIA GeForce GTX 970 or AMD Radeon™ RX 480 o migliore e almeno 8 GB RAM. Va collegato tramite un DisplayPort 1.2 e un USB 3.0

**Controlli:** con il visore sono stati annunciati dei particolari controller chiamati Knuckles Controller che, grazie alla loro impugnatura, consentono di essere



usati senza per forza doverli stringere tra le mani, ma bensì appoggiandoli sulle nocche (da cui il nome). I controllers hanno delle batterie a ioni di litio.

**Altre caratteristiche:** il visore ha uno slot frontale dove sarà possibile inserire dei dispositivi per aumentare le capacità del visore (come ad esempio un sensore per l'hand tracking). I controller sono utilizzabili con altri visori che fanno uso di Steam VR. I controller consentono inoltre un finger tracking approssimativo in base a quante dita sono appoggiate sul controller.

**Prezzo:** il visore costa 540 Euro, ma è sprovvisto di controllers e base station (acquistabili per 300 e 160 Euro rispettivamente). Un kit completo costa 1080 Euro. La disponibilità di varie configurazioni ha lo scopo di dare la libertà a chi già è in possesso di un sistema Steam VR di scegliere i componenti a cui è davvero interessato senza obbligare l'utente all'acquisto del kit completo.

## 2.6 Sony PSVR



Il PSVR (PlayStation VR<sup>[17]</sup>) è l'alternativa più economica in assoluto per quanto riguarda l'esperienza cablata di realtà virtuale in quanto il costo complessivo è composto dal kit del visore e dalla console a cui è collegato (PlayStation 4).

Ovviamente questo visore è esclusivamente legato al consumo di esperienze virtuali legate al mondo dei videogames, in particolare alla piattaforma Sony. Questo visore è stato incluso nella lista (ma verrà comunque descritto in maniera rapida) proprio per evidenziare la sua natura elementare e al contempo funzionale.

**Schermo:** un pannello OLED da 5.7", con FOV di 100° e risoluzione 1080p (960 × 1080 per occhio) e un refresh rate variabile di 90–120Hz.

**Sistema di tracking:** il visore fa uso di una telecamera (PS Camera) che rileva 9 sezioni luminose dal visore e dalla punta dei controllers. Con questo semplice setup riesce comunque ad ottenere tracking posizionale e 6DoF.

**Audio:** uscita tramite il jack headphone e un microfono integrato

**Peso ed Ergonomia:** il visore pesa 600g circa ed è assicurato alla testa con una halo headband regolabile in larghezza.

**Ecosistema:** il sistema si appoggia sulla piattaforma PlayStation e funziona esclusivamente in quell'ambiente.

**Requisiti minimi/consigliati:** i requisiti minimi consistono nell'acquisto di una PS4 e del Kit VR.

**Controlli:** il visore si appoggia su due controller PlayStation Move, due periferiche provenienti dall'epoca PS3 nella quale i controlli di movimento erano esplosi grazie all'avvento di Nintendo Wii. Per giochi più complessi o veloci si ricade sull'uso del controller Dualshock 4 (joystick base di PS4).

**Prezzo:** 400 Euro per il kit VR e almeno altri 300 per la console.

## Capitolo 3

# Campi d'uso

La realtà virtuale è uno strumento estremamente potente e flessibile, queste caratteristiche costituiscono un immenso potenziale inespresso che può essere espresso in qualsiasi settore, portando notevoli vantaggi in termini di costo e permettendo risultati finora inarrivabili. In questo capitolo verranno mostrati i principali settori in cui la VR viene applicata, specificando per ognuno dei campi specifici.

### 3.1 Intrattenimento

Siccome il visore proietta l'utente in uno scenario totalmente diverso da quello in cui si trova, le possibilità concesse in campo artistico, ludico e cinematografico sono illimitate.

Il settore in cui la realtà virtuale ha avuto maggiore impatto è sicuramente quello videoludico. Il già citato Steam conta ormai svariate migliaia di giochi nella sua sezione "VR Enabled"<sup>[26]</sup> molti dei quali hanno una player base vibrante ed entusiasta. Il loro successo è dato principalmente dalla dinamicità consentita nella VR, altri settori, come quello cinematografico non riescono a farne uso in maniera altrettanto brillante, difatti tendono a limitare l'interazione se non ad escluderla del tutto. Anche nel settore artistico vanta un elevato tasso di adozione, sia a livello di esperienze che cercano di "ampliare i sensi" dell'utente ponendolo in un contesto surreale,<sup>[27]</sup> sia in quelle dove viene invece ampliata la conoscenza dell'utente, ovvero veri e propri tour dinamici nell'arte e nella storia.

Anche nel settore artistico vanta un elevato tasso di adozione, grazie a questo strumento si è in grado di comunicare coinvolgendo i sensi nella realtà proposta dall'artista coinvolgendoli nel suo mondo. Queste esperienze vengono di solito realizzate basandosi sulle implicite contraddizioni nel cercare di creare una realtà alternativa.

Anche all'interno di parchi divertimento e musei vengono usate tecnologie VR, talvolta insieme a feedback aptici per aumentare l'immersione dell'attrazione. Nei musei e nelle mostre alla loro funzione di intrattenimento viene aggiunto un valore istruttivo sfruttando la capacità comunicativa del mezzo. Uno degli approcci più comuni consiste nel proiettare il visitatore in un'altra epoca e consentirgli di esplorare l'ambiente.

Tra le varie forme espressive, troviamo le sfilate virtuali, una dimensione creativa in grado di unire Fashion, Arte e 3D. Grazie alle nuove tecnologie è possibile catalizzare una pluralità di linguaggi per coinvolgere ed emozionare un pubblico sempre più esigente. L'innovazione non è soltanto nel prodotto, il modo in cui lo si presenta diventa fondamentale per differenziarsi dalla massa.

## 3.2 Formazione e Simulazione

È stato già accennato nella sezione precedente come la realtà virtuale possa essere impiegata in contesti istruttivi. In maniera del tutto simile, in campo professionale vengono sviluppate esercitazioni immersive per operazioni impegnative e altamente specifiche, che richiederebbero condizioni particolari non sempre avverabili. Ad esempio in campo chirurgico si insegnano operazioni rare e complesse su dei manichini che simulano le reazioni dell'organismo;<sup>[28]</sup> oppure in campo militare vengono proposti degli addestramenti che rievocano battaglie in scenari realistici o creano delle nuove situazioni da cui apprendere tattiche specifiche. Attraverso la simulazione è possibile ricreare ambienti complessi nei quali l'emergenza evolve, al passare del tempo, proprio come avviene nella realtà, incluso le conseguenti ripercussioni su cose e persone. La realtà virtuale costituisce così uno strumento utile ed economico nell'addestramento del personale di gestione delle situazioni di emergenza.

I training in realtà virtuale hanno vantaggi sia economici, sia di efficacia formativa, altrimenti irraggiungibile attraverso gli insegnamenti teorici.<sup>[29]</sup> La capacità di rendere attiva e partecipe l'audience a cui fa riferimento è il punto di forza dei tutorial in VR: l'utente può reagire a degli stimoli, collaborare con altre persone, e soprattutto mettere in pratica gli insegnamenti prima della loro effettiva necessità e senza mettere a repentaglio la propria incolumità o quella degli altri.

Un'ulteriore approccio istruttivo consiste nell'uso di simulatori.<sup>[30]</sup> Questi sistemi servono appunto a "simulare" dispositivi (o situazioni) che richiedono di un addestramento specifico: tra i più famosi i simulatori di volo, di guida o di mezzi di soccorso. I simulatori possono essere esclusivamente software o trovare l'aggiunta di hardware specifico come joystick o intere piattaforme molto simili al dispositivo da simulare. Questa specializzazione costituisce forse una tra le più primitive applicazioni del termine realtà virtuale.

I simulatori vengono impiegati anche nel contesto ingegneristico, dove i prototipi sono testati in anteprima all'interno della realtà virtuale, dando al team di sviluppo un punto di vista unico della loro creazione, questo può essere il cosiddetto POV (ovvero Point of View), che identifica la prima persona dalla prospettiva del dispositivo. Inoltre seguendo un approccio simile, alcuni dispositivi vengono controllati facendo immedesimare l'utente nel mezzo, tramite l'uso del visore. È il caso dei droni e dei robot destinati ad operazioni di soccorso e ricognizione. È possibile anche l'inverso, ovvero tracciare e renderizzare dispositivi reali all'interno della realtà virtuale e renderli interagibili con l'utente. Sistemi come Lighthouse sono già predisposti per questo scopo: basta integrare un ricevitore nei dispositivi per abilitarne il tracking con 6DoF.

### 3.2.1 Gamification e Serious Games

Negli approcci descritti è solito trovare esempi di gamification,<sup>[31]</sup> ovvero l'introduzione di un sistema ludico per aumentare l'engagement dell'utente verso un task particolarmente dispendioso in termini di energie mentali e/o fisiche. Tali tecniche di solito consistono in sistemi a punti con obiettivi che, connessi all'aspetto sociale e ad altre attività competitive, trasformano molte attività di tutti i giorni verso il divertimento e un maggiore rendimento.

Nel caso della VR, una semplice trasposizione dell'esperienza nella realtà virtuale, costituisce una forma diversa di gamification. Parliamo in questo caso di serious games: giochi digitali concepiti come strumenti formativi dove idealmente gli aspetti seri e ludici sono in equilibrio. L'attenzione è posta nel creare un'esperienza formativa efficace e piacevole. Lo scopo fondamentale è quello di sviluppare abilità e competenze da applicare nel mondo reale attraverso l'esercizio in un ambiente simulato e protetto. Diversamente dalla Gamification, che contiene solo alcuni elementi mutuati dai giochi, quali l'assegnazione di punti o il raggiungimento di livelli, il Serious Game è gioco a tutti gli effetti.<sup>[32]</sup>

### 3.3 Retail

Il marketing immersivo è uno dei nuovi termini utilizzati per descrivere un contenuto che include video, eventi o esperienze. Altri termini simili sono marketing esperienziale e multisensoriale. Queste tecnologie cambiano il modo in cui l'utente interagisce con il contenuto. Ossia, al posto di guardare o visualizzare il contenuto in maniera passiva, grazie ad esse l'utente interviene nel processo e si sente parte dell'azione. I negozianti possono usare la VR come estensione della brand experience per portare i clienti in ambienti immersivi. La personalizzazione dell'esperienza d'acquisto è una delle maggiori innovazioni del retail nell'era della digitalizzazione, e in questo contesto rivoluzionario non può certo mancare la realtà virtuale.<sup>[33]</sup> L'esperienza si fa immersiva e il cliente può godere dalla comodità casalinga di nuovi mondi. Ad esempio, consentire di "provare" abiti (e accessori di ogni tipo) prima di acquistarli e senza la necessità di passare dal camerino. Inoltre, visto il crescente sviluppo di interesse verso gli aspetti più social della realtà virtuale, si potrà disporre della possibilità di girare per vetrine insieme ai propri amici.

Il virtual retailing è ideale per creare legami positivi tra il cliente e il brand: la personalizzazione è considerata un'arma fondamentale, mette in chiaro la necessità dell'utente di essere protagonista di tutte le esperienze grazie alla tecnologia che permette di gestire questo parametro. La realtà virtuale garantisce che le aziende creino contenuto personalizzato. Un brand che usa il marketing immersivo ed è interattivo, emergerà rispetto alla concorrenza. Anche a livello di accesso linguistico si può trovare facilmente una soluzione visuale al problema. Il mondo dei video è trasversale e non richiede traduzione. I brand e i negozianti saranno in grado di creare il proprio spazio senza l'incombenza di dover fornire una struttura fisica, con tutto quello che ne consegue, riuscendo inoltre ad attrarre un pubblico molto più vasto rispetto a quello locale.

Un esempio di quest'ultimo caso è dato da Netflix, che ha deciso di proporre una sorta di spazio virtuale con tanto di scaffali dove selezionare i film ma con possibilità di vederne i trailer e di ulteriori informazioni, reminescente della catena Blockbusters.

Altre applicazioni per il settore dell'e-commerce comprendono quello edile. Grazie a software di rendering 3D, ad esempio, si potranno costruire ed esplorare palazzi e abitazioni virtuali, così da verificarne eventuali deficit progettuali e dare modo ai committenti di "abitarle" (anche se per pochi minuti) prima della posa della prima pietra

La Realtà Aumentata, grazie alla diffusione del mobile, trova facilmente supporto per le brand-experience digitali, ma grazie alla tecnologia WebVR, la rete inizia ad accettare anche la Realtà Virtuale, avvicinandoci così ancora di più alla possibilità di navigare internet nella realtà virtuale. I crescenti investimenti nel settore hanno segnato la nascita di divisioni specifiche come MozillaVR ed il conseguente arrivo di a-frame, uno dei framework che consente di sviluppare applicazioni 3D fruibili in VR su qualsiasi browser web, senza dover installare alcun plugin o app dedicate.

Le sperimentazioni, hanno ormai ceduto il passo ai primi esercizi pilota che, se funzioneranno, potranno effettivamente offrire ai retailer online una carta in più da giocare per la conquista del cliente. Ancora ai primi passi, la VR non sarà la panacea per ottenere una shopping experience con tutti i vantaggi, ma sicuramente sta entrando nella vita di molti, se uniamo questo ad una maggiore diffusione dell'offerta, lo shopping immersivo, nel prossimo futuro, potrebbe rappresentare l'ennesima alternativa/sfida al retail tradizionale.

## Capitolo 4

# Concetti teorici

Quando si parla di realtà virtuale vanno tenuti in considerazione diversi aspetti legati alla percezione dell'utente per far sì che l'illusione avvenga con successo.

Abbiamo già anticipato alcuni concetti come immersione e presenza, in questo capitolo verranno introdotti altri fattori incisivi nell'esperienza dell'utente, sia positivamente che negativamente, e verranno ampliate le definizioni già date.

Inoltre verrà presentata una lista di linee guida alla progettazione di una "buona" applicazione VR priva di effetti collaterali.

### 4.1 Immersione

È possibile ampliare quanto noto riguardo l'immersione attraverso la seguente definizione alternativa e andando successivamente ad espandere le singole parti: "L'immersione è il grado oggettivo al quale un sistema in VR progetta gli stimoli sui recettori sensoriali dell'utente, in modo tale che siano quanto più estensivi, congruenti, circondanti, vividi, interattivi, e informanti della trama".<sup>[34]</sup>

**Estensibilità:** la quantità di tipi di stimoli sensoriali forniti all'utente (es. stimoli visivi, uditivi, fisici, ecc...);

**Congruenza:** corrispondenza tra stimoli forniti dall'utente e gli stimoli che arrivano al cervello (es. quando l'utente sposta la visuale, anche l'immagine sul visore deve muoversi);

**Circondanti:** gli stimoli sono di tipo panoramico (es. FOV ampia, un audio stereo, un tracciamento a 360°...);

**Vividezza:** la qualità degli stimoli trasmessi (es. una buona risoluzione, la luce, il frame rate, ecc...);

**Interattività:** la capacità dell'utente di poter cambiare il mondo virtuale in cui è immerso, le risposte di entità virtuali ad azioni dell'utente, e l'abilità dell'utente di influenzare eventi futuri;

**Trama:** la rappresentazione coerente di un messaggio o di un'esperienza, lo svolgersi della sequenza dinamica degli eventi, e il comportamento del mondo e delle sue entità.

## 4.2 Presenza e illusione

Ricordiamo che la presenza è il senso di essere dentro uno spazio, sebbene fisicamente si ci trovi altrove, per cui è uno stato psicologico interno e completamente soggettivo.<sup>[35]</sup>

La seguente definizione ci aiuta ad esprimere al meglio questo sentimento viscerale: "Presenza è quello stato psicologico o percezione soggettiva nella quale sebbene parte o tutta la corrente esperienza dell'individuo sia generata e/o filtrata attraverso una tecnologia artificiale, parte o tutta la percezione dell'individuo fallisce ad appurare accuratamente il ruolo che la tecnologia abbia avuto nell'esperienza" (International Society for Presence Research, 2000).

Da tale definizione si può comprendere come i concetti di immersione e presenza siano legati. La presenza di fatto, è in funzione sia dell'utente che dell'immersione, più elevato è il livello, più alta è la possibilità per il sistema di generare nell'utente la sensazione di presenza. Al contrario, al rompersi dell'immersione, l'illusione si spezza, si perde la presenza e l'utente riapprende coscienza della situazione, ricordandosi di indossare un HMD.

La presenza può essere categorizzata in base all'illusione che va a creare:

**Stabilità spaziale:** è la sensazione di essere in un ambiente fisico in cui tutti gli stimoli sensoriali dell'utente sono congrui ai feedback;

**Impersonificazione:** è la percezione di avere un proprio corpo all'interno dell'applicazione, molte esperienze di VR non soddisfano questo requisito, non essendo effettivamente fondamentale, ma migliorativo. In applicazioni che fanno uso di questa tecnica, il controllo del movimento del corpo virtuale è fondamentale per non spezzare la presenza e di conseguenza l'illusione creata, siccome questo aspetto è molto complesso da gestire ed esula dallo scopo della tesi, è stato del tutto evitato rappresentando solamente le mani e i controllers (come è solito fare in questi casi).

**Interazione Fisica:** quando la mera esplorazione visiva non è abbastanza per credere di essere in un mondo diverso, occorre ricorrere ai feedback fisici (es: vibrazioni del controller, audio di sottofondo, feedback tattile, etc...);

**Comunicazione sociale:** l'illusione di stare veramente comunicando, sia verbalmente che attraverso il linguaggio del corpo, con altri interlocutori, siano essi controllati dal computer o da utenti. Questo aspetto è particolarmente importante e verrà tenuto molto in considerazione nell'ideazione dell'assistente vocale.



## 4.3 Percezione dell'utente

In questa sezione verranno descritte le varie modalità di percezione che l'utente sperimenta durante l'uso di un'applicazione VR.<sup>[36]</sup> Seppur questa serie di sensazioni risultino essere del tutto naturali, non è altrettanto semplice riuscire a simularle e darne una rappresentazione effettiva all'interno della realtà virtuale.

Basta pensare al classico esempio riguardante la sensazione di movimento percepita dal cervello quando guarda uno schermo. Questo "inganno" è costituito da immagini statiche mostrate in rapidissima sequenza che creano l'illusione di movimento. Non è affatto semplice e tuttora è materia di studi e continuo campo di sfida sul fronte tecnologico. In generale, lo scopo di chi sviluppa questi sistemi è alterare le percezioni della mente, ai fini di creare un'illusione abbastanza plausibile così da far immergere completamente l'utente.

### 4.3.1 Percezione dello spazio

Normalmente più unità sensoriali cooperano per darci una percezione spaziale. La vista certamente è la modalità spaziale preferita in quanto è più precisa, veloce, accurata e efficiente, specie nelle lunghe distanze.

Fanno parte di questo senso i giudizi esocentrici, ovvero il senso di dove gli oggetti si trovano relativamente ad altri oggetti o riferimenti esterni. I giudizi esocentrici comprendono il senso di gravità, di direzione geografica (per esempio, nord), e la distanza tra due oggetti. La percezione della posizione degli oggetti comprende anche la direzione e distanza rispetto al proprio corpo.

I giudizi egocentrici, invece, sono il senso di dove gli oggetti sono rispetto all'osservatore. Diverse variabili influenzano la percezione egocentrica, sicuramente più stimoli sono disponibili, più i giudizi sono stabili. Per meglio comprendere come percepiamo lo spazio circostante, esso può essere diviso, a seconda della distanza in tre regioni egocentriche:

**Spazio personale:** è considerato come il volume tra il corpo dell'utente e la massima estensione delle braccia, esso raggiunge solitamente i due metri cubi. Questo spazio include anche l'area sottostante l'utente. Lavorare all'interno di spazio personale offre vantaggi rispetto il sistema propriocettore, si può ottenere un mapping più diretto tra movimento della mano e il movimento oggetto, una più forte visione binoculare e una precisione di movimento della testa e del corpo.

**Spazio di azione:** consiste nello "spazio pubblico" in cui l'utente si può muovere, parlare o interagire a distanza con la realtà simulata. Questo spazio va dai due metri ai venti metri dall'utente.

**Spazio di vista:** oltre i venti metri di distanza dall'utente, gli stimoli percettivi sono solo accennati, e utilizzati, con il solo scopo di creare un contesto e migliorare l'immersione dell'applicazione.

La percezione naturale dello spazio all'interno di un ambiente in VR rimane ancora una sfida, come anche la creazione di contenuti che ne facciano saggio uso. Per esempio, mentre la percezione della distanza egocentrica può essere abbastanza buona per il mondo reale, in VR la distanza egocentrica è fin troppo compressa. Sebbene la percezione spaziale veritiera non sia essenziale, da parecchio valore ad alcune esperienze in determinati campi come: architettura, automobilismo, militare e medico.

### 4.3.2 Percezione del tempo

Il tempo può essere espresso come la percezione personale o l'esperienza di cambiamenti successivi. La percezione che abbiamo di questo fenomeno è manipolabile e alterabile sotto determinate circostanze. Ad esempio, la distorsione temporale, o percezione ritardata, avviene quando il tempo dal ricevimento dello stimolo alla reazione soffre di una latenza oltre i 200 millisecondi.

Nella creazione di un'esperienza VR si tende a cercare di mantenere questo valore più basso e costante possibile, per garantire una fluidità complessiva tale da non rompere l'illusione. Un'eccessiva latenza rischia di causare la cosiddetta motion sickness, che verrà discussa ampiamente nel capitolo successivo.

### 4.3.3 Percezione del movimento

La percezione del movimento è un processo complesso che coinvolge un numero elevato di sistemi sensoriali e componenti fisiologiche. Questo tipo di percezione non è statica, ma varia continuamente nel tempo. Si è in grado di percepire il movimento anche quando gli stimoli visivi non si muovono sulla retina, come ad esempio durante l'inseguimento di un oggetto in movimento con gli occhi. Altre volte invece, gli stimoli sono in movimento ma essi non vengono percepiti (ad esempio quando gli occhi si muovono per esaminare oggetti diversi).

Un mondo percepito senza movimento può essere inquietante e pericoloso: chi soffre di achinetopsia (cecità al movimento) ha la sensazione che gli oggetti in movimento scompaiano per poi riapparire improvvisamente. Tale sensazione può essere indotta anche dalla VR nel caso in cui il sistema soffra di pesanti latenze, ciò ovviamente può comportare problemi di coerenza al sistema vestibolare dell'utente con conseguenti nausea e danni fisici.

## 4.4 Effetti nocivi della VR

Già più volte è stato accennato come alcune caratteristiche influenzino l'utente introducendo diversi tipi di disturbo. In questa sezione verranno approfondite i principali disturbi di carattere psicofisico legati alla fruizione di esperienze in realtà virtuale.

Questi problemi sono legati principalmente a caratteristiche personali degli utenti (soprattutto fisiche), per cui le soluzioni finora adottate non sono immediate, né quantomeno universali. Seppur non sarà mai possibile eliminare del tutto gli aspetti

negativi della VR, andando a capire i problemi e la loro genesi, è possibile progettare sistemi e applicazioni capaci di ridurre la loro gravità e la loro durata.

È importante notare, in ogni caso, come la risoluzione di tali problemi parta sempre dal visore. Se la base su cui le applicazioni lavorano non è buona, è quasi del tutto impossibile sviluppare un'esperienza che non soffra di tali problemi. Non è vero il viceversa

### 4.4.1 Motion Sickness

Non esiste una definizione specifica di motion sickness, in generale può essere espressa come l'effetto collaterale con segni osservabili e documentabili associati all'esposizione ad un movimento reale o apparentemente reale.<sup>[35]</sup> È classificato come uno degli effetti collaterali più comuni legati alla realtà virtuale. I sintomi sono molteplici: malessere generale, nausea, capogiri, mal di testa, disorientamento, vertigini, pallore, sudorazione, sonnolenza e nel peggiore dei casi vomito. Per rendere più semplice la comprensione di questo fenomeno, basta riportare la motion sickness al più comune "mal di mare".

Diverse cause possono essere ricondotte alla motion sickness:

**Movimento della scena:** è il movimento visuale dell'intero ambiente che normalmente non accade nella realtà, ciò può essere dovuto dall'applicazione (a volte addirittura voluto) o può essere causato dalla tecnologia quando non la macchina non ha potenza sufficiente o viene mal calibrato il sistema di tracciamento.

**Movimento di se:** il proprio movimento all'interno dello spazio virtuale di solito non dovrebbe causare motion sickness. Il malessere avviene quando la scena si muove in modo non contestuale con il movimento fisico dell'utente (in altre parole, c'è una discrepanza fra ciò che esso vede e ciò che prova), in quanto si genera una accelerazione virtuale che gli organi otoliti non provano, e, non trovando nulla con cui confrontare la velocità osservata, generano sintomi come capogiri o nausea.

Esistono diverse teorie sulla ragione per cui avviene la motion sickness, vediamo alcune:

**Teoria del conflitto sensoriale:** è la teoria maggiormente accettata, essa afferma che la motion sickness sia dovuta ad una tale alterazione dell'ambiente da rendere incompatibile ogni informazione in arrivo verso gli organi sensoriali principali (vista, udito e sistema vestibolare). Infatti i conflitti sono principalmente causati da incongruenze tra sistema vestibolare e sistema visivo. Nella maggior parte delle applicazioni in VR dove si prova motion sickness il sistema visivo percepisce un movimento ma il sistema vestibolare non lo prova.

**Teoria dell'avvelenamento:** offre una ragione per cui il movimento fa star male, è fondamentale per la sopravvivenza percepire correttamente il movimento del nostro corpo e il movimento di ciò che ci circonda.

Se vengono ricevute informazioni contrastanti dai centri sensoriali, significa che si stanno sperimentando problemi col sistema percettivo e motorio. Il corpo umano si è evoluto in maniera da proteggersi, riducendo al minimo i disturbi fisiologici prodotti da tossine assorbite, praticando l'espulsione del veleno tramite sudorazione, vomito o causando nausea e malessere, al fine di scoraggiare il ripetersi di comportamenti dannosi in futuro. La risposta associata alla motion sickness può verificarsi perché il cervello interpreta le mancate corrispondenze sensoriali come un segno di intossicazione, e innescando nausea o vomito come programma di autodifesa.

## 4.4.2 Effetti fisici e igiene

L'uso di un'apparecchiatura di realtà virtuale può portare a innumerevoli effetti fisici,<sup>[37]</sup> tra cui:

**Gorilla arms:** fuori dalla VR gli utenti mantengono una posizione rilassata delle braccia, nelle applicazioni di realtà virtuale dove si necessita un qualche tipo di controllo, spesso occorre tenere le mani frontalmente causando affaticamento e pesantezza (da qui il nome). Le applicazioni perciò dovrebbero essere progettate per riuscire ad evitare questo sintomo facendo riposare l'utente ad intervalli regolari.

**Traumi e ferite:** con l'uso dell'HMD, l'utente è ignaro di ciò che si trova in sua prossimità, e rischia continuamente collisioni con oggetti vicini e cadute dovute all'inciampare in eventuali cavi, oppure dalla perdita di equilibrio, ecc... Questi spiacevoli inconvenienti vengono comunemente evitati delimitando a priori una area sicura priva di ostacoli. Gli HMD in commercio permettono inoltre di visualizzare i confini all'interno dell'applicazione proiettando un reticolato quando l'utente si avvicina troppo ai bordi. Un altro sistema comune ai visori dotati di telecamere consiste nella visualizzazione di un'immagine in trasparenza di ciò che è presente nelle vicinanze qualora esso si avvicini troppo. Quest'ultima tecnica è molto utile per attirare l'attenzione dell'utente immerso, direttamente dall'esterno.

**Igiene:** l'HMD è un dispositivo fisico capace di trasportare agenti patogeni come batteri, virus, o funghi, trasmissibili quando sono presenti svariati utilizzatori per lo stesso HMD. La soluzione consigliata dai produttori, è la pulizia costante dell'HMD e la sostituzione dei cuscinetti che vanno a contatto col volto.

## 4.5 Linee guida alla progettazione in VR

L'unione dei concetti appresi dai questo capitolo e dai precedenti, costituisce una buona base per stabilire come un'applicazione in VR debba essere progettata:

- Focalizzare le risorse sulla User Experience piuttosto che sui tecnicismi;
- Semplificare e armonizzare la comunicazione tra utente e tecnologia (ad esempio per mezzo di interfacce diegetiche o controlli semplici);
- Scegliere quale forma di realtà si vuole creare e quale sistema di input/output usare;
- Creare una storia concettualmente forte per immergere l'utente nell'applicazione (ove necessaria);
- Massimizzare l'immersione, così da poter abbassare la soglia di inizio della sensazione di presenza dell'utente;
- Minimizzare le rotture della presenza;
- Porre particolare attenzione alla stabilità dell'ambiente virtuale;
- Concentrarsi sull'interazione fisica;
- Gestire la comunicazione sociale;
- Pensare in termini di Spazio personale, di azione e di vista quando si modellano gli ambienti e le interazioni;
- Usare diverse tecniche di indicazione della profondità, come l'aggiunta di ombre o tecniche di occlusione per aumentare il riconoscimento dello spazio dell'utente e di conseguenza la presenza;
- Usare interfacce diegetiche laddove possibile;
- Evitare qualsiasi tipo di ritardo nei movimenti o discrepanza temporale per evitare effetti quali motion sickness.



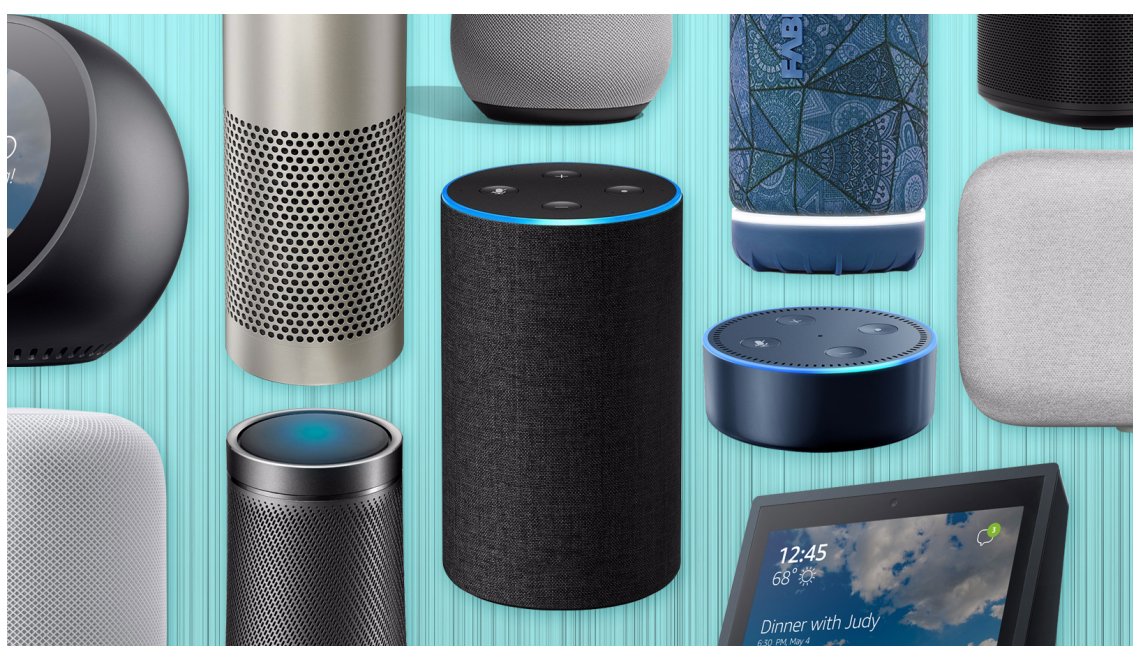
**Parte II**  
**Virtual Assistant**





## Capitolo 5

# Stato dell'arte e campi d'uso



L'assistente virtuale (anche detto IVA, o Intelligent Virtual Assistant) è un software che interpreta il linguaggio naturale e dialoga con degli interlocutori umani per fornire loro informazioni o compiere determinate operazioni, riuscendo a riconoscere l'utente come entità unica, con delle preferenze associate e delle informazioni che caratterizzano l'utilizzo. L'assistente virtuale può essere usato in vari ambiti, come all'interno di siti Internet per facilitare la navigazione degli utenti o come assistente personale in dispositivi mobili, il primo in quest'ultima categoria è stato Siri, introdotto come feature per l'iPhone 4S nel 2011. L'assistente virtuale è attualmente considerato uno dei più importanti avanzamenti tecnologici recenti. Gli esperti dicono che queste "esperienze digitali" potrebbero avere un'importanza pari se non addirittura maggiore a quelle reali. Il trend è confermato da un incremento sia nell'adozione che nell'uso da parte degli utenti (nel 2017, il numero di utenti complessivo a livello globale era stimato attorno al miliardo). Aumenta inoltre in numero di settori nei quali queste tecnologie vengono impiegate (tra cui i già citati automobilistico, telecomunicazioni, educazione e, ovviamente, retailing). Il mercato

per le tecnologie di voce recognition è stimato crescere del 34.9% globalmente nel periodo tra 2016-2024, raggiungendo un valore di 7.5 miliardi di Dollari, secondo uno studio della Ovum.<sup>[38]</sup>

Esistono diverse tipologie di dialogo con i virtual assistants:

**Testo (online chat):** principalmente bot di messaggistica

**Voce:** Amazon Alexa<sup>[39]</sup>, Siri<sup>[40]</sup> di Apple, Google Assistant<sup>[41]</sup> per Android, Snips<sup>[42]</sup> o Mycroft<sup>[43]</sup> (aziende open e indipendenti)

**Immagini scattate o caricate:** Samsung Bixby e Google Assistant per Android

Alcuni di questi assistenti utilizzano più metodi (come ad esempio Google Assistant con cui possiamo dialogare a voce o via testo). Gli assistenti utilizzano un processo chiamato Natural Language Processing (NLP) per convertire l'input ricevuto dall'utente in comandi eseguibili. Alcuni di essi continuano ad imparare attraverso l'uso di tecniche di AI (tra cui Machine Learning) per affinare la capacità di ascolto, talvolta tarandosi su uno specifico utilizzatore. Per attivare un assistente virtuale usando la voce dovremo dapprima usare una "parola d'ordine" detta Wake Word. Questa servirà per avviare l'ascolto senza incorrere in violazioni della privacy, alcuni esempi: "Hey Siri", "OK Google", "Alexa", etc...

Gli assistenti possono essere integrati in diversi tipi di piattaforme:

**Smart Speakers:** dispositivi con diverse capacità da usare in casa tramite comandi vocali (Amazon Echo, Google Home o Apple HomePod) App di Messaggistica, dalle più generiche (Facebook M Virtual Assistant) a quelle più specifiche per compagnie o organizzazioni (Aerobot di Aeromexico, WeChat Secretary o Dom di Domino's)

**Sistemi Operativi:** diversi ora offrono servizi di voce assistant integrati direttamente nel sistema per garantire più funzionalità (di nuovo Google Assistant su Android, Cortana per Windows o Siri nell'ecosistema Apple)

**Dispositivi generici:** ad esempio case produttrici di auto, wearable o smart home devices offrono servizi integrati e personalizzati (spesso molto limitati alle stesse funzioni del dispositivo ospite)

Gli assistenti offrono svariati servizi, tra cui: raccolta di informazioni, meteo, fatti (Wikipedia, IMDb, ect), ricerche on-line, news, azioni di pianificazione, creare appuntamenti, impostare sveglie, liste della spesa, riprodurre musica o video da servizi di streaming. In tutt'altra categoria rientrano gli assistenti per le pubbliche amministrazioni o che offrono un servizio complementare o sostitutivo per il customer service umano (es. risponditorie autonome nei call center) oppure per il conversational Commerce: una forma di e-commerce basato sull'iterazione tra un utente e un mezzo di dialogo, esso potrebbe essere ad esempio un commesso (o un bot) in live chat con il cliente ma ai fini della trattazione è più interessante l'impiego del voice assistant.

Gli utenti usano gli assistenti principalmente per l'intrattenimento, attraverso l'audio, la musica, ma anche come acceleratore per le tecnologie smart home e la gestione della casa attraverso la home automation. Altre attività spaziano dall'aggiunta di un timer in cucina all'ascolto di informazioni e news, all'aggiunta di

qualcosa alla propria lista della spesa. Anche i bambini interagiscono con Alexa: ci sono skills pensate proprio per la famiglia. L'interfaccia vocale diventa è un momento di inclusione collettiva, a differenza di altre tecnologie (come lo smartphone) più "divisive" e solitarie.

A questo punto della discussione, è interessante capire come l'utente percepisce uno smart assistant. In generale, quando si vuole misurare il livello di accettazione di una determinata tecnologia si interroga un campione di utenti usando il TAM (Technology Acceptance Model). Grazie a questo modello è possibile capire quali fattori influenzano l'accettazione e come l'utente percepisce l'utilità di un sistema e a sua semplicità d'uso.

Nell'interesse della trattazione sono stati analizzati degli studi<sup>[44][45]</sup> riguardanti gli smart assistant: essi affermano che, mediamente, l'atteggiamento dell'utente, la percezione di quanto utile possa risultare nell'uso e quindi la volontà d'uso risultano essere alti e positivamente influenzati tra di loro. Influenzano negativamente pochi fattori legati principalmente ad aspetti sociali della tecnologia ed altri fattori esterni.

Nei risultati viene evidenziato come non venga fatta una assoluta preferenza rispetto ai normali sistemi di interazione uomo-macchina, ma questo non presuppone che questa visione sia infondata o insensata. Al migliorare delle tecnologie (specie della loro precisione) è possibile che questa diventi il principale metodo d'interfaccia, una volta raggiunto un sistema d'uso detto "naturale" o "attuale".

## 5.1 Voice Commands vs Voice Assistant

Un'importante distinzione va fatta tra le due tecnologie che solitamente vengono confuse o usate in maniera intercambiabile. Il voice assistant è estremamente più complesso, tant'è che i comandi vocali sono solo una componente dell'intero sistema, più precisamente è il mezzo che interfaccia l'utente con l'assistente. La vera potenza dell'assistente virtuale si esprime quando diverse informazioni vengono richieste dall'utente, in particolare quando esse provengono dall'esterno.

Grazie a complessi algoritmi di machine learning, che fanno uso di dati raccolti durante le iterazioni con l'utente, l'assistente vocale riesce a portare in evidenza risultati più pertinenti, sfruttando quanto appreso tramite algoritmi predittivi. Inoltre, un'altra caratteristica saliente del VA è la capacità di intuire quando un'interazione può portare naturalmente ad un'altra, ad esempio: richiedere informazioni per un volo può suggerire la prenotazione di un albergo, oppure l'acquisto di un articolo può portare ad altri articoli correlati.

Il comando vocale, invece, richiede che tutta la logica sia gestita internamente all'applicazione, perdendo del tutto la centralità delle informazioni e aumentando il carico computazionale dell'host dell'esperienza, inoltre, ciò che può essere portato a compimento dal comando vocale è estremamente limitato, non mantiene il contesto e non risolve problemi all'atto pratico, bensì metterà in moto quello che poi si occuperà di svolgere l'azione.

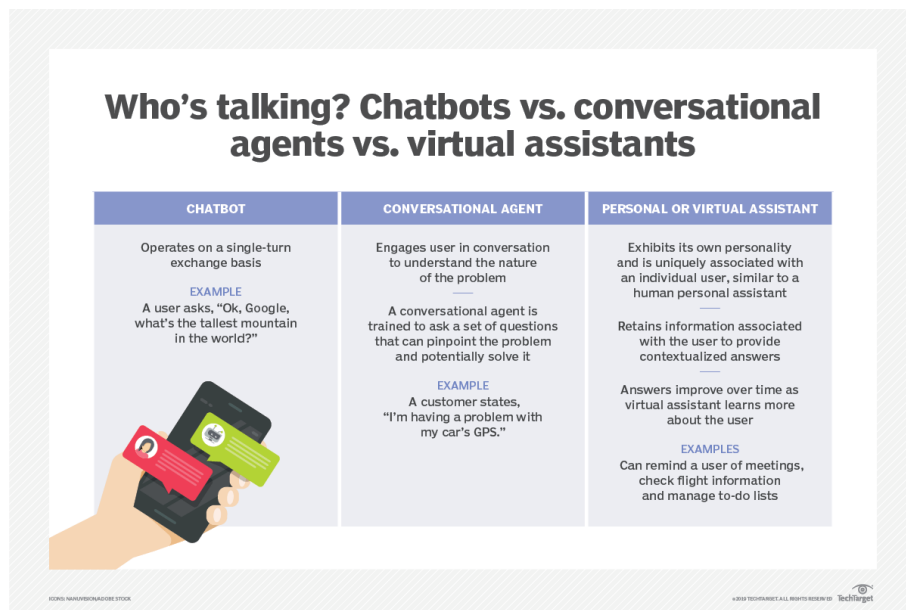


Figura 5.1: Infografica riassuntiva delle principali caratteristiche riguardanti le varie tipologie di assistente

## 5.2 VR + AR

La domanda che viene naturale porsi è se sia già stato tentato di unire le due tecnologie. Dalla letteratura in materia, è possibile ricondurre studi inerenti all'unione di assistenti virtuali ad ambiti immersivi ad articoli risalenti al 2007,<sup>[46]</sup> in questo articolo chiaramente il contesto tecnologico è diverso in quanto al tempo non erano disponibili in commercio soluzioni compatte per la VR. Oltre questo e altri pochi casi la letteratura a riguardo è piuttosto scarsa.

Tutto ciò che si è riuscito a trovare di simile, sono state: proposte di assistenti nella XR (mixed reality, un termine per indicare la generica esperienza che include AR, VR, etc...) senza uso di algoritmi di IA, per cui non classificabili come smart assistant, esperienze di VR che fanno uso di comandi vocali come mezzo di interazione o, ancora, AI in ambito VR con il solo scopo di sostenere un dialogo con un utente. In ogni caso non sembrerebbe esserci soluzione assolutamente identica a quella proposta in questo documento.

Anche nel mercato abbiamo diverse iterazioni di applicazioni che cercano di aprire alla possibilità di unire VA e VR, quella che si avvicina di più è senz'altro Amazon Sumerian, anche se con le sue differenze: in questo caso abbiamo la presenza di un assistente (con forma umana) in ambiente virtuale, esso riesce a rispondere a specifiche richieste tramite feedback vocale, ma di base non riesce a ricevere comandi vocali e non ha funzioni di smart assistant.

Volendo fare un parallelo tra mercato e ricerca, uno sta cercando di comprendere cosa la clientela si aspetta e cosa vorrebbe ottenere da un'esperienza VR cercando di andare in contro alle sue esigenze, mentre la tendenza dell'altro invece è quella di cercare di comprendere quale possa essere la prossima direzione e quanto in là è possibile spingersi con queste tecnologie.

## Capitolo 6

# Panoramica delle alternative

In questo capitolo si introdurranno le principali alternative disponibili in questa categoria.

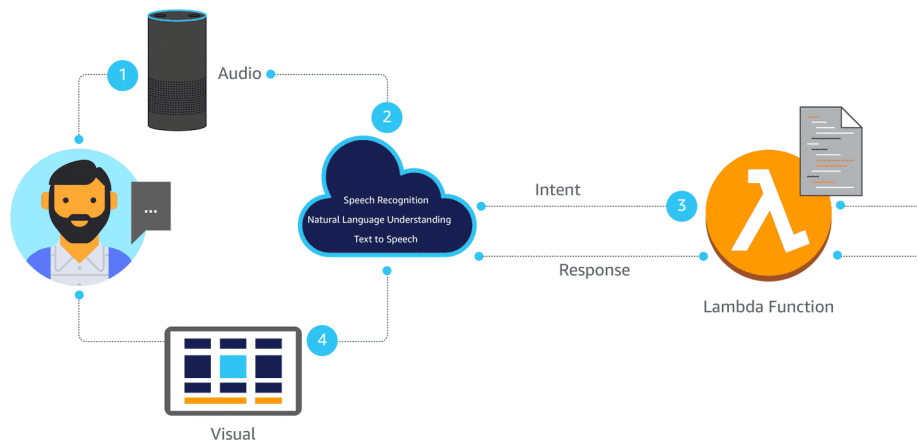
Siccome non esistono metriche precise per valutare la bontà di un assistente virtuale, solitamente ci si orienta nella scelta in base alle necessità specifiche e in base all'impiego che si intende fare dell'assistente virtuale. Un altro fattore di rilevante importanza risiede nel dispositivo che ospita l'assistente virtuale, in quanto molti sono dipendenti dalla piattaforma.

Nella lista sono presenti le offerte dei principali competitors del settore, accompagnate infine da alcune delle più rilevanti alternative open source. Gli assistenti virtuali open source si contraddistinguono per avere una maggiore libertà nella progettazione dell'esperienza utente, un buon numero di piattaforme compatibili e maggiore personalizzazione. Di contro hanno una scarsa adozione, potrebbero costituire un costo maggiore per gli sviluppatori quando si necessita di integrare altri servizi, mentre per gli utenti finali potrebbero risultare troppo complessi o poco pratici, se non addirittura sprovvisti di alcune funzionalità (come il feedback vocale).

D'altro canto i voice assistant proprietari hanno come vantaggio di essere molto più precisi e di avere un'ampia suite di servizi integrabili, così da diminuire il carico di lavoro per lo sviluppatore. Inoltre, la presenza più ampia di questi assistenti nel mercato implica una migliore conoscenza del mezzo, in quanto sono maggiori le possibilità che un utente ne abbia già fatto uso, risultando quindi più familiare. Come svantaggi abbiamo una maggiore rigidità nell'approccio e scarsa personalizzazione. Inoltre, per garantire maggiori prestazioni questi assistenti non sono decentralizzati, ovvero non tutta la logica e i dati sono in locale, ma vengono salvati nei server dei produttori. Questo approccio ha fatto scaturire diversi dibattiti in materia di privacy.<sup>[47]</sup> Ad esempio: recentemente alcune grandi multinazionali hanno ammesso di ascoltare a campione alcune registrazioni degli utenti per valutare e migliorare il servizio, questa notizia ha fatto talmente scandalo da aver reso possibile per l'utente di negare il consenso all'uso dei dati raccolti dall'assistente per finalità legate allo sviluppo.

Per gli utenti interessati a queste tematiche, sicuramente sono più indicati gli assistenti open source, che operano principalmente on-device, ovvero sul dispositivo, garantendo un maggiore livello protezione a scapito magari di precisione e performance.

## 6.1 Alexa



**Figura 6.1:** Il modello di interazione di Alexa

Alexa è lo strumento che Amazon ha pensato per offrire agli utenti la possibilità di interagire tramite comandi vocali con algoritmi in grado di aprire ad una lunga serie di servizi.<sup>[39]</sup> L'idea di Amazon è quella di un servizio in continua evoluzione, che può essere continuamente arricchito e che, proprio tramite le interazioni con l'utente, è in grado di apprendere nuove capacità, nuovi tipi di risposte e nuove sfumature dialettiche in grado di aumentare l'empatia tra la macchina e la persona.

Alexa è dunque l'insieme di algoritmi e funzioni che stanno dietro i dispositivi con cui Amazon ha dato corpo al proprio assistente, la cosiddetta linea Echo. Questa serie di prodotti si arricchisce e migliora periodicamente, attualmente è arrivata a comprendere smart glasses, speakers, cornici digitali e prese intelligenti. Inoltre, attraverso l'uso dell'apposita app è possibile utilizzare Alexa su qualsiasi smartphone e sistema operativo (Windows ha la sua app dedicata mentre macOS e Linux hanno un client da terminale).

Una caratteristica molto importante di Alexa è la possibilità di integrare delle Skills, app appositamente sviluppate su tramite le tecnologie disposte da Amazon. Ottenibili come le comuni app attraverso il loro specifico market, le Skills consentono di ampliare le capacità collegandovi servizi esterni (musica, notizie, video, etc...) o creando esperienze uniche basate sull'uso della voce, tra cui integrazioni domotiche.

Questo voice assistant, come già accennato, si appoggia ai servizi in cloud di Amazon, tra cui Polly per la sintesi vocale e lo stack Cloud Formation per la base funzionale.

Al momento, secondo le statistiche,<sup>[48]</sup> gli Echo device sono gli assistenti più venduti, vantando un market share del 69% (con quasi 100 milioni di unità vendute), dovuto principalmente alla loro distribuzione su svariate fasce di prezzo.

Dopo le recenti polemiche in materia di privacy, Amazon applica una policy chiamata "opt-out privacy", ovvero se l'utente decide di non dare il consenso all'ascolto della propria voce, allora ciò non avverrà.

## 6.2 Google Assistant



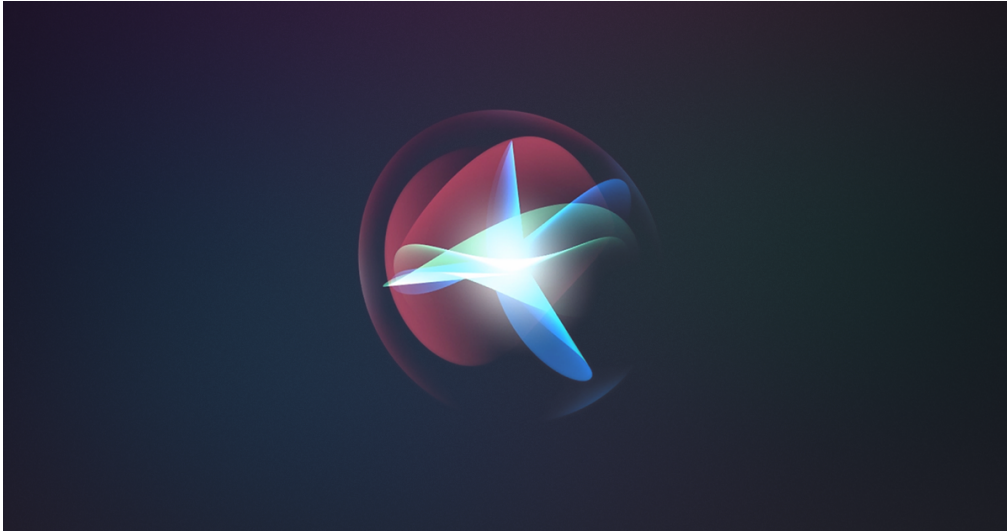
Molto simile ad Alexa, Google Assistant è un assistente con comandi vocali, integrato su più dispositivi.<sup>[41]</sup> In questo caso, l'assistente è integrato in tutti gli smartphone Android e in tutti i prodotti hardware e software a marchio Google. In base alla versione dell'assistente, ma anche a seconda del dispositivo, Google Assistant ha capacità diverse. Questa scelta rende le versioni in circolazione molto frammentate, ma al contempo, consente anche ai dispositivi più datati di godere di funzionalità di voice assistant.

L'assistente di Google è anche capace di ottenere informazioni tramite testo (come un chatbot), tramite immagini e nelle versioni più recenti integra un servizio chiamato Google Duplex, capace di sostenere una conversazione di base con un operatore umano (ad esempio per prenotare in un ristorante). Un'altra interessante integrazione, stavolta nel campo dell'AR, consente Google Assistant di ottenere informazioni dalle vicinanze tramite videocamera.

Un grande selling point è senz'altro l'integrazione nativa con gli algoritmi di ricerca di Google e tutte le app native degli smartphone. Al momento non consente di sviluppare comandi o app personalizzate. Le uniche integrazioni che si hanno tramite il "Google Assistant SDK" sono API per applicazioni domotiche e robotiche (ancora non validato per la commercializzazione).



## 6.3 Siri



Siri è il voice assistant responsabile per il rilancio dell'interesse riguardante questa tecnologia.<sup>[40]</sup> Nel 2011 è stato lanciato come servizio integrato nell'iPhone 4S e da allora viene reso disponibile con ogni nuova iterazione dei prodotti Apple. Di recente è stato aggiunto all'ecosistema Apple lo smart speaker HomePod, che usa esclusivamente Siri per ricevere comandi e dialogare con gli altri dispositivi.

A differenza degli altri assistenti, Siri è stata sviluppata esclusivamente per i prodotti Apple, senza possibilità di essere integrata in altri devices. Un'altra caratteristica che differenzia Siri è la volontà da parte degli sviluppatori di lasciare maggior controllo sui dati raccolti dall'assistente, delegando parte dell'elaborazione on-device, grazie ad appositi coprocessori specializzati in machine learning.

Siri non consente di sviluppare applicazioni vere e proprie come Alexa, ma da la possibilità di integrare comandi personalizzati sia agli sviluppatori, che agli utenti, che con un po' creatività potranno usare il visual scripting dell'app Comandi per creare le proprie shortcut ed abilitarne l'uso con l'assistente vocale. Inoltre, l'integrazione domotica è particolare rispetto i suoi competitors, perché per motivi di sicurezza e qualità legati al brand, non tutti possono sviluppare applicazioni compatibili con dispositivi Apple (e quindi con Siri) se non aderendo ad HomeKit, protocollo proprietario per la domotica.

## 6.4 Mycroft AI



# MYCROFT AI

Mycroft è il primo assistente virtuale open source, è scritto interamente in Python ed è disponibile gratuitamente su più piattaforme: desktop, Raspberry Pi (un mini computer per applicazioni IoT) ed un dispositivo proprietario chiamato Mark.<sup>[43]</sup>

In quanto open source il software dell'assistente può essere modificato, esteso e migliorato. Ogni componente dell'assistente virtuale è documentata ed utilizzabile ai fini dello sviluppo (dal sistema di voice recognition a quello di machine-learning). Mycroft può essere usato per qualsiasi tipo di progetto, sia esso scolastico, scientifico o aziendale. Difatti, questo assistente riesce a soddisfare tutte le necessità dell'utente, con la capacità di aggiungere Skills da per avere un'esperienza più completa e personalizzata. A differenza di Alexa, le skills non vengono viste dall'assistente come applicazioni esterne da richiedere espressamente, ma vere e proprie estensioni dell'assistente.

Ovviamente la privacy è uno dei punti cardine del progetto, esso applica una policy chiamata "opt-in privacy", ovvero se l'utente non decide di dare il consenso all'ascolto della propria voce, allora ciò non avverrà. Questa impostazione verrà modificata solo ed esclusivamente tramite consenso esplicito dell'utente.

Attualmente Mycroft ha un'ottima reputazione ed un discreto seguito, con rilasci continui e nuovi progetti legati al prodotto.

## 6.5 Snips

Snips è un progetto open source molto particolare. Si tratta di un assistente vocale che può funzionare interamente in locale, senza necessità di connessione internet. In quanto tutta l'elaborazione della richiesta viene eseguita on-device, senza alcuna connessione, aumentando di molto la privacy dell'utente e diminuendo i requisiti che il prodotto deve avere per abilitare funzionalità vocali.<sup>[42]</sup>

Questo perché il business model di Snips è molto più puntato all'integrazione hardware, piuttosto che all'uso diretto come assistente da parte dell'utente. Ovvero l'assistente può essere introdotto su dispositivi generici per abilitare comandi vocali custom, ma che non necessitano di altre funzionalità comuni negli assistenti vocali. Un classico esempio consiste in una macchina del caffè che può funzionare tramite comandi vocali.

Per la prototipazione e sviluppo il software è disponibile su tutte le piattaforme, in particolare è possibile appoggiarsi all'uso di una Raspberry Pi per testare devices IoT.

Offre la personalizzazione e la flessibilità di un progetto open source, con un'ottima documentazione a supporto e la possibilità di registrare il proprio dispositivo (a pagamento) per integrarvi Snips e commercializzarlo.

Chiaramente la scelta di lasciare tutto on-device ha un peso in termini di capacità dell'assistente: tutto è lasciato allo sviluppatore a cui vengono messe a disposizione delle semplici API, sarà poi suo il compito di sviluppare le iterazioni, anche quelle più basilari.

Inoltre, come si poteva intuire, Snips non analizza né trasferisce dati nel cloud, rendendo l'applicazione estremamente sicura ed in regola con le regolamentazioni più stringenti nel campo della privacy.

In conclusione, Snips non è un vero e proprio assistente vocale nel senso comune del termine, ma è un'applicazione interessante e degna di nota della tecnologia, che dimostra come essa possa venir applicata in contesti diversi da quelli che siamo soliti usare.

## Capitolo 7

# Concetti teorici

In questo capitolo si descriverà la tecnica adottata dai vari assistenti virtuali per il riconoscimento e la sintesi vocale. Si entrerà nello specifico di ogni parte del processo, seguendolo nel suo ordine naturale.

Laddove necessario, per motivi di chiarezza, verranno mostrati esempi di frasi o parti di esse.

In seguito si offrirà una panoramica sulla percezione dell'utente riguardo queste tecnologie, e una breve discussione sulle implicazioni che queste tecnologie hanno nel quotidiano.

### 7.1 Anatomia di un comando vocale

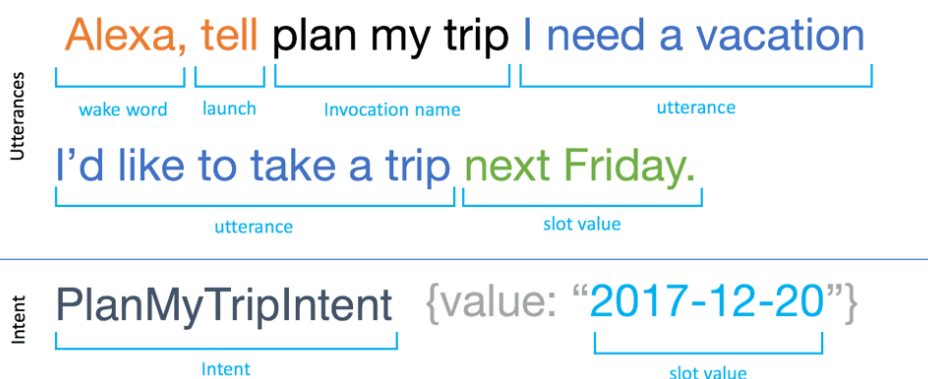


Figura 7.1: Un esempio di comando vocale

Prima di addentrarsi nel flusso seguito dagli assistenti vocali per trasformare la nostra voce in un comando, è importante capire come una voce line viene interpretata dall'assistente. Per comprendere questa struttura è bene partire da un esempio e successivamente scomporlo.

La frase presa in osservazione è: *"Riproduci musica rilassante da Spotify"*.

Un comando vocale è suddiviso in tre parti, secondo il name scheme di Alexa:<sup>[49]</sup>

**Intent:** l'obiettivo di un'interazione vocale. L'intento racchiude in se la motivazione che ha portato l'utente ad interagire con l'assistente.

L'intent può essere più o meno specifico in base alla necessità o all'azione richiesta. Nel design di una VUI è molto semplice inquadrare quali siano gli intent specifici. Molto meno semplice è inquadrare gli intent meno specifici in quanto possono risultare estremamente poco chiari e ambigui da decifrare, spesso in questo caso come risposta viene esposta una domanda per chiarire le intenzioni dell'utente. Nell'esempio l'intent è chiaramente voler ascoltare della musica.

**Utterance:** come l'utente formula un comando vocale, in italiano corrisponde ad un enunciato. Questa parte può risultare essere la più complessa, in quanto richiede un impegno notevole nel prevedere in che maniera l'utente può porre la richiesta. Nell'esempio l'utente ha detto "*Riproduci musica...*", ma avrebbe potuto dire "*Metti della musica...*", una buona previsione di questi casi aumenta la precisione complessiva dell'IA, che si occuperà di comprendere le variazioni minime della frase come accenti, plurali, errori, etc...

**Slots:** sono quelle parti variabili della richiesta che consentono di identificare specifiche risorse a cui si vuole accedere. Quando l'intento da solo non basta per raggiungere lo scopo voluto, l'utente aggiunge spontaneamente o viene portato a fornire queste informazioni. Possiamo vedere gli slot come dei campi da compilare in un modulo. Gli slot possono anche essere completamente opzionali a tal punto da non aggiungere nulla ai fini del risultato finale, ma al contempo conferisce un valore aggiuntivo al prodotto dal punto di vista dell'utente, come avviene nel caso dell'esempio con "*...rilassante...*".

Nella definizione di un comando vocale bisogna concentrarsi su:

1. Identificazione dell'intent: lo scopo preciso che l'utente vuol raggiungere, va definito precisamente e senza ambiguità.
2. Identificazione delle utterances: quali sono le parole e le frasi che gli utenti potrebbero usare per raggiungere l'intent
3. Dare la possibilità di correggere gli errori: le conversazioni di solito non sono perfette, bisogna dare all'utente la possibilità di correggersi o modificare la loro richiesta.
4. Gestire gli errori: dare una risposta che indichi quando un'informazione non è reperibile o non è stato possibile comprendere la richiesta.

## 7.2 Processo di riconoscimento vocale

Durante il ciclo di vita di un assistente vocale, il microfono del dispositivo è costantemente attivo, in attesa della cosiddetta Wake Word, una frase specifica che indica all'assistente che si sta per iniziare un dialogo con l'utente (es. "Hey Siri"). La richiesta dell'utente viene immediatamente analizzata dalle componenti che si occupano di SLU (Spoken Language Understanding<sup>[50]</sup>), un processo che permette una traduzione uomo-macchina. Per primo il motore di ASR (Automatic Speech Recognition<sup>[51]</sup>) ascolta la frase e la trasla in semplice testo, questa tecnologia è ciò che si occupa, in tutti i moderni smartphone, della dettatura vocale. In seguito, il risultato di questa elaborazione viene dato in pasto all'algoritmo di NLU (Natural Language Understanding<sup>[52]</sup>) che classifica e trascrive il linguaggio in categorie di significato. Questo processo è reso particolarmente difficile e complesso a causa delle caratteristiche intrinseche di ambiguità del linguaggio umano.

Per questo motivo il processo di elaborazione viene suddiviso in fasi diverse, tuttavia simili a quelle che si possono incontrare nel processo di elaborazione di un linguaggio di programmazione:

**Analisi lessicale:** scomposizione di un'espressione linguistica in "parole"

**Analisi grammaticale:** associazione delle parti del discorso a ciascuna parola

**Analisi sintattica:** arrangiamento dei token in una struttura sintattica

**Analisi semantica:** assegnazione di un significato (semantico) alla struttura sintattica e, di conseguenza, all'espressione linguistica. Nell'analisi semantica la procedura automatica che attribuisce all'espressione linguistica un significato tra i diversi possibili è detta disambiguazione.

Una volta ottenuto un comando comprensibile dalla macchina esso viene elaborato dalla business logic della Skill. Alla fine di questa elaborazione viene fornito un output, che solitamente, viene presentato all'utente sotto forma di risposta vocale, tramite un algoritmo di TTS (Text to Speech<sup>[53]</sup>). Programmi appositi di sintesi vocali riescono a fornire un dettaglio di pronuncia molto vicino a quello naturale. I sistemi di TTS più avanzati riescono ad ampliare le capacità del linguaggio grazie ad appositi linguaggi di markup che modulano aspetti della voce quali ad esempio tono, pronuncia, fonemi o addirittura riprodurre suoni, tracce, musica etc... Questo ultimo step è opzionale e sostituibile da un comune output testuale. Ovviamente una risposta ottenuta tramite voce e suoni conferisce un valore molto superiore all'esperienza utente.

## 7.2.1 NLU

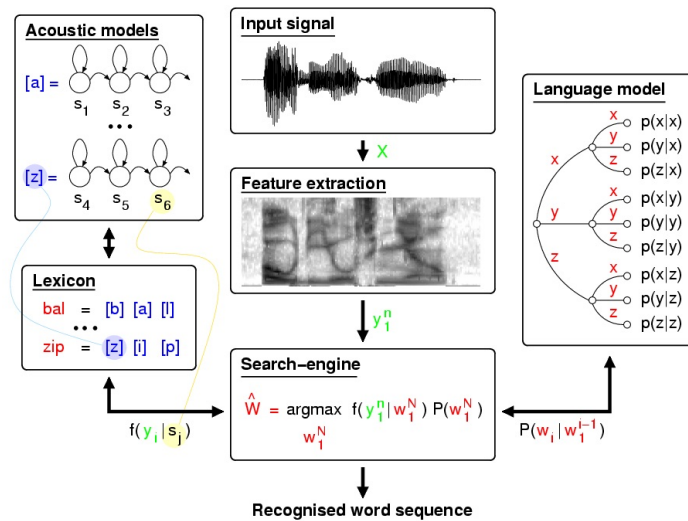


Figura 7.2: Un esempio di modello per voce recognition

Grazie al Natural Language Understanding (NLU<sup>[52]</sup>), una macchina può dedurre il significato di cosa l'interlocutore vuole, non solo di cosa sta dicendo. Difatti, anche la comunicazione tra umani è un costante esercizio di decifrazione del significato, difatti spesso si usano le parole sbagliate, volendo intendere altro. Lo scopo principale dell'NLU è fornire alle macchine il contesto necessario a comprendere di cosa l'utente ha bisogno, con la giusta flessibilità che consenta di coprire le possibili variazioni. Questa parte consente al VUI di trasformare quello che l'utente chiede in una richiesta effettivamente comprensibile dal programma.

L'NLU è basato su intelligenze artificiali incentrate sul riconoscimento di pattern e significati del linguaggio umano. Questo approccio di machine learning, consente all'assistente di migliorarsi col tempo: l'algoritmo è in grado di comprendere più variazioni del dialogo e inferire maggiormente il contesto, senza fare richieste specifiche all'utente, rendendo la comunicazione più naturale. L'apprendimento avviene in maniera costante attraverso una memoria condivisa di apprendimento, grandi sistemi centralizzati come quello di Alexa riescono ad elaborare milioni di richieste al giorno affinando di volta la precisione del modello.

## 7.3 Percezione dell'utente

In questa sezione si descriverà qual'è la percezione che l'utente ha di queste tecnologie. Si elencheranno i fattori positivi che conferiscono valore ad una VUI, e successivamente si discuteranno gli aspetti negativi che rallentano l'approcciarsi di queste tecnologie.

### 7.3.1 Fattori positivi

Il vantaggio principale di una VUI è che permette di avere le mani libere e gli occhi liberi, in cui gli utenti possono interagire con un prodotto e concentrare la loro attenzione altrove. Allo stesso modo, la costruzione di un "rapporto" di comunicazione contribuirà a migliorare la soddisfazione nei confronti dell'esperienza, quando la VUI, conoscendo sempre più la voce dell'oratore, lo ricompensa con risposte più accurate. Non c'è bisogno di avere il messaggio scritto, dato che già ne comprendiamo il contesto.

Un ulteriore vantaggio è che chiunque può usare le skills, dal bambino all'anziano, proprio perché l'unico requisito è saper parlare, non è richiesta la capacità di apprendere un'interfaccia o saper leggere.

Nel loro libro sull'interazione vocale, *Wired for Speech*,<sup>[54]</sup> i ricercatori di Stanford Clifford Nass e Scott Brave sostengono che gli utenti in una certa misura si relazionano alle interfacce vocali nello stesso modo in cui si relazionano con altre persone. Più le VUI riescono a diventare naturali e conversazionali, più risulteranno essere accessibili, veloci ed efficienti.

Considerare l'uso di un'assistente vocale nel quotidiano una comodità non è immediato, questa tecnologia si è evoluta a tal punto da rendere "sconveniente" premere due pulsanti sul microonde piuttosto che dire a voce "Computer, imposta un timer per due minuti".

### 7.3.2 Privacy e Sicurezza

Nel Maggio 2018, i ricercatori dell'Università della California, Berkeley, hanno pubblicato un articolo<sup>[55]</sup> che dimostrava come comandi audio non rilevabili dall'orecchio umano potessero essere direttamente incorporati nella musica o nel testo parlato, manipolando così gli assistenti virtuali per eseguire determinate azioni senza il consenso dell'utente. I ricercatori hanno apportato piccole modifiche ai file audio nei quali sono stati inseriti suoni che, ingannando i sistemi di interpretazione del sistema, consentono di comporre numeri di telefono, aprire siti web o anche trasferire denaro. Questa possibilità è nota dal 2016, e riguarda i dispositivi Apple, Amazon e Google, solo di recente sono state applicate delle patch migliorative.

Oltre alle azioni non intenzionali e alla registrazione vocale, un altro rischio per la sicurezza e la privacy associato agli assistenti virtuali intelligenti è rappresentato dai comandi vocali dannosi: un aggressore che impersona un utente ed emette comandi vocali dannosi, per esempio sbloccare una porta intelligente per ottenere l'accesso non autorizzato a casa o in garage o ordinare articoli online senza che l'utente ne sia a conoscenza. Anche se alcune IVA forniscono una funzione di formazione vocale per prevenire tale impersonificazione, può essere difficile per il sistema distinguere tra voci simili. Così, una malintenzionato in grado di accedere a un dispositivo abilitato all'IVA potrebbe essere in grado di ingannare il sistema facendogli credere di essere il vero proprietario e di compiere atti criminali o maliziosi.

Un ulteriore aspetto che coinvolge la privacy dell'utente è strettamente collegato alla natura centralizzata di alcuni assistenti. In particolare, nell'ultimo anno, dopo



la manifestazione di perplessità riguardante il trattamento dei dati da parte dei colossi come Google o Amazon, è stato reso noto come molte delle richieste degli utenti venissero registrate e rese disponibili agli sviluppatori per effettuare una "peer review", ovvero una sorta di verifica incrociata, utile a comprendere se l'NLU riuscisse a fare le giuste rilevazioni.<sup>[47]</sup>

Chiaramente queste dichiarazioni hanno generato parecchio scandalo, in quanto, potenzialmente, non solo è possibile ascoltare il comando dall'audio, ma anche qualsiasi altro suono o discorso presente in background. Per non parlare poi dei cosiddetti "false trigger", ovvero rilevazioni erronee della wake word, che potrebbero registrare interi dialoghi senza il minimo sforzo. Inoltre, lo stesso problema è rilevabile anche a livello personale, se un dispositivo è usato da più persone tutti gli utenti che hanno accesso alla dashboard possono ascoltare le registrazioni passate che l'assistente ha effettuato, nonché quali richieste sono state effettuate.

Per dovere di cronaca, a seguito di questi scandali Amazon ha integrato un'opzione per permettere a chi vuole di togliere il consenso all'uso dei propri dati per finalità di riascolto, e Google ha, nell'ultimissima versione di Assistant, minimizzato le dimensioni dell'IVA a pochi mega e ha spostato tutta la logica on-device, senza mandare nulla ai server. Seppur molto apprezzabili questi sforzi risolvono solo parzialmente il problema, e tuttora questo aspetto è il più grande turn-over legato a questa tecnologia. Una vera risposta sarebbe passare all'uso di assistenti open, ma la loro scarsa pubblicizzazione e difficoltà di reperimento non convincono gli utenti a cui solitamente queste tecnologie sono indirizzate.

A questi problemi di sicurezza si aggiungono alcune problematiche più "d'immagine", ovvero l'utente ha ancora la percezione che l'assistente vocale sia qualcosa di estraneo e raramente viene usato all'infuori della propria abitazione, in quanto la sensazione di "parlare al vento" è molto accusata in pubblico.<sup>[56]</sup> Inoltre, in molte situazioni l'uso della voce non è possibile o è sconsigliabile (ad esempio nel mezzo di una riunione, in treno o nelle attività sportive...).

Inoltre, questa è la tipica tecnologia che o la si ama o la si odia, e tutto dipende dalle prime esperienze. Se già dai primi utilizzi l'assistente non capta bene i comandi (talvolta per motivi non inerenti all'IVA in se) gli utenti finiscono per provare frustrazione, spezzando qualsiasi tipo di immersione e diventando poco collaborativi.



**Parte III**  
**Realizzazione di ViVi**



## Capitolo 8

# Analisi delle problematiche

In questo capitolo si introducono i vari problemi, la vision e l'obiettivo da raggiungere.

Come già accennato, nel virtual retailing la realtà virtuale trova svariati casi d'uso: dallo store design alle camminate interattive all'interno del negozio stesso, il suo principale scopo è quello di mostrare i dati dei prodotti nel contesto dello shop, piuttosto che con le sole pagine statiche. Ciò che interessa maggiormente è la possibilità per l'utente di interagire con la scena e completare acquisti all'interno di essa. Durante l'esperienza di shopping possiamo imbatterci in una serie di problematiche legate alla struttura dei negozi che visitiamo ma anche all'esperienza stessa di visita e acquisto. Da notare che i problemi saranno esposti partendo dalla realtà per poi essere trasposti nell'ambito virtuale.

### 8.1 Sfide strutturali

Questi problemi sono legati alla forma del negozio, se è troppo grande e dispersivo c'è il rischio che l'utente si smarrisca tra gli scaffali, se invece è troppo piccolo si potrebbe trovare una proposta troppo esigua di prodotti. In generale questi problemi vengono studiati e approcciati in maniera differente a seconda del brand che si vuol rappresentare.<sup>[57]</sup> Ma quando questa realtà deve entrare in contatto con lo shopping online e in particolare con la realtà virtuale subentrano nuove sfide e molto spesso è necessario cambiare totalmente approccio. Ambienti molto piccoli o articoli esposti in una specifica maniera per cercare di rendere più semplice l'uso dell'applicazione vanno a limitare la presentazione e l'unicità dell'esperienza che l'utente si aspetta da determinati brand. Basti prendere catene normalmente associate all'idea di un negozi molto grandi, con capi disposti con un criterio, costretti a dover modificare il proprio formato per non essere troppo dispersivo rischiando di sembrare più un museo o una boutique che uno store. A questo si affianca anche un fattore di asetticità e solitudine derivante dall'ideazione stessa dell'ambiente: un cliente quasi sicuramente sarà a disagio nel trovarsi in un ambiente sconosciuto da solo con i prodotti.<sup>[58]</sup> In sintesi se un utente cerca di replicare l'esperienza di un tradizionale acquisto in negozio non vuole imbattersi in qualcosa di sconosciuto e meno familiare, ma in una replica quanto più fedele con un certo margine di miglioramento.

## 8.2 Carenza di informazioni

Questi problemi sono legati alla presentazione e alla preparazione su alcune tematiche importanti al cliente. Spesso le informazioni (es. assistenza, dettagli di un prodotto, conversione di taglie e misure, etc...) che si cercano non sono immediatamente disponibili e chi è addetto purtroppo potrebbe esserne carente o addirittura non disporne affatto, in quanto basate su una conoscenza estremamente specifica o personale. Un ulteriore fattore di smarrimento potrebbe essere la primissima visita al negozio e la poca esperienza da parte dell'utente. Di nuovo nella trasposizione in realtà virtuale anche quelle poche informazioni che possiamo mettere a disposizione nelle maniere tradizionali vanno riviste nella loro presentazione, risultando in problemi di UI/UX ben noti come l'interface overload o altre scelte poco intuitive per concentrare una mole maggiore di dati.<sup>[59]</sup> Tuttavia maggiori sono i dettagli nell'applicazione e meno si perde sul fronte dell'immersione, in quanto l'utente non sarà costretto a togliere il visore per recuperarli.

È un peccato, inoltre, che si sia persa la figura dell'assistente/avatar nel mondo della UI/UX, in favore delle eteree voci senza forma che combaciano maggiormente con l'innalzarsi delle aspettative verso queste tecnologie che cercano di essere onniscenti e infallibili.<sup>[60]</sup> Esso portava una sottile distrazione e un aiuto concreto in molte situazioni; grazie a questo utenti poco esperti potevano usare il programma con più "leggerezza" e ricevere un'introduzione all'utilizzo dei servizi, rendendo l'esperienza in generale meno ostica. Da questa riflessione l'ispirazione che ha dato il via a questa tesi.



## Capitolo 9

# Approccio e soluzioni

Nel capitolo saranno spiegati i principi seguiti nell'implementazione così da dare una panoramica del prodotto finale e del risultato che si vuole ottenere: l'obiettivo dello sviluppo è rendere l'esperienza quanto più piacevole all'utente, implementando le proposte e i principi in discussione nei capitoli precedenti.

La proposta consiste principalmente in un avatar che si ponga come VUI (Voice User Interface), ed in un'esperienza incentrata intorno ad esso. È importante che l'assistente abbia un aspetto rassicurante e amichevole per stimolare l'interazione. Esso dovrà essere sempre nel campo visivo dell'utente, se non specificato il contrario, per cui è essenziale che l'interfaccia sia poco invasiva. Fornirà istruzioni sull'utilizzo della piattaforma di shopping, comandi e funzionalità. Ovviamente sarà in grado di finalizzare l'acquisto dei prodotti selezionati.

Verrà fatto uso di una serie di comandi vocali e input fisici, piuttosto che dell'immissione di testo in dei campi tramite tastiere virtuali o selettori. Ci si avvarrà principalmente dell'uso della voce per dare e ricevere informazioni, a supporto verrà anche una schermata a scomparsa per presentare informazioni specifiche o visive di alcuni elementi (sia riguardanti l'applicazione che gli articoli). Adottando queste semplici linee guida per l'interfaccia, l'utente si sentirà più portato ad interagire nella realtà virtuale in maniera naturale, ed inoltre si libererà l'uso di una mano, delegando all'altra funzioni base come lo spostamento e la selezione (diminuendo così anche l'effetto gorilla arms).

L'utente, già avvezzo all'uso di assistenti intelligenti (come Siri, Google Assistant o Alexa per l'appunto), non vedrà introdotte ulteriori difficoltà nell'utilizzo, bensì sfrutterà due tecnologie coesistenti per una esperienza innovativa e di immediata comprensione. Al contempo il cliente avrà la consapevolezza di poter ricevere un aiuto esperto in un ambito specifico, che da sempre è motivo di brand affection.

Cambia inoltre la presentazione del dato in favore di un'esperienza più "umana" e familiare. Migliora anche l'aspetto legato alla solitudine nella realtà virtuale. L'introduzione di un'avatar amichevole all'utente e sempre presente può rendere l'esperienza piacevole ed attenuare il senso di alienazione nei confronti di uno store asettico e privo di personale.

Questa applicazione è da considerarsi un prototipo che potrà essere poi ampliato e migliorato in seguito.



## 9.1 Panoramica

Innanzitutto descriviamo la struttura delle varie componenti dell'applicazione, sono presenti più ambienti che dovranno riuscire a dialogare tra loro: abbiamo la scena interna alla VR, che si occupa della presentazione visiva e un di fornire un primo metodo di interazione; ad essa dovrà essere collegata la business logic comprensiva di voice assistant, il quale ascolterà in presa diretta le richieste dell'utente.<sup>[36]</sup>

È importante tenere in considerazione che l'applicazione deve essere accessibile ad utenti con diversa esperienza e demografica, per cui deve risultare semplice ed intuitiva. Per i medesimi motivi i testi devono essere ben leggibili e le interfacce chiare ed intuitive.<sup>[59]</sup> È chiaro che, per i fini della trattazione, non è necessario soffermarsi troppo sui particolari riguardanti un ambito complesso come UI/UX (la disciplina che tratta di interfacce grafiche e esperienza utente), per cui il prodotto finale potrebbe risultare grezzo, ma questo non influisce pesantemente sul risultato finale.

La scena, sviluppata in Unity,<sup>[61]</sup> dovrà essere di dimensioni discrete con una prima area che consentirà l'utente di familiarizzare con i comandi, in particolare quelli di movimento. Spostandosi nella struttura principale si potrà scegliere tra diversi articoli esposti. Si avrà a disposizione un'interfaccia per visualizzare le informazioni sull'articolo (nome, costo, descrizione, etc...) e per acquistare.

L'utente all'interno della scena deve poter eseguire diverse azioni:

**Movimento:** tramite teletrasporto hold and release, una tecnica che consiste nella pressione prolungata del tasto centrale, per indicare la posizione (puntando il controller), e il successivo rilascio del pulsante che farà avvenire lo spostamento, quando possibile. Per evitare che l'utente si spinga troppo "fuori dai binari" verranno limitate le sue possibilità di movimento nei termini di area attiva per il teletrasporto. Il movimento è l'unico caso di azione esclusiva alla VR. Nella progettazione di tutti i comandi si deve sempre considerare la possibilità dell'uso della voce, quindi ogni operazione deve essere eseguibile sia dai controllers, sia tramite comandi vocali.

**Selezione:** tramite la pressione del trigger (il pulsante sottostante) l'utente potrà selezionare l'oggetto desiderato indicato tramite crosshair (un puntatore fisso al centro della visuale che evidenzierà cosa è selezionabile). Questo schema comportamentale è seguito anche dai pulsanti dell'interfaccia. La scelta di questo meccanismo è data dal fatto che attualmente la maggior parte degli utenti è nuova alla VR e può venire facilmente spaesata da comandi troppo complessi. Questa implementazione consente all'utente di fare selezioni senza utilizzare il controller come puntatore, scomodo e poco intuitivo per un neofita.

**Ottenimento informazioni:** le informazioni dei prodotti verranno visualizzate tramite interfacce che compariranno in sovrimpressione dell'articolo selezionato (queste schermate vengono anche dette HUD - Heads Up Display). Le schermate seguiranno la visuale dell'utente, in modo che rimangano sempre visibili. Esse comprendono: nome dell'articolo, prezzo e pulsante di aggiunta al carrello. Queste informazioni vengono integrate da una descrizione, presente nel HUD dedicato all'assistente. Premendo il pulsante "Aggiungi" l'articolo verrà inserito nel carrello e il relativo contatore incrementato.

**Aggiunta dell'articolo nel carrello:** toccando l'apposito tasto sarà possibile aggiungere l'articolo selezionato al carrello, un contatore apposito nell'HUD verrà incrementato. In qualsiasi momento si potrà richiedere all'assistente cosa è presente nel carrello.

**Checkout:** raggiungendo il registratore di cassa posizionato nella scena sarà possibile visualizzare il totale speso e finalizzare l'acquisto tramite una normale selezione, in alternativa si potrà usare un comando vocale. Per indicare che l'acquisto è avvenuto con successo e per gratificare l'utente, verrà riprodotto un suono di registratore di cassa e apparirà una ricompensa visiva.

Tenendo in mente queste necessità l'esperienza è strutturata come segue: una sezione all'esterno, in cui l'utente familiarizzerà con lo spostamento, ed una sezione interna dove, varcato l'ingresso, si avrà accesso ai diversi articoli che l'utente potrà selezionare e alla cassa.

## 9.2 L'assistente

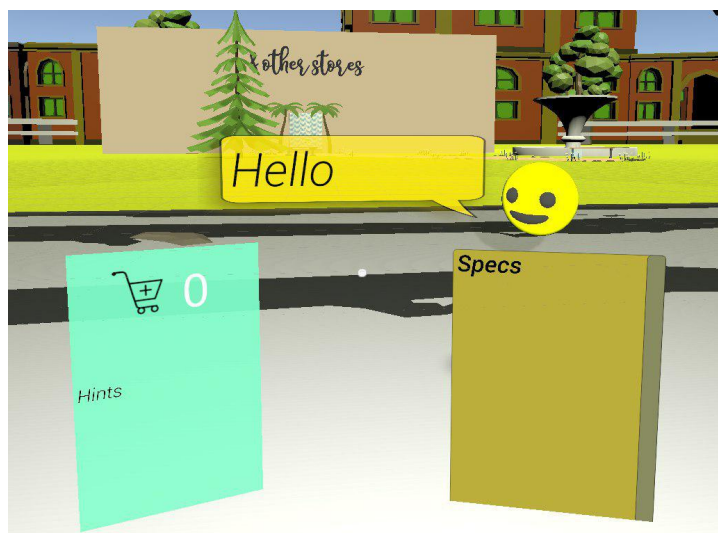


Figura 9.1: Schermata inattiva che ritrae ViVi

Per il voice assistant è stato scelto il nome ViVi, da "VirtuAl assistant in VirtuAl reality", semplice da dire e sufficientemente simile a nomi di persona realmente esistenti.

L'assistente dovrà assumere una forma amichevole, che stimoli il contatto e non risulti invasivo. Potremo rivolgergli la parola ogni volta che vorremo e ottenere una risposta tramite l'uso dell'interfaccia visiva e soprattutto sotto forma di feedback vocale. Durante le nostre azioni l'assistente sarà sempre presente in una posizione poco invasiva all'interno del FOV dell'utente.

Seguendo questi principi, ViVi assumerà la forma di un semplice smile e sarà posto in alto a destra in un layer dedicato dell'HUD. Sotto di esso avremo la schermata di descrizione dell'articolo.

In alto al centro, sarà presente un fumetto con ciò che l'assistente sta dicendo in quel momento (così da rendere più accessibile l'applicazione nel caso in cui non sia possibile ascoltare la voce) e sulla sinistra una lista di esempi di comandi possibili nella situazione attuale. Quando l'assistente non sarà disponibile, non verrà visualizzato.

### 9.2.1 Interazione vocale

Per invocare l'assistente basterà pronunciare la Wake Word di Alexa e il comando di avvio della Skill "Avvia ViVi", a seguire l'assistente darà il benvenuto nell'applicazione e compariranno a video le possibili richieste disponibili.

Quando interagiscono con una VUI, gli utenti non hanno indicazioni chiare su cosa può fare l'interfaccia o quali sono le loro opzioni. Quando si progettano le azioni VUI, è quindi importante che il sistema indichi chiaramente le possibili opzioni di interazione, dica all'utente quali funzionalità sta utilizzando e limiti la quantità di informazioni che fornisce a una quantità che gli utenti possono ricordare. Le possibili frasi da chiedere all'assistente verranno per cui poste sotto forma di esempi in un apposita finestra, per rendere più chiaro cosa è possibile dire senza dover occupare troppo spazio, e lasciando che l'esperienza si spieghi da sola man mano che l'utente interagisce con essa. A questo proposito cambieranno le frasi disponibili in base alla fase dell'acquisto.

Di seguito verranno elencate alcune possibili utterances. Non è sensato elencarle tutte, in quanto molte sono semplicemente ridondanti e utili a comprendere i vari modi in cui un utente potrebbe invocare un intent (es. "acquista articolo" e "voglio comprare articolo"). Notare inoltre che, quando si usano termini generici (come "articolo") è solo per indicare un cosiddetto placeholder, ovvero uno spazio da riempire con la specifica richiesta (ad esempio, per articolo potrebbe essere "scarpe"). Oltre a questi accorgimenti, è buona norma costruire ogni intent come continuazione della frase "Chiedi a ViVi...". Per ogni punto verrà indicato nell'ordine: Intent, Utterance d'esempio e una rapida spiegazione del suo compito.

**intoCart:** *"Cosa c'è nel mio carrello"*, ponendo questa domanda riceveremo in risposta ciò che si trova nel carrello con le relative quantità. Risponderà vuoto altrimenti.

**addToCart:** *"Aggiungi articolo al carrello"*, con questa voce line l'articolo scelto, se trovato nel database, verrà aggiunto al carrello e il prezzo aumentato di conseguenza.

**Price:** *"Quanto costa articolo"*, a questa domanda ViVi risponderà col prezzo dell'articolo scelto, se trovato nel database.

**Hide:** *"Nasconditi"*, questo comando serve a liberare la visuale da tutte le interfacce collegate all'assistente.

**Buy:** *"Acquista tutto"*, con questa breve frase procediamo al checkout di tutti gli articoli del carrello. L'assistente eseguirà e comparirà la ricompensa (audio e video). Questo intent richiede conferma.

**Built-In Intents:** questi sono già presenti in ogni skill, consistono in un comando per annullare ("Annulla"), uno per chiedere istruzioni ("Come funziona") e uno per fermare l'esecuzione della skill ("Stop").

Poiché gli utenti normalmente associano la voce alla comunicazione interpersonale piuttosto che all'interazione persona-tecnologia, a volte non sono sicuri del livello di comprensione che la VUI può ottenere. Quindi, perché una VUI abbia successo, non solo è richiesta un'eccellente capacità di comprendere la lingua parlata, ma anche di avere la competenza ad addestrare gli utenti a capire che tipo di comandi vocali possono usare e che tipo di interazioni possono eseguire. La natura intricata del colloquio di un utente con una VUI necessita molta attenzione in quanto è molto facile deludere le aspettative. Questo è il motivo per cui progettare il prodotto in una forma semplice è importante per rendere l'utente consapevole che una conversazione "umana" a due vie è impossibile.

### 9.3 Business Logic

*NB: è stato incluso tra parentesi il nome dei prodotti di AWS usati per ogni specifica risorsa.*

Questa parte dell'implementazione avrà il compito principale di gestire tutte le risorse condivise tra Alexa e l'applicazione e metterle in comunicazione, rifacendosi al design pattern dell'MVC (Model View Controller<sup>[62]</sup>). Molto brevemente questo principio di progettazione consiste nel separare la presentazione del dato dalla rappresentazione del modello (ovvero delle strutture dati) attraverso l'uso di un controllore intermediario che si occupa di elaborare il dato e compiere operazioni complesse.

Per seguire questo approccio avremo bisogno principalmente di un database (DynamoDB<sup>[63]</sup>), che si occuperà di mantenere i dati e lo stato con due tabelle: una per salvare le informazioni sulla sessione (id dell'utente Alexa, id del client di Unity, contenuto del carrello, prezzo e stato della sessione) e una per emulare il reperimento dei vari articoli, questo perché l'idea di base è che grazie ad uno smart assistant è possibile reperire informazioni direttamente dalla rete per poi aggregare e filtrare i risultati in base alle necessità.

Siccome questo aspetto potrebbe richiedere troppo tempo (per essere implementato), è stata adottata questa soluzione che descrive uno stato finale della ricerca, ovvero si suppone di aver precedentemente richiesto all'assistente una determinata categoria di prodotti (ad esempio: tutti i prodotti di un marchio preciso) per poi essere visualizzata nell'interfaccia. Oltre a questo verranno messe a disposizione delle funzioni (Lambda<sup>[64]</sup>) che si occuperanno delle operazioni su questi dati, quindi aggiunta dell'articolo al carrello, aggiornamento del prezzo e gestione della sessione, nonché business logic per Alexa. Infine saranno configurati dei websockets (ApiGateway<sup>[65]</sup>). Tutte le risorse sono configurate tramite uno YAML (un file che contiene tutte le informazioni necessarie a costruire un'infrastruttura) e deployato tramite Serverless,<sup>[66]</sup> un programma che consente di usare il suddetto YAML per formare lo stack di servizi che ospiteranno il tutto (CloudFormation<sup>[67]</sup>).

Al momento dello sviluppo del progetto, Unity non ha supporto ad Alexa, in quanto non è considerato "utile" nei termini di domanda e offerta sviluppare tale integrazione. A peggiorare la situazione, da pochi mesi, Amazon dispone della sua piattaforma (Sumerian<sup>[68]</sup>) per progettare esperienze AR/VR direttamente sul cloud. Per questi motivi è stato necessario escogitare un sistema per permettere ai due ambienti di dialogare. Il meccanismo congegnato consiste nell'impiego dei websockets,<sup>[69]</sup> una tecnologia che fornisce canali di comunicazione full-duplex attraverso una singola connessione TCP. Grazie ad essi si può stabilire un canale comunicativo che rimane in ascolto per tutta la sessione, e all'interno del quale passano dei messaggi sotto forma di JSON (particolari oggetti del linguaggio JavaScript) contenenti le informazioni da scambiare tra le varie componenti dell'applicazione.

All'avvio dell'applicazione viene stabilita la connessione tramite un messaggio contenente i dati di sessione del client Unity. Essi vengono successivamente memorizzati nell'apposita tabella, dopodiché vengono scaricati dall'altra tabella i dati degli articoli e fatti combaciare i dati con i rispettivi oggetti all'interno della scena. Successivamente per ogni operazione che comporta una modifica delle variabili di sessione verrà inviato un messaggio con i dati da salvare in maniera da avere sempre allineate le informazioni in possesso da entrambi gli ambienti. Lo stesso viene applicato per le operazioni lato Alexa. Questo approccio consente un workflow interattivo, un esempio può essere: selezionare un articolo con la VR e chiedere all'assistente di concludere l'acquisto, esso saprà già cosa è presente nel carrello e concluderà l'acquisto, a sua volta verranno scatenati cambiamenti anche all'interno della VR (aggiornamento del prezzo e comparsa della ricompensa).

Un aspetto meno tecnico ma altrettanto interessante e problematico riguarda l'impostazione dell'interfaccia per l'assistente e per la visualizzazione delle informazioni. Trovare il giusto posizionamento di una finestra in un ambiente virtuale con diversi livelli di profondità aggiunge un ulteriore livello di complessità. Per semplificare, posizioneremo le informazioni in due livelli diversi: il prodotto, a distanza, e l'HUD, sempre in sovraimpressione alla visuale dell'utente. Questo per dare il giusto spazio alle informazioni in base alla loro rilevanza e dimensione complessiva, il prezzo e il nome dell'articolo saranno distanti quanto voluto dall'utente, basterà avvicinarsi o allontanarsi a sufficienza, mentre la descrizione dell'articolo, il dialogo dell'assistente e le opzioni di dialogo saranno sempre fisse, dando la possibilità di concentrarsi maggiormente su ciò che si legge.

In una fase successiva per migliorare l'immersione si potrebbe risiedere nell'uso delle interfacce diegetiche.

## 9.4 Limitazioni

Sono state individuate altre problematiche, non inerenti alla struttura del progetto, ma più alla complessità delle tecnologie di AWS e alle loro relative "particolarità", che hanno cambiato il corso dello sviluppo, per cui degne di nota:

- non è assolutamente possibile avviare la skill su Alexa, se non con l'uso della Wake Word impostata dall'utente proprietario del dispositivo. Per cui non è possibile, ad esempio, sfruttare la sintesi vocale per leggere del testo o dare indicazioni quando un determinato comando viene lanciato dalla VR
- una sessione di una skill di terze parti dura 30 secondi, se entro quella tempistica non si risponde o non si effettuano richieste la sessione cade e bisognerà invocare nuovamente la skill
- i tempi di risposta sono molto limitati by design: ci si aspetta che l'esperienza sia sufficientemente fluida da consentire un'iterazione bot e risposta, per cui non si dispone che di 8 secondi per rispondere. Al termine di questo lasso di tempo, verrà ripetuta la domanda e dopo questo secondo richiamo verrà chiuso il dialogo
- al momento dello sviluppo il client Alexa non è fruibile direttamente sulla stessa macchina che esegue l'applicazione (almeno in maniera agevole). Questo comporta l'obbligo di acquistare un Echo device o una RaspberryPi per consentire l'uso di Alexa
- rimane a carico dello sviluppatore trovare un modo per unire i flussi audio (quello utilizzato in questo caso è stato collegare PC e Echo agli stessi speaker tramite un adattatore)
- non è stato possibile far interagire le Lambda con Alexa se non flaggate come "Production" invece che "Development"
- Alexa non riconosce l'endpoint dei servizi di AWS a meno che ogni risorsa non venga deployata in determinate regioni sulla base della lingua impostata nel dispositivo (nel nostro caso la più conveniente eu-west-1, Ireland)
- la versione Alexa-Hosted (sperimentale ma al tempo stesso consigliata da Amazon) consente di porre tutta la logica all'interno della relativa skill, ma allo stato dello sviluppo questa funzionalità risulta essere ancora acerba: diverse funzionalità di base dei servizi AWS, senza alcuna indicazione, vengono disabilitate o oscurate rendendo impossibile questo approccio

Considerando che alcune di queste caratteristiche hanno limitato la riuscita finale del prodotto, si invita in futuro ad ampliare l'implementazione tenendo in considerazione i nuovi possibili sviluppi dello stack. Si rimanda alla sezione Future Migliorie per maggiori dettagli.



## Capitolo 10

# Tecnologie impiegate

Per l'implementazione di ViVi sono state selezionate le seguenti tecnologie, ad ognuna di esse corrisponderà una breve motivazione che inquadra il criterio di scelta adottato durante la fase di analisi progettuale.

### 10.1 Visore

Il requisito principale che ha indirizzato la scelta è la necessità del supporto completo per Steam VR. Questo framework vanta un'adozione estensiva e consente uno sviluppo semplice e completo anche a chi vi si avvicina per la prima volta. Grazie all'uso di questo software è possibile lavorare sul prodotto senza preoccuparsi dell'implementazione "a basso livello" concentrandosi puramente sulla progettazione dell'esperienza.

Un altro fattore indiscutibilmente importante è la qualità del visore stesso. Oltreché evitare problematiche durante lo sviluppo, vogliamo anche evitare problematiche relative all'uso del visore da parte dell'utente finale. Per cui è stato deciso di andare sul sicuro e di adottare HTC Vive,<sup>[10]</sup> che vanta sia il pieno supporto di Steam VR, sia un hardware all'altezza del compito. Per maggiori ragguagli riguardo questa scelta si riporta ai capitoli precedenti.

### 10.2 Voice Assistant

Per questa tecnologia la tentazione di andare verso l'open source è stata molta, ma alla fine si è preferito puntare su Alexa,<sup>[39]</sup> in quanto la praticità di sviluppo complessiva e il supporto offerto da Amazon, unito alla maggior certezza che il software non venga dismesso e soprattutto una miglior resa a livello di riconoscimento vocale hanno avuto la meglio sulla maggior libertà concessa allo sviluppatore. Grazie all'uso di Alexa non è necessario preoccuparsi della parte di riconoscimento vocale, ML, sintesi vocale ed altri aspetti tecnici che intralcerrebbero il raggiungimento dello scopo di questa trattazione. Alexa dà la possibilità di integrare i servizi di AWS, abbassando ulteriormente il costo di sviluppo e integrazione delle API. La disposizione di questo ambiente comporta alcune limitazioni (delle quali abbiamo accennato e che verranno ridiscusse nei capitoli conclusivi).



Inoltre non è secondario il fatto che Alexa sia naturalmente predisposta all'acquisto di articoli. Difatti essendo lo smart assistant di Amazon, come potrebbe non mancare un collegamento ad una delle più grandi piattaforme di e-commerce globali.

### 10.3 Setup in laboratorio (VarLab)



Ubicato nell'aula studio presso il Laboratorio Ranzani,<sup>[70]</sup> il laboratorio Varlab<sup>[71]</sup> dispone di area e strumenti sufficienti all'uso di Lighthouse.<sup>[22]</sup> Il PC impiegato è stato scelto appositamente per sostenere workload intensivi come machine learning e sviluppo di applicazioni 3D e VR.

Le specifiche sono le seguenti:

**Processore:** Intel Core i7-6850K

**Scheda Video:** Doppia scheda NVIDIA GeForce 1080 in SLI

**Memoria:** DDR4 da 32 GB a 2133 MHz

**Archivio:** Unità SSD principale da 512 GB e HDD 4 TB

**Sistema Operativo:** Windows 10 Pro (64 bit)

Ben oltre superiori a ciò che è richiesto dal Vive e dall'applicazione, ma è notorio che l'ambiente di sviluppo (quindi non compilato ed ottimizzato) risulti sempre essere più pesante del prodotto finito. Il computer è provvisto di due monitor, uno per l'uso e l'altro più grande per fare mirroring dell'applicazione dal punto di vista dell'utente, così da rendere visibile dall'esterno quello che sta avvenendo all'interno della VR.

Il setup è relativamente semplice, si indica un area delimitandola tramite le base-station, poste appositamente su due aste telescopiche all'altezza di 3m circa. È stata scelta questa tecnica di installazione così da evitare di fissarle al muro tramite tasselli e lasciare la possibilità di spostarle facilmente in base alle necessità. Il

visore va collegato al ViveLink, un apposito adattatore che si occupa di instradare i vari tipi di segnali tra cui l'alimentazione dell'HMD. Una volta collegato il tutto all'alimentazione e al PC (inclusi i controller bluetooth), basta seguire la procedura guidata di Steam VR ed è tutto pronto.

## 10.4 Ambiente di sviluppo 3D

La scelta ovvia è ricaduta su Unity:<sup>[61]</sup> è un consolidato IDE per la prototipazione e produzione rapida di esperienze nel campo del gaming, film making e animazione. Supporta tutte le principali tecnologie necessarie all'implementazione del VR (tra cui nativamente le librerie di Steam VR) e la community è molto attiva e piena di riferimenti utili. Queste caratteristiche lo rendono lo strumento migliore quando si richiede rapidità di sviluppo senza compromettere la qualità del risultato finale.

### 10.4.1 Assets e plugins

Per asset si intendono tutte quelle parti che compongono il progetto (siano esse funzionali o estetiche), prese dall'esterno sotto forma di pacchetti o plugins. Tra i plugin di Unity i più importanti sono sicuramente Steam VR<sup>[72]</sup> (il pluricitato framework per creare VR enabled experiences), Pro Builder<sup>[73]</sup> e Pro Grids<sup>[74]</sup> per semplificare il lavoro di costruzione della scena. Questi due plugin servono per velocizzare la modellazione 3D di base, uno per la costruzione di strutture low-poly (ovvero costituite da pochi poligoni) e l'altro invece blocca le coordinate su specifiche griglie consentendo agli oggetti di ottenere le giuste proporzioni e il corretto posizionamento.

Sono stati scaricati dallo Unity Asset Store<sup>[75]</sup> alcuni modelli tridimensionali per abbellire la scena, quali alberi, abiti, scarpe, etc...

I restanti assets, quelli funzionali alla VR, sono forniti direttamente da Steam VR, con tanto di esempi e configurazioni.

### 10.4.2 Suoni e musica

Sono stati usati degli audio samples dallo Shop di Ski Safari 2<sup>[76]</sup> creare un'atmosfera tranquilla e accogliente e un effetto sonoro royalty-free per simulare un registratore di cassa.

## 10.5 Ambiente di sviluppo Backend

Per sviluppare la cosiddetta business logic (ovvero tutto quel codice che si occupa di elaborare i dati, salvarli e farli comunicare) verrà fatto uso di Amazon Web Services (AWS<sup>[77]</sup>) come provider di servizi infrastrutturali serverless. Il vantaggio di questa piattaforma risiede nel fatto che non si possiedono direttamente delle macchine, ma esse vengono astratte e messe a disposizione per lo sviluppo direttamente da Amazon (in base alle necessità i costi potrebbero variare). L'uso di queste tecnologie permette di avere tutta la logica sempre online e al contempo consente di non dover curare l'ambito infrastrutturale. Infine, offre l'integrazione nativa di Alexa, permettendo così di sfruttare ciò che verrà prodotto tramite l'aggiunta di un'apposita proprietà. Il linguaggio principalmente impiegato è NodeJS.

Per semplificare il deploy sulle macchine Amazon verrà usato Serverless, software essenziale per rendere più agibile l'uso dello stack: scrivendo un unico file .YAML è possibile configurare tutte le risorse necessarie ed effettuare rilasci con il minimo sforzo.



# Parte IV

## Risultati



## Capitolo 11

# Somministrazione del test

Nel capitolo verranno trattate le modalità di test impiegate per raccogliere dati qualitativi riguardo l'applicazione. Questi risultati verranno analizzati al fine di discutere la tesi.

### 11.1 Svolgimento

L'esperimento inizierà con la presentazione di ciò che verrà proposto, e la compilazione della sezione preliminare del questionario, riguardante le informazioni demografiche di chi partecipa.

In rapida successione verranno introdotte delle meccaniche di base, in quanto la scarsa familiarità con la realtà virtuale e Alexa potrebbe influenzare le valutazioni sull'usabilità dell'esperienza.

Una volta consolidate tali conoscenze si è passati all'esplorazione di una prima applicazione (Fashion Island di Marco Valerio Manzini<sup>[3]</sup>) per approfondire la familiarità col mezzo e dare un metro di paragone utile in fase di valutazione.

Questa esperienza ritrae l'utente con fattezze umane in un'isola deserta di fronte ad uno specchio e diversi capi di abbigliamento tutti intorno. L'utente avrà la possibilità di selezionare il proprio outfit, visualizzare il suo avatar nello specchio e muoversi liberamente per la piccola isola tramite teletrasporto.

Fashion Island è più mirata all'immersione nella realtà virtuale in contesto moda, ovvero cerca di rendere viva la presenza, facendo vedere l'avatar che risponde ai movimenti dell'utente nella scena. In contrasto, ViVi non fa uso di un modello umano per dare un senso di immersione, in quanto esso è direttamente derivante dalla costruzione della scena, dalla presenza dell'assistente, nonché dalla semplicità e fluidità d'utilizzo. Difatti, come già spiegato in precedenza, un'esperienza più risulta essere naturale e fluida, più da l'impressione di essere realistica nonostante l'apparente irrealtà della situazione.

Alla fine di questo primo approccio si sottopone l'utente ad un questionario qualitativo.

Al termine di questa sezione l'utente inizierà la vera e propria esperienza entrando nella simulazione del negozio. La scena sarà composta in maniera tale da rappresentare fedelmente un negozio, con dei reparti composti da scaffali con vari prodotti. Il cliente si potrà muovere liberamente all'interno di esso tramite la tecnica

del teletrasporto. Potrà interagire con gli oggetti tramite il sistema di puntamento così da poter visualizzare le informazioni o aggiungerli al carrello. Alla stessa maniera l'utente avrà il pieno controllo dei comandi vocali. All'utente sarà chiesto di svolgere azioni utili nella loro sequenza naturale ispirandosi ai suggerimenti a schermo.

Concluso l'acquisto e visualizzata la ricompensa, l'utente sarà invitato a rispondere alle restanti domande del questionario.

## 11.2 Valutazioni e metriche

Per permettere la valutazione ai partecipanti ci si baserà sulla scala Likert,<sup>[78]</sup> una tecnica per la misura dell'atteggiamento. Tale tecnica consiste principalmente nel mettere a punto un certo numero di affermazioni (tecnicamente definite item) che esprimono un atteggiamento positivo e negativo rispetto ad uno specifico oggetto. La somma di tali giudizi tenderà a delineare in modo ragionevolmente preciso l'atteggiamento del soggetto nei confronti dell'oggetto. Per ogni item si presenta una scala di accordo/disaccordo, generalmente a 5 o 7 modalità. Ai rispondenti si chiede di indicare su di esse il loro grado di accordo o disaccordo con quanto espresso dall'affermazione.

L'attribuzione dei punteggi avviene attraverso un processo che Likert definì il "metodo semplice", divenuto uno standard nella codifica dei giudizi espressi su affermazioni.

Si procede alla suddivisione degli item favorevoli all'oggetto di cui si vuole misurare il giudizio dagli item sfavorevoli. Si attribuiscono dei punteggi (ad esempio, su una scala a 5 passi, da 1 sul polo "fortemente in disaccordo", a 5 sul polo "fortemente d'accordo"), a volte per ottenere una maggiore granularità verrà associato un 7 sulla valutazione positiva.

A questo punto, il punteggio totale di un soggetto, indicherà la posizione sul continuum dell'atteggiamento (in questo caso ipotizzato come unidimensionale). Esso risulterà quantificato, pur non esistendo un punto zero nella sua misura, e confrontabile con altri dati. Un metodo alternativo alla somma consiste nella media dei punteggi dei singoli item, anch'essa ottenuta dopo aver invertito i punteggi delle affermazioni negative.

Gli item in studio sono indicati nelle tabelle seguenti, ognuno è identificato da un codice per semplicità. Per motivi di rappresentazione alcune domande potrebbero essere scalate da 5 punti a 7 e viceversa. Un altro dettaglio importante sulle L7 e L8, le domande sono strutturate in maniera multidimensionale, ovvero abbiamo la solita scala in "orizzontale", ma "verticalmente" troviamo più dimensioni, riguardanti se l'esperienza sia stata interessante, divertente, eccitante e gradevole, nel caso delle L7, e soddisfacente, piacevole, esaltante e attraente, per la L8. Nel caso della L7 una valutazione positiva viene posizionata al valore più basso, per motivi rappresentativi verrà successivamente invertita.



Valutazione dell'interfaccia	
Codice	Item
L1	Ho trovato l'interfaccia semplice da usare
L2	Una volta che ho imparato ad usare l'interfaccia questa mi è sembrata intuitiva e semplice
L3	Una volta che ho imparato ad usare l'interfaccia la manipolazione degli oggetti mi è sembrata intuitiva e semplice
L4	Una volta che ho imparato ad usare l'interfaccia questo mi ha permesso di essere creativo e di focalizzarmi meno sulle tecniche dell'interfaccia
L5	Preferisco la nuova interfaccia al mouse o alla tastiera
L6	Sono stato in grado di creare gli oggetti o le scene che desideravo
L7	Sentimento generale riguardo l'esperienza di VR

**Tabella 11.1**

Apprezzamento generale	
Codice	Item
L8	Sentiment riguardo l'esperienza di VR nel campo della moda
L9	Assumi di essere interessato ad acquisti on-line, quanto ti piacerebbe comprare tramite un'app di fashion VR?
L10	Visiteresti un VR retailer, simile alle applicazioni che hai provato?
L11	In futuro, pensi che potresti comprare da un VR retailer?

**Tabella 11.2**

Per le valutazioni specifiche (A-D), verrà usata la scala Likert su 5 punti

Valutazioni specifiche delle esperienze	
Codice	Item
A	Mi piace la VR perché mi estranea dall'ambiente circostante
B	Trovo che il visore di VR sia molto pesante
C	L'uso dei comandi vocali rende più semplice l'utilizzo del sistema
D	Mi piacerebbe poter fare tutto coi comandi vocali, liberando le mani dai controllers

**Tabella 11.3**

Per le domande in tabella 12.4 verrà adottata una scala percentuale, dove la risposta "Si" indica un 100% e la risposta "No" uno 0%.

Domande sull'usabilità	
Codice	Domanda
Q1	Utilizzeresti queste applicazioni?
Q2	Utilizzeresti queste applicazioni per acquisti online?
Q3	Utilizzeresti queste applicazioni per divertimento?
Q4	Utilizzeresti queste applicazioni per comprare un maglione?
Q5	Utilizzeresti queste applicazioni per comprare delle scarpe?
Q6	Utilizzeresti queste applicazioni per comprare un costume da mare?
Q7	Utilizzeresti queste applicazioni per provare abbinamenti?
Q8	Utilizzeresti queste applicazioni per conoscere nuove collezioni?
Q9	Utilizzeresti queste applicazioni per pubblicizzare i tuoi prodotti?

Tabella 11.4

Il questionario somministrato termina con una sezione a domande aperte, che non danno luogo ad una valutazione analitica ma danno spazio ad eventuali riflessioni dei partecipanti. Le domande sono le seguenti:

- Cosa ti è piaciuto maggiormente di questa esperienza VR? Perché?
- Cosa non ti è piaciuto delle app di VR? Perché?
- A livello di tecnologia, dopo queste esperienze, apprezzi la VR? Perché?
- Prima di oggi, avevi già provato le tecnologie di VR? Se sì, in quale contesto?
- Conosci qualche esperienza di VR che è stata fatta nel campo della moda? Se sì, quali?
- Riesci ad immaginare un'applicazione VR di forte interesse per la moda? Che caratteristiche dovrebbe avere?

La sessione in totale dura circa 30 minuti a persona. Aumentare ulteriormente il tempo d'impiego del visore potrebbe causare problemi di nausea e stress fisico, un'esperienza di breve durata aiuta a ridurre questi fattori. Tenere i tempi compatti viene incontro all'ulteriore esigenza di dover far provare diverse applicazioni a più utenti facendo a turno su di un unico visore.

## 11.3 Contesto moda

Per fornire un caso più specifico abbiamo scelto di inscenare un negozio di vestiti, in quanto il settore moda è attualmente uno tra i più interessati alle tecnologie di Realtà Virtuale. L'ambiente rappresenta una boutique di modeste dimensioni, un luogo familiare e rilassante in cui l'utente è portato dimenticarsi di essere in un mondo virtuale, immergendosi completamente nell'applicazione.

Abbiamo scelto il settore moda in quanto, nello scorso decennio ha cercato di svilupparsi nei mercati emergenti, aprendo centinaia di negozi, ma soprattutto aumentando la loro presenza online. In questo contesto di crescita, la realtà virtuale e aumentata costituisce un potenziale per accrescere la presenza nel digitale dei brand, addirittura portandoli ad un livello superiore a quello raggiunto grazie al mobile.

La crescita della clientela relativa alla VR è stimata a 1000 nuovi utenti al giorno. Nonostante i tassi di adozione siano lenti a causa del costo di ingresso, è difficile fare stime in quanto ci si trova di fronte ad una tecnologia emergente il cui andamento nel mercato non è ancora definito stabile. Di sicuro l'adozione di questa tecnologia nel campo della moda non sarà rapidissima, ma si pensa richiederà una ventina di anni, come per l'e-commerce.

Idealisticamente, l'utilizzo di un virtual assistant porta dei vantaggi relativi all'usabilità e alla completezza del software orientato al fashion shopping. Piuttosto che riempire un campo testuale o usare un'interfaccia complessa per fare la ricerca, l'utente potrà semplicemente usare la voce per fare richieste. E grazie alle informazioni che l'assistente possiede sull'utente ottenere risultati coerenti con la propria persona. Un esempio immediato potrebbe consistere nella ricerca di un articolo per determinato brand, con la frase "ViVi mostrami pantaloni della Nike", non solo l'assistente compirà una ricerca per quel determinato brand ma, se si suppone che lo stile dell'utente sia sportivo, mostrerà nello store solo pantaloni da tuta, posizionandoli in maniera strategica all'interno del negozio, e così via.

Grazie a questo ed altri fattori è ormai consolidato che, se la realtà virtuale riuscisse ad entrare nel giornaliero, allora potrebbe costituire un potente canale per la vendita e comunicazione di informazione di articoli d'abbigliamento e simili. In attesa che VR e AR rappresentino la fascia mainstream dei consumatori, l'industria della moda inizia a creare gli strumenti utili ad un futuro ingresso sul mercato. Come anticipato, le simulazioni virtuali possono portare vantaggi, oltre che al cliente, anche ai negozianti permettendo un confronto tra due ambienti diversi, ad esempio permettendo di sperimentare con la VR il layout dello store o l'esposizione degli articoli, senza la necessità di impiegare risorse nel mondo reale, prima della messa in posa delle fondamenta. Alcuni grandi brand, hanno già creato delle esperienze per permettere ai clienti dei loro store di assistere alle sfilate. Altri invece hanno implementato camerini per provare vestiti nel mondo digitale.

Per i motivi sopracitati, il campione è composto principalmente da un gruppo di studenti del Master in Design and Technology for Fashion Communication dell'Università di Bologna.<sup>[5]</sup> I nove partecipanti, di età tra 20 e 25 anni, sono 8 ragazze e un ragazzo, con background diversi per quanto riguarda l'esperienza con le tecnologie in studio. In generale la conoscenza tecnica, il tempo speso al computer e, di

conseguenza, l'esperienza con tecnologie informatiche degli studenti è molto scarsa. È possibile classificare il campione come "Technically Unaware", ovvero coloro che fanno uso di mezzi tecnologici, ma non sono particolarmente interessati all'avanzare del progresso o ad aumentare della loro conoscenza tecnica. Questo tipo di utenza è quella più difficile da "convertire" al presentarsi di una nuova tecnologia nel mercato. Inoltre questi studenti ricadono in una nicchia molto difficile da soddisfare a livello qualitativo. Il campione è altamente consapevole dell'ampio spettro di problematiche legate al settore ed è attento ad ogni piccolo dettaglio.

In particolare, a fronte di una piccola interrogazione preliminare il gruppo risultava molto scettico nei confronti della proposta, in quanto queste esperienze soffrono spesso di diverse problematiche, maggiormente collegate alla preferenza dell'utente di poter toccare i capi, o verificare la conformazione degli articoli.

Questi problemi sono solitamente risolvibili aumentando la qualità degli asset degli articoli o includendo particolari tecnologie che consentono all'utente di abilitare altri sensi all'infuori della vista. Nel caso di questa trattazione, tali problemi non risultano essere interessanti. Verrà infatti posta maggiore attenzione all'usabilità, per cui si richiederà ai soggetti di immaginare una situazione ideale, così da evitare eventuali errori nella valutazione.

Si fa notare come non sia stato preso alcun accorgimento specifico per adattare le funzionalità dell'assistente al caso specifico, a dimostrazione della sua flessibilità applicativa.

## Capitolo 12

# Risultati ottenuti

In questo capitolo verranno raccolti e discussi i risultati della sperimentazione in laboratorio. Verranno trattati per primi i dati analitici e qualitativi raccolti con le domande in scala, successivamente verranno analizzate le opinioni più incisive dei partecipanti.

### 12.1 Valutazioni qualitative

Il processo di test si è svolto senza particolari inconvenienti, in due giornate distinte. I partecipanti, seppur senza alcuna esperienza, non hanno provato nausea o trovato eccessiva difficoltà con l'uso dei mezzi, se non per casi isolati in cui gli utenti si sono dichiarati affetti da cause esterne (stanchezza derivante dalla giornata passata, caldo, problemi personali, etc...).

La funzione che è stata più problematica è stata senz'altro quella di riconoscimento vocale: purtroppo la stanza del laboratorio è direttamente comunicante con la sala studio, che nelle giornate di test si è rivelata essere abbastanza rumorosa da disturbare Alexa. Questa situazione unita all'acustica della stanza e al fatto che Alexa è comunque tarata sulla voce dell'utente principale ha creato qualche singhiozzo durante l'esperienza, ma nulla di invalidante, specie se si pensa che l'applicazione di per se sarebbe destinata all'uso personale e non condiviso.

In generale le valutazioni si sono mantenute mediamente intorno al 4, segno che le considerazioni fatte durante la pianificazione e sviluppo dell'applicazione sono state accolte positivamente. In particolare gli utenti si trovavano molto meno disorientati dal metodo di spostamento e puntamento adottato in ViVi.

L'unica risposta che risulta relativamente negativa è L5, ovvero "Preferisco questa nuova interfaccia a mouse e tastiera", alla quale i partecipanti hanno dato parere indifferente (ovvero a metà della scala) o inferiore. È comprensibile come un completo cambio di paradigma come questo possa suscitare scetticismo, specie se visto in questo stato embrionale, piuttosto che nel vivo delle sue potenzialità.

In ogni caso l'esperienza con comandi vocali ha raccolto maggiore consenso rispetto a quella "tradizionale" di realtà virtuale raccontata da Fashion Island.

Valutazione dell'interfaccia		
	Fashion Island	ViVi
L1	4	4,33
L2	4,44	3,89
L3	3,89	4
L4	4	3,89
L5	2,67	2,89
L6	3,78	3,78
Totale su 6	3,78	3,78
L7	3,92	3,74
Totale su 7	3,8	3,78

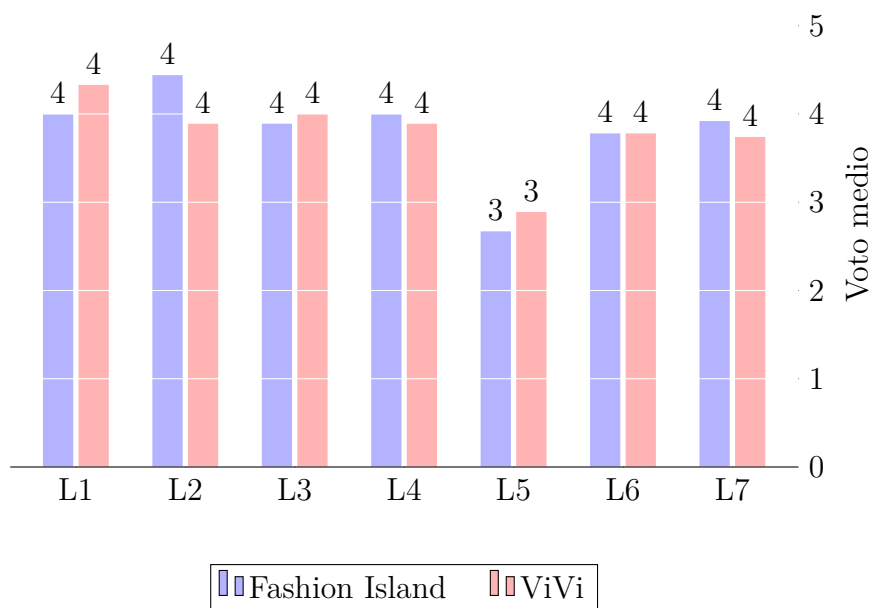


Figura 12.1

Il dato più rilevante è senza dubbio L1, "Ho trovato l'interfaccia semplice da usare", che dimostra un ampio apprezzamento verso ViVi, nonostante la valutazione già alta data a Fashion Island.

Proseguendo, le domande relative alla sezione "VR e fashion" (14.2 e 14.3) riportano una visione molto concorde riguardo le questioni proposte: circa tre quarti dei partecipanti utilizzerebbero l'applicazione per fare acquisti online.

Le domande Q4-7 (riguardanti cosa l'utente comprerebbe con questo sistema a disposizione) riportano anch'esse un responso quasi unanime sui vari prodotti e, come anticipato, gli articoli su cui il campione è riluttante hanno necessità specifiche che vengono soddisfatte solo nel mondo reale. Difatti il più basso numero di "Si" riguarda le scarpe e il costume da mare, due indumenti che necessitano di essere provati prima dell'acquisto per avere una certezza sul loro impatto estetico o sulla loro comodità. In generale c'è da dire che la problematica è riscontrata anche negli odierni sistemi di shopping online dove c'è una relativa riluttanza a comprare abbigliamento.

L'unica considerazione che possiamo fare è che se c'è una possibilità di superare questo ostacolo apparentemente insormontabile, allora questa risiede unicamente

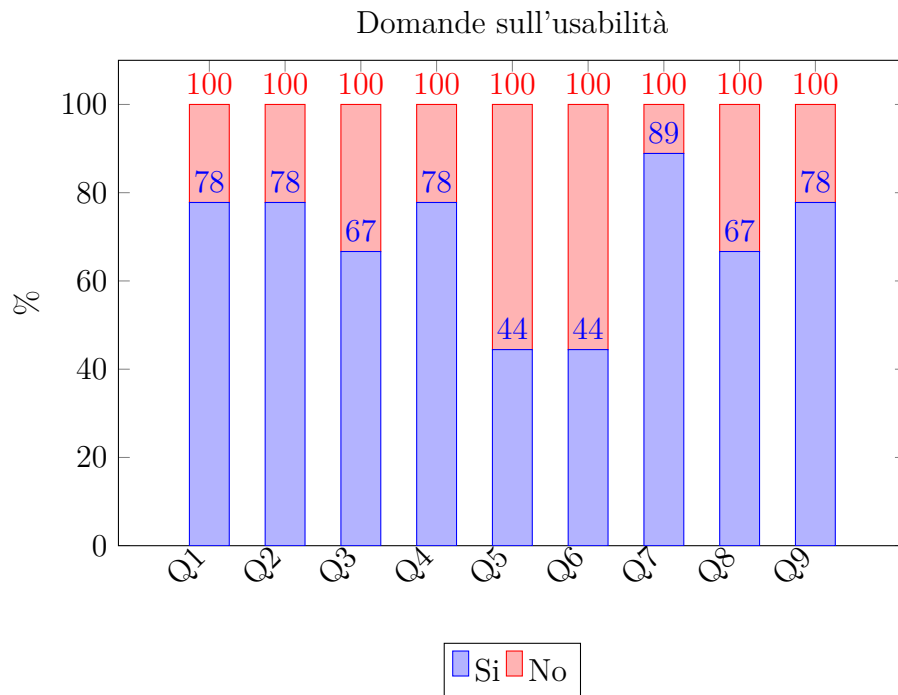


Figura 12.2

nella realtà virtuale. A quel punto ViVi potrebbe rispondere ad un'esigenza diffusa di supporto all'interno di queste esperienze.

Molto più apprezzata la prospettiva di poter pubblicizzare o visualizzare collezioni e prodotti direttamente nella realtà virtuale, giustificando, tra l'altro, l'impiego che le grandi marchi finora hanno fatto di questa tecnologia (domande Q8-9).

Affiancate alle precedenti domande troviamo le L8-11, ovvero le domande in scala Likert riguardanti le esperienze VR nel campo della moda. Nel grafico relativo (14.3) vengono incluse anche le L7 relative alle specifiche esperienze per paragonarle all'apprezzamento generale. Anche qui è stato ottenuto un risultato positivo, mediamente la risposta è stata superiore al 5, per cui anche l'esperienza sviluppata è assolutamente in pari con l'apprezzamento generale del campione.

## Apprezzamento generale

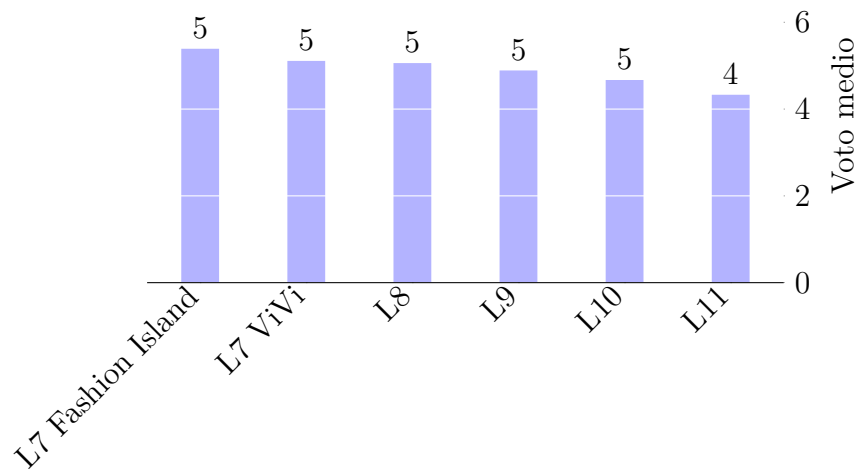


Figura 12.3

Nelle valutazioni specifiche alle applicazioni (A-D), si è ottenuto un risultato più che soddisfacente alla domanda C, riguardante l'uso dei comandi vocali all'interno di un'esperienza di realtà virtuale. Un risultato altrettanto positivo è stato ottenuto per la domanda D, riguardante la possibilità di sostituire tutti i comandi tramite controllers con alternative vocali. Chiaramente è una richiesta molto radicale, da cui può scaturire dell'attrito, in quanto alcuni comandi risultano ancora più semplici via controllers (l'esempio più lampante: lo spostamento). Per cui anche un risultato sopra il mediano risulta essere positivo.

## 12.2 Domande aperte

In questa sezione si discuteranno gli spunti più interessanti lasciati dai partecipanti. Nella prima domanda, relativa agli aspetti positivi della VR, i partecipanti hanno indicato come principali l'immersione, seguito poi dal fatto di essere divertente ed innovativa. Nella domanda riguardante i fattori negativi invece troviamo proprio l'estraneazione, l'uso macchinoso dei controllers e soprattutto il fatto che "sembra di essere all'interno di un videogame", una frase molto spesso ripetuta, che fa capire come questa tecnologia nell'immaginario comune sia ancora relegata all'ambito videoludico e dell'intrattenimento. Difatti anche tra le risposte positive, "divertente" ha legami con questa motivazione. Nello specifico un partecipante non ha apprezzato lo stile low poly dell'app, ma si aspettava qualcosa di più fotorealistico. Un altro studente ha esternato come la VR, nella sua opinione, sia fuori contesto in questo ambito e che sarebbe più sensata in altri ambiti.

In ogni caso, nonostante queste critiche, la VR ha riscosso particolare successo con 5 partecipati che, generalmente, la preferirebbero rispetto ad altre tecnologie e sarebbero interessati ad altri impieghi futuri.



### Valutazioni specifiche delle esperienze

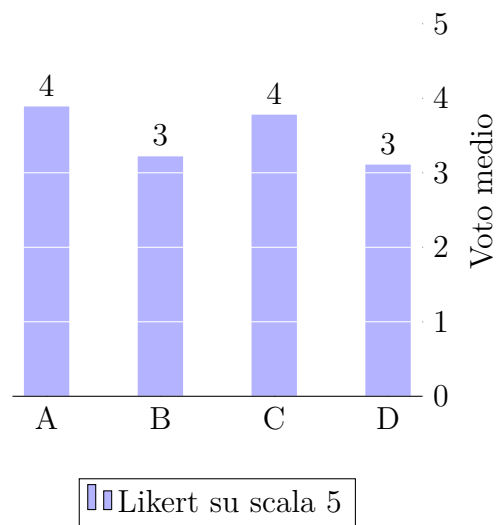


Figura 12.4

Nelle sezioni relative ai suggerimenti, i partecipanti hanno espresso molto interesse in alcune specifiche esperienze che vorrebbero veder realizzate in un futuro prossimo. Tra le tante proposte, spicca maggiormente per numero di interessati, quella di un camerino/armadio virtuale che consenta di vedere come determinati abiti ed abbinamenti stanno indosso all'utente, sia esso avatar che in carne ed ossa. A seguire la possibilità di utilizzare queste tecnologie per progettare e visualizzare spazi ed eventi, tra cui: showroom, sfilate, allestimenti, etc... Una proposta fuori dal coro, ma decisamente interessante, consiste nel creare un'applicazione che consenta di visualizzare il processo di realizzazione di un capo d'abbigliamento.

## Capitolo 13

# Future migliorie

A seguito delle valutazioni fatte sia in corso di sviluppo che in fase sperimentale, sono state evidenziate una serie di possibili migliorie, sia tecniche che funzionali.

Siccome le principali problematiche in fase di sviluppo sono date dalle strategie di implementazione che Amazon ha scelto di sfruttare per il suo prodotto potrebbe essere interessante sostituire Alexa con un altro assistente, magari facente parte delle alternative open.<sup>[42][43]</sup> Inizialmente si era pensato all'impiego di Mycroft, che risulta essere il più completo e compatibile col nostro progetto. Purtroppo le motivazioni che hanno spinto a scegliere Alexa sono molto più legate ad un fattore di costo piuttosto che a motivazioni di natura strettamente tecnica: la sfida principale nello sviluppo è stata creare una struttura che permettesse di supportare il voice assistant all'interno di Unity, andare ad aggiungere ulteriore lavoro dalla parte dell'assistente avrebbe comportato un notevole allungamento dei tempi, che si è evitato scegliendo una soluzione "già pronta". Un'implementazione con un assistente open è sicuramente più flessibile e personalizzabile, ma non altrettanto semplice da realizzare e sicuramente con molto meno supporto. Tuttavia, abbattuta la barriera del dialogo tra le parti, è ora fattibile e altamente consigliato adottare un'altro tipo di assistente.

Questa scelta potrebbe portare a diversi miglioramenti e dare vita a nuove opportunità, tra cui:

- Cambio della Wake word
- Deploy del backend in locale
- Integrazione diretta con desktop (senza necessità di dispositivi esterni)
- Integrazione di funzionalità collegate a Mycroft

Altri importanti ampliamenti possono riguardare l'aggiunta di nuove funzioni per rendere l'esperienza ancora più personalizzabile e legata alle potenzialità che l'assistente stesso offre. Uno dei punti focali potrebbe consistere nell'avvalersi di un profilo utente generato tramite un'apposita interfaccia. Tale profilo informerà l'assistente di ogni tipo di preferenza e necessità che potrà poi essere usata per rendere più smart l'esperienza di shopping (ad esempio indicate le allergie l'assistente sarà in grado di filtrare in anticipo i prodotti o avvertire mentre l'utente compie una scelta considerata "rischiosa").

Alcuni esempi di filtraggio potrebbero essere:

- Caratteristiche fisiche (sesso, taglie, colore della pelle, etc...)
- Eticità dei prodotti (child labour-free, equo solidale, riciclabile, etc...)
- Allergie, preferenze alimentari (vegan, gluten-free, etc...)
- Composizione del nucleo familiare (preferenza di altri componenti del nucleo, presenza di bambini, gravidanze, etc...)
- Caratteristiche dell'ambiente casalingo (dimensioni della casa, presenza di animali domestici, stile, etc...)
- Preferenze stilistiche (classico piuttosto che moderno, formale piuttosto che informale, etc...)
- Storico degli acquisti (acquisti passati o oggetti posseduti, etc...)

Questa feature potrà aiutare l'utente a compiere scelte molto più informate senza intaccare l'interfaccia o la scena. Il filtro posto dall'assistente permetterà un acquisto che richieda un minor sforzo da parte dell'utente nell'accertarsi della compatibilità ed adeguatezza del prodotto scelto. L'applicazione e l'assistente rimarrebbero la base per visualizzare e ricevere le dettagli riguardanti gli articoli, ma a differenza dell'app sviluppata, che simula la parte finale dell'operazione, le informazioni potrebbero venire cercate, filtrate e raggruppate da Amazon (o, in generale, dalla rete) attraverso apposite API. All'aumentare dei dati a sua disposizione l'assistente riuscirà ad agire come un personal shopper (figura tanto desiderata quanto irraggiungibile dato il costo). Infatti, data la fama attuale di queste tecnologie, si potrebbe pensare di integrare proprio in questa funzionalità un meccanismo di Machine Learning e AI per rendere l'assistente ancora più preciso.

Sempre ai fini dell'esperienza personalizzata l'utente dovrebbe essere in grado di modificare alcuni aspetti del negozio, come potrebbero essere la musica riprodotta o lo stile, questa funzione in particolare è molto semplice da ottenere ma di grande impatto.

L'assistente potrebbe avere maggiori capacità di supporto su alcuni aspetti di solito associati alla figura del commesso nei negozi, come indicare la posizione di determinati prodotti, spiegare le politiche del negozio o simili e, nel caso in cui non riesca ad essere d'aiuto, ci metterà in contatto con una persona reale (tramite chiamata vocale/video, ad esempio). Diminuendo ulteriormente il divario tra realtà e virtuale.

Sempre riguardo l'ambito assistenziale, si potrebbe spiegare come ottenere la posizione di un determinato articolo o di un settore dello store. In particolare questa funzione potrebbe essere elaborata in maniera tale da risultare efficace in molteplici campi, specie in campo assistenziale o istruttivo (per esempio basta pensare ad un addestramento militare o di sicurezza sul lavoro).

Un'aspetto assolutamente migliorabile della presentazione: si potrebbe adottare un approccio che fa uso di interfacce diegetiche per presentare le informazioni e i comandi. Quindi implementare visualizzazioni che non intaccano l'immersione nell'esperienza.

Volendo cambiare la piattaforma di sviluppo 3D, ma mantenere i servizi Amazon, di recente è stata annunciata Sumerian,<sup>[68]</sup> un prodotto cloud che comprende un vero e proprio ambiente di sviluppo 3D molto simile a Unity. La potenzialità di questo software risiede nella capacità nativa di includere prodotti che abilitano sia AI che sintesi vocale, allo specifico scopo di creare assistenti virtuali in un ambiente 3D. Ovviamente nel tipico stile AWS, tutte le risorse sono affidate ad una macchina di proprietà di Amazon, per cui non è necessario avere un PC estremamente prestante per sviluppare esperienze di VR e AR.

Un'ulteriore possibilità che si apre mantenendo la piattaforma Alexa, è la possibilità di implementare un sistema di acquisto diretto tramite l'integrazione nativa di Amazon per lo shopping.

## Capitolo 14

# Conclusione

In conclusione, dall'analisi dei dati raccolti, si evince che l'esperienza ha raccolto un discreto successo.

Grazie all'applicazione ViVi, è stato possibile analizzare il punto di vista dell'utente riguardo questo particolare tipo di esperienza e capire la validità dell'idea proposta. Seppur il prodotto finale sia tutt'altro che rifinito, è risultato essere piacevole all'uso e pienamente immersivo, permettendo un'ampia libertà d'azione e movimento, e al contempo risultando intuitivo e utile grazie all'uso dell'assistente e della voce.

Come già espresso nell'introduzione, il progetto è stato concepito non solo per dimostrare concetti teorici, ma anche per offrire un ambiente di sviluppo base estensibile, modificabile e replicabile da altri che vogliono contribuire a questa idea. Si può vedere tale sistema come un primo passo verso lo sviluppo di applicazioni in realtà virtuale più complesse, che si concentrano maggiormente su altri aspetti o che introducono un livello di dettaglio maggiore nel comparto grafico, o che cerchino di sormontare le limitazioni che si sono presentate in fase di sviluppo.

Per cui si auspica che questo progetto non rimanga stagnante, bensì sia frutto di una continua e costante ricerca, nell'interesse del progresso e dello studio nel campo applicativo della moda e del retail in abito VR e AR.



# Bibliografia

- [1] Edwin Jiang. *Virtual Reality: Growth Engine for Fashion?* URL: <https://www.businessoffashion.com/articles/fashion-tech/virtual-reality-growth-engine-for-fashion> (visitato il 30/10/2019).
- [2] Jane Price. *A brand manager's guide to virtual assistants*. URL: <https://martechtoday.com/a-brand-managers-guide-to-virtual-assistants-218933> (visitato il 30/11/2019).
- [3] Marco Valerio Manzini. «Progettazione e realizzazione di una applicazione di Realtà Virtuale immersiva utilizzando la tecnologia HTC Vive.» Tesi di dott. Alma Mater Studiorum · Università di Bologna, 2016.
- [4] Sonia di Vaio. «Exploiting Immersive Virtual Reality for Fashion Gamification». In: *2018 IEEE 29th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)* (2018).
- [5] *Master in Design and Technology for Fashion Communication*. URL: <https://master.unibo.it/fashion-communication/it> (visitato il 31/11/2019).
- [6] Charalambos D. Aliprantis e Kim C. Border. *Infinite Dimensional Analysis*. Berlin: Springer, 1994.
- [7] David; Sheridan Thomas; Slater Mel Barfield Woodrow; Zeltzer. «Presence and Performance Within Virtual Environments». In: *Virtual Environments and Advanced Interface Design*. Oxford University Press (1995).
- [8] Brian Thomson Carol; Goldiez. «Predicting presence: Constructing the Tendency toward Presence Inventory». In: *International Journal of Human Computer Studies* (2009).
- [9] Michael Abrash. «What VR could, should, and almost certainly will be within two years». In: Valve VR. 2014.
- [10] *Vive product page*. URL: <https://www.vive.com/eu/product/> (visitato il 31/11/2019).
- [11] *Vive Pro product page*. URL: <https://www.vive.com/eu/product/vive-pro/> (visitato il 31/11/2019).
- [12] *Vive Cosmos product page*. URL: <https://www.vive.com/eu/product/vive-cosmos/features/> (visitato il 31/11/2019).
- [13] *Steam Index page*. URL: <https://www.valvesoftware.com/en/index/> (visitato il 31/11/2019).
- [14] Oculus. *Oculus Rift S product page*. URL: <https://www.oculus.com/rift-s/> (visitato il 30/11/2019).

- [15] Oculus. *Oculus Quest product page*. URL: <https://www.oculus.com/quest/> (visitato il 30/11/2019).
- [16] Pimax landing page. URL: <https://www.pimax.com/> (visitato il 31/11/2019).
- [17] Playstation VR page. URL: <https://www.playstation.com/en-us/explore/playstation-vr/> (visitato il 31/11/2019).
- [18] VR Ready Specifications. URL: <https://www.geforce.co.uk/hardware/technology/vr/system-requirements> (visitato il 02/12/2019).
- [19] Google Cardboard. URL: <https://arvr.google.com/cardboard/> (visitato il 03/12/2019).
- [20] MultiMedia LLC. *MS Windows NT Kernel Description*. URL: <https://developers.google.com/vr/discover/degrees-of-freedom> (visitato il 30/09/2019).
- [21] Steam VR Tracking. URL: <https://partner.steamgames.com/vrlicensing#Tracking/> (visitato il 03/12/2019).
- [22] Lighthouse Tracking. URL: <https://xinreality.com/wiki/Lighthouse/> (visitato il 03/12/2019).
- [23] Facebook. URL: <https://tech.fb.com/the-story-behind-oculus-insight-technology> (visitato il 30/09/2010).
- [24] Facebook AI. URL: <https://ai.facebook.com/blog/powered-by-ai-oculus-insight> (visitato il 30/09/2010).
- [25] Daniel Starkey. *THE RACE TO DEVELOP VR CONTROLS*. URL: <https://www.polygon.com/a/the-race-for-vr/the-race-for-vr-controls> (visitato il 31/11/2019).
- [26] Steam: VR Enabled. URL: <https://store.steampowered.com/vr/> (visitato il 03/12/2019).
- [27] *Artribune: new media section*. URL: <https://www.artribune.com/category/progettazione/new-media/> (visitato il 03/12/2019).
- [28] J Ferris. «Simulated ocular surgery: cataract, glaucoma, strabismus and VR surgery». In: *Acta Ophthalmologica* 92 (set. 2014). DOI: [10.1111/j.1755-3768.2014.3223.x](https://doi.org/10.1111/j.1755-3768.2014.3223.x).
- [29] Patrick Connolly. «Virtual Reality & Immersive Technology in Education». In: gen. 2008. DOI: [10.4018/9781599049496.ch158](https://doi.org/10.4018/9781599049496.ch158).
- [30] *What are VR simulators*. URL: <https://whatis.techtarget.com/definition/360-degree-VR-360-degree-virtual-reality> (visitato il 03/12/2019).
- [31] Kai Huotari e Juho Hamari. «Defining Gamification - A Service Marketing Perspective». In: ott. 2012. DOI: [10.1145/2393132.2393137](https://doi.org/10.1145/2393132.2393137).
- [32] *Classifying Serious Games: the G/P/S model*. URL: [http://www.ludoscience.com/files/ressources/classifying\\_serious\\_games.pdf](http://www.ludoscience.com/files/ressources/classifying_serious_games.pdf) (visitato il 03/12/2019).
- [33] Selcen Ozturkcan. «Virtual Reality Retailing». In: mar. 2018. DOI: [10.1007/978-3-319-08234-9\\_163-1](https://doi.org/10.1007/978-3-319-08234-9_163-1).



- [34] Jacquelyn Morie. «Virtual reality, immersion and the unforgettable experience». In: (feb. 2006). DOI: [10.1117/12.660290](https://doi.org/10.1117/12.660290).
- [35] Nobuhisa Tanaka e Hideyuki Takagi. «Design system of Virtual Reality Environment Based on VR Presence and VR Sickness». In: (giu. 2006), pp. 301–311. DOI: [10.18974/tvrsj.11.2\\_301](https://doi.org/10.18974/tvrsj.11.2_301).
- [36] Yipaer Aierken. «Perceptions of Virtual Reality: Virtual Reality in Perceptions». In: (lug. 2018).
- [37] Oona Kivela et al. «Study on the Motivational and Physical Effects of Two VR Exergames». In: set. 2019, pp. 1–2. DOI: [10.1109/VS-Games.2019.8864544](https://doi.org/10.1109/VS-Games.2019.8864544).
- [38] *Virtual digital assistants to overtake world population by 2021*. URL: <https://ovum.informa.com/resources/product-content/virtual-digital-assistants-to-overtake-world-population-by-2021> (visitato il 03/12/2019).
- [39] *Alexa Devs*. URL: <https://developer.amazon.com/en-US/alexa> (visitato il 03/12/2019).
- [40] *Siri*. URL: <https://www.apple.com/siri/> (visitato il 03/12/2019).
- [41] *Google Assistant*. URL: <https://assistant.google.com/> (visitato il 03/12/2019).
- [42] *Snips*. URL: <https://snips.ai/> (visitato il 03/12/2019).
- [43] *Mycroft site*. URL: <https://mycroft.ai/> (visitato il 03/12/2019).
- [44] Arttu Kääriä. *TECHNOLOGY ACCEPTANCE OF VOICE ASSISTANTS: ANTHROPOMORPHISM AS A FACTOR*. URL: <https://pdfs.semanticscholar.org/072e/f2fc80cfe07e48517ce361482f40b9512d5a.pdf>.
- [45] Robbert Willem de Kruijff. *Technology Acceptance of the Smart Speaker Exploring factors affecting the Use Intention of an emerging technology*. URL: <http://www.scriptiesonline.uba.uva.nl/document/668183>.
- [46] A. Schmeil e W. Broll. «MARA - A Mobile Augmented Reality-Based Virtual Assistant». In: *2007 IEEE Virtual Reality Conference*. 2007, pp. 267–270. DOI: [10.1109/VR.2007.352497](https://doi.org/10.1109/VR.2007.352497).
- [47] Dorian Lynskey. 'Alexa, are you invading my privacy?' – the dark side of our voice assistants. URL: <https://www.theguardian.com/technology/2019/oct/09/alexa-are-you-invading-my-privacy-the-dark-side-of-our-voice-assistants> (visitato il 30/11/2019).
- [48] Greg Sterling. *Alexa devices maintain 70% market share in U.S. according to survey*. URL: <https://marketingland.com/alexa-devices-maintain-70-market-share-in-u-s-according-to-survey-265180> (visitato il 31/10/2019).
- [49] *Intents, Utterances, and Slots: The New Vocabulary Needed To Develop For Voice*. URL: <https://medium.com/screenmedia-lab/utterances-slots-and-skills-the-new-vocabulary-needed-to-develop-for-voice-7428bff4ed79> (visitato il 03/12/2019).
- [50] Michael Metear, Zoraida Callejas e David Griol. «Spoken Language Understanding». In: mag. 2016, pp. 161–185. ISBN: 978-3-319-32965-9. DOI: [10.1007/978-3-319-32967-3\\_8](https://doi.org/10.1007/978-3-319-32967-3_8).

- [51] Xugang Lu, Sheng Li e Masakiyo Fujimoto. «Automatic Speech Recognition». In: gen. 2020, pp. 21–38. ISBN: 978-981-15-0594-2. DOI: [10.1007/978-981-15-0595-9\\_2](https://doi.org/10.1007/978-981-15-0595-9_2).
- [52] Nishith Pathak. «Using Natural Language Understanding». In: ago. 2017, pp. 71–92. ISBN: 978-1-4842-2948-4. DOI: [10.1007/978-1-4842-2949-1\\_4](https://doi.org/10.1007/978-1-4842-2949-1_4).
- [53] Yoshinori Shiga et al. «Text-to-Speech Synthesis». In: gen. 2020, pp. 39–52. ISBN: 978-981-15-0594-2. DOI: [10.1007/978-981-15-0595-9\\_3](https://doi.org/10.1007/978-981-15-0595-9_3).
- [54] Clifford Nass e Scott Brave. «Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship». In: (dic. 2019).
- [55] *Hidden Voice Commands*. URL: <https://people.eecs.berkeley.edu/~daw/papers/voice-usenix16.pdf> (visitato il 03/12/2019).
- [56] *USER PERCEPTION VS. REALITY IN VOICE CONTROLLED USER INTERFACES*. URL: <https://bunnyfoot.com/2018/01/user-perception-vs-reality-voice-controlled-user-interfaces/> (visitato il 03/12/2019).
- [57] Bethan Alexander e Daniela Alvarado. «Convergence of Physical and Virtual Retail Spaces». In: gen. 2017. DOI: [10.4018/978-1-5225-1865-5.ch008](https://doi.org/10.4018/978-1-5225-1865-5.ch008).
- [58] Yipaer Aierken. «Perceptions of Virtual Reality: Virtual Reality in Perceptions». In: (lug. 2018).
- [59] A. McNamara et al. «Information Placement in Virtual Reality». In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2019, pp. 1078–1079. DOI: [10.1109/VR.2019.8797910](https://doi.org/10.1109/VR.2019.8797910).
- [60] R. Jacobs, J. Karreman e J. Botma. «Beyond Clippy’s Counsel: Word Processor Feature Underuse among the Digital Generation». In: *2019 IEEE International Professional Communication Conference (ProComm)*. 2019, pp. 145–153. DOI: [10.1109/ProComm.2019.00033](https://doi.org/10.1109/ProComm.2019.00033).
- [61] *Unity landing page*. URL: <https://unity.com/> (visitato il 03/12/2019).
- [62] A. Syromiatnikov e D. Weyns. «A Journey through the Land of Model-View-Design Patterns». In: *2014 IEEE/IFIP Conference on Software Architecture*. 2014, pp. 21–30. DOI: [10.1109/WICSA.2014.13](https://doi.org/10.1109/WICSA.2014.13).
- [63] *DynamoDB main page*. URL: <https://aws.amazon.com/dynamodb/> (visitato il 03/12/2019).
- [64] *Lambda main page*. URL: <https://aws.amazon.com/lambda/> (visitato il 03/12/2019).
- [65] *Api Gateway main page*. URL: <https://aws.amazon.com/api-gateway/> (visitato il 03/12/2019).
- [66] *Serverless main page*. URL: <https://serverless.com/> (visitato il 03/12/2019).
- [67] *Cloudformation main page*. URL: <https://aws.amazon.com/cloudformation/> (visitato il 03/12/2019).
- [68] *Sumerian landing page*. URL: <https://aws.amazon.com/sumerian/> (visitato il 03/12/2019).
- [69] *Glossary: WebSockets*. URL: <https://developer.mozilla.org/en-US/docs/Glossary/WebSockets> (visitato il 03/12/2019).

- [70] *Pagina del laboratorio Ranzani*. URL: <https://disi.unibo.it/it/servizi-e-strutture/laboratori> (visitato il 03/12/2019).
- [71] *Pagina del Varlab*. URL: <https://site.unibo.it/varlab/en/about> (visitato il 03/12/2019).
- [72] *Steam VR*. URL: [https://support.steampowered.com/kb\\_article.php?ref=1131-WSFG-3320](https://support.steampowered.com/kb_article.php?ref=1131-WSFG-3320) (visitato il 03/12/2019).
- [73] *ProBuilder*. URL: <https://unity3d.com/unity/features/worldbuilding/probuilder/> (visitato il 03/12/2019).
- [74] *ProGrids*. URL: <https://www.procore3d.com/progrids/> (visitato il 03/12/2019).
- [75] *Unity Asset Store*. URL: <https://assetstore.unity.com/> (visitato il 03/12/2019).
- [76] *Rich Brian - History (Sample used in song) Ski Safari 2 Shop Theme*. URL: <https://www.youtube.com/watch?v=dRSwRowZb-o> (visitato il 30/11/2019).
- [77] *AWS*. URL: <https://aws.amazon.com/> (visitato il 03/12/2019).
- [78] *Enciclopedia Britannica. Liker scale*. URL: <https://www.britannica.com/topic/Likert-Scale> (visitato il 30/10/2019).