



**ΑΛΕΞΑΝΔΡΕΙΟ Τ.Ε.Ι. ΘΕΣΣΑΛΟΝΙΚΗΣ**  
**ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ**  
**ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ**



**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

**«Φωνητικές Πύλες-Τεχνολογία και Εφαρμογές»**



Του φοιτητή  
Γκουγκούλη Ανδρέα  
Αρ. Μητρώου: 03/2370

Επιβλέπων καθηγητής  
Αρσλάνογλου Αχιλλέας

Θεσσαλονίκη 2011

## Πρόλογος

Η παρούσα πτυχιακή εργασία εκπονήθηκε κατά το ακαδημαϊκό έτος 2010-2011 στα πλαίσια του προγράμματος σπουδών του τμήματος πληροφορικής της σχολής τεχνολόγων εφαρμογών του Αλεξάνδρειου Τεχνολογικού Ιδρύματος Θεσσαλονίκης. Για την ολοκλήρωση της απαιτήθηκε εκτενής έρευνα κυρίως στο διαδίκτυο, σε πηγές επί το πλείστον ξένης προέλευσης, αλλά και Ελληνικής . Η εργασία αυτή γράφτηκε με σκοπό να βοηθήσει τους αναγνώστες της να κατανοήσουν την έννοια των Φωνητικών Πυλών, να γνωρίσουν τις τεχνολογίες που χρησιμοποιούνται και να δουν τη χρησιμότητά τους μέσω του πλήθους υπηρεσιών που προσφέρουν.

## Περίληψη

Στην ανθρώπινη ιστορία η ομιλία ήταν το κυρίαρχο μέσο επικοινωνίας μεταξύ των ανθρώπων, γι' αυτό και η χρήση φωνής στην επικοινωνία του ανθρώπου με τον υπολογιστή αποτελεί αντικείμενο εντατικής έρευνας για πολλά χρόνια. Με τη ραγδαία ανάπτυξη της πληροφορικής και των τηλεπικοινωνιών και τη σύγκλισή τους δημιουργήθηκαν μια σειρά τεχνολογιών όπως η αναγνώριση και η σύνθεση ομιλίας, το αυτόματο σύστημα αλληλεπίδρασης φωνής καθώς και η VoiceXML μέσω των οποίων αναπτύχθηκαν οι Φωνητικές Πύλες. Μπορούμε να παρομοιάσουμε μία φωνητική πύλη με μια κοινή δικτυακή πύλη με τη διαφορά όμως ότι το περιβάλλον αλληλεπίδρασης είναι αποκλειστικά φωνητικό και προσβάσιμο μέσω μίας απλής τηλεφωνικής γραμμής. Όλοι πλέον, οπουδήποτε κι αν βρίσκονται και οποιαδήποτε στιγμή, μπορούν μέσω του τηλεφώνου και της φωνής τους να έχουν πρόσβαση σε όποια πληροφορία τους ενδιαφέρει, φυσικά και γρήγορα, χωρίς να είναι απαραίτητο να γνωρίζουν τη χρήση ηλεκτρονικών υπολογιστών και να περιμένουν στα τηλεφωνικά κέντρα. Οι φωνητικές πύλες προσφέρουν πλήθος υπηρεσιών από απλή πληροφόρηση μέχρι υπηρεσίες υγείας, υπηρεσίες αγοράς προϊόντων και τραπεζικές συναλλαγές.

ΛΕΞΕΙΣ-ΚΛΕΙΔΙΑ: Τηλεματική, Φωνητικές Πύλες, Τεχνολογία αυτόματης Αναγνώρισης Ομιλίας ASR, Τεχνολογία Μετατροπής Κειμένου Σε Ομιλία TTS, VoiceXML.

## Περίληψη στα Αγγλικά (Abstract)

In human history, the speech was the dominant mean of communication between people, so the use of voice in communication between man and computer is the subject of intensive research for many years. With the rapid development and convergence of informatics and telecommunications created technologies such as Speech Recognition *and* Synthesis, interactive voice response and the VoiceXML through which developed the Voice Portals. We can liken a voice portal with a common web portal, but the difference is that the interactive environment is exclusively vocal and accessible through a single telephone line. Everybody now anytime and wherever they are can access via phone and voice in whatever information they want, quickly and naturally, without needing to know the use of computers and wait for call centers. Voice portals offer many services from simple information to health care services, product market services and voice banking.

KEYWORDS: Telematics, Voice Portals, Automatic Speech Recognition ASR, Text to Speech technology TTS, VoiceXML.

## Ευχαριστίες

Σε αυτό το σημείο θα ήθελα να ευχαριστήσω ιδιαίτερα τον επιβλέποντα της εργασίας μου, Καθηγητή Κο Αχιλλέα Αρσλάνογλου, που μου έδωσε την ευκαιρία να εκπονήσω την παρούσα πτυχιακή εργασία καθώς η βασική ιδέα της ερευνας του ανήκει. Τον ευχαριστώ για τις πολύτιμες συμβουλές και για τη καθοδήγησή του σε όλη την διάρκεια της εκπόνησης αυτής της διπλωματικής εργασίας καθώς επίσης και για την εμπιστοσύνη και τη κατανόηση που έδειξε στο πρόσωπο μου παρόλες τις αργοπορίες μου.

Ευχαριστίες επίσης, θα ήθελα να εκφράσω στον Κο Κωνσταντίνο Σταμέλο για τη βοήθεια που μου προσέφερε σχετικά με το κεφάλαιο που αναφέρεται στη φωνητική πύλη «Γραμμή Για Την Υγεία»

Τέλος, θέλω να πω «ευχαριστώ» στην οικογένειά μου που μου συμπαραστάθηκε και με ενθάρρυνε με κάθε τρόπο σ' όλη τη διάρκεια των σπουδών μου αλλά και για την υπομονή που έδειξαν όλα αυτά τα χρόνια.

## Ευρετήριο Περιεχομένων

<b>ΠΡΟΛΟΓΟΣ</b> .....	<b>1</b>
<b>ΠΕΡΙΛΗΨΗ</b> .....	<b>2</b>
<b>ΠΕΡΙΛΗΨΗ ΣΤΑ ΑΓΓΛΙΚΑ (ABSTRACT)</b> .....	<b>3</b>
<b>ΕΥΧΑΡΙΣΤΙΕΣ</b> .....	<b>4</b>
<b>ΕΥΡΕΤΗΡΙΟ ΠΕΡΙΕΧΟΜΕΝΩΝ</b> .....	<b>5</b>
<b>ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ ΚΑΙ ΣΧΗΜΑΤΩΝ</b> .....	<b>8</b>
<b>ΕΙΣΑΓΩΓΗ</b> .....	<b>10</b>
<b>1. ΕΙΣΑΓΩΓΗ ΣΤΗ ΣΥΓΚΛΙΣΗ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ (ΤΗΛΕΜΑΤΙΚΗ)</b> .....	<b>11</b>
ΕΙΣΑΓΩΓΗ.....	11
1.1 ΙΣΤΟΡΙΑ .....	13
1.2 ΣΥΣΤΗΜΑ ΤΗΛΕΜΑΤΙΚΗΣ.....	16
1.3 ΔΙΚΤΥΑ ΤΗΛΕΜΑΤΙΚΗΣ.....	16
1.4 ΚΑΤΗΓΟΡΙΕΣ ΔΙΚΤΥΩΝ ΤΗΛΕΜΑΤΙΚΗΣ.....	17
1.5 ΤΗΛΕΜΑΤΙΚΕΣ ΥΠΗΡΕΣΙΕΣ .....	18
1.6 ΤΗΛΕΜΑΤΙΚΕΣ ΕΦΑΡΜΟΓΕΣ .....	21
1.6.1 Τηλεργασία (Teleworking ή Telecommuting).....	21
1.6.2 Τηλεκπαίδευση (Telelearning).....	22
1.6.3 Τηλεϊατρική.....	24
1.6.4 Ηλεκτρονικό εμπόριο (electronic commerce, e-commerce).....	24
ΕΠΙΛΟΓΟΣ.....	25
<b>2. ΦΩΝΗΤΙΚΕΣ ΠΥΛΕΣ</b> .....	<b>26</b>
ΕΙΣΑΓΩΓΗ.....	26
2.1 ΟΝΤΟΤΗΤΕΣ ΥΠΗΡΕΣΙΩΝ ΦΩΝΗΤΙΚΩΝ ΠΥΛΩΝ.....	28
2.2 ΠΥΛΗ (PORTAL) .....	28
2.3 ΓΙΑΤΙ ΧΡΗΣΙΜΟΠΟΙΕΙΤΑΙ Η ΦΩΝΗ;.....	29
2.4 ΚΑΤΗΓΟΡΙΕΣ ΦΩΝΗΤΙΚΩΝ ΠΥΛΩΝ.....	29
2.5 ΑΡΧΙΤΕΚΤΟΝΙΚΗ .....	30
ΕΠΙΛΟΓΟΣ.....	31
<b>3. ΤΕΧΝΟΛΟΓΙΕΣ ΦΩΝΗΤΙΚΩΝ ΠΥΛΩΝ</b> .....	<b>32</b>
ΕΙΣΑΓΩΓΗ.....	32
3.1 INTERACTIVE VOICE RESPONSE (IVR) .....	33
3.1.1 Ιστορία.....	34
3.1.2 DTMF.....	35

3.1.3 IVR Συστήματα και Φωνητικές Πύλες.....	36
3.1.4 Πλεονεκτήματα των IVR συστημάτων .....	37
3.1.5 Χαρακτηριστικά της IVR τεχνολογίας .....	37
3.2 ASR AUTOMATIC SPEECH RECOGNITION (ΤΕΧΝΟΛΟΓΙΑ ΑΝΑΓΝΩΡΙΣΗΣ ΟΜΙΛΙΑΣ) .....	39
3.2.1 Ιστορία .....	40
3.2.2 Αυτόματη διαδικασία αναγνώρισης ομιλίας.....	40
3.2.3 Φάσεις Συστήματος Αναγνώρισης Ομιλίας .....	42
3.2.4 Λόγοι παρουσίας σφαλμάτων στα Συστήματα Αναγνώρισης Ομιλίας.....	44
3.2.5 Εφαρμογές της ASR τεχνολογίας .....	45
3.2.6 Οφέλη της ASR τεχνολογίας.....	45
3.2.7 Μελλοντικά Σχέδια .....	45
3.3 TTS TEXT-TO-SPEECH .....	46
3.3.1 Συστατικά ενός TTS συνθέτη .....	47
3.3.2 Ποιότητα ενός TTS σύνθετη.....	48
3.3.3 Αρχιτεκτονική ενός TTS συνθέτη .....	48
3.3.4 Ιστορία της σύνθεσης φωνής .....	51
3.3.5 Εφαρμογές ενός TTS συστήματος.....	54
3.3.6 Γλώσσες σήμανσης σύνθεσης ομιλίας .....	54
3.4 VOICEXML .....	55
3.4.1 Αρχιτεκτονικό μοντέλο.....	58
3.4.2 Εφαρμογές της VoiceXML .....	58
3.4.3 Ιστορία της VoiceXML .....	59
3.4.4 Παραδείγματα VoiceXML .....	60
3.4.5 Γλώσσες που σχετίζονται με τη VoiceXML.....	61
<b>4. ΤΕΧΝΟΛΟΓΙΑ ΑΝΑΓΝΩΡΙΣΗΣ ΣΥΝΑΙΣΘΗΜΑΤΩΝ .....</b>	<b>63</b>
ΕΙΣΑΓΩΓΗ.....	63
4.1 ΔΙΑΔΙΚΑΣΙΕΣ ΕΝΟΣ ΣΥΣΤΗΜΑΤΟΣ ΑΝΑΓΝΩΡΙΣΗΣ ΣΥΝΑΙΣΘΗΜΑΤΩΝ .....	65
4.2 ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ ΣΥΝΑΙΣΘΗΜΑΤΙΚΗΣ ΟΜΙΛΙΑΣ .....	67
4.3 ΑΡΧΙΤΕΚΤΟΝΙΚΗ .....	70
4.4 ΣΤΡΑΤΗΓΙΚΕΣ ΣΥΝΔΙΑΛΛΑΓΗΣ .....	72
4.5 ΑΝΑΓΝΩΡΙΣΗ ΣΥΝΑΙΣΘΗΜΑΤΩΝ ΣΤΙΣ ΦΩΝΗΤΙΚΕΣ ΠΥΛΕΣ.....	73
4.6 ΕΦΑΡΜΟΓΕΣ ΤΕΧΝΟΛΟΓΙΑΣ ΑΝΑΓΝΩΡΙΣΗΣ ΣΥΝΑΙΣΘΗΜΑΤΩΝ .....	76
ΕΠΙΛΟΓΟΣ.....	77
<b>5. ΠΛΕΟΝΕΚΤΗΜΑΤΑ ΦΩΝΗΤΙΚΩΝ ΠΥΛΩΝ.....</b>	<b>79</b>
<b>6. ΥΠΗΡΕΣΙΕΣ ΦΩΝΗΤΙΚΩΝ ΠΥΛΩΝ .....</b>	<b>80</b>
ΕΙΣΑΓΩΓΗ.....	80
6.1 ΥΠΗΡΕΣΙΕΣ ΠΛΗΡΟΦΟΡΗΣΗΣ .....	80
6.2 ΥΠΗΡΕΣΙΕΣ VOICE MESSAGING .....	80
6.3 ΥΠΗΡΕΣΙΕΣ ΔΙΕΚΠΕΡΑΙΩΣΗΣ ΣΥΝΑΛΛΑΓΩΝ. ....	81
ΕΠΙΛΟΓΟΣ.....	81
<b>7. ΠΡΟΪΟΝΤΑ ΑΓΟΡΑΣ.....</b>	<b>82</b>

ΕΙΣΑΓΩΓΗ.....	82
7.1 1-800-555-TELL.....	82
7.2 ΦΩΝΗΤΙΚΗ ΠΥΛΗ ΤΟΥ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΚΡΗΤΗΣ .....	83
7.3 ΦΩΝΗΤΙΚΗ ΠΥΛΗ ΚΤΕΟ.....	86
7.4 ΜΥCOSMOS .....	86
7.5 ΦΩΝΗΤΙΚΗ ΠΥΛΗ ΔΡΟΜΟΛΟΓΙΩΝ ΠΛΟΙΩΝ 14789 .....	87
7.6 INFO TRIP .....	88
7.7 PHONE BANKING ΕΘΝΙΚΗΣ ΤΡΑΠΕΖΑΣ .....	89
7.7.1 Υπηρεσία Phone Banking.....	90
7.7.2 Ασφάλεια.....	91
7.7.3 Πλεονεκτήματα.....	92
7.7.4 Οδηγός Χρήσης.....	93
7.7.5 Συναλλαγές με φωνητικές εντολές!.....	94
<b>8. CASE STUDY «ΓΡΑΜΜΗ ΓΙΑ ΤΗΝ ΥΓΕΙΑ 1535» .....</b>	<b>97</b>
ΕΙΣΑΓΩΓΗ.....	97
8.1 ΥΠΗΡΕΣΙΕΣ .....	98
8.2 ΤΕΧΝΙΚΗ ΠΕΡΙΓΡΑΦΗ .....	99
8.3 ΛΕΙΤΟΥΡΓΙΑ .....	100
8.4 ΠΑΡΑΔΕΙΓΜΑ.....	101
ΕΠΙΛΟΓΟΣ.....	103
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ .....</b>	<b>104</b>



## Ευρετήριο Εικόνων Και Σχημάτων

Εικόνα 1 Τηλεματική : Πληροφορική + Τηλεπικοινωνίες .....	12
Εικόνα 2 Ηλεκτρικός τηλεγράφος .....	13
Εικόνα 3 Αυτόματη τηλεφωνική ανταλλαγή Strowger, 1891.....	14
Εικόνα 4 Η γενική μορφή ενός δικτύου τηλεματικής .....	16
Εικόνα 5 Ταξινόμηση υπηρεσιών τηλεματικής.....	20
Εικόνα 6 Τηλεργασία.....	21
Εικόνα 7 Σύστημα Τηλεκπαίδευσης.....	23
Εικόνα 8 Τηλεϊατρική .....	24
Εικόνα 9 Ηλεκτρονικό εμπόριο .....	25
Εικόνα 10 Φωνητικές Πύλες .....	26
Εικόνα 11 Αρχιτεκτονική Φωνητικών Πυλών .....	30
Εικόνα 12. Αρχιτεκτονική Φωνητικών Πυλών .....	31
Εικόνα 13 Συνομιλία ανθρώπου - υπολογιστή.....	32
Εικόνα 14 Αρχιτεκτονική IVR.....	34
Εικόνα 15 DTMF Dual Tone Multi frequency .....	35
Εικόνα 16 Χαρακτηριστικά της IVR τεχνολογία .....	38
Εικόνα 17 ASR Τεχνολογία Αναγνώρισης Ομιλίας.....	39
Εικόνα 18 Αρχιτεκτονική ενός συστήματος αναγνώρισης ομιλίας .....	41
Εικόνα 19 Οι φάσεις λειτουργίας ενός ολοκληρωμένου συστήματος αναγνώρισης ομιλίας .....	43
Εικόνα 20 Ακρίβεια αναγνώρισης ομιλίας.....	44
Εικόνα 21 Τεχνολογία Text-to-speech TTS.....	47
Εικόνα 22 Φάσεις ενός text-to-speech συνθέτη.....	47
Εικόνα 23 Επεξεργασία Φυσικής Γλώσσας.....	48
Εικόνα 24 Οι ακουστικοί ταλαντωτές του Christian Kratzenstein .....	51
Εικόνα 25 Μηχανή ομιλίας του Kempelen .....	52
Εικόνα 26 Ο συνθέτης φωνής VODER.....	53
Εικόνα 27 Χρονικό της σύνθεσης φωνής.....	53
Εικόνα 28 VoiceXML.....	56
Εικόνα 29 Διαδικτυακή εφαρμογή.....	57
Εικόνα 30 Φωνητική εφαρμογή.....	57
Εικόνα 31 Αρχιτεκτονικό μοντέλο VoiceXML.....	58
Εικόνα 32 Εξέλιξη της VoiceXML.....	60
Εικόνα 33 Διαφορετικές χρήσεις της συναισθηματικής επεξεργασίας σε συστήματα ηλεκτρονικών υπολογιστών .....	65
Εικόνα 34 διαδικασία εκπαίδευσης στο αυτόματο σύστημα αναγνώρισης συναισθημάτων.....	66
Εικόνα 35 Διαδικασία κανονικής λειτουργίας στο αυτόματο σύστημα αναγνώρισης συναισθημάτων .....	67
Εικόνα 36 Ένταξη της συναισθηματικής επεξεργασίας στην αρχιτεκτονική των φωνητικών πυλών .....	70
Εικόνα 37 Μεταβολή χαρακτηριστικών προσωπείας και ποιότητας φωνής σε σχέση με τα συναισθήματα.....	71
Εικόνα 38 Φωνητική Πύλη του Πανεπιστημίου Κρήτης.....	85

Εικόνα 39 Phone Banking.....	89
Εικόνα 40 Παράδειγμα Υπολοίπου .....	95
Εικόνα 41 Παράδειγμα μεταφοράς ποσού.....	96
Εικόνα 42 Αρχιτεκτονική φωνητικής πύλης 1535 Γραμμή για την Υγεία .....	99

## Εισαγωγή

Στο πρώτο κεφάλαιο γίνεται μια εισαγωγή στη σύγκλιση τηλεπικοινωνιών και πληροφορικής, αναφέρονται οι κατηγορίες των δικτύων τηλεματικής καθώς και οι υπηρεσίες και οι εφαρμογές που προσφέρονται. Στο επόμενο κεφάλαιο δίνεται ο ορισμός και η αρχιτεκτονική των φωνητικών πυλών και η χρησιμότητα της φωνής στην επικοινωνία ανθρώπου-μηχανής. Στο τρίτο κεφάλαιο γίνεται μια επισκόπηση στις τεχνολογίες που χρησιμοποιούνται για να υλοποιηθεί μια φωνητική πύλη, τεχνολογίες όπως η σύνθεση και αναγνώριση φωνής το αυτόματο σύστημα αλληλεπίδρασης φωνής αλλά και το πρότυπο VoiceXML. Στο τέταρτο κεφάλαιο εξετάζεται η τεχνολογία αναγνώρισης συναισθημάτων που κάνει πιο φιλική και αληθοφανή τη φωνητική πύλη στους χρήστες, και αναφέρονται οι εφαρμογές και η αρχιτεκτονική της τεχνολογίας αυτής. Στη συνέχεια γίνεται μια αναφορά στα πλεονεκτήματα και στις υπηρεσίες που προσφέρουν οι φωνητικές πύλες. Στα τελευταία δύο κεφάλαια παρουσιάζονται ορισμένες πρακτικές εφαρμογές, με περισσότερη ανάλυση του προϊόντος Γραμμή Για Την Υγεία.

# 1. Εισαγωγή στη Σύγκλιση Τηλεπικοινωνιών και Πληροφορικής (Τηλεματική)

## Εισαγωγή

Κυρίως την τελευταία εικοσαετία η χρήση συστημάτων ηλεκτρονικών υπολογιστών (Η/Υ) εξαπλώθηκε στον ευρύτερο γεωγραφικό χώρο . Η ραγδαία ανάπτυξη του τομέα της πληροφορικής δημιούργησε την ανάγκη αυτά τα συστήματα να μπορέσουν να συνδεθούν και να επικοινωνήσουν μεταξύ τους. Αυτό επετεύχθη με τη χρήση τηλεπικοινωνιακών μέσων τα οποία είχαν αρχικά κατασκευαστεί για διαφορετικού είδους επικοινωνία όπως για παράδειγμα η τηλεφωνία.

Ωστόσο ,παράλληλα, το ίδιο χρονικό διάστημα υπήρξαν και σημαντικές εξελίξεις και στο τομέα των τηλεπικοινωνιών, που χαρακτηρίζονται κυρίως από την εφαρμογή της ψηφιακής τεχνολογίας, η οποία αποτελεί τη γέφυρα ολοκλήρωσης της τηλεπικοινωνιακής τεχνολογίας και της τεχνολογίας των υπολογιστών.

Η σύγκλιση και η σύζευξη της Πληροφορικής και των Τηλεπικοινωνιών οδήγησε στην εμφάνιση της Τηλεματικής ή Τηλεπληροφορικής.

Πιο συγκεκριμένα τηλεματική είναι η επιστήμη της αποστολής, της λήψης και της αποθήκευσης των πληροφοριών μέσω των συσκευών τηλεπικοινωνιών. Περιλαμβάνει κάθε είδους μεταφορά δεδομένων μεταξύ υπολογιστικών συστημάτων που συνδέονται μεταξύ τους με τηλεπικοινωνιακές γραμμές υποστηρίζοντας εντυπωσιακές και χρήσιμες εφαρμογές στις ιδιωτικές και δημόσιες επιχειρήσεις.

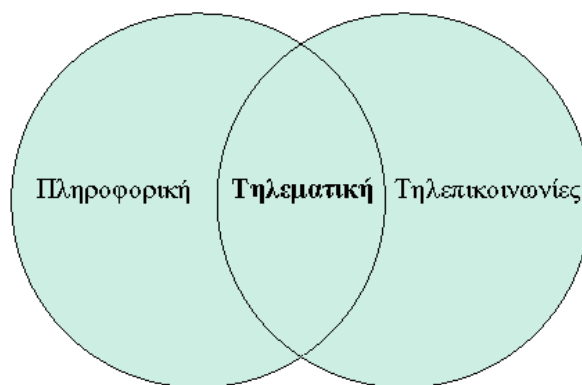
Ο όρος τηλεματική αρχικά αναφερόταν στην τεχνολογία συστημάτων παγκόσμιας πλοήγησης GPS που ενσωματώνεται με την τεχνολογία υπολογιστών και κινητής επικοινωνίας στα συστήματα πλοήγησης αυτοκινήτων.

Η τηλεματική περιλαμβάνει αλλά δεν περιορίζεται στη τεχνολογία GPS. Το πιο χαρακτηριστικό παράδειγμα τηλεματικής είναι το ίδιο το Internet. Οι υπηρεσίες τηλεματικής τα τελευταία χρόνια έχουν αναπτυχθεί ιδιαίτερα. Με την εξάπλωση του διαδικτύου έχει δοθεί νέα ώθηση στην ανάπτυξη και παροχή νέων υπηρεσιών. Υπηρεσίες όπως η μετάδοση πολυμέσων και η τηλεδιάσκεψη μέσω του διαδικτύου γνωρίζουν μεγάλη αποδοχή. Μερικά παραδείγματα τηλεματικής είναι τα συστήματα επικοινωνίας μέσω του δικτύου υπολογιστών (π.χ., ηλεκτρονικό

ταχυδρομείο e-mail), τα δίκτυα καλωδιακής τηλεόρασης και ψηφιακής τηλεόρασης, το Internet, τα συστήματα αγορών από απόσταση, τα συστήματα ανάληψης χρημάτων από απόσταση, τα συστήματα videoδιάσκεψης μέσω του Internet που επιτρέπουν ομαδικές συζητήσεις όπου οι συνομιλητές βλέπουν και ακούν ο ένας τον άλλο σε πραγματικό χρόνο. Άλλες εφαρμογές μελετώνται ή αναπτύσσονται για την παρακολούθηση της ρύπανσης των υδάτων και του αέρα, για την ιατρική πληροφορική και την υγειονομική περίθαλψη.

Η χρήση της τηλεματικής έχει προκαλέσει σημαντικές αλλαγές στην κοινωνία και στην καθημερινή ζωή. Οι πολίτες έχουν καλύτερη πρόσβαση σε κάθε είδους πληροφορία που μπορεί να χρειαστούν. Οι καταναλωτές μπορούν να αγοράζουν προϊόντα από απόσταση. Οι μαθητές σε όλους τους τομείς εκπαίδευσης παρακολουθούν εκπαίδευση από απόσταση, χρησιμοποιούν το Internet και τον Παγκόσμιο Ιστό για να βρίσκουν πληροφορίες και τα σχολεία έχουν νέες ευκαιρίες για επικοινωνία και συνεργασία με άλλα σχολεία ή κοινωνικούς φορείς. Οι εργαζόμενοι μπορούν να εργάζονται από το σπίτι μέσω της τηλεργασίας. Οι επιχειρήσεις μπορούν να παρέχουν νέες υπηρεσίες από απόσταση με χαμηλό κόστος και να διευρύνουν το αγοραστικό κοινό τους.

Ο όρος Τηλεματική δημιουργήθηκε από τους Γάλλους Simon Nora και Alain Minc το 1976 και προέρχεται από τις αγγλικές λέξεις : Telecommunication + Informatics.



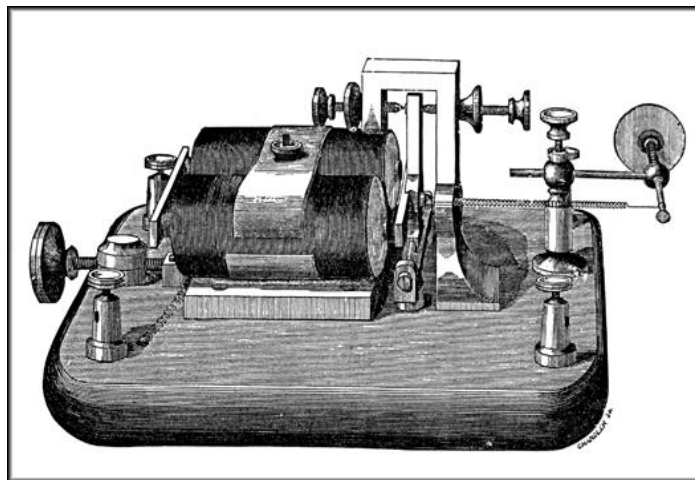
Εικόνα 1 Τηλεματική : Πληροφορική + Τηλεπικοινωνίες

## 1.1 Ιστορία

Η αφετηρία για την τηλεματική θεωρείται το 1832, όταν πραγματοποιήθηκε η εφεύρεση του ηλεκτρικού τηλέγραφου. Παρά το γεγονός ότι διαφορετικά είδη οπτικών τηλέγραφων και εξοπλισμών σηματοδότησης είχαν εφευρεθεί μαζί την πορεία της ανθρώπινης ιστορίας, ο ηλεκτρικός τηλέγραφος θεωρείται το πρώτο σύγχρονο σύστημα επικοινωνίας δεδομένων. Ο ηλεκτρικός τηλέγραφος περιέχει διάφορα βασικά χαρακτηριστικά που είναι κοινά στα περισσότερα σύγχρονα τηλεματικά συστήματα επικοινωνιών:

- Στηρίχθηκε σε ένα ηλεκτρομαγνητικό σήμα για να μεταβιβάσει τις πληροφορίες
- Βασίστηκε στη μετάδοση των συμβόλων από ένα ιδιαίτερο (αλφαριθμητικό) αλφάβητο.
- Ενσωμάτωσε την έννοια της βέλτιστης κωδικοποίησης πηγής
- Από 1845, χρησιμοποίησε τους κώδικες συμπίεσης: κώδικες Μορς και άλλους

Η στοιχειώδης χρήση της ψηφιακής επεξεργασίας, της κωδικοποίησης πηγής, της συμπίεσης στοιχείων, και των κωδίκων κρυπτογράφησης έχει οδηγήσει στους αντίστοιχους τομείς της τηλεματικής σήμερα: θεωρία της πληροφορίας, βέλτιστη κωδικοποίηση πηγής, αλγόριθμοι και τεχνικές συμπίεσης, θεωρίες κρυπτογράφησης και ασφάλειας των δικτύων.



Εικόνα 2 Ηλεκτρικός τηλέγραφος

Το 1891 σηματοδοτεί τη δεύτερο σχετικό κύριο σημείο της Τηλεματικής, με την εφεύρεση της αυτόματης τηλεφωνικής ανταλλαγής από τον Strowger. Η αυτόματη ανταλλαγή, πρώτα ηλεκτρομηχανικά και αργότερα πλήρως ηλεκτρονικά, ενσωμάτωσε την έννοια της σηματοδότησης μεταξύ των συστημάτων. Η έννοια της μεταγωγής (switching) είναι θεμελιώδης για την Τηλεματική, καθώς είναι βασικό στοιχείο στο σχεδιασμό και την εφαρμογή του δικτύου. Η ανάπτυξη του switch άνοιξε νέους τομείς έρευνας που οδήγησαν στη γέννηση νέων στοιχείων όπως της αρχιτεκτονικής μεταγωγής κυκλωμάτων και των τεχνικών δρομολόγησης.



Εικόνα 3 Αυτόματη τηλεφωνική ανταλλαγή Strowger, 1891.

Η σύγκλιση μεταξύ της μετάδοσης στοιχείων και αυτόματης μεταγωγής επήλθε μέσω της ανάπτυξης των συστημάτων τηλετύπου. Τα δίκτυα τηλετύπου άρχισαν στη δεκαετία του '20 σαν βελτιώσεις του ηλεκτρικού τηλεγράφου αρχικά στον τερματικό εξοπλισμό και αργότερα στις τεχνικές μετάδοσης. Στα τέλη της δεκαετίας του 1930 η ενσωμάτωση της αυτόματης μεταγωγής, των συστημάτων σηματοδότησης και της αποθήκευσης μηνυμάτων είχαν ως αποτέλεσμα το πρώτο πραγματικά Τηλεματικό σύστημα επικοινωνίας. Τα συστήματα τηλετύπου στη δεκαετία του '30 βασίστηκαν στην ηλεκτρομηχανική επεξεργασία και αποθήκευση. Η δημιουργία του πρώτου ψηφιακού ηλεκτρονικού υπολογιστή, του ENIAC, το 1945 πυροδότησε τη σύγκλιση μεταξύ του υπολογισμού και της επικοινωνίας. Η ανάπτυξη των τεχνολογιών της ηλεκτρονικής επεξεργασίας και αποθήκευσης έδωσαν τη δυνατότητα για τη δημιουργία ενός μεγάλου αριθμού τηλεματικών υπηρεσιών επικοινωνίας.

Η σύγκλιση μεταξύ των υπολογιστών και των επικοινωνιών ξεκίνησε με την ανάγκη να αποκτήσουν απομακρυσμένη πρόσβαση σε εγκατάσταση υπολογιστών( τηλεπεξεργασία). Η ανάγκη των μηχανών να επικοινωνήσουν μεταξύ τους οδήγησαν στην ανάπτυξη «γλωσσών» οι οποίες δεν είναι τίποτε άλλο από τα πρωτόκολλα επικοινωνίας, που βασίζονται στη διαμόρφωση της ανταλλαγής πληροφοριών μεταξύ των ανθρώπων με τέτοιο τρόπο ώστε να μπορεί να εφαρμοστεί σε επίπεδο μηχανής. Η ανάπτυξη πρωτοκόλλων επικοινωνίας σηματοδοτεί μια ριζική τομή στην ιστορία της τηλεματικής καθώς οι αρχιτεκτονικές πρωτοκόλλου και η θεωρία πρωτοκόλλου είναι τώρα μια από τις πιό θεμελιώδεις έννοιές της.

Στο τέλος της δεκαετίας του 60 αναπτύχθηκε και επεκτάθηκε ο πρώτος υπολογιστής βασισμένος σε εξοπλισμό τηλεφωνικής μεταγωγής. Το 1968 η Ericsson AKE 12, εγκαταστάθηκε στη Tumba (Σουηδία). Η κατευθυντήρια ιδέα ήταν να εφαρμοστεί η ευελιξία και η μεταβλητότητα των υπολογιστών στη διαχείριση και τον έλεγχο των συστημάτων τηλεπικοινωνιών. Η επιτυχία αυτής της αρχικής εφαρμογής οδήγησε αργότερα σε έναν πλήρη επανασχεδιασμό της τηλεφωνικής σηματοδότησης βασισμένη στα δίκτυα πακετων και την επικοινωνία Πρωτόκολλων. Η εφαρμογή των υπολογιστών εισήχθη επίσης για να αυτοματοποιήσει με έναν πολύ εύκαμπτο τρόπο τις διοικητικές λειτουργίες των συστημάτων επικοινωνιών.

Το 1969 το ARPANET δημιουργήθηκε χρησιμοποιώντας το τότε επαναστατικό 32bit μικροεπεξεργαστή , ήταν δυνατό να εκτελεσθεί η μετατροπή, η σηματοδότηση και η διαχείριση στοιχείων ολόκληρου του δικτύου με έναν ολοκληρωμένο τρόπο. Το ARPANET ενσωμάτωσε συγχρόνως μια άλλη επαναστατική έννοια: τις επικοινωνίες μεταγωγής πακέτων. Το ARPANET ήταν το πρώτο σύγχρονο δίκτυο που έχει σχεδιαστεί ειδικά μετά τις θεωρίες και τις αρχές του επιστημονικού κλάδου της Τηλεματικής. Εξελίχθηκε για να γίνει αυτό που σήμερα θεωρείται ως η βάση μιας παγκόσμιας κοινωνικο-οικονομικής επανάστασης: το Διαδίκτυο.

Χρησιμοποιώντας το πρωτόκολλο διαδικτύου Internet Protocol (IP) ως τη κινητήρια δύναμη υπηρεσίας δικτύου, το 1971 είδε την αρχή της έκρηξης των βασισμένων σε υπολογιστή τηλεματικών υπηρεσιών επικοινωνίας. Ο όρος Τηλεματική ορίστηκε για πρώτη φορά από τους Γάλλους Simon Nora και Alain Minc το 1976.



## 1.2 Σύστημα Τηλεματικής

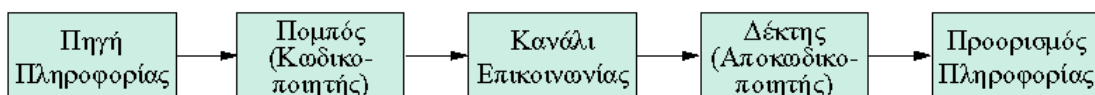
Ένα σύστημα τηλεματικής είναι ένα σύστημα διαχείρισης πληροφοριών το οποίο αξιοποιεί στο έπακρο τις δυνατότητες των τηλεπικοινωνιακών δικτύων και των συστημάτων Η/Υ (λογισμικό και υλικό). Οι πληροφορίες ακολουθούν χρονικά και λογικά έναν κύκλο.

Ο κύκλος αυτός περιλαμβάνει τις εξής φάσεις:

- Δημιουργίας – εισόδου.
- Επεξεργασίας.
- Αρχαιοθέτησης.
- Αναζήτησης – ανάκτησης.
- Τηλεπικοινωνιακής αποστολής – μετάδοσης – λήψης.
- Εξόδου.

## 1.3 Δίκτυα Τηλεματικής

Πάνω στα δίκτυα τηλεματικής πραγματοποιούνται οι διάφορες υπηρεσίες τηλεματικής. Περιλαμβάνουν τα κανάλια επικοινωνίας, τους πομπούς και τους δέκτες όπου γίνεται η κωδικοποίηση και η αποκωδικοποίηση της επικοινωνίας και τις τερματικές συσκευές των χρηστών. Η επικοινωνία μπορεί να είναι αμφίδρομη και ο πομπός μπορεί να είναι ταυτόχρονα και δέκτης.



Εικόνα 4 Η γενική μορφή ενός δικτύου τηλεματικής

Για να μπορέσει να επιτευχθεί η μεταβίβαση πληροφοριών από έναν πομπό σε έναν δέκτη χρειάζονται τα εξής:

- Ένας κώδικας που εξασφαλίζει ότι οι πληροφορίες του πομπού γίνονται κατανοητές από το δέκτη.
- Ένα σύστημα μετάδοσης.
- Ένα κανάλι (δίαυλος) επικοινωνίας.
- Ένα σύστημα λήψης.

Η χρήση κοινού κώδικα είναι απαραίτητη για την πραγματοποίηση της επικοινωνίας.

#### **1.4 Κατηγορίες Δικτύων Τηλεματικής**

Υπάρχουν πέντε κατηγορίες δικτύων τηλεματικής, τα δίκτυα μεταγωγής, τα δίκτυα διανομής, τα δίκτυα συλλογής, τα αναλογικά και τα ψηφιακά δίκτυα. Τα δίκτυα αυτά διακρίνονται από τη δυνατότητα μεταφοράς μεγάλου όγκου πληροφοριών στη μονάδα του χρόνου, το μέγεθος του δικτύου που επηρεάζει τη δομή του (τοπικά δίκτυα, μητροπολιτικά δίκτυα, δίκτυα ευρείας περιοχής), τα χρησιμοποιούμενα μέσα μετάδοσης (κοινά καλώδια, ομοαξονικά καλώδια, οπτικές ίνες, δορυφόροι, κ.λπ.), τη κυριότητα του δικτύου (ιδιωτικό, δημόσιο).

- Δίκτυα μεταγωγής- Αποτελούνται από τις γραμμές μετάδοσης και από μια ιεραρχία κέντρων μεταγωγής. Είναι δυνατή η αποκλειστική επικοινωνία μεταξύ δύο οποιωνδήποτε χρηστών που είναι συνδεδεμένοι στο δίκτυο. Τέτοια δίκτυα είναι το τηλεφωνικό δίκτυο, το τηλετυπικό δίκτυο, τα δίκτυα μεταγωγής δεδομένων.
- Δίκτυα διανομής ή εκπομπής: Από ένα κεντρικό σταθμό αποστέλλονται μηνύματα προς όλους τους χρήστες που συνδέονται με τον κεντρικό σταθμό. Η σύνδεση μπορεί να είναι ενσύρματη ή ασύρματη. Τα ασύρματα ραδιοηλεκτρονικά δίκτυα, και τα δίκτυα ενσύρματης τηλεόρασης ανήκουν σ' αυτή τη κατηγορία.
- Δίκτυα συλλογής: Κάθε χρήστης συνδέεται με ένα κεντρικό σταθμό στον οποίο έχει τη δυνατότητα να αποστέλλει μηνύματα. Συνυπάρχουν συνήθως με τα δίκτυα διανομής.
- Αναλογικά δίκτυα: Μεταβιβάζουν αναλογικά σήματα. Χαρακτηριστικό παράδειγμα είναι το συμβατικό τηλεφωνικό δίκτυο. Είναι φτιαγμένο να μεταδίδει τα αναλογικά σήματα ομιλίας.
- Ψηφιακά δίκτυα: Η μεταγωγή και η μετάδοση γίνονται με βάση την ψηφιακή τεχνική. Τα δίκτυα μεταγωγής είναι σε μεγάλο βαθμό (ή και ολοκληρωτικά) ψηφιακά.

## 1.5 Τηλεματικές υπηρεσίες

Σύμφωνα με τη Διεθνή Ένωση Τηλεπικοινωνιών (International Telecommunications Union, ITU) οι υπηρεσίες τηλεματικής ορίζονται ως εξής:

*«Υπηρεσίες τηλεματικής είναι το σύνολο των υπηρεσιών οι οποίες είναι διαφορετικές από τις συμβατικές τηλεφωνικές και τηλεγραφικές υπηρεσίες και οι οποίες είναι δυνατό να παρέχονται στους συνδρομητές ενός τηλεπικοινωνιακού δικτύου. Οι υπηρεσίες αυτές επιτρέπουν να στέλνονται ή να λαμβάνονται ενημερωτικές πληροφορίες δημόσιες ή ιδιωτικές στις οποίες συμπεριλαμβάνονται λήψη στοιχείων από αρχεία, εμπορικές ή τραπεζικές εργασίες, κλείσιμο θέσεων, κ.λπ.»*

Παρακάτω περιγράφονται μερικές από τις πιο σημαντικές υπηρεσίες τηλεματικής

- Τηλεομοιοτυπία (Fax) το γνωστό σε όλο μας φαξ μέσω του οποίου μπορούν να σταλούν κείμενα ή γραφικά σε χαρτί. Οι συσκευές Fax λειτουργούν συγχρόνως σαν σαρωτές και modem. Πρώτα γίνεται η σάρωση του περιεχομένου του χαρτιού και στη συνέχεια κωδικοποιείται για την αποστολή του.
- Τηλεκειμενογραφία (Teletext). Η συγκεκριμένη υπηρεσία μεταδίδεται μέσω της τηλεόρασης με δεδομένα κειμένου και πληροφορεί τους χρήστες σε πάρα πολλούς τομείς. Χρησιμοποιεί το τηλεοπτικό δίκτυο για να εκπέμπει ταυτόχρονα με τις τηλεοπτικές εκπομπές τις υπηρεσίες που προσφέρει. Ένας κοινός δέκτης τηλεόρασης μπορεί να λάβει τα σήματα της τηλεκειμενογραφίας, αρκεί να εφοδιαστεί με έναν αποκωδικοποιητή.
- Τηλεηχοπληροφόρηση (Audiotext). Η υπηρεσία αυτή παρέχει πληροφορίες με ήχο που είναι μαγνητοφωνημένες ή τεχνικά παραγμένες και καλύπτουν θέματα ψυχαγωγίας, ενημέρωσης και επιστήμης.
- Εικονοτηλεφωνία (VideoPhone). Δίνει τη δυνατότητα στους χρήστες που συνομιλούν μέσω τηλεφώνου να έχουν οπτική επαφή.
- Τηλεδιάσκεψη (Video Conference). Με την υπηρεσία αυτή μπορούν να είναι σε οπτική και ακουστική επαφή ταυτόχρονα αρκετοί άνθρωποι από διαφορετικές γεωγραφικές περιοχές. Χρησιμοποιείται από ομάδες που επικοινωνούν μεταξύ τους για ανταλλαγή ιδεών, απόκτηση πληροφοριών, σύγχρονη εκπαίδευση και διαχείριση οργανισμών.

- Τηλεειδοποίηση (Paging): Πρόκειται για μια συσκευή ειδοποίησης που είναι ένας δέκτης ηχητικού σήματος τον οποίο μπορούμε να καλέσουμε με μια οποιαδήποτε κοινή τηλεφωνική συσκευή. Για την υπηρεσία αυτή λειτουργούν ειδικά κέντρα τηλεειδοποίησης τα οποία είναι συνδεδεμένα με τηλεφωνικά κέντρα.
- Κινητή Τηλεφωνία (mobile communication): στη κινητή τηλεφωνία παρέχονται ένα σύνολο τηλεματικών υπηρεσιών όπως ο αυτόματος τηλεφωνητής, η αποστολή μηνυμάτων από άλλο κινητό ή από υπολογιστή, κλπ.
- Ηλεκτρονικό Ταχυδρομείο (e-mail): δίνει τη δυνατότητα αποστολής μηνύματος κειμένου, και με συνημμένο τρόπο, οποιουδήποτε αρχείου κάθε μορφής.
- Τηλεξυπηρέτηση (Teleservice) Στην κατηγορία αυτή υπάγονται πάρα πολλές υπηρεσίες που παρέχονται από απόσταση. Οι πληροφορίες αυτές μπορεί να προέρχονται από Κρατικές ή ιδιωτικές υπηρεσίες και εταιρείες.
- Τηλεαγορές, τηλετραπεζικές εργασίες (για αύξηση της εξυπηρέτησης και της ασφάλειας των συναλλαγών).

Στο σχήμα που ακολουθεί οι διάφορες υπηρεσίες τηλεματικής ταξινομούνται ανάλογα με το είδος του δικτύου (ενσύρματα, ασύρματα, διανομής και μεταγωγής) και ανάλογα με το είδος της επικοινωνίας (φωνής, κειμένου, δεδομένων κλπ).

Απαιτήσεις σε Δίκτυο		Είδος Επικοινωνίας	Ονομασία-μορφή Υπηρεσίας
Υπάργοντα Δίκτυα	Ενσύρματα Δίκτυα	Επικοινωνία φωνής	Προηγμένη τηλεφωνία
		Επικοινωνία κειμένου	Ταχυτηλετυπία (teletex)
			Τηλεκειμενογραφία (Videotex εκπομπής ή teletext)
			Τηλεεικονογραφία (Διαλογικό videotex)
		Επικοινωνία ακίνητης εικόνας	Τηλεομοιοτυπία Ηλεκτρονική εφημερίδα
		Επικοινωνία δεδομένων	Επικοινωνία μεταξύ Η/Υ Μεταφορά αρχείων
		Συνδυασμός επικοινωνίας φωνής, κειμένου, εικόνας και δεδομένων	Ηλεκτρονικό ταχυδρομείο Παγκόσμιος ιστός (WWW)
	Τηλεπιδραση	Τηλεπίβλεψη Τηλέλεγχος	
Τηλεσυνεδρίαση	Τηλεφωνική συνεδρίαση Τηλεφωνική συνεδρίαση με γραφικά		
Ασύρματα Δίκτυα, Δίκτυα Κινητών Υπηρεσιών	Ασύρματη, κινητή	Τηλεειδοποίηση Κινητή τηλεφωνία Ασύρματη επικοινωνία φωνής, κειμένου, εικόνας, δεδομένων	
Νέα Δίκτυα Ευρείας Ζώνης	Δίκτυα διανομής	Επικοινωνία κινούμενης εικόνας (Διανομή)	Καλωδιακή τηλεόραση (με ή χωρίς κύκλωμα επιστροφής) Δορυφορική τηλεόραση
	Δίκτυα μεταγωγής	Επικοινωνία κινούμενης εικόνας (Μεταγωγή)	Εικονοτηλέφωνο Τηλεεικονοδιάσκεψη Κλήση κινούμενης εικόνας

Εικόνα 5 Ταξινόμηση υπηρεσιών τηλεματικής

## 1.6 Τηλεματικές εφαρμογές

Ο όρος τηλεματικές εφαρμογές αναφέρεται στις υπηρεσίες που προσφέρει η σύγχρονη τεχνολογία μέσω των οποίων μπορούν να αποσταλούν και να ληφθούν κάθε φύσης πληροφορίες. Οι πληροφορίες μπορεί να είναι ακουστικές, οπτικές, εικόνας ή κειμένου και μεταδίδονται μέσω τηλεόρασης, υπολογιστή ή άλλων ειδικών συσκευών. Ο συνδυασμός δύο ή περισσότερων υπηρεσιών τηλεματικής για την κάλυψη πιο σύνθετων πληροφοριακών και επικοινωνιακών αναγκών ονομάζεται εφαρμογή τηλεματικής. Οι τηλεματικές εφαρμογές κερδίζουν συνεχώς έδαφος στο σύγχρονο κόσμο αλλάζοντας ριζικά τους τρόπους επικοινωνίας και μετάδοσης πληροφοριών.

Οι εφαρμογές τηλεματικής κάνουν χρήση πολλών τεχνολογιών και διαφόρων τεχνολογικών μέσων. Πλέον φαίνεται όλο και περισσότερο η προσπάθεια σύγκλισης και ολοκλήρωσης όλων των υπηρεσιών με κεντρικό άξονα τα δίκτυα υπολογιστών.

Στα παρακάτω υποκεφάλαια αναφέρονται και περιγράφονται οι κυριότερες τηλεματικές εφαρμογές .

### 1.6.1 Τηλεργασία (Teleworking ή Telecommuting)

Η τηλεργασία είναι ένας νέος τρόπος οργάνωσης της εργασίας μέσω της χρήσης της Πληροφορικής και των νέων Τεχνολογιών για την επεξεργασία και ανταλλαγή πληροφοριών ή εγγράφων όπως κειμένων, εικόνων, γραφικών και ήχων.

Μια εργασία που θα μπορούσε να εκτελεστεί στις εγκαταστάσεις του εργοδότη εκτελείται κανονικά εκτός αυτών των εγκαταστάσεων. Οι υπάλληλοι οργανισμών καθώς και οι ελεύθεροι επαγγελματίες θα μπορούν να εργάζονται κατ' οίκον ή σε αποκεντρωμένα γραφεία, έτσι ώστε να μην απαιτείται πλέον να διανύουν αποστάσεις από και προς τον τόπο εργασίας τους .



Εικόνα 6 Τηλεργασία

Η τηλεργασία έχει αρκετά πλεονεκτήματα, μερικά από τα οποία αναφέρονται παρακάτω.

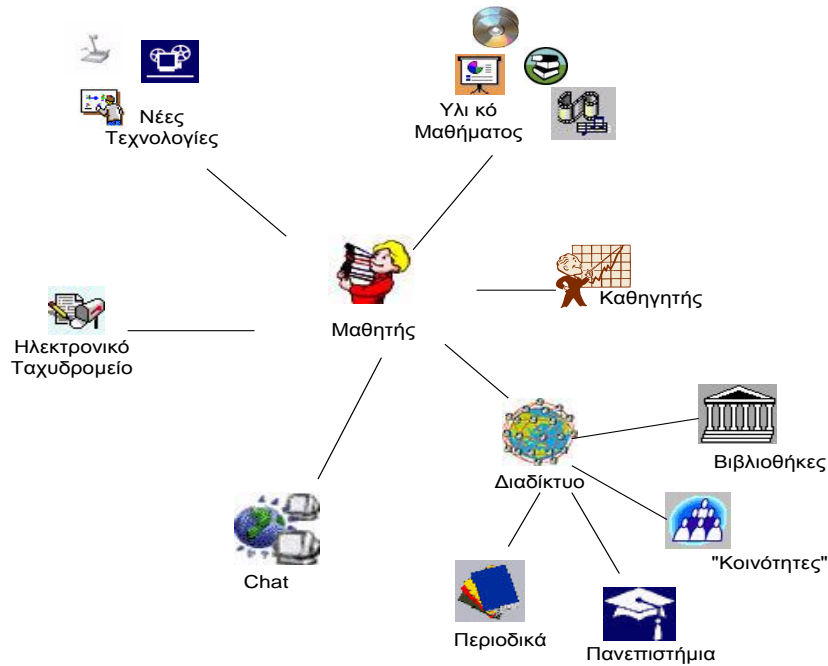
- Καταπολεμά τον κοινωνικό αποκλεισμό
- Διαφυλάσσει την προστασία του περιβάλλοντος
- Ενδυναμώνει την οικονομία
- Αντιστέκεται στο φαινόμενο της αστυφιλίας και προάγει την αποκέντρωση
- Βοηθά τον εργαζόμενο να διαλέξει τον τόπο και το χρόνο εργασίας του
- Προάγει μια πιο ολοκληρωμένη εικόνα του ανθρώπου (όχι μόνο ως επαγγελματία, αλλά και ως γονέα, ως μέλους της κοινωνίας κ.λ.π.)
- Εξοικονομεί χρόνο και για τον εργαζόμενο και για την εταιρεία στην οποία απασχολείται
- Εξοικονομεί χώρο στην εταιρεία
- Εξαπλώνεται παντού και γρήγορα
- Εξοικειώνει το σύγχρονο άνθρωπο με τις απαιτήσεις του μέλλοντος

#### *1.6.2 Τηλεκπαίδευση (Telelearning)*

Η ραγδαία ανάπτυξη της πληροφορικής και των τηλεπικοινωνιών που χαρακτηρίζουν τη σύγχρονη κοινωνία είχε ως αποτέλεσμα τον εκσυγχρονισμό του περιεχομένου, της μεθοδολογίας και των διαδικασιών της εκπαίδευσης και την εμφάνιση της τηλεκπαίδευσης .

Στόχος είναι η εκπαίδευση από απόσταση σε εκπαιδευτικά ιδρύματα, φορείς, επιχειρήσεις, άτομα με ειδικές ανάγκες, προβληματικές γεωγραφικές περιοχές από άποψη πρόσβασης κλπ.

Η εκπαίδευση από απόσταση είναι μια τυπική εκπαιδευτική διαδικασία κατά την οποία το μεγαλύτερο μέρος της εκπαίδευσης πραγματοποιείται όταν ο σπουδαστής και ο εκπαιδευτής δε βρίσκονται στον ίδιο χώρο. Ο εκπαιδευτής επικοινωνεί με τους εκπαιδευόμενους με κάποιο μέσο αμφίδρομης επικοινωνίας σύγχρονης ή ασύγχρονης. Η εκπαίδευση από απόσταση αποτελεί αυτή τη στιγμή μια από τις πιο ταχύτερες αναπτυσσόμενες μορφές εκπαίδευσης. Απευθύνεται σε ανθρώπους όλων των ηλικιών και όλων των βαθμίδων εκπαίδευσης. Ωστόσο τα άτομα τα οποία που βοηθά περισσότερο η ανάπτυξη της μάθησης από απόσταση είναι τα άτομα με ειδικές ανάγκες τα οποία μπορούν πλέον να έχουν ισότιμη πρόσβαση στην εκπαίδευση



Εικόνα 7 Σύστημα Τηλεκπαίδευσης

Υπάρχουν τρεις τύποι τηλεκπαίδευσης

- *Ακουστική τηλεκπαίδευση (audio teletraining)* που υποστηρίζει την ακουστική επικοινωνία ανάμεσα σε δάσκαλο και μαθητή όπου χρησιμοποιείται το τηλεφωνικό δίκτυο.
- *Ακουστική τηλεκπαίδευση με χρήση γραφημάτων (audiographic teletraining)* όπου εκτός από την ακουστική επικοινωνία υπάρχει και η μετάδοση γραφημάτων μέσω ενός ηλεκτρονικού υπολογιστή ή μιας συσκευής τηλεομοιοτυπίας.
- *Τηλεκπαίδευση με χρήση πολυμέσων (multimedia teletraining)* η οποία υποστηρίζει και ακουστική και οπτική επικοινωνία (τηλεεικονοδιάσκεψη) κάνοντας χρήση υπολογιστικών συστημάτων τα οποία επιτρέπουν την ανταλλαγή διαφόρων τύπων πληροφοριών (δεδομένα, ήχος, φωνή, κινούμενη και ακίνητη εικόνα).



### 1.6.3 Τηλεϊατρική



Εικόνα 8 Τηλεϊατρική

Τηλεϊατρική είναι η παροχή ιατρικών υπηρεσιών ακόμα και σε περιπτώσεις όπου παρεμβάλλεται απόσταση μεταξύ ασθενούς, ιατρού και άλλων εξειδικευμένων πληροφοριών και γνώσεων. Η Τηλεϊατρική χρησιμοποιεί τηλεματικές τεχνολογίες δηλαδή συνδυασμό υπολογιστών και επικοινωνιών προκειμένου να υποστηριχτούν οι ιατρικές υπηρεσίες. Η διαδικασία μπορεί να είναι τόσο απλή όπως ένα τηλεφώνημα μεταξύ δυο ιατρών μελετώντας μια υπόθεση, ή και περίπλοκη χρησιμοποιώντας δορυφορική τεχνολογία ούτως ώστε να διευκολύνει μια διαβούλευση μέσω video-conferencing. Η εφαρμογή της τηλεϊατρικής μειώνει τον χρόνο μετάβασης του ασθενούς στον ιατρό και μπορεί να βοηθήσει σημαντικά τον άνθρωπο και ιδιαίτερα τους κατοίκους των μικρών κέντρων και της υπαίθρου. Η τηλεϊατρική μπορεί να πάρει τις εξής μορφές: Τηλεδιάγνωση, Τηλεχειρουργική και Τηλεεικονοδιάσκεψη μεταξύ ιατρικών ομάδων

### 1.6.4 Ηλεκτρονικό εμπόριο (*electronic commerce, e-commerce*).

Ηλεκτρονικό εμπόριο είναι οποιοδήποτε είδος εμπορικής συναλλαγής, η οποία εκτελείται αποκλειστικά σε ηλεκτρονικό επίπεδο, δηλαδή με τη χρήση ηλεκτρονικών υπολογιστών που συνδέονται μέσω τηλεφωνικών γραμμών. Ο όρος χρησιμοποιείται για να περιγράψει την εμπορική δραστηριότητα στον Διαδίκτυο είτε μεταξύ επιχειρήσεων είτε μεταξύ μιας επιχείρησης και του καταναλωτικού κοινού.

Υπάρχουν δύο τύποι ηλεκτρονικού εμπορίου.

- Το Έμμεσο ηλεκτρονικό εμπόριο το οποίο περιλαμβάνει την ηλεκτρονική παραγγελία υλικών αγαθών τα οποία όμως παραδίδονται με

παραδοσιακούς τρόπους, όπως ταχυδρομικά ή μέσω ιδιωτικών υπηρεσιών διανομής και

- το Άμεσο ηλεκτρονικό εμπόριο στο οποίο υπάρχει η ηλεκτρονική παραγγελία, πληρωμή και παράδοση άυλων αγαθών και υπηρεσιών, όπως λογισμικό υπολογιστών



Εικόνα 9 Ηλεκτρονικό εμπόριο

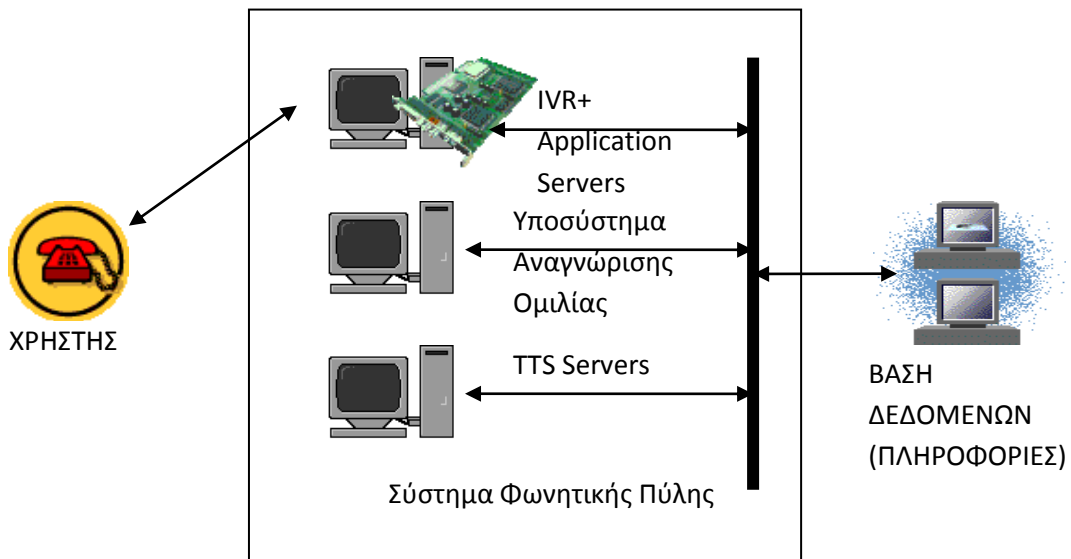
Τα είδη του ηλεκτρονικού εμπορίου είναι τέσσερα και είναι τα εξής :

- επιχείρηση - επιχείρηση
- επιχείρηση - καταναλωτής
- επιχείρηση - δημόσια διοίκηση
- καταναλωτής - δημόσια διοίκηση

## Επίλογος

Τα επόμενα σύνορα στη σύγκλιση και τη τομή του Διαδικτύου και των τηλεπικοινωνιών, θολώνοντας τις διακρίσεις μεταξύ της φωνής και των δεδομένων, των υπολογιστών και των τηλεφώνων είναι οι Φωνητικές Πύλες. Οι φωνητικές τεχνολογίες είναι αυτές οι οποίες θα μεταμορφώσουν τον κόσμο της τηλεματικής κι αυτό γιατί προσφέρουν απόδοση και επεκτασιμότητα. Επίσης μέσω των φωνητικών τεχνολογιών η εξυπηρέτηση των συνδρομητών-πελατών συνήθως γίνεται μέσω των υπολογιστών και μόνο αν κατάσταση κριθεί επιτακτική τότε η εξυπηρέτηση θα γίνει από εκπαιδευόμενο άτομο . Οι Φωνητικές πύλες καθώς και οι τεχνολογίες οι οποίες τις υλοποιούν αναλύονται στα επόμενα δύο κεφάλαια διεξοδικά.

## 2. Φωνητικές Πύλες



Εικόνα 10 Φωνητικές Πύλες

### Εισαγωγή

Η εξέλιξη της τεχνολογίας, η ανάπτυξη των τηλεπικοινωνιών, οι προσδοκίες των χρηστών κι ο ανταγωνισμός των επιχειρήσεων έδωσαν στη φωνή μεγάλη δύναμη και οδήγησαν στην εμφάνιση μιας νέας πλατφόρμας, της Φωνητικής Πύλης. Μέσω της οποίας οποιοσδήποτε ,από οποιοδήποτε τηλέφωνο και από οπουδήποτε μπορεί να έχει πρόσβαση στο περιεχόμενο του Internet και να κάνει συναλλαγές ηλεκτρονικού εμπορίου χρησιμοποιώντας άπλες φωνητικές εντολές.

Τρεις αλλαγές που έχουν γίνει τα τελευταία χρόνια στις τηλεπικοινωνίες είναι ιδιαίτερα σημαντικές:

- Η εμφάνιση του Διαδικτύου ως πρωταρχική δεξαμενή πληροφοριών, περιεχομένου, ψυχαγωγίας, και εμπορικών συναλλαγών
- Η τεράστια ανάπτυξη και αύξηση της διαθεσιμότητας και της δημοτικότητας των συσκευών πρόσβασης(wireless, PCs, Laptops)
- Η αναδυόμενη ανάγκη για μια εξατομικευμένη εμπειρία στην πρόσβαση σε περιεχόμενο ή την επικοινωνία.

Οι φωνητικές πύλες είναι τα επόμενα σύνορα στη σύγκλιση του Διαδικτύου και των τηλεπικοινωνιών, που θολώνουν τις διαφορές μεταξύ της φωνής και των στοιχείων, των υπολογιστών και των τηλεφώνων.

Οι φωνητικές πύλες δίνουν αξία στο βασικότερο προϊόν της Τηλεφωνίας που είναι η φωνή. Έχουν ως στόχο να απλουστεύσουν με τη βοήθεια της υψηλής τεχνολογίας την καθημερινή ζωή των ανθρώπων. Το τηλέφωνο βρίσκεται μέσα σε όλες τις επιχειρήσεις αλλά και τα νοικοκυριά και η σωστή αξιοποίηση του προσφέρει σημαντικές λύσεις τόσο στις επιχειρήσεις όσο και σε φορείς του δημοσίου. Εμφανίστηκαν στα τέλη της δεκαετίας του 1990 και οι πρώτοι πάροχοι ήταν οι εταιρίες Tellme Networks και Quack.com

Η τεράστια ανάπτυξη του διαδικτύου έδωσε τη δυνατότητα στον άνθρωπο να αντλεί οποιαδήποτε πληροφορία όποτε θέλει με το πάτημα ενός κουμπιού. Για αυτούς τους οποίους δεν έχουν υπολογιστές το τηλέφωνο και οι φωνητικές πύλες προσφέρουν μια φυσική πρόσβαση στον νέο ψηφιακό κόσμο της πληροφορίας. Οι χρήστες απλώς καλούν ένα τηλεφωνικό αριθμό, ο οποίος είναι διαφορετικός για κάθε φωνητική πύλη, και χρησιμοποιώντας άπλες φωνητικές εντολές έχει πρόσβαση σε οποιαδήποτε πληροφορία χρειάζεται ή μπορεί να διαχειριστεί τις προσωπικές του υπηρεσίες επικοινωνιών. Είναι απλό, γρήγορο και αποτελεσματικό. Ένα πρόγραμμα περιήγησης στο Internet μπορεί να παρέχει πρόσβαση στον παγκόσμιο Διαδίκτυο από έναν υπολογιστή, αλλά όχι από το τηλέφωνο. Μια φωνητική πύλη είναι ένας τρόπος να γίνει αυτό.

Μία φωνητική πύλη μπορεί να παραλληλισθεί με μια κοινή δικτυακή πύλη με ουσιαστική διαφορά ότι το περιβάλλον αλληλεπίδρασης είναι αποκλειστικά φωνητικό και προσβάσιμο μέσω μία απλής τηλεφωνικής γραμμής, έτσι ανοίγεται ο δρόμος για έναν ευκολότερο τρόπο πλοήγησης στο Διαδίκτυο, βασισμένο στη φωνή, που επιτρέπει στο χρήστη να ζητάει αυτό που θέλει και να λαμβάνει άμεση απάντηση από το σύστημα, χωρίς αναμονή σε τηλεφωνικά κέντρα εξυπηρέτησης και πληροφόρησης πελατών, από οποιοδήποτε μέρος και οποιαδήποτε στιγμή της ημέρας. Οι φωνητικές πύλες βασίζονται στη IVR τεχνολογία και απαιτούν τηλεφωνικές τεχνολογίες όπως τη διαχείριση κλήσεων. Αυτό που είναι νέο με τις φωνητικές πύλες είναι η εκτενής χρήση της αναγνώρισης και σύνθεσης ομιλίας και ο συνδυασμός τεχνολογιών τηλεφωνίας με μοντέλα πρόσβασης στο Internet.

Μια πύλη φωνής ενσωματώνει δύο διαφορετικούς τύπους δικτύων:

- το τηλεφωνικό δίκτυο
- ένα δίκτυο δεδομένων, ή το Διαδίκτυο βασισμένο στην IP

## 2.1 Οντότητες υπηρεσιών φωνητικών πυλών

Η φωνητική πύλη αναγνωρίζει δύο οντότητες υπηρεσιών: Τους παρόχους και τους χρήστες.

- Οι πάροχοι αντιπροσωπεύουν τις επιχειρήσεις που συνδέονται με την καταγραφή και την παράδοση πληροφοριών. Κατασκευάζουν και φιλοξενούν τις υπηρεσίες φωνητικών πυλών που απευθύνονται σε συγκεκριμένα ακροατήρια . Κάθε μια από αυτές τις φωνητικές πύλες παρέχει άμεσα ένα διαφορετικό μίγμα των υπηρεσιών πληροφοριών στις επιχειρήσεις και τους καταναλωτές . Ένα δεύτερο είδος προμηθευτή είναι μια παραδοσιακή πύλη Διαδικτύου που θέλει να επεκτείνει την προσιτότητά της μέσω του τηλεφώνου.
- Οι χρήστες συνδέονται με συγκεκριμένους παρόχους, και μπορούν να έχουν πρόσβαση σε υπηρεσίες μιας φωνητικής πύλης μόνο μετά τον επιτυχή έλεγχο ταυτότητας με τη χρήση κωδικών και προσωπικών αριθμών αναγνώρισης (κωδικών PIN).Ο τελικός χρήστης δε βλέπει τη πολυπλοκότητα των τεχνολογιών πληροφορικής και τηλεπικοινωνιών που αναπτύσσεται, απλώς χρησιμοποιεί τον πιο συνηθισμένο και απλό τρόπο επικοινωνίας, τον φυσικό διάλογο.

## 2.2 Πύλη (Portal)

Στην εποχή που ζούμε των πληροφοριών μια πύλη είναι μια ιστοσελίδα που παρέχεται στους επισκέπτες, η οποία χρησιμεύει ως πύλη προς άλλους δικτυακούς τόπους στο Διαδίκτυο. Οι πύλες παρέχουν γενικά μηχανές αναζήτησης, βιβλία διευθύνσεων, ηλεκτρονικό ταχυδρομείο, ηλεκτρονικό εμπόριο, περιήγηση στο Διαδίκτυο ,chat κτλ. Γενικά, μια πύλη μπορεί να υποστηρίξει ποικίλες μορφές εισαγωγής και παραγωγής που μπορεί να είναι ακουστικές, οπτικές, ή ακόμα και αφής. Μια πύλη μπορεί επίσης να χρησιμοποιηθεί σε ποικίλα διαφορετικά περιβάλλοντα χρηστών όπως ένα γραφείο, ένα αυτοκίνητο, ή σε ένα δρόμο.

Οι Φωνητικές Πύλες φέρνουν την ιδέα της Πύλης στα δίκτυα τηλεφωνίας, επιτρέποντας στον τελικό χρήστη να έχει πρόσβαση σε δίκτυο ανταλλαγής μηνυμάτων, σε υπηρεσίες πληροφόρησης και στο περιεχόμενο του Διαδικτύου. Νέες προηγμένες υπηρεσίες, όπως η περιήγηση στο Internet, η οποία είναι δύσκολο να επιτευχθεί με ένα πληκτρολόγιο τηλεφώνου, προσφέρονται πλέον, λόγω των πρόσφατων βελτιώσεων στη φυσική αναγνώριση ομιλίας (ASR) και στο σχεδιασμό και την ανάπτυξη των φωνητικών διεπαφών χρήστη (VUI)

### **2.3 Γιατί χρησιμοποιείται η φωνή;**

Γιατί η φυσική ομιλία είναι η μορφή που χρησιμοποιεί πιο συχνά ο άνθρωπος κατά την επικοινωνία με άλλους ανθρώπους και μαζί με τον γραπτό λόγο αποτελούν τα δύο κύρια μέσα ανταλλαγής απόψεων, ιδεών, γνώσης και πολιτισμού. Αυτό καθιστά ευκολότερο για έναν χρήστη να μάθει τη λειτουργία των ενεργοποιημένων φωνητικά υπηρεσιών. Με τη χρήση της φυσικής ομιλίας αποφεύγεται να χρησιμοποιούνται τα χέρια και τα μάτια. Η ομιλία δεν εμποδίζει το χρήστη να κάνει άλλες εργασίες ταυτόχρονα όπως το να οδηγεί, επιτρέπει την εύκολη ενσωμάτωση μέσων που βασίζονται στη φωνή (φωνητικό ηλεκτρονικό ταχυδρομείο), επίσης οι πρόοδοι στη text-to-speech (TTS) τεχνολογία σημαίνουν ότι οι πληροφορίες κειμένων μπορούν να διαβιβαστούν εύκολα στο χρήστη. Με το κατάλληλο σχεδιασμό, οι εντολές φωνής μπορούν να απομνημονευθούν από το χρήστη. Η φυσική ομιλία μπορεί να χρησιμοποιηθεί για τη δημιουργία διεπαφής που είναι εύκολη στη χρήση και απαιτεί το ελάχιστο της προσοχής των χρηστών. Μπορεί να χρησιμοποιηθεί από τους τυφλούς ή άτομα που είναι ανάκανα να χρησιμοποιήσουν διεπαφές αφής. (Schmandt, 1994). Με την απλή ομιλία μέσω της χρήσης μιας διασύνδεσης συνομιλίας οι συνδρομητές αποφεύγουν τα ιεραρχικά μενού και τα δέντρα απόφασης και έχουν πρόσβαση άμεσα στην υπηρεσία που θέλουν ή μετατοπίζονται από μια υπηρεσία σε άλλη πολύ εύκολα.

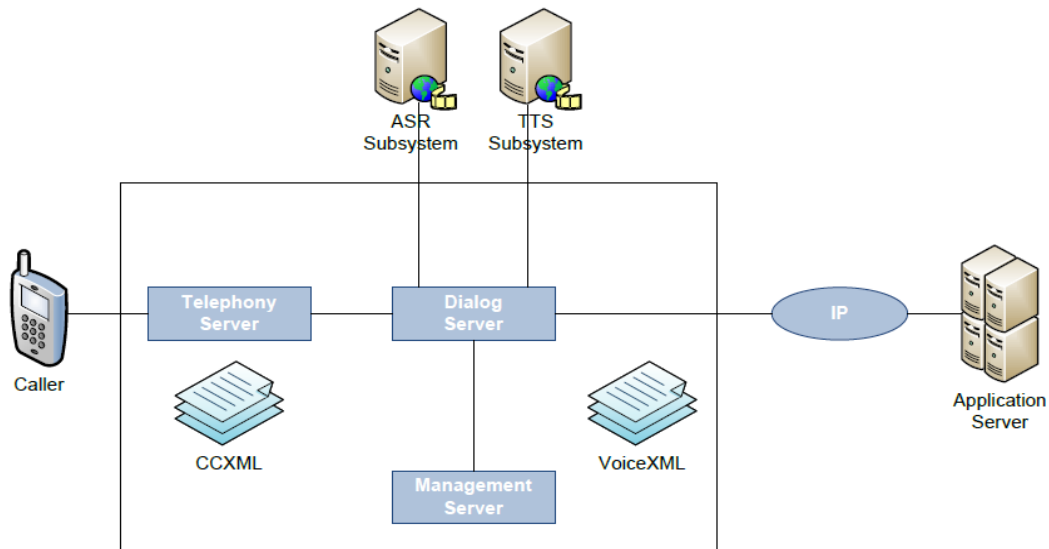
### **2.4 Κατηγορίες Φωνητικών Πυλών**

Υπάρχουν δύο κύριες κατηγορίες φωνητικών πυλών:

- Φωνητικές Πύλες καταναλωτών που παρέχουν γενική πρόσβαση σε πληροφορίες.

- Φωνητικές Πύλες επιχειρήσεων που παρέχουν προσαρμοσμένη πρόσβαση στην υποστήριξη πελατών

## 2.5 Αρχιτεκτονική



Εικόνα 11 Αρχιτεκτονική Φωνητικών Πυλών

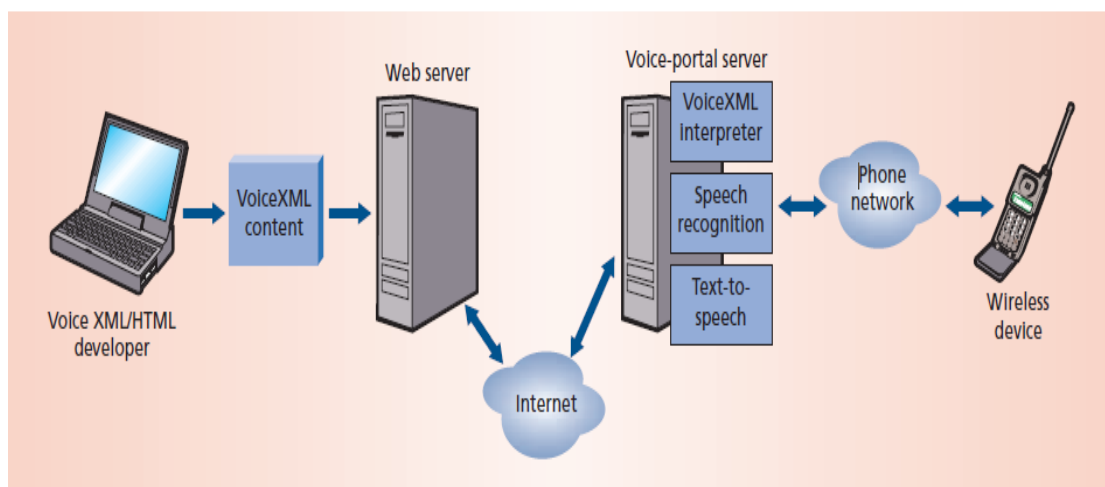
Οι φωνητικές Πύλες αποτελούνται από πέντε βασικά υποσυστήματα:

- Το διακομιστή τηλεφωνίας που είναι ένα ανεξάρτητο υποσύστημα λογισμικού, το οποίο ελέγχει τη λειτουργία του τηλεφωνικού υλικού. Ένας ή περισσότεροι διακομιστές εφαρμογών επικοινωνούν με τον διακομιστή τηλεφωνίας μέσω ενός δικτύου δεδομένων TCP/IP, και λαμβάνουν ειδοποιήσεις για ένα σύνολο γεγονότων, όπως: εισερχόμενες κλήσεις, ολοκλήρωση και μεταφορά κλήσης, εκκίνηση / τερματισμό αναπαραγωγής ήχου, εκκίνηση / τερματισμός εγγραφής ήχου, Χρήση DTMF. Ο διακομιστής τηλεφωνίας υποστηρίζει τα ακόλουθα πρωτόκολλα: T1 / E1 , ISDN PRI , SS7 ,VoIP (SIP/RTP) και είναι συνδεδεμένος με το τηλεφωνικό δίκτυο. Περιλαμβάνει έναν διερμηνέα CCXML, ο οποίος μπορεί να χρησιμοποιηθεί για τον έλεγχο των εισερχόμενων και των εξερχόμενων τηλεφωνικών συνδέσεων και των φωνητικών εφαρμογών που συνδέονται με αυτές.

- Το διαλογικό διακομιστή που ενεργεί ως κέλυφος για όλες τις μονάδες διαλόγου που υποστηρίζουν την αλληλεπίδραση του τελικού χρήστη. Ο διακομιστής τηλεφωνίας προωθεί όλα τα γεγονότα τηλεφωνίας στο διακομιστή διάλογο.
- Το υποσύστημα διαχείρισης χειρίζεται τη διαμόρφωση, την προετοιμασία, την παρακολούθηση και τον τερματισμό όλων των υποσυστημάτων της φωνητικής πύλης.
- Το υποσύστημα ASR
- Το υποσύστημα TTS

## Επίλογος

Οι φωνητικές πύλες βασίζονται στη τεχνολογία λογισμικού IVR(interactive voice response) αυτόματο σύστημα αλληλεπίδρασης φωνής που επιτρέπει στο χρήστη να «συνομιλήσει» με ένα υπολογιστή(φωνητικό σύστημα). Κατά τη συνομιλία χρήστη - φωνητικού συστήματος χρησιμοποιούνται οι τεχνολογίες αναγνώρισης ομιλίας ASR και σύνθεσης φωνής TTS. Επίσης αρκετές φορές αντί για την τεχνολογία TTS χρησιμοποιούνται προηχογραφημένα μηνύματα. Το φωνητικό σύστημα αντλεί τις πληροφορίες από μια βάση δεδομένων όπου είναι αποθηκευμένες. Στις φωνητικές πύλες υπάρχουν ειδικευμένοι περιηγητές φωνής οι οποίοι λαμβάνουν σελίδες από το διακομιστή εφαρμογής μέσω ενός πρότυπου που ονομάζεται VoiceXML.



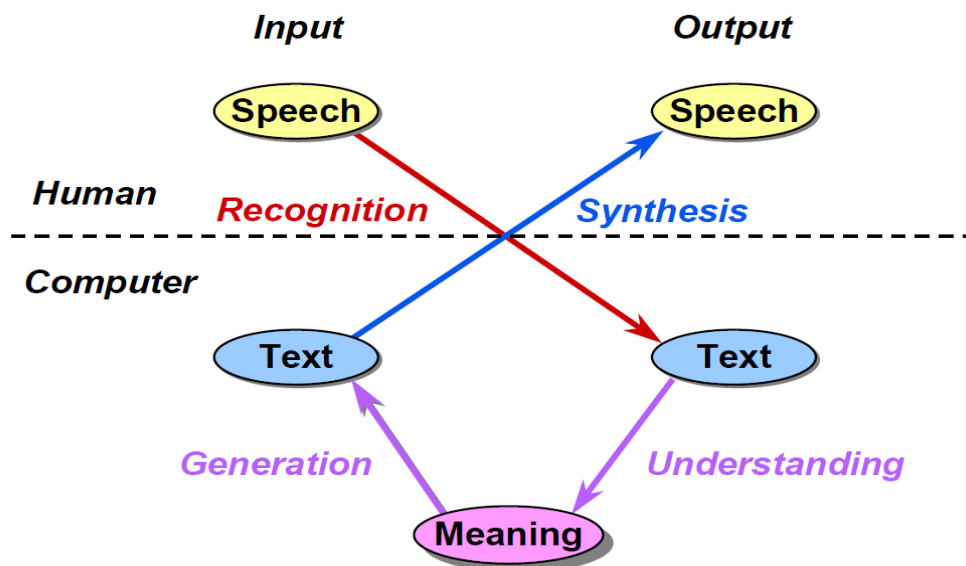
Εικόνα 12. Αρχιτεκτονική Φωνητικών Πυλών



### 3. Τεχνολογίες Φωνητικών Πυλών

#### Εισαγωγή

Στο συγκεκριμένο κεφάλαιο παρουσιάζονται οι τεχνολογίες που χρησιμοποιούνται για να επιτευχθεί η συνομιλία μεταξύ ανθρώπου – υπολογιστή και κατ' επέκταση οι τεχνολογίες που χρησιμοποιούνται στις Φωνητικές Πύλες. Η τεχνολογία Interactive Voice Response (IVR) χρησιμοποιείται στα αυτόματα συστήματα τα οποία μπορούμε να καλέσουμε και να συνδιαλλαγούμε είτε πιέζοντας κάποια πλήκτρα στο τηλέφωνο είτε δίνοντας φωνητικές εντολές. Τα τελευταία χρόνια, η χρήση των υπολογιστών στις τεχνολογίες σύνθεσης ομιλίας και αναγνώρισης ομιλίας έχει γίνει ένας σημαντικός τομέας μελέτης μεταξύ των επιστημόνων της πληροφορικής και της φωνής. Τα αρχικά κίνητρα είναι να παρέχεται στους χρήστες μια φιλική φωνητική διεπαφή με τον υπολογιστή και να επιτρέπει σε ανθρώπους με αναπηρίες να χρησιμοποιούν τους υπολογιστές. Επίσης χρησιμοποιείται η VoiceXML η οποία είναι μια markup γλώσσα για ανάπτυξη εφαρμογών φωνητικού διαδικτύου.



Εικόνα 13 Συνομιλία ανθρώπου - υπολογιστή

### 3.1 Interactive Voice Response (IVR)

Interactive Voice Response (IVR) ή Φωνο-Αλληλεπιδραστική Απόκριση, ονομάζεται η αμφίδρομη επικοινωνία με ένα αυτόματο σύστημα. Η τεχνολογία IVR χρησιμοποιείται εκτενώς στις τηλεπικοινωνίες, αλλά εισάγεται επίσης στα αυτοκινητικά συστήματα για τη λειτουργία hands-free. Στον τομέα των τηλεπικοινωνιών, η IVR επιτρέπει στους χρήστες να έχουν πρόσβαση στη βάση δεδομένων μιας επιχείρησης μέσω ενός τονικού τηλεφώνου (DTMF) με την χρήση των πλήκτρων ή με αναγνώριση ομιλίας. Τα συστήματα IVR μπορούν να αποκριθούν στους χρήστες με προηχογραφημένα μηνύματα ή με δυναμικά δημιουργημένα μηνύματα για το πώς να προχωρήσουν. Επιτρέπουν τη παράλληλη διαχείριση ενός μεγάλου αριθμού ταυτοχρόνων κλήσεων.

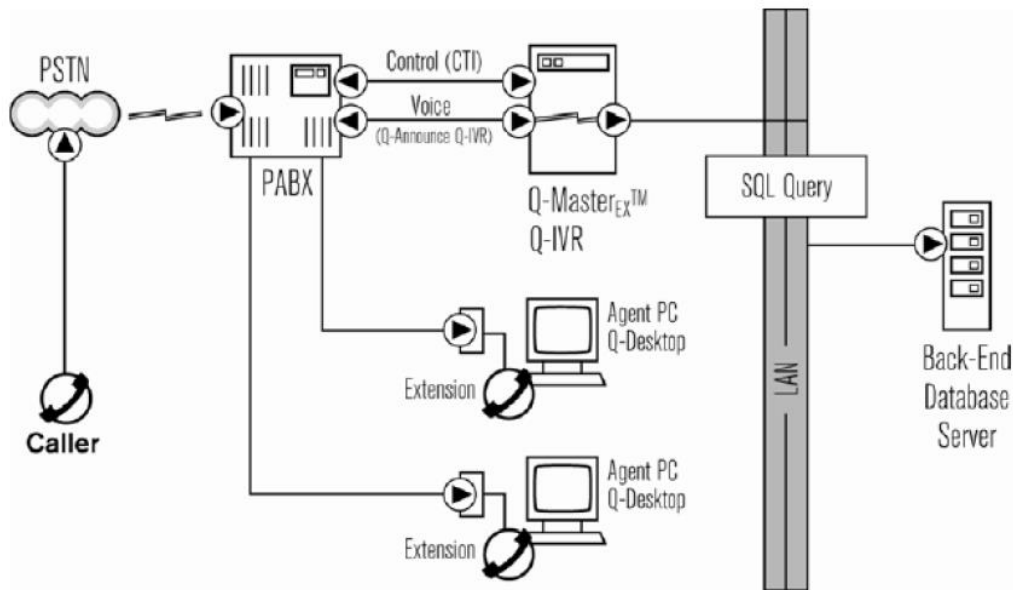
Οι επιχειρήσεις στρέφουν το ενδιαφέρον τους στη τεχνολογία IVR για να μειώσουν το κόστος των πωλήσεων και των υπηρεσιών τους. Τα συστήματα IVR παίρνουν τη θέση των τηλεφωνητών στα τηλεφωνικά κέντρα πραγματοποιώντας από τις πιο απλές εργασίες, όπως η ταυτοποίηση του καλούντα και η δρομολόγηση κλήσεων, μέχρι τις πιο σύνθετες, όπως η σύνδεση με τα μηχανογραφικά συστήματα ενός οργανισμού και η πραγματοποίηση πολλαπλών συναλλαγών. Είναι μια τεχνολογία ανοιχτού προτύπου η οποία διαχειρίζεται και εξυπηρετεί ένα σύνολο τηλεφωνικών κλήσεων και συνδέει το σύστημα με μια ποικιλία από πηγές πληροφορίας όπως μια βάση δεδομένων ή ένα ιστότοπο.

Μπορούμε να παρομοιάσουμε ένα σύστημα IVR με ένα φωνητικό ηλεκτρονικό υπολογιστή. Όπου ένας υπολογιστής χρησιμοποιεί πληκτρολόγιο για να εισάγει πληροφορίες, ενώ ένα IVR σύστημα χρησιμοποιεί σήματα DTMF ή τη φωνή του καλούντα (χρησιμοποιώντας τη τεχνολογία αναγνώρισης φωνής ASR). Ενώ για την προβολή των αποτελεσμάτων ο υπολογιστής έχει την οθόνη το IVR σύστημα χρησιμοποιεί τη φωνή για να διαβάσει την οθόνη στον απομακρυσμένο χρήστη (με ηχογραφημένα μηνύματα ή με τη τεχνολογία σύνθεσης φωνής TTS). Οτιδήποτε μπορεί να κάνει ένας υπολογιστής, μπορεί να το κάνει και ένα IVR σύστημα.

Τα IVR συστήματα χρησιμοποιούνται συνήθως για την εξυπηρέτηση μεγάλου όγκου κλήσεων, μείωση του κόστους και τη βελτίωση της εμπειρίας του πελάτη. Μια σημαντική εφαρμογή της IVR τεχνολογίας είναι οι τηλεφωνικές τραπεζικές συναλλαγές.

Τα ενεργοποιημένα φωνητικά συστήματα (VAD) χρησιμοποιούνται πλέον και αντικαταστούν τους χειριστές τηλεφωνικών κέντρων (PABX) και χρησιμοποιούνται σε πολλά νοσοκομεία και μεγάλες επιχειρήσεις για να μειώσουν τον χρόνο αναμονής των χρηστών.

Οι γλώσσες προγραμματισμού που χρησιμοποιούνται για τη δημιουργία των IVR εφαρμογών είναι η VoiceXML που συνδυάζει την XML με την αναγνώριση φωνής και η SALT(Speech Application language Tags).



Εικόνα 14 Αρχιτεκτονική IVR

### 3.1.1 Ιστορία

Τα IVR συστήματα ξεκίνησαν το 1941 όταν η εταιρία Bell System ανέπτυξε μια νέα μεθοδολογία στο τονικό σχηματισμό . Η ίδια εταιρία παρουσίασε το πρώτο τηλέφωνο που θα μπορούσε να σχηματίσει κώδικες χρησιμοποιώντας την τεχνολογία DTMF στην παγκόσμια έκθεση του Σιάτλ το 1962.

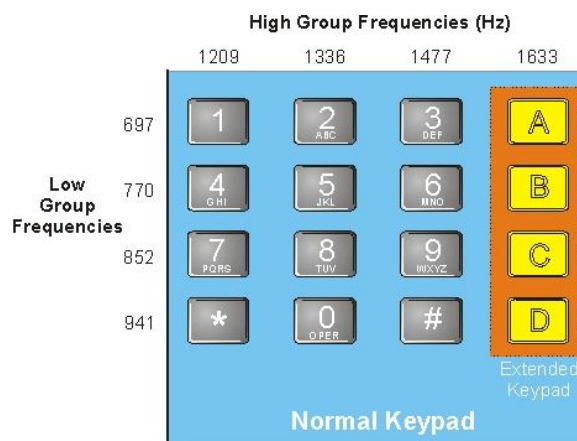
Παρά το γεγονός ότι όλο και περισσότερες εταιρείες άρχισαν να χρησιμοποιούν το σύστημα IVR στη δεκαετία του 1970 για την αυτοματοποίηση των εργασιών στα τηλεφωνικά κέντρα, η τεχνολογία ήταν ακόμη δαπανηρή και περίπλοκη πράγμα το οποίο έκανε δύσκολη τη διείσδυση του στην αγορά. Ωστόσο μέχρι τη δεκαετία του

80 ο ανταγωνισμός των επιχειρήσεων ήταν τόσο μεγάλος με αποτέλεσμα η χρήση της IVR τεχνολογίας να αυξάνεται.

Παράλληλα αναπτύχθηκαν οι τεχνολογίες αναγνώρισης και σύνθεσης φωνής και τα πρότυπα VXML τα οποία ενσωματώθηκαν αμέσως στη IVR τεχνολογία και την απλοποίησαν.

### 3.1.2 DTMF

Αρχικά για την πρόσβαση σε πληροφορίες μέσω του τηλεφώνου χρησιμοποιούταν το πληκτρολόγιο του τηλεφώνου DTMF (Dual Tone Multi frequency). Ωστόσο το πληκτρολόγιο του τηλεφώνου μπορεί να δώσει στους χρήστες μόνον έναν περιορισμένο αριθμό επιλογών. Επίσης, ένα τηλέφωνο δεν παρέχει εύκολο τρόπο για να περιηγηθεί ο χρήστης από τη μια ερώτηση ή πληροφορία στην άλλη, παρά μόνο αν πάει πίσω σε ένα μενού εκκίνησης, κάτι το οποίο είναι πολύ χρονοβόρο.



Εικόνα 15 DTMF Dual Tone Multi frequency

Η τεχνολογία αυτή δεν είναι η βέλτιστη για τον χρήστη γιατί :

- η αποστήθιση των αριθμών τηλεφώνου, των εντολών επιλογών, και των ακολουθιών καθιστούν πολλές εφαρμογές δυσκίνητες και άκαμπτες.
- Επίσης η ασφάλεια αποτελεί όλο και περισσότερο μια σημαντική ανησυχία με τη χρήση του κινητού τηλεφώνου, καθώς μπορεί να αποσπαστεί η προσοχή ενός χρήστη κατά την οδήγηση ή την εκτέλεση άλλων εργασιών του.
- Οι φυσικές ανικανότητες μπορούν να κάνουν τη χρήση αδύνατη ή δύσκολη

- Οι διαφορετικές εφαρμογές, οι στόχοι και τα περιβάλλοντα απαιτούν μια επιλογή των συσκευών εισόδου.
- Ένα τηλεφωνικό αριθμητικό πληκτρολόγιο δεν είναι μια κατάλληλη μορφή εισαγωγής για να έχει κανείς πρόσβαση στο Διαδίκτυο, ακριβώς όπως ένα ποντίκι δεν ταιριάζει στην καταχώριση αριθμών

Σε αντίθεση με αυτή τη τεχνολογία οι φωνητικές πύλες προσφέρουν μια πιο φυσική διεπαφή στο χρήστη και αποτελούν μια εναλλακτική λύση που συνδυάζουν την ιδέα μιας «πύλης» με τη φυσικότερη μέθοδο προσπέλασης, τη φωνή, επιτρέποντας στον χρήστη να χρησιμοποιήσει ένα ολοκληρωμένο σύνολο υπηρεσιών με ένα πιο ευέλικτο και ευπαρουσίαστο τρόπο.

### *3.1.3 IVR Συστήματα και Φωνητικές Πύλες*

Τα IVR συστήματα έχουν χρησιμοποιηθεί με επιτυχία για πολλά χρόνια, αλλά με τις φωνητικές τεχνολογίες, νέας γενιάς, απαιτείται μια πιο φιλική προς το χρήστη και ευέλικτη λύση για αυτοματοποιημένες υπηρεσίες. Στις μέρες μας, τα παραδοσιακά συστήματα IVR δεν είναι πλέον ο καλύτερος τρόπος για την παροχή αυτοματοποιημένων υπηρεσιών, γι 'αυτό και πολλές εταιρείες μετακινούνται από τη IVR τεχνολογία στις φωνητικές Πύλες. Η κύρια διαφορά μεταξύ μιας φωνητικής πύλης και ενός IVR συστήματος είναι η ιδέα - ενώ ένα IVR σύστημα είναι μια αυτόνομη εφαρμογή ,ενσωματωμένη στο λογισμικό των επιχειρήσεων ,που παρέχει στους πελάτες τις απαραίτητες πληροφορίες, μία φωνητική πύλη δεν παρέχει απλώς πρόσβαση στα επιχειρησιακά δεδομένα αλλά και μια φωνητική διασύνδεση στην δικτυακή πύλη και σε άλλες εφαρμογές της επιχείρησης.

Η διαφορά μεταξύ των παραδοσιακών IVR συστημάτων και φωνητικών πυλών δεν έχει να κάνει τόσο με τη τεχνολογία αλλά με τις εφαρμογές που μπορούν να τρέξουν. Οι Φωνητικές Πύλες χαρακτηρίζονται από δύο σημαντικά επιπλέον χαρακτηριστικά σε σχέση με τα παραδοσιακά IVR συστήματα: επεκτασιμότητα σε μεγαλύτερη ποικιλία διαδικτυακών, υπηρεσιών μέσα από πρότυπα, όπως η VoiceXML , και χρήση της τεχνολογίας της αναγνώρισης ομιλίας ως μια μέθοδο εισόδου.

Μερικοί από τους βασικούς λόγους για τους οποίους οι εταιρίες στράφηκαν από τα παραδοσιακά IVR συστήματα στις Φωνητικές Πύλες είναι οι εξής. Οι Φωνητικές

Πύλες προσφέρουν μεγαλύτερη ευελιξία στις επιλογές ανάπτυξης. Η δυνατότητα που έχουν οι φωνητικές πύλες να χρησιμοποιήσουν τη δύναμη του λόγου με την ένταξη των προηγμένων δυνατοτήτων που προσφέρει τεχνολογία αναγνώρισης ομιλίας. Όταν εφαρμοστεί σωστά, ανοίγει την πόρτα σε ένα ευρύτερο φάσμα του τομέα παροχής πληροφοριών που καθιστά την αλληλεπίδραση των πελατών με την εταιρία πλουσιότερη και πιο ικανοποιητική.

Τα ανοικτά πρότυπα που έχουν αναπτυχθεί, πρότυπα όπως VXML έχουν επεκτείνει δραστικά τις ευκαιρίες στους προγραμματιστές να δημιουργήσουν βαθιά λειτουργικές εφαρμογές σε πολύ χαμηλότερο κόστος και σε πολύ λιγότερο χρόνο σε σχέση με τα παλαιού τύπου IVR συστήματα.

#### *3.1.4 Πλεονεκτήματα των IVR συστημάτων*

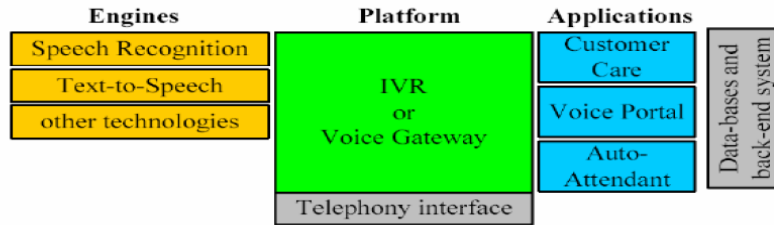
Οι επιχειρήσεις μπορούν να αποκομίσουν διάφορα σημαντικά οφέλη από την εισαγωγή ενός συστήματος IVR:

- Αύξηση της παραγωγικότητας
- Βελτιωμένα επίπεδα ικανοποίησης των πελατών
- Μείωση του χρόνου αναμονής των πελατών
- Επέκταση του χρόνου λειτουργίας χωρίς την αύξηση του προσωπικού ή ωρών εργασίας
- Χαμηλότερες δαπάνες δικτύων
- Απλή και εύκολη λειτουργία
- Τα συστήματα IVR επιτρέπουν επίσης στους χρήστες να λάβουν τα στοιχεία ανώνυμα.

#### *3.1.5 Χαρακτηριστικά της IVR τεχνολογίας*

- Ενοποίηση με τη τεχνολογία αναγνώρισης ομιλίας (ASR).  
Η IVR τεχνολογία μπορεί έξυπνα να αναγνωρίσει φράσεις και φωνητικές εντολές από τον καλούντα σε συνδυασμό με την Touch-tone εισαγωγή.
- Ενοποίηση με τη τεχνολογία σύνθεσης φωνής (Text To Speech).  
Μετατρέπει τα στοιχεία που ανακτώνται από εξωτερικές πηγές σε ομιλία και τα μεταφέρει σε μορφή φωνής στον χρήστη-καλούντα
- Ενοποίηση με τις εφαρμογές φαξ και ηλεκτρονικού ταχυδρομείου
- Πλήρως αυτοματοποιημένη επεξεργασία κλήσεων

- Δυναμικός έλεγχος ροής κλήσεων
- Πολύγλωσση υποστήριξη
- Δυνατότητα να αφήσει μηνύματα



Εικόνα 16 Χαρακτηριστικά της IVR τεχνολογία

Μια επιχείρηση για να αναπτύξει μια IVR τεχνολογία χρειάζεται τα εξής:

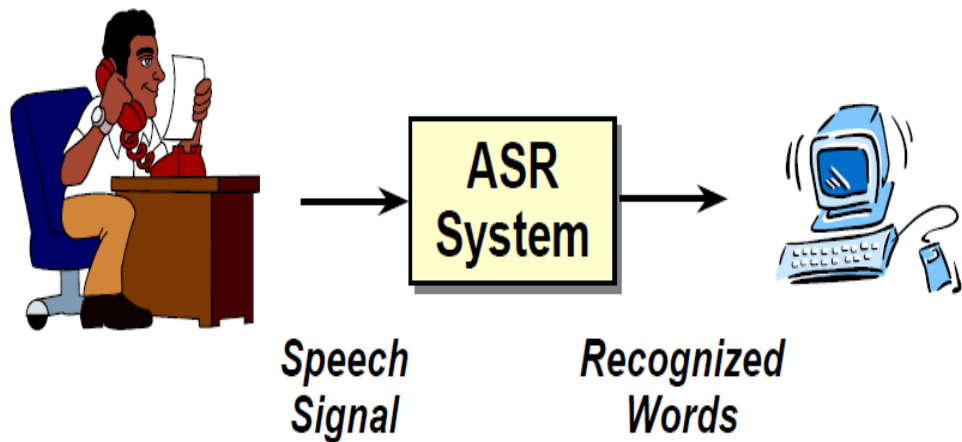
- IVR Πλατφόρμες  
Είναι οι πλατφόρμες υλικού(Hardware) και λογισμικού πάνω στις οποίες θα τρέξει η IVR τεχνολογία. Παρέχουν τη δυνατότητα να παίζουν και να ηχογραφούν μηνύματα και να συλλέγουν touch-tone εισαγωγές . Επίσης μια IVR πλατφόρμα προσφέρει τη δυνατότητα να αναγνωρίσει τη φωνή των χρηστών(ASR), να μετατρέψει το κείμενο σε φωνή (TTS) και να μεταφέρει τις IVR κλήσεις σε οποιοδήποτε τηλέφωνο ή τηλεφωνητή.
- IVR εφαρμογές  
Οι IVR εφαρμογές είναι προγράμματα τα οποία ελέγχουν και ανταποκρίνονται στις κλήσεις των IVR πλατφόρμων.
- Back-end servers  
Είναι κεντρικοί υπολογιστές στους οποίους είναι αποθηκευμένα τα απαραίτητα δεδομένα. Συνήθως είναι αποθηκευμένα σε βάσεις δεδομένων
- Τηλεφωνική υποδομή  
Η τηλεφωνική υποδομή περιλαμβάνει τις τηλεφωνικές γραμμές, τον εξοπλισμό μετατροπής κλήσεων και τους αυτόματους διανομείς κλήσεων τηλεφωνικών κέντρων (ACDs).  
Οι τηλεφωνικές γραμμές για IVR μπορούν να είναι απλές αναλογικές γραμμές, ψηφιακές T1, ή ψηφιακές γραμμές ISDN. Οι γραμμές αυτές συνδέονται από τη μία πλευρά στην πλατφόρμα IVR και από τη άλλη στον εξοπλισμό μετατροπής κλήσεων που περιλαμβάνουν PBX, VoIP πύλες ή

μερικές φορές συνδέονται απευθείας στα τηλεφωνικά κέντρα μέσω ενός ACD.

- Ειδικούς στη IVR τεχνολογία

Είναι υπάλληλοι και σύμβουλοι οι οποίοι γνωρίζουν τη IVR τεχνολογία και τις προκλήσεις παρά πολύ καλά

### 3.2 ASR Automatic Speech Recognition (Τεχνολογία Αναγνώρισης Ομιλίας)



Εικόνα 17 ASR Τεχνολογία Αναγνώρισης Ομιλίας

Μετά από σχεδόν εξήντα χρόνια ερευνών η τεχνολογία αναγνώρισης ομιλίας έχει φθάσει σε ένα σχετικά υψηλό επίπεδο και αντιμετωπίζεται ως ένα αναπόσπαστο τμήμα των μελλοντικών διεπαφών μεταξύ ανθρώπου-υπολογιστή, είναι μια τεχνολογία που επιτρέπει σε έναν υπολογιστή να προσδιορίσει τις λέξεις που λέει ένα πρόσωπο σε ένα μικρόφωνο ή ένα τηλέφωνο και να τις μετατρέψει σε κείμενο. Μια μηχανή που κατανοεί την ανθρώπινη ομιλία. Κάποιος μπορεί να γράψει ένα έγγραφο, να συνθέσει ένα μήνυμα ηλεκτρονικού ταχυδρομείου, και να ανοίξει προγράμματα χρησιμοποιώντας απλώς τη φωνή του. Μερικές φορές η τεχνολογία αναγνώρισης ομιλίας μπορεί να κάνει λάθος ωστόσο ο χρόνος διόρθωσης των λαθών αυτών είναι πολύ μικρότερος από το χρόνο που θα χρειαστεί για να δακτυλογραφηθεί το ίδιο μήνυμα.

Η κατοχή μιας μηχανής η οποία μπορεί να καταλάβει με ευφράδεια τη προφορική ομιλία έχει οδηγήσει σε έρευνες για περισσότερο από 50 χρόνια. Ο απώτερος σκοπός της έρευνας ASR είναι να επιτρέπει στον υπολογιστή να αναγνωρίζει σε



πραγματικό χρόνο, με 100% ακρίβεια, όλες τις λέξεις που ομιλούνται από οποιοδήποτε πρόσωπο, ανεξάρτητα από το μέγεθος του λεξιλογίου, το θόρυβο, τα χαρακτηριστικά του ομιλητή ή τη προφορά του. Αν και ακόμα δε έχει επιτευχθεί αυτός ο στόχος 100% χρησιμοποιείται σε καθημερινή βάση σε ποικίλες εφαρμογές και υπηρεσίες.

### *3.2.1 Ιστορία*

Η πρώτες προσπάθειες για να αναπτυχθούν συστήματα για την αυτόματη αναγνώριση ομιλίας έγιναν στη δεκαετία του 1950. Ένα μεγάλο μέρος των πρώτων ερευνών πάνω στη ανάπτυξη της αναγνώρισης φωνής χρηματοδοτήθηκε από το από το Εθνικό Ίδρυμα Επιστημών National Science Foundation (NSF) και το DARPA . Το πρώτο λεκτικό σύστημα αναγνώρισης εμφανίστηκε το 1952 και αποτελούνταν από μια συσκευή για την αναγνώριση των ενιαίων προφορικών ψηφίων. Μια άλλη πρόωρη συσκευή ήταν η IBM Shoebox, που εκτέθηκε στην παγκόσμια έκθεση της Νέας Υόρκης του 1964.

Η τεχνολογία αναγνώρισης ομιλίας σχεδιάστηκε αρχικά για να βοηθήσει τα άτομα με κινητικές αναπηρίες.

Κατά τη διάρκεια των πρώτων χρόνων της δεκαετίας του 1990 προέκυψαν τεράστιες ευκαιρίες στην αγορά για την τεχνολογία αναγνώρισης ομιλίας .Ωστόσο οι πρώτες εκδόσεις αυτών των προϊόντων ήταν δύσκολο να χρησιμοποιηθούν και είχαν πολλούς περιορισμούς : είτε ήταν "συντονισμένα" να αναγνωρίζουν ένα συγκεκριμένο ομιλητή, είτε είχαν μικρό λεξιλόγιο, είτε χρησιμοποιούσαν μια πολύ τυποποιημένη και άκαμπτη σύνταξη. Εντούτοις, στη βιομηχανία υπολογιστών, τίποτα δεν μένει το ίδιο για πάρα πολύ καιρό και προς το τέλος της δεκαετίας του '90 υπήρξε μια έκρηξη στα εμπορικά πακέτα λογισμικού λεκτικής αναγνώρισης που ήταν ευκολότερα να χρησιμοποιηθούν και αποτελεσματικότερα από τους προκατόχους τους.

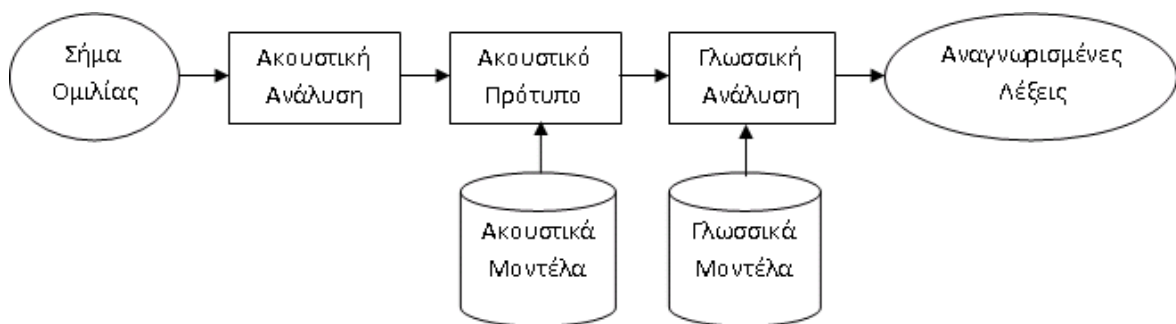
Τα τελευταία χρόνια, η τεχνολογία αναγνώρισης ομιλίας έχει προχωρήσει σε σημείο που να χρησιμοποιείται από εκατομμύρια ατόμων για την αυτόματη δημιουργία έγγραφων μόνο με τη φωνή.

### *3.2.2 Αυτόματη διαδικασία αναγνώρισης ομιλίας*

Η διαδικασία αρχίζει όταν ένας ομιλητής αποφασίζει να μιλήσει μέσω ενός τηλεφώνου ή μικροφώνου . Στη συνέχεια το λογισμικό παράγει μια μορφή

λεκτικών κυμάτων, στα οποία ενσωματώνονται οι λέξεις της πρότασης καθώς και οι ξένοι ήχοι και οι παύσεις της προφορικής εισαγωγής(ακουστικό μοντέλο). Έπειτα, το λογισμικό προσπαθεί να αποκωδικοποιήσει την ομιλία. Πρώτα μετατρέπει το λεκτικό σήμα σε μια ακολουθία διανυσμάτων που μετριοούνται σε όλη τη διάρκεια του λεκτικού σήματος(Γλωσσικό μοντέλο) και μετά χρησιμοποιώντας έναν συντακτικό αποκωδικοποιητή παράγει μια έγκυρη ακολουθία αντιπροσωπεύσεων.

Στην εικόνα 18 παρουσιάζεται η αρχιτεκτονική ενός τυπικού συστήματος αναγνώρισης ομιλίας.



Εικόνα 18 Αρχιτεκτονική ενός συστήματος αναγνώρισης ομιλίας

Το πρώτο βήμα σε ένα σύστημα αυτόματης αναγνώρισης ομιλίας είναι η ακουστική ανάλυση. Η διαδικασία αυτή έχει ως στόχο να εξάγει ακουστικά χαρακτηριστικά από την κυματομορφή ομιλίας. Το αποτέλεσμα της ακουστικής ανάλυσης είναι ένα συμπαγές, αποτελεσματικό σύνολο παραμέτρων που διαπιστώθηκαν από εισαγόμενα σήματα ομιλίας, οι οποίοι θα χρησιμοποιηθούν μεταγενέστερα από την ακουστική μοντελοποίηση.

Οι σημαντικότερες τεχνικές επεξεργασίας ακουστικής ανάλυσης είναι η *linear predictive coding (LPC)*, η *mel-frequency cepstral coefficients (MFCC)* και η *perceptual linear prediction (PLP)*

Στη συνέχεια ακολουθεί το ακουστικό πρότυπο αναγνώρισης κατά το οποίο γίνεται έλεγχος της ομοιότητας μεταξύ μιας εισαγόμενης ομιλίας με ένα μοντέλο το οποίο έχει ληφθεί κατά τη διάρκεια της εκπαίδευσης. Μια προσέγγιση για το ταίριασμα ομιλίας και μοντέλων είναι η *dynamic time warping (DTW)* μια μέθοδος η οποία μετρά την απόσταση μεταξύ των πλαισίων των εισροών και τα συγκρίνει με τα πλαίσια αναφορών. Μια άλλη προσέγγιση είναι βασισμένη σε στατιστικά

μοντέλα. Η πιο επιτυχημένη και αυτή που χρησιμοποιείται ευρέως είναι η Hidden Markov Model που χαρακτηρίζει σήματα ομιλίας χρησιμοποιώντας μια προ-εκπαιδευμένη αλυσίδα. Στο στάδιο της εκπαίδευσης, ένα ή περισσότερα HMMs αντιστοιχούν σε ήχους ομιλίας της ίδιας κατηγορίας (π.χ., φθόγγους, λέξεις, φράσεις). Στο στάδιο της αναγνώρισης, λαμβάνονται μέτρα πιθανοτήτων που χρησιμοποιούνται για να υπολογίσουν πόσο μια άγνωστη εισαγόμενη ομιλία ταιριάζει με ένα συγκεκριμένο σύνολο HMMs. Τα HMMs χρησιμοποιούνται στην αναγνώριση ομιλίας, διότι ένα σήμα ομιλίας μπορεί να θεωρηθεί σαν τμηματικά στάσιμο σήμα ή βραχυπρόθεσμο στάσιμο σήμα. Ένας άλλος λόγος για τον οποίο HMMs είναι δημοφιλή είναι επειδή μπορούν να εκπαιδεύονται αυτόματα και είναι απλά και υπολογιστικά εφικτό να χρησιμοποιηθούν.

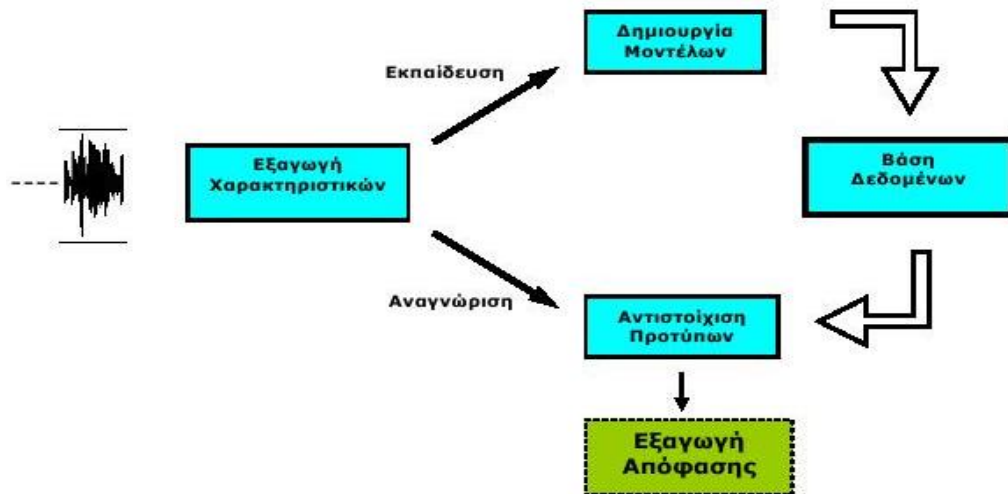
Τα τελευταία χρόνια η γλωσσική ανάλυση γίνεται όλο και περισσότερο σημαντική στην αναγνώριση ομιλίας, ειδικά όταν υπάρχει μεγάλο λεξιλόγιο και συνεχής αναγνώριση ομιλίας. Η διαδικασία αποκωδικοποίησης ομιλίας θα πρέπει να επικαλεστεί τις γνώσεις σχετικά με τη προφορά, το λεξιλόγιο, τη σύνταξη, και τη πραγματολογία, προκειμένου να παράγουν μια ικανοποιητική σειρά κειμένων.

Τα ακουστικά μοντέλα περιλαμβάνουν τις ακουστικές ιδιότητες, τις φωνητικές ιδιότητες, τη μεταβλητότητα του μικροφώνου και του περιβάλλοντος, καθώς και τις διαλεκτικές διαφορές μεταξύ των ομιλητών. Τα γλωσσικά μοντέλα περιλαμβάνουν τη σύνταξη, τη σημασιολογία και γενικότερα όλα τα χαρακτηριστικά μιας γλώσσας. Τα μοντέλα αυτά μπορούν να τροποποιηθούν δυναμικά σύμφωνα με τα χαρακτηριστικά της ομιλίας η οποία είχε αναγνωριστεί προηγουμένως.

### *3.2.3 Φάσεις Συστήματος Αναγνώρισης Ομιλίας*

Οι φάσεις που λειτουργεί Ένα ολοκληρωμένο σύστημα αναγνώρισης ομιλίας λειτουργεί είναι δύο, η φάση της εκπαίδευσης και η φάση της αναγνώρισης.

- Εκπαίδευση συστήματος: Κατά την οποία γίνεται εισαγωγή στο σύστημα φωνητικών δειγμάτων γνωστής προέλευσης, από τα οποία, με ειδική επεξεργασία, θα δημιουργηθούν τα μοντέλα πιθανότητας.
- Αναγνώριση: Κάθε νέο, άγνωστο, φωνητικό δείγμα που εισάγεται στο σύστημα συγκρίνεται με τα ήδη υπάρχοντα, γνωστά δείγματα προκειμένου να επιτευχθεί η αναγνώριση, επιβεβαίωση ή ταυτοποίηση του ομιλητή.



Εικόνα 19 Οι φάσεις λειτουργίας ενός ολοκληρωμένου συστήματος αναγνώρισης ομιλίας

Τα ASR συστήματα ενσωματώνουν τρία χαρακτηριστικά γνωρίσματα:

- Μεγάλα λεξιλόγια. Είναι σαφές ότι όσο μεγαλύτερο είναι το λεξιλόγιο τόσο μεγαλώνει και η πιθανότητα το σύστημα να κάνει λάθη.
- Συνεχή αναγνώριση ομιλίας η οποία είναι πολύ πιο φυσική και φιλική στο χρήστη. Το σύστημα είναι σε θέση να αναγνωρίσει μια σειρά από λέξεις μέσα σε μια πρόταση, όμως η ακρίβεια αναγνώρισης είναι στη πραγματικότητα μικρότερη από την αναγνώριση μεμονωμένων λέξεων. Αυτό μπορεί να οφείλεται σε διάφορους παράγοντες. Τα όρια των λέξεων δεν είναι απαραίτητα σαφή ή η προφορά των ομιλητών δεν είναι προσεκτική.
- Συστήματα αναγνώρισης καθορισμένου και ακαθόριστου κειμένου. Τα συστήματα καθορισμένου κειμένου απαιτούν από τον χρήστη να επαναλάβει κάποιες συγκεκριμένες εκφράσεις, που συνήθως περιέχουν το ίδιο κείμενο με τα δεδομένα της εκπαίδευσής του. Αυτά τα συστήματα έχουν μεγαλύτερη απόδοση εφόσον περιορίζεται η διαδικασία και υπάρχει αυξημένη ασφάλεια καθώς είναι δύσκολο να εξαπατηθεί το σύστημα. Ωστόσο όμως απαιτούν συγκεκριμένη και ακριβή άρθρωση του λόγου προκειμένου να γίνει η ταυτοποίηση. Όσον αφορά την αναγνώριση ομιλίας ακαθόριστου κειμένου, το σύστημα δεν γνωρίζει, εξ' αρχής, το κείμενο που προφέρεται από τον χρήστη, είναι πιο ευέλικτο και τα μοντέλα πιθανότητας γίνονται πιο πολύπλοκα.

Σήμερα, υπάρχουν πολλά συστήματα που περιλαμβάνουν αυτούς τους συνδυασμούς.

Η απόδοση ενός συστήματος ASR αποτελείται κυρίως από τρία βασικά μέρη: την ακρίβεια της αναγνώρισης, την πολυπλοκότητα και την ευρωστία. Τα διαθέσιμα συστήματα αναγνώρισης φωνής που κυκλοφορούν στο εμπόριο απαιτούν συνήθως μόνο μια σύντομη περίοδο εκπαίδευσης και μπορούν επιτυχώς να αναγνωρίσουν τη συνεχή ομιλία με ένα μεγάλο λεξιλόγιο σε πραγματικό χρόνο με μια πολύ υψηλή ακρίβεια. Οι περισσότερες εμπορικές επιχειρήσεις υποστηρίζουν ότι το λογισμικό αναγνώρισης είναι δυνατόν να επιτύχει ακρίβεια μεταξύ 98% και 99% εάν λειτουργεί υπό ιδανικές συνθήκες. «Οι βέλτιστες συνθήκες» συνήθως προϋποθέτουν οι χρήστες να έχουν χαρακτηριστικά ομιλίας που συμφωνούν με τα δεδομένα εκπαίδευσης και να βρίσκονται σε περιβάλλον χωρίς θορύβους Αυτό εξηγεί γιατί μερικοί χρήστες, ειδικά εκείνοι των οποίων η ομιλία είναι βαριά, μπορεί να επιτύχουν ποσοστά αναγνώρισης πολύ χαμηλότερα από το αναμενόμενο.

Στόχος	Ακρίβεια
Ναι / Όχι	> 99%
Ταχυδρομικός Κώδικας	> 97%
Τηλέφωνο	> 95%
Ημερομηνία	> 95%
Λέξη (μέγεθος λεξιλογίου 2.500)	> 98%
Φράση με τεχνική endpointing (4.000 – closed grammar)	> 95%
Natural Language Understanding (open grammar)	> 92%
Speak Freely	> 90%

Εικόνα 20 Ακρίβεια αναγνώρισης ομιλίας

### 3.2.4 Λόγοι παρουσίας σφαλμάτων στα Συστήματα Αναγνώρισης Ομιλίας

- Φράσεις που έχουν διαβαστεί ή εκφραστεί λανθασμένα
- Ακραίες συναισθηματικές συνθήκες.
- Ο τρόπος έκφρασης της φράσης είναι διαφορετικός από τον τρόπο που χρησιμοποιήθηκε στη καταγραφή.
- Αλλαγή θέσης λήψης του μικροφώνου.
- Ασυμβατότητα καναλιού (π.χ. χρήση διαφορετικών μικροφώνων για καταγραφή και επιβεβαίωση).
- Διαφορετική ταχύτητα προφοράς κατά την διάρκεια της επιβεβαίωσης σε σύγκριση με τα δεδομένα της εκπαίδευσης.

- Ασθένεια (π.χ. το κρύωμα μπορεί να αλλοιώσει τη φωνή).
- Η ηλικία.

### 3.2.5 Εφαρμογές της ASR τεχνολογίας

- Σε άτομα με ειδικές ανάγκες
- Αυτόματη μετάφραση
- Στη τηλεματική (GPS)
- Στη IVR τεχνολογία
- Στη κινητή τηλεφωνία, συμπεριλαμβανομένου του κινητού ηλεκτρονικού ταχυδρομείου
- Στη ρομποτική
- Στα βιντεοπαιχνίδια
- Στο αυτοκίνητο
- Στα μαχητικά αεροσκάφη
- Στις προσομοιώσεις

### 3.2.6 Οφέλη της ASR τεχνολογίας

Υπάρχουν ουσιαστικά τρεις βασικοί λόγοι για τους οποίους έχει γίνει τόση πολύ έρευνα και προσπάθεια για να μπορέσουν να διδάξουν σε μηχανές να αναγνωρίζουν και να κατανοούν την ομιλία:

- Προσβασιμότητα για τα άτομα με προβλήματα ακοής
- Μείωση των δαπανών μέσω της αυτοματοποίησης
- Ικανότητα αναζήτησης κειμένων

### 3.2.7 Μελλοντικά Σχέδια

- Το MIT της Βοστώνης αναπτύσσει ένα λογισμικό το οποίο θα ενσωματωθεί σε ένα smart phone και θα μετατρέπει τα ηχητικά κύματα σε δόνηση που θα γίνεται αισθητή από τον κωφό και φυσικά κάθε μία θα διαφοροποιείται ώστε να γίνεται διακριτή η λέξη, η έννοια.

- Το YouTube Ανακοίνωσε ότι τα εκατομμύρια βίντεο που φιλοξενεί, θα αποκτήσουν αυτόματους υπότιτλους, κάτι που θα τα καταστήσει πιο προσβάσιμα σε κωφούς και άτομα με προβλήματα ακοής. Η πρωτοβουλία της βασίζεται στις τελευταίες εξελίξεις στην τεχνολογία αναγνώρισης φωνής. Αποτελεί του μεγαλύτερο online πείραμα του είδους του στον κόσμο.
- Αναζήτηση πληροφοριών στο διαδίκτυο με χρήση φωνητικών εντολών  
1-800-GOOG-411 Google

### 3.3 TTS text-to-speech

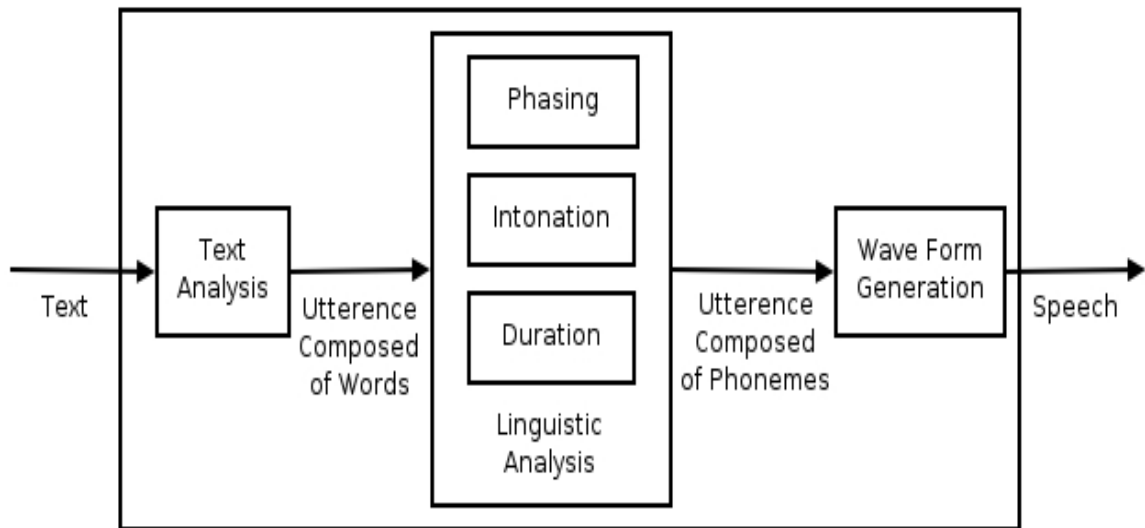
Το όνειρο πολλών επιστημόνων και όχι μόνο ήταν να μπορέσουν να κατασκευάσουν μια μηχανή η οποία θα μπορούσε να μιλά με ανθρώπινη φωνή. Για να επιτευχθεί αυτό χρειάστηκαν πολλά χρόνια μελετών ώσπου φτάσαμε σήμερα με την εξέλιξη της τεχνολογίας να έχουμε μηχανές που να συνθέτουν φωνή που να μοιάζει όλο και περισσότερο στην ανθρώπινη φωνή.

Ένα text-to-speech (TTS) μετατροπής κειμένου-σε-ομιλία σύστημα μετατρέπει ένα κανονικό γλωσσικό κείμενο σε ανθρώπινη ομιλία με υψηλή ποιότητα ,παρέχοντας στους ανθρώπους πληροφορίες κειμένου μέσω φωνητικών μηνυμάτων . Ένα TTS λογισμικό μπορεί να "διαβάσει" το κείμενο από ένα έγγραφο, μια ιστοσελίδα ή ένα ηλεκτρονικό βιβλίο. Η σύνθεση ομιλίας μπορεί να δημιουργηθεί από τη σύνδεση ηχογραφημένων κομματιών ομιλίας τα οποία είναι αποθηκευμένα σε μια βάση δεδομένων.

Η ποιότητα ενός συνθέτη ομιλίας εξαρτάται από την ομοιότητα της με την ανθρώπινη φωνή και από την ικανότητά της να γίνει κατανοητή από τον άνθρωπο. Το κείμενο το οποίο θα μετατραπεί σε ομιλία είτε εισάγεται από κάποιο χρήστη απευθείας στον υπολογιστή είτε σαρώνεται και στην συνέχεια περνάει από ένα σύστημα οπτικής αναγνώρισης χαρακτήρων (Optical Character Recognition system – OCR) .

Η τεχνολογία σύνθεσης ομιλίας περιλαμβάνει την αλγοριθμική μετατροπή των δεδομένων εισόδου κειμένου σε φωνητικές κυματομορφές, επιτρέπει το μετασχηματισμό μιας σειράς φωνητικών και προσωδιακών συμβόλων σε ένα σήμα συνθετικής ομιλίας .Οι λεκτικοί συνθέτες χαρακτηρίζονται από τη μέθοδο που χρησιμοποιείται για την αποθήκευση, την κωδικοποίηση και τη σύνθεση της

ομιλίας. Η μέθοδος σύνθεσης καθορίζεται από το μέγεθος του λεξιλογίου, δεδομένου ότι όλες οι πιθανές εκφράσεις της γλώσσας πρέπει να διαμορφωθούν. Οι τομείς εφαρμογής της σύνθεσης ομιλίας μπορούν να περιλαμβάνουν αναγνώστες ηλεκτρονικού ταχυδρομείου ή ειδήσεων, διαλογικά συστήματα, αυτόματα συστήματα μετάφρασης.

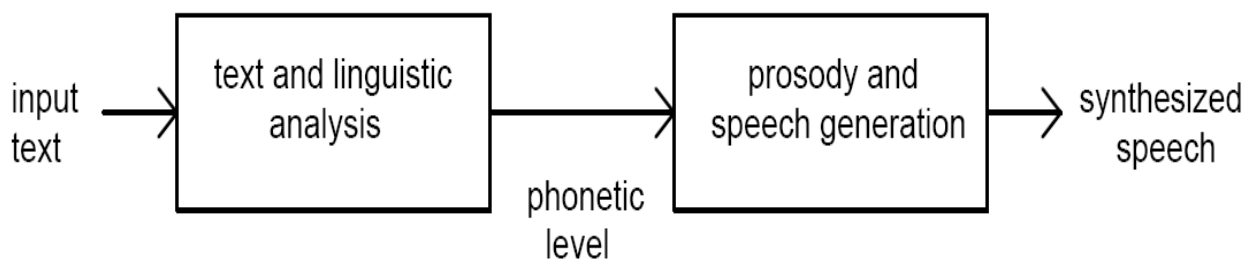


Εικόνα 21 Τεχνολογία Text-to-speech TTS

### 3.3.1 Συστατικά ενός TTS συνθέτη

Ένας TTS συνθέτης αποτελείται από δύο φάσεις:

- Μια υψηλού επιπέδου φάση σύνθεσης η οποία δέχεται σαν είσοδο κείμενα και εξάγει μια συμβολική γλωσσική αναπαράσταση.
- Μια υψηλού επιπέδου φάση σύνθεσης που δέχεται τη συμβολική γλωσσική αναπαράσταση και εξάγει παραγόμενη ομιλία σε κυματομορφή



Εικόνα 22 Φάσεις ενός text-to-speech συνθέτη



### 3.3.2 Ποιότητα ενός TTS σύνθετη

Η ποιότητα ενός TTS συνθέτη εξαρτάται από δύο κριτήρια :

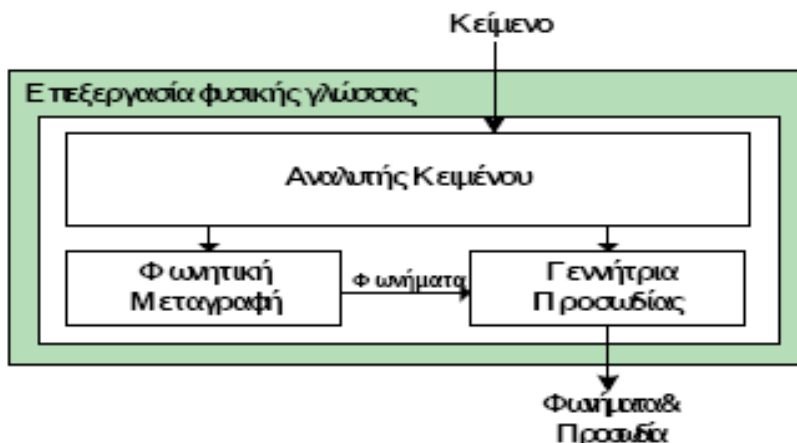
- **Σαφήνεια** - αναφέρεται στο πόσο εύκολα η παραγόμενη ομιλία μπορεί να γίνει κατανοητή. Μπορεί να μετρηθεί λαμβάνοντας υπόψη διάφορα είδη μονάδων (φωνήματα, συλλαβές, λέξεις, φράσεις, κλπ..).
- **Φυσικότητα** – αναφέρεται στο πόσο μια παραγόμενη ομιλία μοιάζει με την ομιλία ενός πραγματικού προσώπου.

Οι συνθέτες ομιλίας μπορούν αν χρειαστεί να μιλήσουν πιο γρήγορα από τους ανθρώπους. Αυτό μπορεί να είναι χρήσιμο μερικές φορές ωστόσο όμως η ποιότητα της ομιλίας που παράγεται να είναι κατά πολύ χαμηλότερη από την ανθρώπινη ομιλία.

### 3.3.3 Αρχιτεκτονική ενός TTS συνθέτη

Πώς ένας υπολογιστής διαβάζει; Αυτό το φαινομενικά δύσκολο ερώτημα μπορεί να απαντηθεί αν κανείς εξετάσει την αρχιτεκτονική ενός τυπικού TTS συνθέτη. Σε αυτή την ενότητα εξετάζονται τα βασικά χαρακτηριστικά-λειτουργίες ενός TTS συνθέτη.

Ένα Text-to-Speech σύστημα αποτελείται από τέσσερις βασικές λειτουργίες : την ανάλυση κειμένου, την αυτόματη φωνηματική μεταγραφή , τη παραγωγή προσωδίας και τέλος η ψηφιακή σύνθεση φωνής .Ο αναλυτής κειμένου η φωνηματική μεταγραφή και ο υπολογισμός της Προσωδίας αποτελούν αντικείμενα της Επεξεργασίας Φυσικής Γλώσσας (Natural Language Processing).



Εικόνα 23 Επεξεργασία Φυσικής Γλώσσας

- Κατά τη ανάλυση κειμένου ένα μοντέλο προεπεξεργασίας οργανώνει τις προτάσεις εισαγωγής σε εύχρηστους καταλόγους λέξεων, προσδιορίζει τους αριθμούς, τις συντομογραφίες, και τα αρκτικόλεξα και τα μετατρέπει σε πλήρες κείμενο. Γίνεται διαχωρισμός των προτάσεων και εν συνεχεία κατάτμηση των προτάσεων σε λέξεις, συμβολοσειρές κλπ. Επίσης υπάρχει ένα μορφολογικό μοντέλο ανάλυσης που προτείνει όλες οι πιθανές κατηγορίες ομιλίας για κάθε λέξη που λαμβάνεται χωριστά, βάσει της ορθογραφίας τους, κι ένα μοντέλο ανάλυσης βασισμένο στα συμφραζόμενα που εξετάζει τις λέξεις από τα συμφραζόμενα τους. Γίνεται φραστική απόδοση δειγμάτων που δεν ανήκουν στη φυσική γλώσσα (@, &) και τέλος γίνεται μορφολογική ανάλυση για το προσδιορισμό γραμματικών χαρακτηριστικών των δειγμάτων (μέρος του λόγου, γένος) και συντακτική ανάλυση για την εύρεση της συντακτικής δομής μιας πρότασης
- Το μοντέλο αυτόματης φωνηματικής μεταγραφής είναι υπεύθυνο για τον αυτόματο προσδιορισμό της φωνητικής μεταγραφής του εισερχόμενου κειμένου, μετατρέπει δηλαδή τους χαρακτήρες του κειμένου εισόδου σε φωνήματα. Ωστόσο συνήθως η προφορά μιας λέξης διαφέρει από την ορθογραφία της. Όπως για παράδειγμα ο χαρακτήρας «ξ» αντιστοιχεί σε δύο φωνήματα /ks/ και μια σειρά από ένα ή περισσότερους χαρακτήρες π.χ. τα «εί», «οί», «υί» να αντιστοιχούν σε ένα φώνημα στο /i/. Η λέξη <<σβέλτος>> θα πρέπει να μετατραπεί σε /zveltos/ και όχι σε /sveltos/. Επίσης το μοντέλο αυτόματης φωνηματικής μεταγραφής πρέπει να χειρίζεται σωστά τις περιπτώσεις όπου λέξεις που περιέχουν ίδιους φθόγγους προφέρονται διαφορετικά όπως για παράδειγμα ενώ η λέξη <<πιεστικός>> γίνεται /πι-ε-στικός/ η λέξη <<πιες>> γίνεται /πχές/. Τα λεξικά προφοράς δε μπορούν πάντα να βοηθήσουν αυτό το μοντέλο γιατί αναφέρονται μόνο σε ρίζες λέξεων και οι λέξεις που ενσωματώνονται στις προτάσεις δεν προφέρονται σαν ήταν απομονωμένες. Τέλος το συγκεκριμένο μοντέλο θα πρέπει να χειρίζεται ειδικούς χαρακτήρες, όπως οι συντομογραφίες και τα νούμερα, τα οποία προφανώς, πρέπει πρώτα να μετατραπούν σε λέξεις στην σωστή κλίση, και στην συνέχεια σε φωνή.

Παράδειγμα λεξικού προφοράς με μετατροπή λέξεων σε ήχους :

A	ah
A	ax
A	ey
CALL	k ao l
DIAL	d ay ax l
EIGHT	ey t
PHONE	f ow n
SEVEN	s eh v n
TO	t ax
TO	t uw
ZERO	z ia r ow

- Η προσωδία ορίζεται ως το σύνολο των χαρακτηριστικών της φωνής που προσδίδουν φυσικότητα στο σήμα και χαρακτηρίζονται από χρονική διάρκεια μεγαλύτερη από αυτή των φωνημάτων, είναι το αποτέλεσμα της κινητοποίησης ενός φυσικού μηχανισμού στον οποίο συμμετέχουν ποικίλα όργανα. Τα χαρακτηριστικά της είναι οι παύσεις, η τονικότητα, η ένταση και η διάρκεια. Με τη προσωδία ένας ομιλητής μπορεί να εκφράσει ερώτηση, άρνηση ή κατάφαση ακόμα και να εκφράσει τη συναισθηματική του κατάσταση. Επίσης μπορεί να δείξει ειρωνεία και σαρκασμό καθώς επίσης και άλλα στοιχεία της γλώσσας που δεν μπορούν να κωδικοποιηθούν από γραμματική και λεξιλόγια. Αποτελεί σημαντικό κομμάτι της καθημερινής λεκτικής επικοινωνίας των ανθρώπων άρα και του συστήματος TTS. Ωστόσο η εκπαίδευση μιας μηχανής στο πώς θα υπολογίζει τις προσωδικές παραμέτρους ενός κειμένου είναι αρκετά δύσκολη αν και τα τελευταία χρόνια παρουσιάζεται τεράστια βελτίωση.
- Η τελευταία λειτουργία ενός TTS συνθέτη είναι η Ψηφιακή Σύνθεση Φωνής. Οι τεχνικές Ψηφιακής Σύνθεσης Φωνής, έχουν μελετηθεί και βελτιωθεί περισσότερο, απ' ό,τι οι αλγόριθμοι Επεξεργασίας Φυσικής Γλώσσας επειδή ο μηχανισμός παραγωγής φωνής είναι ο ίδιος για όλους τους Ανθρώπους, ενώ οι γλώσσες διαφέρουν.

Υπάρχουν διάφορες τεχνικές σύνθεσης φωνής:

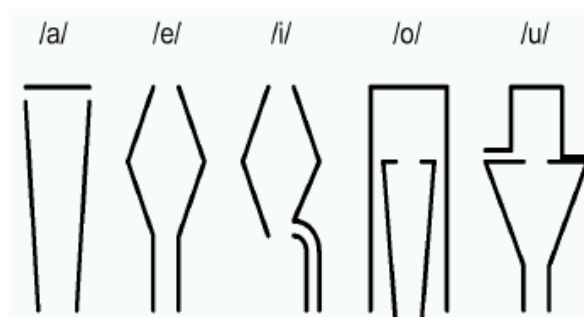
- Συνθέτες φωνής που χρησιμοποιούν προηχογραφημένα κομμάτια φωνής τα οποία συνδέονται αλυσιδωτά δημιουργώντας λέξεις και προτάσεις (concatenation synthesis)
- παραμετρικοί συνθέτες, οι οποίοι χρησιμοποιούν ψηφιακά, χρονομεταβλητά φίλτρα με απόκριση συχνότητας κατάλληλη έτσι ώστε να δημιουργηθεί συνθετική φωνή, παραπλήσια φασματικά με την ανθρώπινη (formant synthesizers).
- Συνθέτες που προσπαθούν να μιμηθούν τον Ανθρώπινο τρόπο παραγωγής φωνής. Σ' αυτή τη κατηγορία, ανήκουν οι Συνθέτες άρθρωσης (articulation synthesizers), οι οποίοι με μαθηματικά μοντέλα προσομοιώνουν τις κινήσεις όλων των αρθρωτήρων του φωνητικού σωλήνα (vocal tract), για την παραγωγή κάθε φθόγγου.

### 3.3.4 Ιστορία της σύνθεσης φωνής

Η ιδέα της δημιουργίας μιας μηχανής η οποία θα έχει τη δυνατότητα να μιλάει δεν είναι νέα. Για αιώνες, οι άνθρωποι έχουν προσπαθούσαν να βρουν τρόπους να κάνουν τις μηχανές να μιλήσουν. Οι πρώτες προσπάθειες για να παραχθεί η ανθρώπινη ομιλία από μια μηχανή έγιναν το δέκατο όγδοο αιώνα.

Οι πρώτες ομιλούσες μηχανές ήταν μηχανικές λεκτικές συσκευές που μιμήθηκαν την ανθρώπινη φωνητική οδό.

Το 1773 ο Christian Kratzenstein στη Αγία Πετρούπολη εξήγησε τις φυσιολογικές διαφορές μεταξύ πέντε φωνήεντων (/ a /, / e /, / i /, / o / και / u /) με μία συσκευή που μπορούσε να τα παράγει τεχνητά. Ένα σύνολο ακουστικών ταλαντωτών οι οποίοι λειτουργούσαν με αέρα μιμήθηκαν το ανθρώπινη φωνητική οδό.

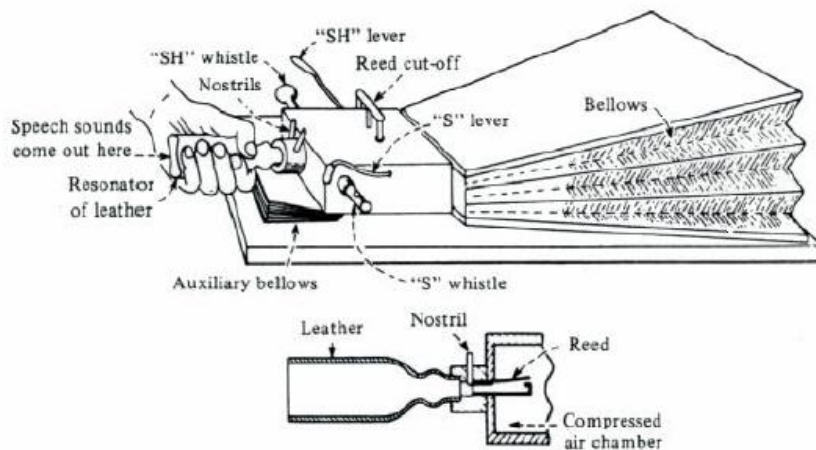


Εικόνα 24 Οι ακουστικοί ταλαντωτές του Christian Kratzenstein

Το 1791 στη Βιέννη ο πρώτος πειραματικός φωνολόγος Wolfgang von Kempelen έγραψε το *Mechanismus der menschlichen Sprache nebst Beschreibung einer*

*sprechenden Maschine*. Το βιβλίο είχε μια λεπτομερή περιγραφή μιας μηχανής μιλάει. Οι μελέτες του Kempelen οδήγησαν στη θεωρία ότι η φωνητική οδός, μια κοιλότητα μεταξύ των φωνητικών χορδών και των χειλιών, είναι η κύρια περιοχή της ακουστικής άρθρωσης. Τα βασικά μέρη της μηχανής ήταν μια περιοχή πίεσης αέρα για τους πνεύμονες, τα βασικά μέρη της μηχανής ήταν μια αίθουσα πίεσης για τους πνεύμονες, ένα παλλόμενο έλασμα που ενεργούσε όπως οι φωνητικές χορδές, και ένας σωλήνας δέρματος αντί των φωνητικών οδών. Η μηχανή λειτουργούσε χειροκίνητα και μπορούσε να παραγάγει ενιαίους ήχους .

Το 1835 η συγκεκριμένη μηχανή ανακατασκευάστηκε από τον Charles Wheatstone και πλέον μπορούσε να παράγει όχι μόνο ενιαίους ήχους αλλά και ολόκληρες λέξεις ακόμα και σύντομες φράσεις.



Εικόνα 25 Μηχανή ομιλίας του Kempelen

Στην αρχή του 20ου αιώνα, η πρόοδος στην ηλεκτρική εφαρμοσμένη μηχανική κατέστησε πιθανό να παραχθούν λεκτικοί ήχοι με ηλεκτρικά μέσα. Η πρώτη συσκευή που θεωρείται ως λεκτικός συνθέτης ήταν ένας ηλεκτρικός συνθέτης αποκαλούμενος VODER που παρουσιάστηκε από Homer Dudley στην παγκόσμια έκθεση στη Νέα Υόρκη το 1939. Η συσκευή είχε ένα πετάλι ποδιού για να ελέγχει τη συχνότητα και ένα σύνθετο πληκτρολόγιο . Τα σήματα, μέσα σε αυτό το συνθεσάιζερ, δρομολογούνται μέσα από δέκα φίλτρα. Ωστόσο, αυτή η συσκευή απαιτεί πολύ μεγάλο χρονικό διάστημα εκπαίδευσης για επιτυχή χρήση.



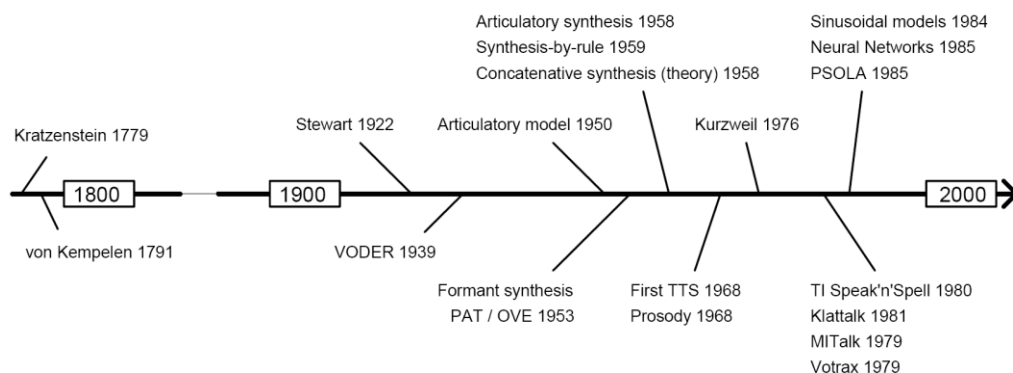
Εικόνα 26 Ο συνθέτης φωνής VODER

Το 1953 ο Walter Lawrence παρουσίασε τον πρώτο συνθέτη με κανόνες (formant) PAT (Parametric Artificial Talker). Ο PAT αποτελείται από τρεις συντονιστές που συνδεόταν παράλληλα. Το σήμα εισόδου ήταν είτε ένας βόμβος είτε θόρυβος. Την ίδια εποχή ο Gunnar Fant παρουσίασε τον συνθέτη OVE (Orator Verbis Electricis) με συντονιστές σε σειρά.

Το 1958 ο πρώτος αρθρωτικός συνθέτης DAVO (Dynamic Analog of the VOcal tract) εισήχθη από τον George Rosen στο Τεχνολογικό Ινστιτούτο της Μασαχουσέτης. Τα σήματα ελέγχου εισαγόταν χειροκίνητα και ήταν ηχογραφημένα.

Ο πρώτος πλήρης TTS συνθέτης για τα αγγλικά αναπτύχθηκε στο τεχνικό εργαστήριο, στην Ιαπωνία το 1968 από το Noriko Umeda και την ομάδα του.

Το 1979 οι Allen, Hunnicutt, και Klatt παρουσίασαν τον MITalk ένα εργαστηριακό TTS συνθέτη που αναπτύχθηκε στο M.I.T.



Εικόνα 27 Χρονικό της σύνθεσης φωνής

### 3.3.5 Εφαρμογές ενός TTS συστήματος

Η σύνθεση ομιλίας μπορεί να χρησιμοποιηθεί σε διάφορες εφαρμογές, μερικές από αυτές περιγράφονται παρακάτω

- **Εφαρμογές στις τηλεπικοινωνίες και τα πολυμέσα**– τα TTS συστήματα μπορούν να παρέχουν στους χρήστες πρόσβαση σε πληροφορίες κειμένου μέσω του τηλέφωνου. Όπως για παράδειγμα η ηχητική αναγνώριση κλήσεων και η ανάγνωση ενός ολοκληρωμένου μηνύματος (μπορεί κάποιος να ακούσει μήνυμα από το ηλεκτρονικό ταχυδρομείο)
- **Εκπαιδευτικές εφαρμογές** – Ένας υπολογιστής με ένα συνθέτη ομιλίας μπορεί να διδάξει 24 ώρες την ημέρα και 365 μέρες το χρόνο. Μπορεί να προγραμματιστεί για ειδικές λειτουργίες όπως η διδασκαλία ορθογραφίας και προφοράς για διαφορετικές γλώσσες. Ένα TTS σύστημα μπορεί να συνδυαστεί με ένα σύστημα εκμάθησης, και να παρέχει ένα χρήσιμο εργαλείο για να μπορέσει κάποιος να μάθει μια νέα γλώσσα.
- **Βοήθεια σε άτομα με ειδικές ανάγκες** - Ίσως το πιο σημαντικό και χρήσιμο πεδίο εφαρμογής της σύνθεσης ομιλίας είναι η ανάγνωση και τα βοηθήματα επικοινωνίας για άτομα με προβλήματα στην όραση και στην ομιλία. Για παράδειγμα αν κάποιος ο οποίος έχει πρόβλημα ομιλίας χρησιμοποιήσει ένα ειδικά σχεδιασμένο πληκτρολόγιο και ένα γρήγορο πρόγραμμα σύνθεσης προτάσεων, η συνθετική ομιλία μπορεί να παραχθεί σε μερικά δευτερόλεπτα και να του προσφέρει μια ανεκτίμητη βοήθεια.
- **Εφαρμογές Πολυμέσων, επικοινωνία ανθρώπου-μηχανής** Η ανάπτυξη υψηλού επιπέδου TTS συστημάτων είναι απαραίτητα για μια πιο πλήρη επικοινωνία μεταξύ ανθρώπου και υπολογιστή.

### 3.3.6 Γλώσσες σήμανσης σύνθεσης ομιλίας

Διάφορες γλώσσες σήμανσης έχουν καθιερωθεί για την απόδοση του κειμένου ως ομιλία σε μορφή XML, η πιο πρόσφατη είναι η γλώσσα σήμανσης σύνθεσης ομιλίας Speech Synthesis Markup Language (SSML). Οι παλαιότερες γλώσσες σήμανσης σύνθεσης ομιλίας περιλαμβάνουν τη Java Speech Markup Language (JSML) και τη SABLE. Αν και κάθε μία από αυτές προτάθηκε ως πρότυπο, καμία δεν έχει υιοθετηθεί ευρέως. Οι Γλώσσες σήμανσης σύνθεσης ομιλίας διαφέρουν

από τις γλώσσες σήμανσης διαλόγου, η VoiceXML, για παράδειγμα εκτός από τις text-to-speech σημάνσεις περιλαμβάνει ετικέτες που σχετίζονται με την αναγνώριση ομιλίας, τη διαλογική διαχείριση και την κλήση touchtone .

### 3.4 VoiceXML

Μέχρι πριν μερικά χρόνια οι Φωνητικές Πύλες δε παρείχαν πρόσβαση σε ολόκληρο το Διαδίκτυο παρά μόνο σε συγκεκριμένα στοιχεία και υπηρεσίες. Τώρα πλέον το Διαδίκτυο είναι διαθέσιμο μέσω φωνητικών διεπαφών.

Η VoiceXML είναι μια ειδικής χρήσης γλώσσα για την περιγραφή του διαδραστικού φωνητικού διαλόγου. Η χαρακτηριστική μηχανή αναζήτησης φωνής VoiceXML έχει σχεδιαστεί για τη δημιουργία ηχητικών διαλόγων που συνδυάζουν την ομιλία, τον ψηφιακό ήχο, τη τεχνολογία αναγνώρισης ομιλίας, τη τεχνολογία DTMF, την ηχογραφημένη ή την παραγόμενη ομιλία και τη τηλεφωνία.

Βασικός στόχος της VoiceXML είναι να φέρει την πλήρη δύναμη της ανάπτυξης του Web και του περιεχομένου του στις φωνητικές εφαρμογές ανταπόκρισης και να απελευθερώσει τους συντάκτες εφαρμογών από τον χαμηλού επιπέδου προγραμματισμό. Δίνει τη δυνατότητα ενσωμάτωσης των υπηρεσιών φωνητικής τηλεφωνίας με υπηρεσίες δεδομένων, χρησιμοποιώντας το γνωστό μοντέλο client-server. Μια φωνητική υπηρεσία αντιμετωπίζεται ως μια ακολουθία διαλόγων αλληλεπίδρασης μεταξύ ενός χρήστη και μιας πλατφόρμας υλοποίησης.

Η VoiceXML

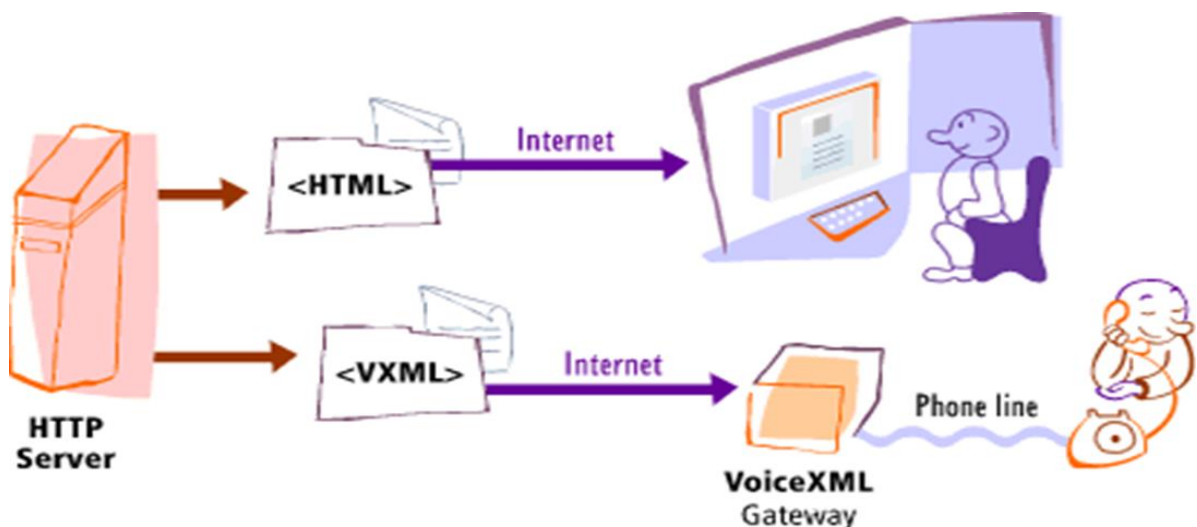
- Απλοποιεί τη ανάπτυξη εφαρμογών
- Ελαχιστοποιεί την αλληλεπίδραση κίνησης του διαδικτύου
- Χωρίζει τον κώδικα αλληλεπίδρασης χρηστών από τον κώδικα λογικής εφαρμογής
- Παρέχει τη φορητότητα υπηρεσιών σε πλατφόρμες υλοποίησης. Η VoiceXML είναι κοινή γλώσσα για τους προμηθευτές, τους προμηθευτές εργαλείων, και τους προμηθευτές πλατφορμών
- Παρέχει την απλότητα, είναι εύχρηστη για τις απλές αλληλεπιδράσεις, και όμως παρέχει τα γλωσσικά χαρακτηριστικά γνωρίσματα για να υποστηρίξει τους σύνθετους διαλόγους.
- Υποστηρίζει τη γρήγορη προτυποποίηση και τον επαναληπτικό καθαρισμό.



Οι εφαρμογές Ιστού παραδίδουν τις πληροφορίες και τις υπηρεσίες αποκλειστικά μέσω των οπτικών διεπαφών στους υπολογιστές με τις οθόνες, τα πληκτρολόγια, και το ποντίκι. Η VoiceXML είναι μια βασισμένη στο διαδίκτυο γλώσσα σήμανσης ακριβώς όπως η HTML. Ενώ, όμως, η HTML προϋποθέτει μια γραφική μηχανή αναζήτησης του διαδικτύου, με οθόνη, πληκτρολόγιο και ποντίκι, η VoiceXML προϋποθέτει φωνητική μηχανή αναζήτησης με έξοδο ήχου (παραγόμενο ή ηχογραφημένο λόγο), καθώς και είσοδο ήχου (φωνή ή τονικά πληκτρολόγια). Η VoiceXML αξιοποιεί το Διαδίκτυο για την ανάπτυξη και τη παράδοση φωνητικών εφαρμογών και δημιουργεί νέες ευκαιρίες. Το συνηθισμένο τηλέφωνο είναι πολύ σημαντικό στην ανάπτυξη της VoiceXML. Η τυπική φωνητική μηχανή αναζήτησης VoiceXML τρέχει σε ένα εξειδικευμένο κόμβο πυλών φωνής που συνδέεται και με το δημόσιο μεταστρεφόμενο τηλεφωνικό δίκτυο και με το Διαδίκτυο.

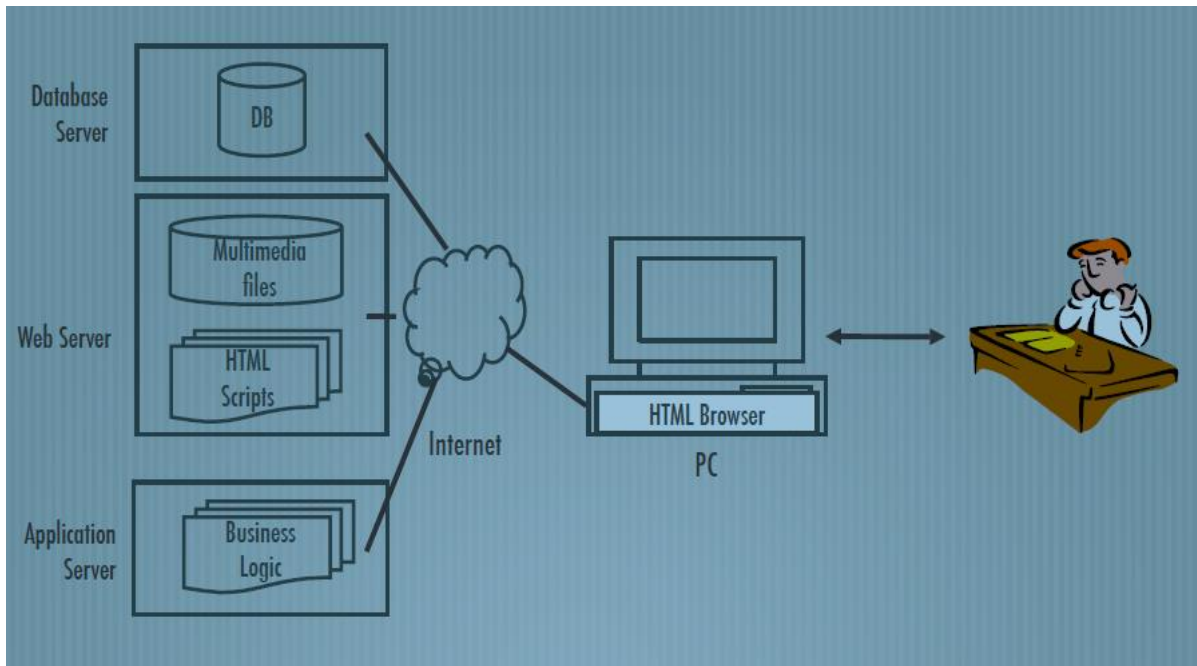
Οι φωνητικές ιστοσελίδες είναι σαν τις ιστοσελίδες για τα τηλέφωνα. Για να χρησιμοποιήσει κάποιος μια φωνητική ιστοσελίδα πρέπει να καλέσει από οποιοδήποτε τηλέφωνο, να ακούσει τις επιλογές που του παρουσιάζονται, και να κάνει την επιλογή του μιλώντας φυσικά. Αποτελούν μια επαναστατική τεχνολογία για τις επιχειρήσεις, τις οργανώσεις και τις κυβερνήσεις να παρέχει για την παροχή οικονομικά αποδοτικών και υψηλής ποιότητας εξυπηρέτησης πελατών μέσω των προηγμένων εφαρμογών αναγνώρισης ομιλίας.

Επίσης οι φωνητικές ιστοσελίδες επιτρέπουν στους καλούντες να πάρουν πληροφορίες, να συνδεθούν με άλλους ανθρώπους ή να αφήσουν μηνύματα από οποιοδήποτε τηλέφωνο.

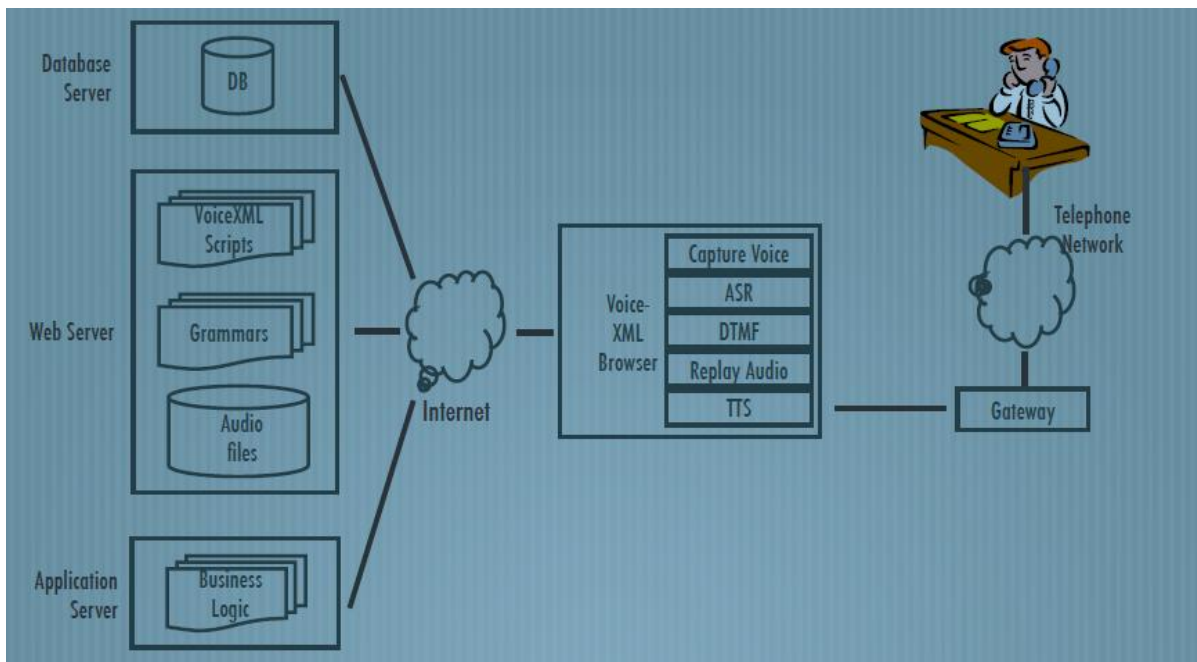


Εικόνα 28 VoiceXML

Στις εικόνες που ακολουθούν βλέπουμε τις διαφορές μεταξύ μιας διαδικτυακής εφαρμογής που χρησιμοποιεί HTML και μιας φωνητικής εφαρμογής που χρησιμοποιεί VoiceXML.

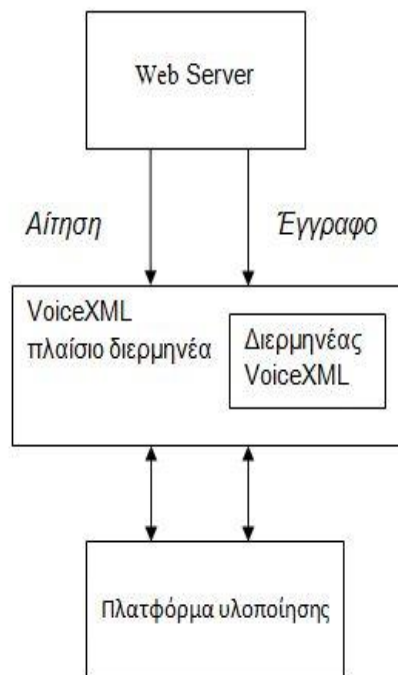


Εικόνα 29 Διαδικτυακή εφαρμογή



Εικόνα 30 Φωνητική εφαρμογή

### 3.4.1 Αρχιτεκτονικό μοντέλο



Εικόνα 31 Αρχιτεκτονικό μοντέλο VoiceXML

Ο Web Server επεξεργάζεται αιτήσεις από μια εφαρμογή πελάτη μέσω του πλαισίου διερμηνέα VoiceXML. Ο διακομιστής παράγει VoiceXML έγγραφα σε απάντηση, τα οποία υποβάλλονται σε επεξεργασία από το διερμηνέα VoiceXML. Το πλαίσιο διερμηνέα VoiceXML μπορεί να το παρακολουθεί τις εισαγωγές του χρήστη παράλληλα με τον διερμηνέα VoiceXML.

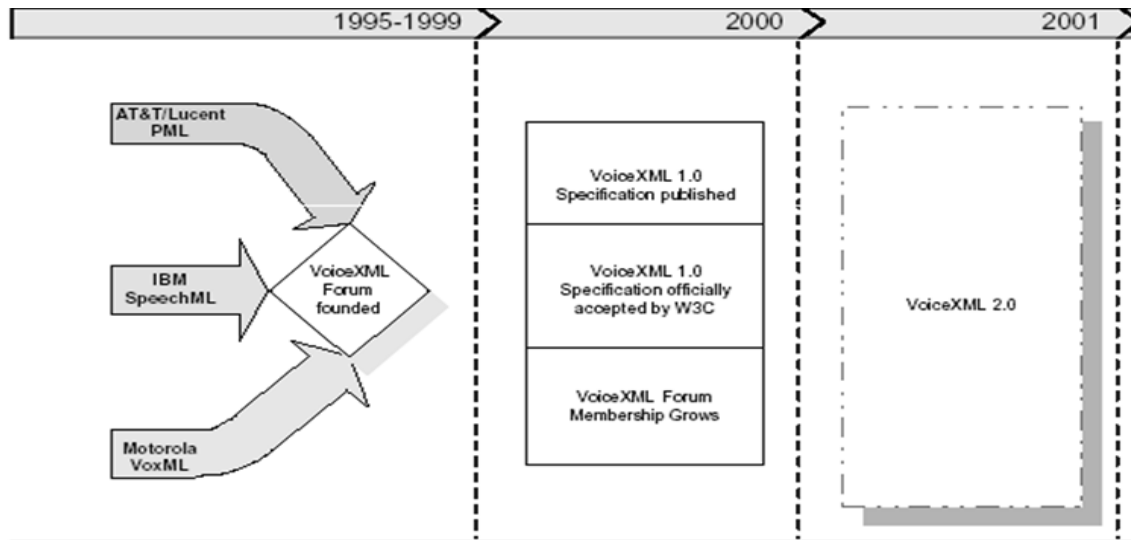
Η πλατφόρμα υλοποίησης ελέγχεται από το πλαίσιο διερμηνέα VoiceXML και από το διερμηνέα VoiceXML. Για παράδειγμα σε μια φωνητική πύλη το VoiceXML πλαίσιο διερμηνέα μπορεί να είναι υπεύθυνο για την ανίχνευση μιας εισερχόμενης κλήσης, αποκτώντας το αρχικό VoiceXML έγγραφο, και στη συνέχεια απαντά τη κλήση, ενώ ο διερμηνέας VoiceXML διεξάγει το διάλογο μετά την απάντηση.

### 3.4.2 Εφαρμογές της VoiceXML

- **Εφαρμογή στις φωνητικές πύλες** - Παρέχει εξατομικευμένες υπηρεσίες πρόσβασης σε πληροφορίες, όπως τιμές μετοχών, τον καιρό, κατάλογοι εστιατορίων, ειδήσεις, κτλ.
- **Υπηρεσίες βασιζόμενες στη τοποθεσία** - Εφαρμογές χρησιμοποιούν τον αριθμό τηλεφώνου από τον οποίο καλείτε.
- **Φωνητικές ειδοποιήσεις (όπως για τη διαφήμιση)** – αποστολή συγκεκριμένων σημάτων σε ένα χρήστη. Ο χρήστης μπορεί να λαμβάνει ειδικές ειδοποιήσεις ενημερώνοντάς τον για τις επερχόμενες εκδηλώσεις.
- **Εμπόριο** - εφαρμογές που επιτρέπουν στους χρήστες να αγοράσουν μέσω του τηλέφωνου. Χρησιμοποιείται κυρίως για προϊόντα που δεν χρειάζονται πολλή περιγραφή (όπως εισιτήρια, CDs, είδη γραφείου, κλπ.)

### 3.4.3 Ιστορία της VoiceXML

Η προέλευση της VoiceXML ξεκίνησε το 1995 ως μια διαλογική και σχεδιαστική γλώσσα βασισμένη στην XML που είχε σαν σκοπό να απλοποιήσει τη διαδικασία ανάπτυξης εφαρμογών λεκτικής αναγνώρισης. Το 1998, η W3C χρησιμοποίησε τη VoiceXML στις φωνητικές μηχανές αναζήτησης. Οι εταιρίες AT&T, IBM, Lucent, και Motorola διαμόρφωσαν VoiceXML φόρουμ το Μάρτιο του 1999 προκειμένου να αναπτυχθεί μια τυποποιημένη γλώσσα σήμανσης για τη διευκρίνιση των διαλόγων φωνής. Τον Σεπτέμβριο του 1999, το Φόρουμ κυκλοφόρησε την έκδοση VoiceXML 0.9 για μέλη και τον Μάρτιο του 2000 δημοσίευσε τη VoiceXML 1.0. Λίγο αργότερα, η VoiceXML 1.0 υποβλήθηκε στο W3C ως βάση για τη δημιουργία ενός νέου διεθνούς προτύπου. Η έκδοση VoiceXML 2.0 είναι το αποτέλεσμα αυτής της εργασίας. Η W3C παρήγαγε διάφορες ενδιάμεσες εκδόσεις VoiceXML 2.0, οι οποίες έφθασαν στο τελικό στάδιο «σύστασης» το Μάρτιο του 2004.



Εικόνα 32 Εξέλιξη της VoiceXML

#### 3.4.4 Παραδείγματα VoiceXML

Το ανώτατου επιπέδου στοιχείο είναι το `<vxml>` το οποίο δέχεται διαλόγους. Υπάρχουν δύο τύποι διαλόγων: φόρμες και μενού. Οι φόρμες παρουσιάζουν τις πληροφορίες και συλλέγουν τις εισαγωγές. Τα μενού προσφέρουν επιλογές για το τι θα κάνουμε στη συνέχεια.

Το πρώτο παράδειγμα έχει μια ενιαία φόρμα, η οποία περιέχει ένα μπλοκ που συνθέτει και παρουσιάζει το "Hello World!" στο χρήστη.

```
<vxml version="2.0" xmlns="http://www.w3.org/2001/vxml">  
  <form>  
    <block>  
      <prompt>  
        Hello world!  
      </prompt>  
    </block>  
  </form>  
</vxml>
```

Το δεύτερο παράδειγμά ζητά από το χρήστη να επιλέξει ένα ποτό και στη συνέχεια το υποβάλλει σε ένα script.

```
?xml version="1.0"?>
<vxml version="2.0" xmlns="http://www.w3.org/2001/vxml">
  <form>
    <field name="drink">
      <prompt>
        Would you like coffee, tea, milk, or nothing
      </prompt>
      <grammar src="drink.grxml" type="application/srgs+xml"/>
    </field>
    <block>
      <submit next="http://www.drink.example.com/drink2.asp"/>
    </block>
  </form>
</vxml>
```

Ένας διάλογος που μπορεί να προκύψει από το παραπάνω παράδειγμα μεταξύ χρήστη και υπολογιστή είναι ο εξής:

**Υπολογιστής:** Would you like coffee, tea, milk, or nothing?

**Χρήστης:** Orange juice.

**Υπολογιστής:** I did not understand what you said.(ειδικό προεπιλεγμένο μήνυμα)

**Υπολογιστής:** Would you like coffee, tea, milk, or nothing?

**Χρήστης:** Tea

**Υπολογιστής:** (συνεχίζει στο αρχείο drink2.asp)

### 3.4.5 Γλώσσες που σχετίζονται με τη VoiceXML

- Η Speech Synthesis Markup Language (SSML) Είναι μια γλώσσα σήμανσης βασισμένη σε XML και χρησιμοποιείται για εφαρμογές σύνθεσης φωνής. Η SSML ενσωματώνεται συνήθως στα VoiceXML scripts για να οδηγήσει τα διαλογικά συστήματα τηλεφωνίας. Ωστόσο, είναι επίσης δυνατό να χρησιμοποιηθεί και μόνη της, όπως για τη δημιουργία ηχητικών βιβλίων.
- Call Control eXtensible Markup Language (CCXML) είναι και αυτή ένα XML πρότυπο. Ενώ η VoiceXML έχει σχεδιαστεί για να παρέχει ένα φωνητική διαχείριση χρήστη σε μια μηχανή αναζήτησης φωνής, η CCXML έχει σχεδιαστεί για να ενημερώνει το πρόγραμμα περιήγησης φωνής για το πώς μπορεί να χειριστεί τον έλεγχο τηλεφωνίας του καναλιού φωνής.

## 4. Τεχνολογία αναγνώρισης συναισθημάτων

### Εισαγωγή

Στο τέταρτο κεφάλαιο περιγράφεται το πώς οι φωνητικές πύλες μπορούν να αναγνωρίσουν τη κατάσταση των συναισθημάτων του ομιλητή. Η κατασκευή μιας μηχανής η οποία θα μπορεί να αναγνωρίζει τα συναισθήματα από την ομιλία δεν είναι νέα ιδέα, οι έρευνες ξεκίνησαν τη δεκαετία του 80 χρησιμοποιώντας στατιστικές ιδιότητες ορισμένων ακουστικών χαρακτηριστικών γνωρισμάτων. Η αυτόματη αναγνώριση συναισθημάτων από την ομιλία έχει κατά την τελευταία δεκαετία μετατοπιστεί από ένα δευτερεύον ζήτημα σε μείζον θέμα στην αλληλεπίδραση ανθρώπου-υπολογιστή και την επεξεργασία του λόγου.

Ο στόχος είναι να καταστεί δυνατή μια πολύ φυσική αλληλεπίδραση με τον υπολογιστή, να μπορεί δηλαδή η μηχανή να κατανοεί όχι μόνο το λεκτικό περιεχόμενο αλλά και των συναισθημάτων που μπορεί να εκφράσει ανά πάσα στιγμή ο ομιλητής. Τα συναισθήματα διαδραματίζουν σημαντικό ρόλο στην ανθρώπινη επικοινωνία και κατά συνέπεια και τα διαλογικά συστήματα ανθρώπου μηχανής μπορεί να επωφεληθούν από τη συναισθηματική επεξεργασία.

Αρχικά για την αναγνώριση συναισθημάτων χρησιμοποιούταν επεμβατικές μέθοδοι μέτρησης όπως η μέτρηση της θερμοκρασίας του σώματος και η μέτρηση των καρδιακών παλμών. Αν και οι συγκεκριμένες μέθοδοι ήταν ιδιαίτερα αποτελεσματικές ήταν και αρκετά δύσχρηστες καθώς ο χρήστης θα έπρεπε να συνδέεται με κάποιον αισθητήρα. Έτσι οι επιστήμονες προσπάθησαν να δημιουργήσουν μη συμβατικές μεθόδους αναγνώρισης συναισθημάτων οι οποίες βασίζονται στη φωνή, στις χειρονομίες και τις εκφράσεις του προσώπου του χρήστη.

Τα ανθρώπινα συναισθήματα περιγράφονται πολύ σπάνια και συνήθως τα συναισθηματικά μηνύματα δεν είναι λεκτικά. Για να μπορέσει κάποιος να ανιχνεύσει τη συναισθηματική κατάσταση του άλλου θα πρέπει να έχει την ικανότητα να διαβάζει τα μη λεκτικά στοιχεία της επικοινωνίας. Όπως το τρέμουλο και το τόνο της φωνής, τις χειρονομίες καθώς και τις εκφράσεις του προσώπου. Η φωνή παρέχει συγκεκριμένα συναισθήματα μέσω ακουστικών ιδιοτήτων όπως για παράδειγμα η διακύμανση του ύψους της φωνής (Pitch) . Οι πληροφορίες σχετικά



με τα συναισθήματα είναι κωδικοποιημένες σε όλες τις πτυχές της γλώσσας, σε ότι λέμε και στο πώς το λέμε, και το «πώς» είναι πολύ πιο σημαντικό από το «τι».

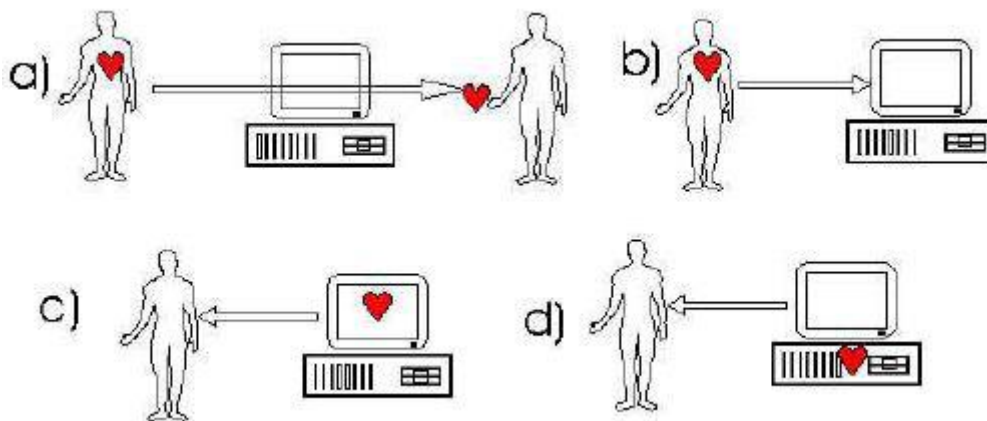
Οι εμπορικές συναλλαγές που δεν απαιτούν προσωπική συνάντηση διεκπεραιώνονται όλο και περισσότερο από τις αυτοματοποιημένες φωνητικές πύλες . Ωστόσο μέχρι πρόσφατα σε ένα κομμάτι της ανθρώπινης επικοινωνίας, αυτό των συναισθημάτων δεν είχε δοθεί ιδιαίτερη βαρύτητα. Έτσι για να αντιμετωπίσει αυτό το πρόβλημα η βιομηχανία άρχισε να αναζητά τρόπους λύσης με αυτοματοποιημένες διαλογικές στρατηγικές. Αναπτύχθηκαν φωνητικές πύλες που βασίζονται σε μια τυποποιημένη αρχιτεκτονική που ενσωματώνει ένα προσωδιακό και σημασιολογικό μοντέλο ανίχνευσης συναισθημάτων.

Η ανάπτυξη μιας εφαρμογής η οποία θα μπορεί να αναγνωρίζει συναισθήματα απαιτεί τα παρακάτω βήματα:

1. Πρέπει να προσδιοριστεί η λειτουργικότητα της εφαρμογής
2. Με βάση τις προδιαγραφές, οι συναισθηματικές καταστάσεις και το πώς αυτές χρησιμοποιούνται θα πρέπει να προσδιορίζονται
3. Επίσης θα πρέπει να περιγράφονται τα συναισθήματα
4. Για να εκπαιδευτούν οι αλγόριθμοι ανίχνευσης κατάλληλα απαιτείται να συλλεχτεί ένα σύνολο δεδομένων δοκιμών είτε από ένα πρωτότυπο, από μια μελέτη ή από μια συγκρίσιμη εφαρμογή
5. Τα δεδομένα θα πρέπει να επισημαίνονται από ανθρώπινους ακροατές ούτως ώστε να αποκτήσουν μια βάση αλήθειας
6. Η αναγνώριση συναισθημάτων μπορεί στη συνέχεια να ενσωματωθεί στην εφαρμογή αλλά θα πρέπει να είναι συντονισμένη και να ανανεώνεται-εκπαιδεύεται όσο το δυνατόν συχνότερα προκειμένου να προσαρμοστεί στο πραγματικό κόσμο.

Παρακάτω αναλύονται οι διαφορετικές χρήσεις της συναισθηματικής επεξεργασίας στα συστήματα ηλεκτρονικών υπολογιστών.

- a) Broadcast: Στις τηλεπικοινωνίες θα ήταν επιθυμητό να παρέχεται ένα ειδικό κανάλι για συναισθηματική επικοινωνία. Ένα δημοφιλές παράδειγμα είναι τα λεγόμενα «emotions » που χρησιμοποιούνται στην επικοινωνία μέσω chat.
- b) Αναγνώριση: Η έκφραση των ανθρώπινων συναισθημάτων μπορεί να αναλυθεί με διάφορους τρόπους, και αυτό μπορεί να χρησιμοποιηθεί για να μεταβάλλει την αντίδραση του συστήματος.
- c) Προσομοίωση: Οι Συναισθηματικές εκφράσεις μπορούν να μιμηθούν από το σύστημα προκειμένου να ενισχυθεί η φυσική επικοινωνία.
- d) Μοντελοποίηση: Τα εσωτερικά υποδείγματα των συναισθηματικών αναπαραστάσεων μπορεί να χρησιμοποιηθούν για να αναπαρασταθεί η κατάσταση του χρήστη ή του συστήματος ή για να χρησιμοποιηθούν ως πρότυπο στην τεχνητή νοημοσύνη.



Εικόνα 33 Διαφορετικές χρήσεις της συναισθηματικής επεξεργασίας σε συστήματα ηλεκτρονικών υπολογιστών

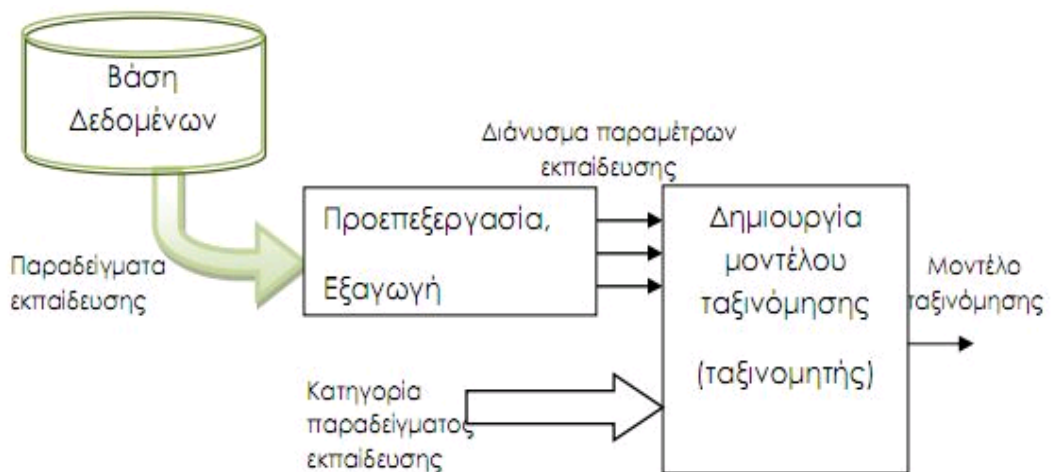
#### 4.1 Διαδικασίες ενός συστήματος αναγνώρισης συναισθημάτων

Για να επιτευχθεί η κατασκευή ενός συστήματος αναγνώρισης συναισθημάτων απαιτούνται δύο διαδικασίες η διαδικασία της εκπαίδευσης και διαδικασία της κανονικής λειτουργίας ταξινόμησης.

Κατά τη διαδικασία εκπαίδευσης το σύστημα αναγνώρισης συναισθημάτων πρέπει να εκπαιδευτεί. Για να γίνει αυτό ένας αρκετά μεγάλος αριθμός αντιπροσωπευτικών παραδειγμάτων τα οποία υπάρχουν αποθηκευμένα σε μια βάση δεδομένων, δίνεται ως είσοδος στο σύστημα αναγνώρισης. Επίσης μαζί με

αυτά τα παραδείγματα δίνεται και η κατηγορία στην οποία ανήκει το κάθε παράδειγμα. Στη συνέχεια τα παραδείγματα αυτά επεξεργάζονται και δημιουργείται ένα γενικευμένο μοντέλο ταξινόμησης σύμφωνα με τον επιλεγμένο αλγόριθμο ταξινόμησης, και την επιθυμητή έξοδο (κατηγορία παραδείγματος) . Στην εικόνα που ακολουθεί φαίνονται τα βήματα που ακολουθούνται κατά τη διαδικασία της εκπαίδευσης.

Αρχικά γίνεται η επιλογή μια κατάλληλης βάσης δεδομένων, τα παραδείγματα της οποίας να είναι ταξινομημένα από πριν, ώστε να γνωρίζουμε την κατηγορία στην οποία ανήκει το καθένα, μετά γίνεται προεπεξεργασία του σήματος εισόδου και εξαγωγή των κατάλληλων παραμέτρων, που είναι ικανές να ξεχωρίσουν τα δεδομένα που μας ενδιαφέρουν. Τέλος γίνεται επιλογή ενός κατάλληλου αλγόριθμου εκπαίδευσης – ταξινόμησης, και δημιουργία του μοντέλου ταξινόμησης, σύμφωνα με το οποίο θα μπορούν να ταξινομηθούν και άγνωστα στο σύστημα παραδείγματα.



Εικόνα 34 διαδικασία εκπαίδευσης στο αυτόματο σύστημα αναγνώρισης συναισθημάτων

Στη συνέχεια ακολουθεί η διαδικασία κανονικής λειτουργίας ταξινόμησης όπου ταξινομούνται οι εισοδοί που μας ενδιαφέρουν. Πρώτα επεξεργάζονται τα σήματα εισόδου και εξάγονται οι επιλεγμένοι παράμετροι (Feature extraction), και μετά ταξινομούνται οι εισοδοί σύμφωνα με τον επιλεγμένο αλγόριθμο αναγνώρισης προτύπων (Classification).

Η επεξεργασία σήματος περιλαμβάνει την ψηφιοποίηση και την ακουστική προεπεξεργασία, όπως το φιλτράρισμα, καθώς και την κατάτμηση του σήματος εισόδου σε σημαντικές μονάδες. Ο υπολογισμός των επιλεγμένων παραμέτρων ασχολείται με τον εντοπισμό ιδιαίτερων χαρακτηριστικών που παρουσιάζει το ακουστικό σήμα όσον αφορά τα συναισθήματα και τέλος ο αλγόριθμος ταξινόμησης. Ταξινομεί τα συναισθήματα σύμφωνα με τη μάθηση από τα παραδείγματα που υπάρχουν στη βάση δεδομένων



Εικόνα 35 Διαδικασία κανονικής λειτουργίας στο αυτόματο σύστημα αναγνώρισης συναισθημάτων

## 4.2 Βάσεις Δεδομένων συναισθηματικής ομιλίας

Οι βάσεις δεδομένων με συναισθηματική ομιλία δεν είναι μόνο ουσιαστικής σημασίας για ψυχολογικές μελέτες, αλλά και για την αυτόματη αναγνώριση συναισθημάτων. Κατά την έναρξη της έρευνας για την αυτόματη φωνητική αναγνώριση συναισθημάτων, η οποία ξεκίνησε σοβαρά στα μέσα της δεκαετίας του '90, χρησιμοποιούνταν ο προσπονητός λόγος ενώ στρέφεται τώρα προς πιο ρεαλιστικά στοιχεία

Οι συναισθηματικές βάσεις δεδομένων δεν είναι επαρκείς γι' αυτό και οι ερευνητές αναγκάζονται κάθε φορά να δημιουργούν από την αρχή μια νέα συναισθηματική βάση ανάλογα με την έρευνά τους. Για να δημιουργηθεί μια βάση δεδομένων η οποία θα χρησιμεύει στην αναγνώριση ομιλίας, χρησιμοποιούνται τρία είδη λόγου.

- Ο προσπονητός λόγος όπου επαγγελματίες ηθοποιοί εκφωνούν προτάσεις και κείμενα με διαφορετικά συναισθήματα κάθε φορά, ο συγκεκριμένος λόγος είναι αυτός που χρησιμοποιείται πιο συχνά, η φυσικότητα όμως των συναισθημάτων πολλές φορές αμφισβητείται. Συνήθως οι ομιλητές λένε τις ίδιες φράσεις με διαφορετικά συναισθήματα κάθε φορά ούτως ώστε να είναι

πιο εύκολη η σύγκριση μεταξύ των συναισθημάτων. Οι βάσεις δεδομένων της προσπονητής ομιλίας δεν αντικατοπτρίζουν τον πραγματικό αυθόρμητο διάλογο και δε μπορούμε να πούμε ότι τα αποτελέσματα είναι αποδοτικά σε καθημερινές εφαρμογές. Χαρακτηριστικά παραδείγματα βάσεων δεδομένων προσπονητού λόγου είναι η βάση δεδομένων συναισθηματικής ομιλίας του Βερολίνου (Berlin database of emotional Speech) και η Danish Emotional Speech corpus (DES).

- Ο φυσικός ή αυθόρμητος λόγος που προκαλείται από φυσικές καταστάσεις. Ο αυθόρμητος λόγος περιέχει τις πιο αυθεντικές συναισθηματικές καταστάσεις αλλά είναι πολύ δύσκολο να συλληχθεί καθώς οι ομιλητές δε θα πρέπει να γνωρίζουν ότι καταγράφονται ,γι' αυτό και πολλοί ερευνητές χρησιμοποιούν υλικό από τα μέσα επικοινωνίας όπως προγράμματα και συνεντεύξεις από το ραδιόφωνο και τη τηλεόραση. Ένα μειονέκτημα αυτού του είδους ομιλίας είναι η δυσκολία στη ταξινόμηση των συναισθημάτων.
- Ο προκλητός λόγος κατά τον οποίο τα συναισθήματα παράγονται με φυσικό τρόπο μέσα από διάφορες διαδικασίες. Τέτοιες διαδικασίες είναι οι εξής : φαντασία και ύπνωση του ομιλητή, δίνοντας υλικό που μπορεί να προκαλέσει συναισθήματα στον ομιλητή όπως μουσική, ταινίες, φωτογραφίες, χρησιμοποιώντας τεχνικά διεγερτικά φάρμακα έτσι ώστε να έρθει ο ομιλητής σε συγκεκριμένη ψυχολογική κατάσταση, καθοδήγηση του ομιλητή σε μία συναισθηματική κατάσταση με παροτρύνσεις προτάσεις, αρνητική και θετική αξιολόγηση.

Η απόδοση ενός ταξινομητή συναισθημάτων στηρίζεται σε μεγάλο βαθμό στη ποιότητα της βάσης δεδομένων που χρησιμοποιείται. Τα συναισθήματα τα οποία έχουν καταγραφεί και εξετασθεί στις περισσότερες βάσεις δεδομένων είναι τα ίδια και είναι τα εξής : θυμός ,χαρά, λύπη, έκπληξη, άγχος και φόβος.

Είναι γνωστό ότι οι άνθρωποι χρησιμοποιούν συγκεκριμένες λέξεις και φράσεις για να εκφράσουν τα συναισθήματα τους γι' αυτό και στις βάσεις δεδομένων σε κάθε αρχείο ήχου αντιστοιχούν και κάποιες φράσεις. Οι πιο συνηθισμένες κατηγορίες φράσεων είναι οι ακόλουθες:

- Εκφράσεις αποδοχής όπως ναι, ευχαριστώ, σωστά, μάλιστα, βεβαίως, ακριβώς, καλά, εντάξει.
- Εκφράσεις άρνησης – απόρριψης όπως όχι, δεν θέλω, λάθος δεν το καταλαβαίνω
- Εκφράσεις επανάληψης όπως επαναλαμβάνετε παρακαλώ, συγγνώμη δεν σας άκουσα
- Εκφράσεις δισταγμού όπως εεε, χμμ, ουφ, πώς

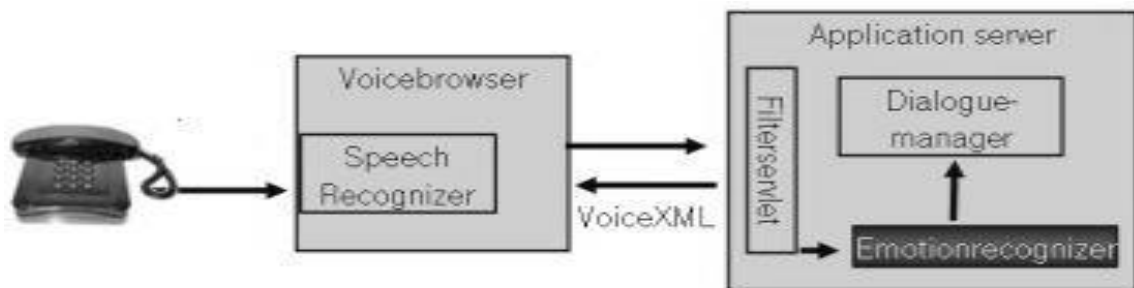
Πέραν των άλλων στις βάσεις δεδομένων συναισθηματικής ομιλίας λαμβάνονται υπόψη και ένα σύνολο μη εκμεταλλεύσιμων δεδομένων τα οποία δεν περιέχουν συναίσθημα, σε αυτά τα δεδομένα εντάσσεται ο θόρυβος που υπάρχει στα αρχεία ήχου που είναι αποθηκευμένα. Ως θόρυβος ορίζεται κάθε ήχος που προέρχεται από τον περιβάλλοντα χώρο όπως τηλεόραση, ραδιόφωνο, θόρυβο αυτοκινήτου κλπ, γάβγισμα σκύλου, από τη χρήση του εξοπλισμού όπως πάτημα ενός πλήκτρου του τηλεφώνου ή κλείσιμο του ακουστικού και από τον ίδιο χρήστη-πελάτη γέλιο, βήχας, φτέρνισμα.

Παρακάτω αναφέρονται κάποια παραδείγματα βάσεων δεδομένων ηχητικών καταγραφών που προέρχονται από κέντρα εξυπηρέτησης πελατών.

- Η βάση How may I help you? Η οποία δημιουργήθηκε από τον Gorin το μέγεθος της οποίας είναι 10 KB. Κάθε φράση έχει επισημειωθεί ορθογραφικά, έχει σημειωθεί το τέλος της και η ποιότητα του ήχου. Το μήκος μίας φράσης μπορεί να είναι από μια λέξη μέχρι ένα σύνολο 183 λέξεων, ενώ ο μέσος όρος είναι 8 λέξεις και ο ρυθμός ομιλίας αντιστοιχεί σε 3 λέξεις ανά δευτερόλεπτο.
- Η βάση DARPA στην οποία οι χρήστες-ομιλητές κάλεσαν συστήματα τα οποία ήταν διαθέσιμα σε διάφορες ηλεκτρονικές διευθύνσεις για να κάνουν κρατήσεις αεροπορικών εισιτηρίων. Οι χρήστες γνώριζαν εξ αρχής ότι οι πτήσεις δε θα πραγματοποιούνταν γι' αυτό και η δυσαρέσκεια και η απογοήτευση τους ήταν χαμηλότερη από το κανονικό. Συλλέχθηκαν 287 διάλογοι και 21899 φράσεις.

- Η βάση SYMPAFLY στην οποία χρησιμοποιήθηκε ένα πλήρως αυτόματο σύστημα διαλόγου για τη κράτηση πτήσεων μέσω τηλεφώνου. Στη συγκεκριμένη βάση χρησιμοποιήθηκε αυθόρμητη ομιλία και περιλαμβάνει 3 υποσύνολα. Στο πρώτο τις κρατήσεις τις έκαναν τα άτομα που ανέπτυξαν το σύστημα. Στο δεύτερο τις έκαναν εθελοντές που δεν είχαν καμία ανάμειξη με την ανάπτυξη του συστήματος και στο τρίτο τις κρατήσεις τις έκαναν εθελοντές που τους επέβλεπαν άτομα που ανέπτυξαν το σύστημα.

### 4.3 Αρχιτεκτονική



Εικόνα 36 Ένταξη της συναισθηματικής επεξεργασίας στην αρχιτεκτονική των φωνητικών πυλών

Στην αρχιτεκτονική των φωνητικών πυλών το μοντέλο ανίχνευσης των συναισθημάτων έχει ενσωματωθεί στη πλευρά του διακομιστή αν και θα ήταν πιο αποτελεσματικό να βρίσκεται μαζί με τη τεχνολογία αναγνώρισης ομιλίας καθώς οι δύο αυτές τεχνολογίες μοιράζονται πολλά βήματα επεξεργασίας. Μέσω ενός φίλτρου “FilterServlet” όλες οι αιτήσεις φιλτράρονται από τον φωνητικό browser κι έτσι το διαλογικό παράθυρο μπορεί να ανταποκρίνεται προς τον χρήστη χωρίς καμία χρονική καθυστέρηση που προκλήθηκε από την ανίχνευση του θυμού. Το μοντέλο αναγνώρισης συναισθημάτων αποτελείται από ένα αντικείμενο, το οποίο τροφοδοτείται από συναισθηματικές αξιολογήσεις από το νήμα ανάλυσης.

Αυτές οι αξιολογήσεις υπολογίζονται από δύο πηγές:

- Την ανάλυση των ηχητικών χαρακτηριστικών και
- Την ανάλυση του σημασιολογικού περιεχομένου.

Η σημασιολογική ανάλυση δεν έχει αναπτυχθεί αρκετά μέχρι σήμερα και συνίσταται απλώς από λέξεις κλειδιά. Η ακουστική ανάλυση γίνεται από ένα στατιστικό ταξινομητή κατά την οποία αναλύονται προσωδιακά χαρακτηριστικά.

Ο ακουστικός ταξινομητής λειτουργεί ως εξής: από την έκφραση του χρήστη ένα σύνολο από προσωδιακά χαρακτηριστικά υπολογίζεται και αυτό το διάνυσμα χαρακτηριστικών κατατάσσεται σε μία, δύο ή τρεις κατηγορίες, αντίστοιχα γίνεται η χρήση ενός αλγορίθμου ο οποίος βασίζεται στις Gaussian πυκνότητες. Σε αυτό το βήμα επεξεργασίας μόνο το σήμα ομιλίας χρησιμοποιείται ως εισροή. Για την ταξινόμηση χρησιμοποιείται ένα σύνολο μέχρι 31 προσωδιακών χαρακτηριστικών. Οι φωνητικές παράμετροι οι οποίες έχουν ερευνηθεί καλύτερα από ψυχολογικές μελέτες σε σχέση με το συναίσθημα είναι η προσωδία (τόνος, ένταση, ρυθμός ομιλίας) και η ποιότητα της φωνής.

Στον παρακάτω πίνακα φαίνεται πως μεταβάλλονται τα χαρακτηριστικά της προσωδίας και η ποιότητα της φωνής σε τέσσερα κύρια συναισθήματα, στο θυμό, τη χαρά τη λύπη και τον φόβο.

	<b>Τόνος</b>	<b>Ένταση</b>	<b>Ρυθμός ομιλίας</b>	<b>Ποιότητα φωνής</b>
<b>Θυμός</b>	υψηλός	αυξάνεται	αυξάνεται	Τρεμάμενη φωνή
<b>Χαρά</b>	αυξάνεται	αυξάνεται	αυξάνεται	Λίγο τρεμάμενη φωνή
<b>Λύπη</b>	Ίδιος ή μειώνεται	μειώνεται	μειώνεται	ψιθυριστή
<b>Φόβος</b>	αυξάνεται	κανονική	αυξάνεται	τρεμάμενη

Εικόνα 37 Μεταβολή χαρακτηριστικών προσωδίας και ποιότητας φωνής σε σχέση με τα συναισθήματα



Για παράδειγμα όταν ένας χρήστης λυπάται τότε η φωνή του γίνεται πιο αργή και πιο ψιθυριστή καθώς και ο τόνος και η ένταση μειώνονται. Ενώ όταν ένας χρήστης βιώνει θυμό ή χαρά τότε μιλά πιο γρήγορα και πιο δυνατά.

#### 4.4 Στρατηγικές συνδιαλλαγής

Στη καθημερινή επικοινωνία τα αρνητικά συναισθήματα προσδιορίζονται με τρεις βασικές λειτουργίες:

- Ενημερώνοντας τον συνομιλητή σχετικά με τη συναισθηματική μας κατάσταση προκειμένου να μπορέσει να κατανοήσει πλήρως τις πληροφορίες που θέλουμε να εκφράσουμε
- Ενημερώνοντας τον συνομιλητή σχετικά με την αντίληψη που έχουμε για τη ποιότητα της σχέσης που υπάρχει μεταξύ μας (για παράδειγμα δυσπιστία)
- Προκαλώντας μια συγκεκριμένη ενέργεια ή συμπεριφορά την οποία θέλουμε να παρατηρήσει ο συνομιλητής (πχ οπισθοχώρηση)

Στις καθημερινές καταστάσεις επικοινωνίας υπάρχει πιθανότητα οι συνομιλητές να αντιδράσουν με θετικές ή αρνητικές συνέπειες στην έκφραση των αρνητικών συναισθημάτων. Μια γρήγορη λύση της αρνητικής συναισθηματικής κατάστασης μπορεί να θεωρηθεί ως ένα κριτήριο για την επιτυχή και επαρκή αντίδραση. Στα λειτουργικά και επαγγελματικά πλαίσια επικοινωνίας όπως για παράδειγμα στα τηλεφωνικά κέντρα η εμφάνιση των αρνητικών συναισθημάτων (ειδικά του θυμού) θεωρούνται ως επί το πλείστον ανεπιθύμητα. Τα ισχυρά αρνητικά συναισθήματα μπορεί να αποσπάσουν τη προσοχή από το πραγματικό θέμα της επικοινωνίας με συνέπεια να οδηγήσει σε απώλεια της αποτελεσματικότητας. Επίσης είναι πιθανόν να οδηγήσουν σε μία δυνατή συναισθηματική πίεση. Για τους δύο αυτούς λόγους οι στρατηγικές συνδιαλλαγής χρησιμοποιούνται κυρίως από επαγγελματίες προκειμένου να αντιμετωπίσουν μια συναισθηματικά φορτισμένη κατάσταση γρήγορα και να επιστρέψουν αμέσως στο κύριο θέμα της επικοινωνίας. Η κύρια πρόθεση ωστόσο δεν είναι απλώς να προσπεράσει τα αρνητικά συναισθήματα, αλλά η ικανοποίηση των πελατών ώστε να έχουν την εντύπωση ότι τα αρνητικά τους συναισθήματα λαμβάνονται σοβαρά υπόψη.

Παρακάτω παραθέτονται παραδείγματα αποτελεσματικών στρατηγικών συνδιαλλαγής για αρνητικά συναισθήματα.

- Διάσπαση της προσοχής / αλλαγή του θέματος χωρίς δείχνουν ενδιαφέρον για τα αρνητικά συναισθήματα
- Παροχή πληροφοριών οι οποίες αντικρούουν τις πιθανώς λανθασμένες εκτιμήσεις που οδήγησαν σε μια αρνητική συναισθηματική κατάσταση.
- Εμφατικές δηλώσεις όπου φαίνεται η συμπάθεια για τα συναισθήματα των ομιλητών
- Περαιτέρω ενθάρρυνση του ομιλητή για να εκφράσει την συναισθηματική κατάσταση του.
- Κάνει χιούμορ και αστειεύεται για να αντιμετωπίσει τα αρνητικά συναισθήματα

Μεταφέροντας τις στρατηγικές συνδιαλλαγές της ανθρώπινης επικοινωνίας στις διαλογικές συμπεριφορές των φωνητικών πυλών αντιμετωπίζουμε δύο βασικούς περιορισμούς.

- Αν και χρησιμοποιείται η τελευταία λέξη της τεχνολογίας για την ανίχνευση των συναισθημάτων δυστυχώς αυτό που δε μπορεί να ανιχνευτεί είναι η αιτία που αυτά προκαλούνται. Για να είναι αξιόπιστα, μπορούν να μεταφερθούν στα διαλογικά συστήματα ανθρώπου-μηχανής μόνο αυτές οι στρατηγικές που είναι γενικού χαρακτήρα έτσι ώστε να μη χρειαστεί να γίνουν αναφορές στην αιτία των αρνητικών συναισθημάτων.
- Η επικοινωνία μεταξύ ανθρώπου και μηχανής ,όταν εμφανιστούν τα αρνητικά συναισθήματα, γίνεται περιπλοκή καθώς ο πελάτης-ομιλητής αναφέρεται σε άσχετα θέματα (πχ ο λόγος για τον οποίο ο πελάτης είναι θυμωμένος)

#### **4.5 Αναγνώριση συναισθημάτων στις φωνητικές πύλες**

Η κακή διαχείριση ενός τηλεφωνικού κέντρου μπορεί οδηγήσει στη δυσαρέσκεια πελατών, σε ζημιές την επιχείρηση και πολλές φορές στη μετακίνηση των πελατών

προς άλλες επιχειρήσεις - ανταγωνιστές. Έτσι η ικανοποίηση των πελατών μέσω της αποτελεσματικής εκπαίδευσης των αντιπροσώπων αποτελεί πρωτεύων ζήτημα στις φωνητικές πύλες. Θα πρέπει να παρέχεται μια αξιολόγηση πραγματικού χρόνου ώστε να ενισχυθεί η σωστή διαχείριση της κλήσης με τον πελάτη και η καλύτερη εξυπηρέτηση του ενώ βρίσκεται σε διάλογο. Τα πρόωρα προειδοποιητικά σημάδια της απογοήτευσης πελατών μπορούν να ανιχνευθούν από τον τόνο της φωνής, τις απότομες και μη ολοκληρωμένες προτάσεις και τις αλλαγές στον τρόπο ομιλίας.

Αν και συναισθηματική ομιλία υπήρξε το επίκεντρο της έρευνας για πολλά χρόνια, τα αποτελέσματα συχνά δεν μπορούν να εφαρμοστούν σε τηλεφωνικές υπηρεσίες για διάφορους λόγους:

- Το σήμα ομιλίας είναι χαμηλού εύρους ζώνης, και διαταράσσεται από το θόρυβο του περιβάλλοντος, έτσι η εξαγωγή χαρακτηριστικών είναι επιρρεπής σε λάθη.
- Το διαλογικό παράθυρο είναι συνήθως σύντομο, και συχνά αποτελείται από μόνο ένα πολύ περιορισμένο σύνολο λέξεων εντολών.
- Οι άνθρωποι δεν χρειάζεται να ακολουθήσουν τους κανόνες ευγένειας που ισχύουν στην ανθρώπινη επικοινωνία, και συνήθως χρησιμοποιούν ένα εχθρικό και αυταρχικό τόνο στη φωνή τους.
- Οι άνθρωποι έχουν την τάση μιλάνε αργά και δυνατά ακόμα και με τρόπο ρομποτικό, πιστεύοντας πως με αυτό τον τρόπο θα διευκολύνουν την αυτόματη αναγνώριση ομιλίας.
- Σε πολλές εφαρμογές οι πελάτες καλούν έχοντας κάποιο είδος καταγγελίας στο μυαλό τους και την τάση να μιλήσουν με αρκετά αρνητικό τόνο, ανεξαρτήτως των προβλημάτων που ενδέχεται να προκύψουν από την αλληλεπίδραση.

Η ανίχνευση του θυμού είναι ένα θέμα το οποίο κερδίζει όλο και περισσότερη προσοχή στις φωνητικές πύλες και μπορεί να συμβάλει στην ενίσχυση της αλληλεπίδραση ανθρώπου υπολογιστή. Αρκετές φορές οι χρήστες των φωνητικών

πυλών είναι απογοητευμένοι που μιλάνε σε μηχανές και δεν έχουν τη δυνατότητα να εκφράσουν το θυμό τους. Για να λυθεί αυτό το πρόβλημα έχουν ενσωματωθεί στις φωνητικές πύλες ολοκληρωμένα συστήματα αναγνώρισης συναισθημάτων. Όταν για παράδειγμα το σύστημα αναγνώρισης συναισθημάτων ανιχνεύσει θυμό στη φωνή του ομιλητή τότε το διαλογικό σύστημα αρχικά προσπαθεί να αντιμετωπίσει τη διάθεση του με κατάλληλες διαλογικές στρατηγικές και εν συνεχεία μεταφέρει τον καλούντα στο τηλεφωνικό κέντρο .

Στην πραγματικότητα, υπάρχουν τρεις εφαρμογές οι οποίες σχετίζονται με την ανίχνευση του θυμού στις φωνητικές πύλες.

- Στέλνουν τον καλούντα να επικοινωνήσει με εκπαιδευμένους υπάλληλους οι οποίοι θα μπορέσουν να αντιμετωπίσουν κατάλληλα το πρόβλημα
- Προσαρμόζουν το διαλογικό σύστημα με κατάλληλες στρατηγικές οι οποίες γίνονται αυτόματα.
- Συλλέγουν στατιστικά στοιχεία, προκειμένου να συγκεντρώσουν πληροφορίες για το πόσο ευχαριστημένοι είναι οι πελάτες κατά την πάροδο του χρόνου.

Στην εξυπηρέτηση πελατών των φωνητικών πυλών είναι πολύ χρήσιμη η ανίχνευση ενδεχόμενων προβλημάτων που προκύπτουν από μια μη ικανοποιητική πορεία της αλληλεπίδρασης, προκειμένου να βοηθήσει τον πελάτη είτε προσφέροντας βοήθεια μέσω ενός ανθρώπινου φορέα είτε προσπαθώντας να αντιδράσει με κατάλληλες διαλογικές στρατηγικές. Ένα σύνολο απαιτήσεων που πρέπει να ικανοποιηθούν για να θεωρηθεί επιτυχής η επικοινωνία στις φωνητικές πύλες χωρίς προβλήματα είναι οι εξής :

- Το μοντέλο ανίχνευσης θυμού πρέπει να ενσωματωθεί στην υφιστάμενη αρχιτεκτονική.
- Η καθυστέρηση που προκλήθηκε από την επεξεργασία δεν πρέπει να εμποδίζει την ροή του διαλόγου.
- Η ταξινόμηση πρέπει να βασίζεται αποκλειστικά στα στοιχεία που έχει αποκτηθεί αυτόματα.

- Ο αλγόριθμος πρέπει να δουλέψει σε μικρές μονολεκτικές εντολές και σε κακή κατάσταση ήχου.

Οι περισσότεροι αλγόριθμοι ταξινόμησης για την ανίχνευση του θυμού βασίζονται σε μια προσέγγιση τριών σταδίων :Πρώτον, μια σειρά ακουστικών και προσωδιακών χαρακτηριστικών υπολογίζεται από το εισερχόμενο σήμα ομιλίας . Δεύτερον διαφορετικοί αλγόριθμοι αναγνώρισης όπως οι Gaussian Mixture Models (GMMs), Support Vector Machines (SVMs) αποφασίζουν αν στον τρέχον διάλογο υπάρχει ή όχι θυμός. Τέλος τεχνολογίες μεταεπεξεργασίας μπορούν να χρησιμοποιηθούν για συνδυασμό των αποτελεσμάτων των ταξινομητών.

#### **4.6 Εφαρμογές τεχνολογίας αναγνώρισης συναισθημάτων**

Η αναγνώριση των συναισθημάτων στην ανθρώπινη ομιλία έχει αποκτήσει ολοένα και μεγαλύτερη προσοχή τα τελευταία χρόνια λόγω της μεγάλης ποικιλίας των εφαρμογών που επωφελούνται από την τεχνολογία αυτή. Παρόλο που τα ανθρώπινα συναισθήματα είναι πολύ δύσκολο να χαρακτηριστούν, η έρευνα για μηχανές οι οποίες θα μπορούν να κατανοούν τα ανθρώπινα συναισθήματα έχει προχωρήσει με ταχείς ρυθμούς. Οι λύσεις για την αναγνώριση συναισθημάτων εξαρτώνται από το ποια συναισθήματα θέλουμε να αναγνωρίζει μια μηχανή και για ποιο σκοπό.

Η τεχνολογία αναγνώρισης συναισθημάτων έχει εφαρμογές σε ομιλούντα και ηλεκτρονικά παιχνίδια , σε τηλεφωνικά κέντρα και φωνητικές πύλες, στη ψυχολογία καθώς και σε ευφυή συστήματα διδασκαλίας. Για παράδειγμα τα Ρομπωτικά κατοικίδια ζώα θα πρέπει να είναι σε θέση να κατανοήσουν όχι μόνο τις προφορικές εντολές, αλλά και άλλες πληροφορίες, όπως η συναισθηματική και η κατάσταση της υγείας του ανθρώπου που επικοινωνούν και να μεταβάλουν τις ενέργειές τους αναλόγως. Στα αυτόματα τηλεφωνικά κέντρα η ένταξη της αναγνώρισης συναισθημάτων κάνει την επικοινωνία χρήστη-υπολογιστή πιο φυσική και αν παρουσιαστεί σοβαρό πρόβλημα τότε ο χρήστης οδηγείται σε ειδικευμένο αντιπρόσωπο. Επίσης στα ευφυή συστήματα διδασκαλίας ο εντοπισμός των συναισθημάτων των μαθητών είναι παρά πολύ σημαντικός και μειώνει το χάσμα μεταξύ πραγματικού δασκάλου και δασκάλου-υπολογιστή.

Μελέτες της εκπαιδευτικής ψυχολογίας επισημαίνουν ότι τα συναισθήματα των μαθητών μπορούν να επηρεάσουν την απόδοσή τους και τη μάθηση. Ακόμη στο τομέα της ψυχολογίας η τεχνολογία αναγνώρισης συναισθημάτων θα μπορούσε να βοηθήσει τους γιατρούς να κατανοήσουν βαθύτερα τη φύση των συναισθημάτων, και να τους βοηθήσει να διαγνώσουν πιο εύκολα την ασθένεια που έχει ο ασθενής όπως της κατάθλιψης.

Τέλος μια σημαντική χρήση της τεχνολογίας αναγνώρισης συναισθημάτων θα ήταν να υπήρχε σύστημα στα αυτοκίνητα το οποίο θα αναγνώριζε την άσχημη συναισθηματική κατάσταση του οδηγού και θα προσπαθούσε είτε να τον ηρεμήσει με κατάλληλα μηνύματα είτε να μειώνει τη ταχύτητα του αυτοκινήτου, καθώς έχει αποδειχθεί ότι η επίδοση του οδηγού στην οδήγηση εξαρτάται σημαντικά από τη συναισθηματική του κατάσταση ,έτσι θα είχαν αποφευχθεί πολλά αυτοκινητιστικά ατυχήματα που οφείλονται στην ασταθή συναισθηματική κατάσταση του οδηγού.

Συνοψίζοντας, μπορούμε να πούμε ότι υπάρχει ένα μεγάλο πλήθος εφαρμογών στα συστήματα που χρησιμοποιούν την τεχνολογία αναγνώρισης συναισθημάτων Παρ' όλα αυτά, η εκτεταμένη χρήση των συστημάτων αυτών εξακολουθεί να παρεμποδίζεται από μεγάλα τεχνικά ζητήματα που πρέπει να επιλυθούν, ειδικά στην απευθείας αναγνώριση συναισθημάτων η οποία απαιτείται στη πραγματικού χρόνου επικοινωνία ανθρώπου με υπολογιστή

## **Επίλογος**

Η ενσωμάτωση της αυτόματης αναγνώρισης συναισθημάτων στην ανθρώπου-υπολογιστή επικοινωνία παρουσιάζει μεγάλες προκλήσεις δεδομένου ότι τα συναισθήματα πρέπει να αναγνωρίζονται σε πραγματικό χρόνο, ενώ ο χρήστης αλληλεπιδρά με μια εφαρμογή. Η συναισθηματική πληροφορία που μεταδίδεται μέσω της ανθρώπινης ομιλίας αποτελεί σημαντικό παράγοντα στην ανθρώπινη επικοινωνία .Στην επικοινωνία ανθρώπου και μηχανής υπήρχε ένα κενό μέχρι την ανάπτυξη της τεχνολογίας αναγνώρισης συναισθημάτων, γι' αυτό και η ένταξη στις φωνητικές πύλες ήταν σημαντική. Οι απογοητευμένοι πελάτες συνεχώς αυξάνονταν επειδή ένα μηχάνημα δε μπορούσε να κατανοήσει άμεσα τις αιτήσεις του χρήστη, κι επειδή δεν υπήρχε τρόπος για να εκφράσει το θυμό του ο

καλλών αισθανόταν όλο και πιο πολύ δυσφορία. Με τη συγκεκριμένη τεχνολογία η αλληλεπίδραση πελάτη και φωνητικής πύλης έγινε πιο φυσική και η δυσαρέσκεια των χρηστών ελαττώθηκε.

## 5. Πλεονεκτήματα φωνητικών πυλών

Για τον χρήστη :

- Λειτουργία σε βάση καθημερινή βάση
- Τα τηλέφωνα μπορούν να χρησιμοποιηθούν οπουδήποτε
- Συντριπτική ανάπτυξη των κινητών τηλεφώνων
- Μηδενική αναμονή χρηστών
- Εξάλειψη ανθρώπινου λάθους
- Συνέπεια και συνέχεια στην εξυπηρέτηση
- Χρήση προσαρμοσμένη στις ικανότητες των χρηστών και στο φυσικό τους περιβάλλον.
- Φιλικότητα στη χρήση, ακόμα και από άτομα χωρίς καθόλου εξοικείωση.
- Μεγαλύτερη δυνατή ασφάλεια σε ευαίσθητα δεδομένα και συναλλακτικές διαδικασίες μέσω τηλεφώνου (πχ αριθμοί πιστωτικών καρτών), καθώς δεν παρεμβάλλεται τρίτος μεταξύ του χρήστη και του φορέα των συναλλαγών.
- Μείωση των άσκοπων μετακινήσεων
- Ευκολότερη πρόσβαση στο διαδίκτυο σε σχέση με τους υπολογιστές

Για τους Πελάτες :

- Μείωση λειτουργικού κόστους
- Αναβάθμιση υπηρεσιών με τη χρήση προηγμένης τεχνολογίας
- Υπεροχή στον ανταγωνισμό
- Καλύτερη κατανομή εργασίας
- Αύξηση παραγωγικότητας
- Αύξηση κερδών



## 6. Υπηρεσίες Φωνητικών Πυλών

### Εισαγωγή

Οι υπηρεσίες που προσφέρουν οι φωνητικές πύλες είναι τριών κατηγοριών, οι υπηρεσίες πληροφόρησης όπου χρήστες ενημερώνονται για θέματα τα οποία τους ενδιαφέρουν, οι υπηρεσίες voice messaging όπου οι εταιρίες πληροφορούν τους χρήστες μαζικά για διάφορα γεγονότα και οι υπηρεσίες διεκπεραίωσης συναλλαγών.

### 6.1 Υπηρεσίες πληροφόρησης

Οι χρήστες χρησιμοποιούν τις συγκεκριμένες φωνητικές πύλες για να ενημερωθούν σχετικά με:

- Πρόγραμμα τηλεόρασης, κινηματογράφων, θεάτρων
- Διευθύνσεις, τηλέφωνα επικοινωνίας και ώρες λειτουργίας δημόσιων οργανισμών, τραπεζών κ.λπ.
- Διανυκτερεύοντα φαρμακεία και νοσοκομεία
- Τυχερά παιχνίδια και λαχεία (π.χ. αριθμοί που κληρώθηκαν)
- Αποτελέσματα ποδοσφαιρικών αγώνων, ιπποδρομιών κ.ο.κ.
- Μετεωρολογικές προβλέψεις
- Νέα της επικαιρότητας
- Χρηματιστηριακά νέα
- Αστρολογικές προβλέψεις
- Δρομολόγια (αεροπορικά, ακτοπλοϊκά, σιδηροδρομικά)
- Πρόσβαση στο προσωπικό ηλεκτρονικό ταχυδρομείο και δυνατότητα ανάγνωσης και αποστολής email
- Προσωπική ατζέντα (ραντεβού, υπενθυμίσεις συναντήσεων ή σημαντικών ημερών κ.λπ.)

### 6.2 Υπηρεσίες voice messaging

Όπου οι εταιρίες και οι δημόσιοι οργανισμοί μπορούν να ενημερώσουν το χρήστη για διάφορα γεγονότα. Έτσι, υπηρεσίες αυτόματης τηλεφωνικής ενημέρωσης,

μέσω φωνητικών μηνυμάτων ή τεχνολογίας σύνθεσης φωνής (TTS- Text To Speech), ενημερώνουν πολίτες, μέλη συλλόγων, πελάτες εταιρειών κλπ. , στην τηλεφωνική συσκευή τους στην ανάπτυξη μιας καμπάνιας. Για παράδειγμα, με τη χρήση μιας τέτοιας υπηρεσίας είναι δυνατό ένας Δήμος να ενημερώσει άμεσα και αυτόματα τους δημότες του για τις πολιτιστικές εκδηλώσεις του Δήμου ή για την διεξαγωγή μιας συνέλευσης.

### **6.3 Υπηρεσίες διεκπεραίωσης συναλλαγών.**

Οι πολίτες χρησιμοποιούν τις συγκεκριμένες φωνητικές πύλες για να διεκπεραιώσουν κάποιες συναλλαγές. Για να πραγματοποιήσει μια τέτοια συναλλαγή θα πρέπει να γίνει η ταυτοποίηση του χρήστη είτε με κάποιο κωδικό είτε με τον αριθμό της ταυτότητας του κ.λπ. είτε αυτόματα με τη βοήθεια φυσικών προσώπων.

Στη κατηγορία αυτή ανήκουν οι συναλλαγές με Δημόσιες Υπηρεσίες – εξυπηρέτηση πολίτη (αυτοματοποίηση δικαιολογητικών και πιστοποιητικών, ενημέρωση τιμολογίων), υπηρεσίες υγείας, υπηρεσίες αγοράς προϊόντων και φυσικά οι τραπεζικές συναλλαγές.

Το προϊόν φωνητικής πύλης «Γραμμή για την Υγεία -1535» εντάσσεται σ' αυτή τη κατηγορία .

### **Επίλογος**

Είναι φανερό λοιπόν το πόσο σημαντικές και αναγκαίες είναι οι φωνητικές πύλες από το πλήθος υπηρεσιών που προσφέρουν. Στο κεφάλαιο που ακολουθεί αναλύονται κάποιες φωνητικές πύλες που προσφέρουν και τα τρία είδη υπηρεσιών που μόλις αναφέρθηκαν .

## 7. Προϊόντα αγοράς

### Εισαγωγή

Στο παρόν κεφάλαιο παρουσιάζονται κάποια προϊόντα αγοράς – πρακτικές εφαρμογές φωνητικών πυλών ελληνικά και ξένα. Οι πληροφορίες που παρέχονται από αυτές τις πύλες μπορούν να προέρχονται από αρχεία κειμένου, βάσεις δεδομένων και από το διαδίκτυο. Αναφέρονται η φωνητική πύλη του Πανεπιστημίου Κρήτης μέσω της οποίας εξυπηρετούνται τα μέλη της ακαδημαϊκής κοινότητας και στα οποία παρέχεται ένα σύνολο πληροφοριών του πανεπιστημίου ,η φωνητική πύλη του ΚΤΕΟ Θεσσαλονίκης όπου οι πελάτες μπορούν να κλείσουν ραντεβού, η Φωνητική Πύλη 14789 που παρέχει πληροφορίες σχετικά με τα δρομολόγια των πλοίων καθώς και η φωνητική πύλη της εθνικής τράπεζας μέσω της οποίας οι χρήστες-πελάτες μπορούν να διαχειριστούν ένα πλήθος τραπεζικών συνδιαλλαγών.

#### 7.1 1-800-555-TELL



Η εταιρία Tellme Network παρουσίασε τη φωνητική πύλη 1-800-555-TELL (1-800-555-8355). Είναι η πρώτη φωνητική πύλη η οποία παρέχει δωρεάν πρόσβαση στους χρήστες σε υπηρεσίες πληροφόρησης και διασκέδασης.

Υπηρεσίες που προσφέρονται:

- Παρέχεται πρόσβαση και πληροφόρηση σε ένα πλήθος επιχειρήσεων
- Οι καλούντες μπορούν να βρουν και να συνδεθούν με μια τοπική υπηρεσία ταξί οπουδήποτε στις Ηνωμένες Πολιτείες
- Η υπηρεσία συνδέει άμεσα τους χρήστες με μια αεροπορική εταιρεία απλά λέγοντας το όνομά της
- Οι χρήστες της συγκεκριμένης φωνητικής πύλης μπορούν να μάθουν ποιες παραστάσεις και ταινίες παίζονται σε κάθε θέατρο και κινηματογράφο

αντίστοιχα και επίσης να δουν τις ώρες και τα σχόλια της παράστασης ή της ταινίας.

- Οι χρήστες μπορούν να ελέγχουν ένα εξατομικευμένο χαρτοφυλάκιο και να παρακολουθήσουν τα νέα της εταιρείας
- Η tellme παρέχει στους χρήστες τις πρώτες ειδήσεις του CNN, και ειδήσεις σε θέματα της τεχνολογίας, του καιρού, της υγείας και της διασκέδασης
- Παρέχεται άμεση πληροφόρηση για την εξέλιξη των αθλητικών αποτελεσμάτων
- Επίσης οι χρήστες μπορούν να μάθουν τα νέα για τα ζώδια τους
- Οι Tellme καλούντες μπορούν να παίξουν blackjack
- Οι χρήστες έχουν τη δυνατότητα να κλείσουν κάποιο ξενοδοχείο και να ενοικιάσουν αυτοκίνητο
- Τέλος οι χρήστες μπορούν να μάθουν πως μπορούν να φτάσουν σε μια συγκεκριμένη κατεύθυνση οδηγώντας

## 7.2 Φωνητική Πύλη του Πανεπιστημίου Κρήτης

Το Πανεπιστήμιο Κρήτης σε συνεργασία με την Cardisoft εγκατέστησε την πρώτη Φωνητική Πύλη εξυπηρέτησης φοιτητών η οποία με τη χρήση της τεχνολογίας αναγνώρισης και σύνθεσης ομιλίας εξυπηρετεί αυτόματα τα μέλη της ακαδημαϊκής κοινότητας. Είναι μία εφαρμογή που επιτρέπει στους χρήστες να ανακτούν, αποθηκεύουν και να διαχειρίζονται πληροφορίες μέσω του τηλεφώνου όλο το 24ωρο και 7 ημέρες την εβδομάδα. Η επικοινωνία μεταξύ του φοιτητή και του ηλεκτρονικού υπολογιστή γίνεται αυτόματα μέσω του φυσικού λόγου από την πλευρά του χρήστη και μέσω σύνθεσης ομιλίας από κείμενο ή τοιχογραφημένων φωνητικών αρχείων από τον ηλεκτρονικό υπολογιστή. Οι φοιτητές μπορούν μέσω μιας τηλεφωνικής συσκευής να καλούν στα τηλέφωνα 2810-394444 ή 28310-77777 του ιδρύματος και να συνδέονται στο αυτοπονημένο σύστημα υπολογιστή όπου τους παρέχονται όλες οι υπηρεσίες της γραμματείας. Σκοπός του συστήματος είναι η μείωση του χρόνου αναμονής του φοιτητή για την εξυπηρέτηση του και η παράλληλη αποσυμφόρηση των τηλεφωνικού κέντρου του πανεπιστημίου και των γραμματειών.

Ο φοιτητής για να μπορέσει να εισέλθει στο σύστημα πρέπει να πει τον αριθμό μητρώου και τον προσωπικό του αριθμό PIN.

Το σύστημα αποτελείται από δύο υποσυστήματα:

- Υποσύστημα Αυτόματης Φωνητικής Απόκρισης (Interactive Voice Response, IVR) Βασίζεται στην πλατφόρμα Cisco IP Interactive Response της εταιρείας Cisco και έχει σχεδιασθεί σύμφωνα με το μοντέλο Cisco AVVID.
- Υποσύστημα αναγνώρισης και σύνθεσης φωνής (ASR - TTS) Η αναγνώριση φωνής βασίζεται στο Nuance OpenSpeech Recognizer 3.x. Το σύστημα έχει την δυνατότητα αναγνώρισης και κατανόησης της ομιλίας διαχειριζόμενο σύνθετες συναλλαγές προκειμένου να εξασφαλιστεί έξυπνος και φιλικός διάλογος του φοιτητή με το σύστημα. Η δε σύνθεση φωνής βασίζεται στο Nuance Realspeak Telecom 4.x προκειμένου να αντιμετωπιστεί ο μεγάλος όγκος της πληροφορίας, μαθήματα όλων των σχολών του Π.Κ τα οποία είναι σημαντικό να τονιστεί ότι μεταβάλλονται συχνά και θα ήταν δύσκολη η προσπάθειά τους μέσω προηχογραφημένων μηνυμάτων.

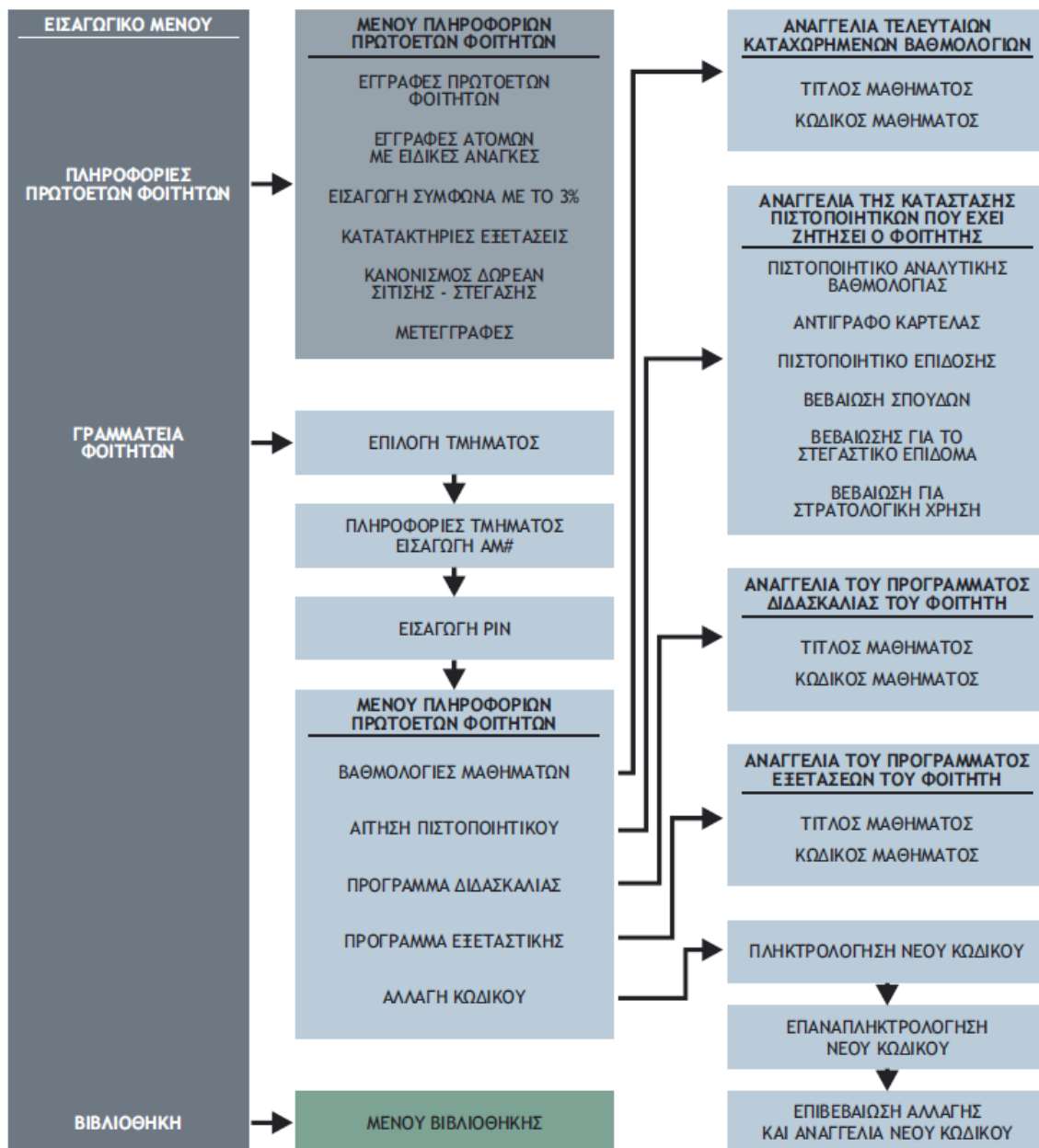
Θα πρέπει να τονιστεί ότι το σύστημα λειτουργεί συγχρόνως και με DTMF (πληκτρολόγηση τηλεφώνου), δυνατότητα που χρησιμοποιήθηκε για την εξυπηρέτηση των ΑΜΕΑ. Η διαδικτυακή πύλη του Πανεπιστημίου Κρήτης λειτουργεί με μεγάλη επιτυχία εδώ και δύο χρόνια -με την τεχνολογική εξέλιξη του συστήματος να συνεχίζεται και σήμερα- κάτι που φαίνεται από τον αριθμό των φοιτητών που τη χρησιμοποιούν, ο οποίος αυξάνεται συνεχώς.

Μερικές από τις υπηρεσίες που προσφέρονται στο φοιτητή είναι οι εξής:

- Πρόσβαση σε γενικές πληροφορίες (για το ίδρυμα, για εγγραφές πρωτοετών, κλπ)
- Δρομολόγηση κλήσεων μέσω αναγνώρισης ομιλίας στο τμήμα επιλογής του χρήστη.
- Πλήρη και δυναμική ενημέρωση για βαθμούς, προγράμματα διδασκαλίας και εξετάσεων, δηλώσεις μαθημάτων

- Αίτηση έκδοσης πιστοποιητικών, αναλυτικής βαθμολογίας, στρατολογίας, κλπ τα οποία οι χρήστες λαμβάνουν μέσω email ή fax
- Διακοπή της εκφώνησης από τον χρήστη για την ταχύτερη εξέλιξη του διαλόγου
- Υπηρεσίες για μεταπτυχιακούς φοιτητές
- Πληροφορίες για πρωτοετείς φοιτητές
- Πληροφορίες για τη βιβλιοθήκη

Η αρχιτεκτονική του συστήματος της φωνητικής πύλης του πανεπιστημίου Κρήτης απεικονίζεται στο ακόλουθο σχήμα



Εικόνα 38 Φωνητική Πύλη του Πανεπιστημίου Κρήτης

### 7.3 Φωνητική Πύλη ΚΤΕΟ

Η εταιρία Cardisoft συνεργασία με τη νομαρχία Θεσσαλονίκης υλοποίησε το 2007 τη Φωνητική πύλη εξυπηρέτησης πελατών ΚΤΕΟ Θεσσαλονίκης .Η Φωνητική Πύλη των ΚΤΕΟ, χρησιμοποιεί όλες τις σύγχρονες τεχνολογίες ανοιχτών πρωτοκόλλων για την επικοινωνία με το σύστημα διαχείρισης ραντεβού των ΚΤΕΟ μέσω XML, Web Services , SOAP, ενώ για τον διάλογο χρησιμοποιείται πρωτόκολλο VoiceXML και γεννήτριες αναγνώρισης ομιλίας. Στις καινοτομίες της Φωνητικής Πύλης, συγκαταλέγονται:

- Ο αλγόριθμος που χρησιμοποιείται για την αναγνώριση του αριθμού πινακίδας των οχημάτων, για την κατανόηση της μικρής ηχητικής πληροφορίας που δίνεται από τον πολίτη (π.χ. XYZ1234)
- Η ενσωμάτωση όλων των κανονισμών για τις κατηγορίες οχημάτων (επιβατικά Ι.Χ, ταξί, φορτηγά, ταξί, λεωφορεία, δίκυκλα, ανάρθημα κτλ.)
- Η δυνατότητα πρότασης ημερομηνίας, ώρας και τόπου (ΚΤΕΟ Α' , Β' ή Γ') από τον ίδιο τον πολίτη
- Οι δυνατότητες αλλαγής ή και ακύρωσης των ραντεβού μέσω αναφοράς στο ticket number που του δίνει το σύστημα.
- Ενημέρωση του πολίτη για την χρέωση των εκπρόθεσμων οχημάτων
- Οδηγίες για τον τρόπο προσέλευσης του πολίτη στα ΚΤΕΟ.
- Η συνεργασία με το υφιστάμενο Call Center των ΚΤΕΟ μέσω τεχνολογίας C.T.I. (Computer Telephony Integration).

Η Φωνητική Πύλη ενσωματώθηκε στην υπάρχουσα τηλεπικοινωνιακή υποδομή των ΚΤΕΟ και λειτουργεί σε βάση (24x7) 24ώρες το 24ωρο, 7 ημέρες την εβδομάδα και αποσκοπεί στο να κλείνει περισσότερα από 500 ραντεβού ημερησίως, ενώ ιδιαίτερα σημαντικό είναι να τονιστεί ότι λαμβάνει υπόψη όλους τους περιορισμούς και τις πολιτικές του συστήματος διαχείρισης ραντεβού.

### 7.4 MyCosmos

Η εταιρία Διάλογος υλοποίησε την πρώτη φωνητική πύλη στην Ελλάδα για λογαριασμό της Cosmote. Η Cosmote προσέφερε υπηρεσίες πληροφόρησης και διασκέδασης στους συνδρομητές της, οι οποίοι καλώντας 14000 από κινητό

Cosmote μπορούν να χρησιμοποιήσουν τη φωνητική πύλη MyCosmos. Οι υπηρεσίες της φωνητικής πύλης καλύπτουν τους παρακάτω τομείς: χρηματιστήριο, ταξίδια (ακτοπλοϊκά και αεροπορικά δρομολόγια), διασκέδαση (πρόγραμμα κινηματογράφων, θεάτρων και τηλεόρασης), αστρολογικές προβλέψεις, παιχνίδια-ψυχαγωγία (παιχνίδι "31", τεστ προσωπικότητας, ερωτική χαρτομαντεία, quiz γνώσεων). Η φωνητική πύλη **εμπλουτίζεται διαρκώς** με νέες υπηρεσίες φωνητικής αναγνώρισης και σύνθεσης φωνής από τη Διάλογος. Ο σχεδιασμός της φωνητικής πύλης επιτρέπει στο χρήστη να αποκτά με τον πιο φυσικό και γρήγορο τρόπο την πληροφορία που τον ενδιαφέρει, αποκλειστικά μέσω της φωνής του. Χωρίς αναμονή σε τηλεφωνικά κέντρα εξυπηρέτησης και πληροφόρησης πελατών, από οποιοδήποτε μέρος και οποιαδήποτε στιγμή της ημέρας οι συνδρομητές της Cosmote έχουν τη δυνατότητα ενημέρωσης και διασκέδασης εξ αποστάσεως μέσω φωνητικών εντολών.

### 7.5 Φωνητική Πύλη Δρομολογίων Πλοίων 14789



Η φωνητική πύλη για τις ακτοπλοϊκές συγκοινωνίες εξυπηρετεί εύκολα, άμεσα και έγκυρα όσους ταξιδεύουν με πλοίο αλλά και επαγγελματίες, όπως ταξιδιωτικούς πράκτορες, ξενοδόχους. Η υπηρεσία λειτουργεί με την τεχνολογία αναγνώρισης φωνής (speech recognition), ώστε να παρέχει στους χρήστες εύκολα, γρήγορα και αξιόπιστα την επιθυμητή πληροφόρηση. Ο χρήστης καλεί από το σταθερό, ή το κινητό του τηλέφωνο στο 14789 και κατευθύνεται με διάλογο ώστε να λάβει άμεσα την ενημέρωση που θέλει, χωρίς πολύπλοκα μενού και χρήση πλήκτρων στο καντράν της τηλεφωνικής του συσκευής.

Το 14789 προσφέρει:



- Πληροφορίες σε πραγματικό χρόνο (real time) για αφίξεις πλοίων.
- Πληροφορίες σε πραγματικό χρόνο (real time) για αναχωρήσεις πλοίων.
- Ενημέρωση σε πραγματικό χρόνο (real time) για το σημείο που βρίσκεται το πλοίο που ενδιαφέρει τον χρήστη.
- Πληροφορίες για προγραμματισμένα δρομολόγια, ώστε ο χρήστης να κανονίσει την ημέρα και ώρα της αναχώρησης του.

## 7.6 INFO TRIP

Η υπηρεσία παρέχει τη δυνατότητα στον καλούντα να ενημερώνεται τηλεφωνικά για την κατάσταση της κυκλοφορίας σε οδικά τμήματα της επιλογής του, καθώς επίσης και για εκδηλώσεις σε επιλεγμένες από το χρήστη περιοχές, μέσω αυτοματοποιημένης υπηρεσίας τηλεφωνικής κλήσης. Οι αριθμοί κλήσης με αστική χρέωση είναι **210-7458450** και **210-7458452**. Όλες οι υπηρεσίες IVR της INFO TRIP είναι διαθέσιμες στα Ελληνικά και στα Αγγλικά. Η επιλογή της γλώσσας γίνεται με την πληκτρολόγηση του αντίστοιχου αριθμού στην έναρξη της τηλεφωνικής σύνδεσης.

Από το κεντρικό μενού, ο χρήστης επιλέγει την υπηρεσία για την οποία θέλει να λάβει πληροφορίες. Οι παρεχόμενες υπηρεσίες παρουσιάζονται παρακάτω:

- Κυκλοφοριακά συμβάντα στην Αθήνα. Ο χρήστης επιλέγει τη περιοχή και ενημερώνεται για το κυκλοφοριακό της.
- Εκδηλώσεις στη Αθήνα Παροχή πληροφοριών για τις εκδηλώσεις στην Αθήνα, με επιλογή υποπεριοχής (κατά τον ίδιο τρόπο όπως στην υπηρεσία κυκλοφοριακών συμβάντων).
- Σημεία ενδιαφέροντος στην Αθήνα (Μετακίνηση, Δημόσια Υπηρεσία ,Πολιτισμός ,Διασκέδαση ,Κοινωνική Πρόνοια ,Διατροφή ,Διαμονή)
- Εύρεση ταξιδιού με μέσα μαζικής μεταφοράς

## 7.7 Phone banking Εθνικής Τράπεζας



Εικόνα 39 Phone Banking

Το Phone Banking της ΕΘΝΙΚΗΣ βασίζεται σε μία από τις αρτιότερες διεθνώς τεχνικές υποδομές, και περιλαμβάνει σύγχρονα λειτουργικά χαρακτηριστικά που ξεπερνούν κατά πολύ τις δυνατότητες των περισσότερων Ευρωπαϊκών τραπεζών. Ο πελάτης εξυπηρετείται μέσα από τον ενιαίο αριθμό πρόσβασης 181818, είτε από σταθερό τηλέφωνο είτε από κινητό. Το Phone Banking λειτουργεί ως ενιαίο σημείο σύνδεσης με όλη την Τράπεζα, προσφέροντας 144 διαφορετικές τραπεζικές συναλλαγές, πληροφορίες, αιτήσεις για νέα προϊόντα και υποστήριξη χρηστών. Επίσης έχει προστεθεί ολοκληρωμένο σύστημα διαχείρισης παραπόνων που υποδέχεται σχόλια και προτάσεις πελατών. Οι επιλογές του πελάτη περιλαμβάνουν αυτόματο σύστημα εξυπηρέτησης με φωνητική αναγνώριση (Interactive Voice Recognition), προσωπική συνομιλία με εξειδικευμένο στέλεχος της Τράπεζας, τηλεφωνικό ραντεβού (voice call-back) και επιβεβαίωση συναλλαγής μέσω φαξ, e-mail ή μηνύματος σε κινητό. Η ευελιξία του συστήματος επιτρέπει τόσο στους πελάτες που χρησιμοποιούν το Phone Banking περιστασιακά και συμπληρωματικά με το Internet Banking, όσο και στους πελάτες που προτιμούν την προσωπική εξυπηρέτηση, να έχουν στη διάθεσή τους ένα εύχρηστο και ασφαλές περιβάλλον με ιδιαίτερα χαμηλούς χρόνους αναμονής που ανταποκρίνονται στις διεθνείς βέλτιστες πρακτικές.

Μερικές από τις ενέργειες που γίνονται με το Phone Banking είναι οι εξής:

- Δήλωση απώλειας της κάρτας σας (χρεωστικής ή πιστωτικής)

- Πληροφορίες για τις πιστωτικές σας κάρτες: υπόλοιπο, κίνηση, τρόποι εξόφλησης κ.λπ.
- Ενημέρωση για προϊόντα / υπηρεσίες, σημεία εξυπηρέτησης (Καταστήματα / ATMs), τιμές συναλλάγματος κ.ά. Η ενημέρωση αυτή παρέχεται ακόμα και αν δεν είστε πελάτης της Τράπεζας.
- Συναλλαγές μέσω Phone Banking: Για χρήστες της υπηρεσίας, μέσω του αυτόματου συστήματος φωνητικών εντολών και μέσω εκπροσώπου.
- Σύνδεση με Εκπρόσωπό μας.

Το τηλεφωνικό κέντρο είναι στη διάθεσή 24 ώρες την ημέρα, 365 μέρες τον χρόνο. Το μόνο που έχετε να κάνετε είναι να καλέσετε στο 181818 από σταθερό ή κινητό τηλέφωνο, αν βρίσκεστε στην Ελλάδα, ή στο +30 210 4848484 αν βρίσκεστε στο εξωτερικό: Μια φιλική φωνή θα σας καθοδηγήσει στο μενού υπηρεσιών και θα περιμένει οδηγίες, προτείνοντάς σας κάθε φορά την κατάλληλη λέξη –κλειδί. Οποιαδήποτε στιγμή μπορείτε να διακόψετε και να ζητήσετε να συνδεθείτε με κάποιον εκπρόσωπο, λέγοντας τη λέξη «Εκπρόσωπος».

#### *7.7.1 Υπηρεσία Phone Banking*

Εάν είστε χρήστης της υπηρεσίας Phone Banking, έχετε 181818 λόγους για να μας καλέσετε:

- Θέλετε να ελέγξετε το υπόλοιπο του λογαριασμού σας, για να δείτε αν σας κατέθεσαν χρήματα οι γονείς σας;
- Θέλετε να πληροφορηθείτε για το υπόλοιπο και την κίνηση της πιστωτικής σας κάρτας;
- Θέλετε να πληρώσετε την πιστωτική σας κάρτα, στην Εθνική ή σε άλλη τράπεζα;
- Θέλετε να εξοφλήσετε λογαριασμό ΟΤΕ, ΔΕΗ, Internet ή κινητής τηλεφωνίας;
- Θέλετε να καταθέσετε εξέταστρα για ξένη γλώσσα σε λογαριασμό άλλης τράπεζας;
- Θέλουν οι γονείς σας να πληρώσουν το φόρο εισοδήματος αλλά δεν έχουν χρόνο;

- Μόλις κλείσατε ξενοδοχείο για διακοπές και θέλετε να στείλετε προκαταβολή σε λογαριασμό στην Εθνική ή σε άλλη Τράπεζα;
- Θέλετε να πληροφορηθείτε τις τιμές συναλλάγματος για να προμηθευτείτε το ποσό που χρειάζεστε από το Κατάστημα;

Όλα αυτά, και ακόμα περισσότερα, γίνονται από το Phone Banking απλά, γρήγορα και χωρίς κόπο. Επιπλέον, οι πληροφορίες που λαμβάνετε μέσω Phone Banking (υπόλοιπα και κινήσεις λογαριασμών, καρτών και δανείων), αλλά και οι αποδείξεις των συναλλαγών σας, αποστέλλονται με fax, e-mail και sms.

Με την εγγραφή σας στο Phone Banking αποκτάτε παράλληλα πρόσβαση και στο Internet Banking και απολαμβάνετε τις υπηρεσίες ηλεκτρονικής εξυπηρέτησης i-bank της Εθνικής Τράπεζας, 24 ώρες το 24ωρο, επιλέγοντας το μέσον που σας εξυπηρετεί κάθε φορά: το Internet, το κινητό σας ή οποιοδήποτε

#### 7.7.2 Ασφάλεια

Οι υπηρεσίες ηλεκτρονικής εξυπηρέτησης i-bank της Εθνικής καινοτομούν και στον τομέα της ασφάλειας, καθώς οι συναλλαγές μέσω Phone Banking διασφαλίζονται με την πλέον σύγχρονη τεχνολογία:

- Η ταυτοποίηση του πελάτη για πληροφόρηση σχετικά με τους λογαριασμούς του ή διενέργεια συναλλαγών Phone Banking πραγματοποιείται με τον Κωδικό Εισόδου (User ID) και ένα κωδικό μίας χρήσης, που παράγεται ηλεκτρονικά από τη συσκευή i-code. Κατά την εγγραφή στην υπηρεσία ο πελάτης λαμβάνει ένα μοναδικό, προσωπικό Κωδικό Εισόδου (User ID) και τη συσκευή παραγωγής ηλεκτρονικών κλειδαριθμών (i-code). Η συσκευή αυτή αντιστοιχίζεται μοναδικά με τον κάτοχό της, οπότε τα κεντρικά μας συστήματα αναγνωρίζουν ότι κάθε κωδικός μίας χρήσης προέρχεται από συγκεκριμένο πελάτη. Έτσι, ακόμα και σε περίπτωση κλοπής ή απώλειας της συσκευής, αυτή δεν μπορεί να χρησιμοποιηθεί από οποιονδήποτε τρίτο. Επιπλέον, η ισχύς κάθε κωδικού είναι μόλις 32", με την παρέλευση των οποίων ακυρώνεται και δεν μπορεί να χρησιμοποιηθεί. Το γεγονός αυτό καθιστά αδύνατη την υποκλοπή από τρίτους, για μεταγενέστερη χρήση.

- Η διασφάλιση των εγχρημάτων συναλλαγών Phone Banking πραγματοποιείται με χρήση ενός επιπλέον κωδικού i-code. Μετά τη διεκπεραίωση της συναλλαγής το σύστημα «επιστρέφει» στον πελάτη ένα τριψήφιο Κωδικό Επιβεβαίωσης, ο οποίος πρέπει να είναι ίδιος με τον τριψήφιο Κωδικό που παράγεται, στη συνέχεια, από τη συσκευή i-code. Η «επιστροφή» του τριψήφιου κωδικού γίνεται είτε από το αυτόματο φωνητικό σύστημα είτε από τον Εκπρόσωπο, ανάλογα με τον τρόπο συναλλαγής που έχει επιλέξει ο πελάτης. Πληροφορηθείτε αναλυτικά για τη συσκευή i-code [εδώ](#).
- Κλείδωμα κωδικών. Σε περίπτωση που ο πελάτης εισάγει τέσσερις φορές λάθος το UserID ή τον κωδικό i-code, τότε αυτόματα απαγορεύεται η πρόσβασή του στο σύστημα. Για να επιτραπεί, εκ νέου, η χρήση των υπηρεσιών Phone Banking θα πρέπει να συνδεθεί με έναν από τους εκπροσώπους της Τράπεζας και να ακολουθήσει τις οδηγίες του.
- Καταγραφή και μαγνητοφώνηση κλήσεων. Όλες οι κλήσεις που πραγματοποιούνται προς το τηλεφωνικό κέντρο καταγράφονται και μαγνητοφωνούνται από το σύστημα.

### 7.7.3 Πλεονεκτήματα

- Ταχύτητα  
Με τη χρήση φωνητικών εντολών, ο χρόνος που χρειάζεστε πλέον για πληροφόρηση και τραπεζικές συναλλαγές μειώνεται στο ελάχιστο.
- Ευελιξία  
Τώρα μπορείτε να διεκπεραιώσετε τις τραπεζικές σας εργασίες με τον πιο εύκολο τρόπο, από οποιαδήποτε μέρος του κόσμου και όποια ώρα εξυπηρετεί εσάς.
- Ασφάλεια  
Η πρόσβαση στα δεδομένα των λογαριασμών σας και σε συναλλαγές διασφαλίζεται με Κωδικό Εισόδου (UserID), καθώς και με κωδικούς μιας χρήσης που παράγονται από τη συσκευή [i-code](#).
- Προνόμια  
Για εσάς τους σπουδαστές και φοιτητές (ανεξαρτήτως ηλικίας!) η δυνατότητα χρήσης του Phone Banking, καθώς και η απόκτηση του i-code

παρέχεται δωρεάν.  
Επιπλέον, με την εγγραφή σας στην υπηρεσία του Phone Banking αποκτάτε αμέσως πρόσβαση και στο Internet/Mobile Banking.

- Οικονομία  
Οι περισσότερες συναλλαγές μέσω Phone Banking παρέχονται δωρεάν, ενώ για δικαιούχους λογαριασμού «ΣΠΟΥΔΑΖΩ» δεν παρακρατείται προμήθεια υπέρ της Τράπεζας για αποστολή εμβασμάτων σε τράπεζες του εσωτερικού και της ΕΕ. Για περισσότερες πληροφορίες πιέστε εδώ.
- Ενημέρωση  
Για τους λογαριασμούς σας (καταθετικούς, δανειακούς και επενδυτικούς) και όχι μόνο, τη στιγμή ακριβώς που τη θέλετε, με δυνατότητα αποστολής sms, e-mail και fax.

#### 7.7.4 Οδηγός Χρήσης

Η χρήση της υπηρεσίας είναι εξαιρετικά απλή. Αυτό συμβαίνει γιατί το ίδιο το σύστημα προτείνει κάθε φορά την κατάλληλη λέξη «κλειδί» για την επιλογή που επιθυμείτε. Ακόμα, όμως, και αν νιώσετε ότι χρειάζεστε κάποια βοήθεια ή διευκρίνιση δεν έχετε παρά να εκφωνήσετε τη λέξη «εκπρόσωπος» και ο εκπρόσωπος της Τράπεζάς θα σας λύσει οποιαδήποτε απορία.

Αν δεν είστε χρήστης Phone Banking: Μπορείτε να χρησιμοποιήσετε τις υπηρεσίες του κέντρου τηλεφωνικής εξυπηρέτησης της Τράπεζας για δήλωση απώλειας κάρτας Εθνικής Τράπεζας, καθώς και πληροφόρηση για προϊόντα και υπηρεσίες, σημεία εξυπηρέτησης (Καταστήματα/ATMs), τιμές συναλλάγματος κ.λπ. Η πληροφόρηση αυτή παρέχεται είτε με φωνητικές εντολές και απαντήσεις (αυτόματα ηχογραφημένα μηνύματα) είτε από εκπρόσωπο. Επιπλέον, μπορείτε να πληροφορηθείτε, επιλέγοντας από το αρχικό μενού «ΠΛΗΡΟΦΟΡΙΕΣ», σχετικά με τη πιστωτική σας κάρτα: τρέχουσα οφειλή και όριο, διαθέσιμο υπόλοιπο, ποσό μελλοντικών δόσεων, τελευταίες κινήσεις της κάρτας, τρόπους εξόφλησης, αλλαγή διεύθυνσης κατοικίας ή επικοινωνίας σας κ.λπ.

Εάν είστε χρήστης Phone Banking: Μπορείτε να λαμβάνετε πληροφόρηση για τους λογαριασμούς σας (καταθετικούς και επενδυτικούς), τα δάνεια και τις πιστωτικές σας κάρτες, και να διενεργείτε συναλλαγές αυτόματα (μέσω

του συστήματος αναγνώρισης φωνητικών εντολών -IVR) ή με τη βοήθεια εκπροσώπου. Δεν έχετε παρά να επιλέξετε «ΣΥΝΑΛΛΑΓΕΣ», ακούγοντας το αρχικό μενού, και να ακολουθήσετε, βήμα προς βήμα, τις οδηγίες που ακολουθούν:

- Για την είσοδό σας στην υπηρεσία Phone Banking απαιτείται ο 6ψήφιος αριθμητικός Κωδικός Εισόδου (User ID), που είναι ο ίδιος με τον Κωδικό Εισόδου του Internet Banking και ένας κωδικός μιας χρήσης που παράγεται από τη συσκευή i-code. Εάν έχετε αλφαριθμητικό Κωδικό Εισόδου (User ID), παρέχεται η δυνατότητα άμεσης παραλαβής του νέου, αριθμητικού κωδικού, που θα σας επιτρέψει την παράλληλη πρόσβαση στο Phone Banking, μέσω του menu «Αιτήσεις» του Internet Banking.
- Για τη διενέργεια εγχρήματων συναλλαγών και συναλλαγών ασφαλείας απαιτείται η εισαγωγή ενός ακόμη κωδικού i-code.

#### *7.7.5 Συναλλαγές με φωνητικές εντολές!*

Είναι πραγματικά απίστευτα εύκολη, γρήγορη και απόλυτα ασφαλής η πραγματοποίηση συναλλαγών μέσω Phone Banking, αφού το μόνο που χρειάζεται είναι ο 6ψήφιος αριθμητικός Κωδικός Εισόδου (UserID), η συσκευή i-code και ένα τηλέφωνο!

Παρακάτω μπορεί κάποιος να δει πόσο απλά, γρήγορα και με απόλυτη ασφάλεια, πραγματοποιούνται ορισμένες από τις διαθέσιμες συναλλαγές στο Phone Banking μέσω του αυτόματου συστήματος αναγνώρισης φωνητικών εντολών (IVR):

Στην παρακάτω εικόνα φαίνεται ένα παράδειγμα για το πως μπορεί ένας πελάτης να πληροφορηθεί σχετικά με το υπόλοιπο ενός συνδεδεμένου λογαριασμού καταθέσεων και πώς θα μπορέσει να επιλέξει να του αποσταλεί η σχετική ενημέρωση με fax, e-mail ή sms

Ενώ στη εικόνα 41 στο δεύτερο παράδειγμα φαίνεται πως ο πελάτης μπορεί να μεταφέρει ένα ποσό μεταξύ δύο συνδεδεμένων λογαριασμών και πώς θα λάβει απόδειξη της συναλλαγής με fax ή e-mail.

## Ενημέρωση για το υπόλοιπο καταθετικού λογαριασμού σας και αποστολή του με sms στο κινητό σας τηλέφωνο

Καλείτε: **181818** (από την Ελλάδα)  
**+30 210 4848484** (από το εξωτερικό)

### Ακούτε

το εισαγωγικό μήνυμα και το Αρχικό Μενού

### Πείτε

«Συναλλαγές»

### Ακούτε

το αυτόματο σύστημα αναγνώρισης φωνητικών εντολών, το οποίο σας ζητά να πείτε ή να πληκτρολογήσετε τον 6-ψήφιο Κωδικό Χρήστη (User ID) και έναν ηλεκτρονικό κλειδάριθμο από τη συσκευή i-code. Αν τα στοιχεία αυτά είναι λάθος, τα επαναλαμβάνει καθοδηγώντας σας να τα επανεισάγετε. Αν είναι σωστά, η διαδικασία συνεχίζεται.

### Πείτε

«Πληροφορίες Λογαριασμών»

### Ακούτε

το Μενού «Πληροφορίες Λογαριασμών»

### Πείτε

«Υπόλοιπα Λογαριασμών»

### Ακούτε

τους συνδεδεμένους λογαριασμούς σας

### Πείτε

«Επιλογή», κατά τη διάρκεια της εκφώνησης του λογαριασμού που επιθυμείτε. Εναλλακτικά πείτε ή πληκτρολογήστε τον αριθμό αυτό.

### Ακούτε

τον αριθμό και το είδος του λογαριασμού που επιλέξατε και στη συνέχεια το διαθέσιμο και το λογιστικό του υπόλοιπο.

### Ακούτε

«Αν θέλετε να σας αποσταλούν οι πληροφορίες που ακούσατε μέσω SMS, fax ή e-mail, πείτε sms, fax ή e-mail»

### Πείτε

«SMS»

### Ακούτε

«Παρακαλώ πείτε ή πληκτρολογήστε τον αριθμό κινητού τηλεφώνου σας».

Αφού πείτε ή πληκτρολογήσετε τον αριθμό του κινητού σας τηλεφώνου,

### Ακούτε

τον αριθμό του κινητού σας τηλεφώνου

### Πείτε

«Σωστά»

Αμέσως λαμβάνετε sms με το υπόλοιπό σας και η συναλλαγή σας έχει ολοκληρωθεί!

Για να επιστρέψετε στο Αρχικό Μενού επιλογών:

### Πείτε

«Αρχικό Μενού»

Για να τερματίσετε την κλήση:

### Πείτε

«Έξοδος»

Εικόνα 40 Παράδειγμα Υπολοίπου



## Πώς θα μεταφέρετε ποσό σε λογαριασμό σας στην ΕΤΕ και θα λάβετε απόδειξη της συναλλαγής με e-mail

**Καλείτε: 181818 (από την Ελλάδα)  
+30 210 4848484 (από το εξωτερικό)**

<b>Ακούτε</b> το εισαγωγικό μήνυμα και το Αρχικό Μενού	<b>Πείτε</b> «Συναλλαγές»
<b>Ακούτε</b> το αυτόματο σύστημα αναγνώρισης φωνητικών εντολών, το οποίο σας ζητά να πείτε ή να πληκτρολογήσετε τον 6-ψήφιο Κωδικό Χρήστη (User ID) και έναν ηλεκτρονικό κλειδαριθμό από τη συσκευή i-code. Αν τα στοιχεία αυτά είναι λάθος, τα επαναλαμβάνει καθοδηγώντας σας να τα επανεισάγετε. Αν είναι σωστά, η διαδικασία συνεχίζεται.	
<b>Ακούτε</b> το Μενού Συναλλαγών	<b>Πείτε</b> «Μεταφορά χρηματικών ποσών»
<b>Ακούτε</b> Το Μενού «Μεταφοράς χρημάτων»	<b>Πείτε</b> «Λογαριασμό ΕΤΕ ίδιου»
<b>Ακούτε</b> Τους συνδεδεμένους λογαριασμούς σας, για να επιλέξετε το λογαριασμό που θα χρεωθεί.	<b>Πείτε</b> «Επιλογή», κατά τη διάρκεια της εκφώνησης του λογαριασμού που επιθυμείτε. Εναλλακτικά πείτε ή πληκτρολογήστε τον αριθμό αυτό.
<b>Ακούτε</b> Τους συνδεδεμένους λογαριασμούς σας, για να επιλέξετε το λογαριασμό που θα πιστωθεί.	<b>Πείτε</b> «Επιλογή», κατά τη διάρκεια της εκφώνησης του λογαριασμού που επιθυμείτε. Εναλλακτικά πείτε ή πληκτρολογήστε τον αριθμό αυτό.
<b>Ακούτε</b> «Πόσα χρήματα θέλετε να μεταφέρετε; Πείτε το ποσό και το νόμισμα.»	<b>Πείτε</b> Π.χ. «1000 Ευρώ»
<b>Ακούτε</b> Τα στοιχεία της συναλλαγής που έχετε εισάγει. • Εάν συμφωνείτε με τα στοιχεία, το σύστημα σας ζητά να πείτε ή να πληκτρολογήσετε έναν αριθμό ηλεκτρονικού κλειδαριθμού. • Σε διαφορετική περίπτωση, σας καθοδηγεί να διορθώσετε τα στοιχεία της συναλλαγής ή να ξεκινήσετε από την αρχή.	
<b>Ακούτε</b> «Αν θέλετε να σας αποσταλούν οι πληροφορίες που ακούσατε μέσω SMS, fax ή e-mail, πείτε sms, fax ή e-mail»	<b>Πείτε</b> «e-mail»

Αμέσως λαμβάνετε e-mail με τα στοιχεία της συναλλαγής σας και η συναλλαγή σας έχει ολοκληρωθεί!

Εικόνα 41 Παράδειγμα μεταφοράς ποσού

## 8. Case study «Γραμμή για την υγεία 1535»



### Εισαγωγή

Μια από τις σημαντικότερες εφαρμογές φωνητικών πυλών που έχουν γίνει στην Ελλάδα είναι η Φωνητική Πύλη 1535 Γραμμή για την Υγεία που δημιουργήθηκε από την εταιρία Διάλογος σε συνεργασία με τον ΟΤΕ.

Η «Γραμμή για την Υγεία» είναι μια ολοκληρωμένη εφαρμογή φωνητικής πύλης που προσφέρει τη δυνατότητα στους χρήστες - ασθενείς να κλείσουν πρωινά και απογευματινά ραντεβού στα εξωτερικά ιατρεία των δημόσιων νοσοκομείων, τα οποία συμμετέχουν σ' αυτή τη φωνητική πύλη, αυτόματα και απλά μέσω ενός τηλεφώνου χρησιμοποιώντας τη φωνή του. Με αυτό τον τρόπο αποφεύγονται οι άσκοπες μετακινήσεις, η τλαιπωρία και οι αναμονές των ασθενών στα νοσοκομεία.

Η λειτουργία της υπηρεσίας «Γραμμή για την Υγεία -1535» ξεκίνησε το Δεκέμβρη του 2003 σε 9 δημόσια νοσοκομεία και μέχρι το Μάρτιο του 2006 ήταν πιλοτική. Όλο αυτό το χρονικό διάστημα έγιναν βελτιώσεις για τη καλύτερη δυνατή εξυπηρέτηση των ασθενών. Το 2007 το μοντέλο λειτουργίας αλλάζει και το σύστημα ενισχύεται με ιατρικό σύμβουλο. Μέχρι και τα τέλη του 2009 προστίθενται συνεχώς νέες μονάδες υγείας με αποτέλεσμα πλέον να είναι ενταγμένες στην έρχεται να συμπληρώσει διαδικτυακά την παροχή υπηρεσιών προς το κοινό υπηρεσία 56 μονάδες υγείας σε ολόκληρη την επικράτεια. Το επόμενο χρονικό διάστημα αναμένεται η αναβάθμιση της υπηρεσίας και του εξοπλισμού.

Επίσης το portal για την υγεία <http://www.healthpages.gr> συμπληρώνει διαδικτυακά τις υπηρεσίες που παρέχει προς το κοινό η φωνητική πύλη «Γραμμή

για την Υγεία –1535». Στο συγκεκριμένο portal ,το οποίο επιμελείται η εταιρεία Ιατρικές Εκδόσεις ΚΑΥΚΑΣ ,ο επισκέπτης μπορεί να βρει πληροφορίες σχετικά με τα εφημερεύοντα νοσοκομεία και φαρμακεία, αναζήτηση ιατρών, τηλέφωνα έκτακτης ανάγκης, κλπ. Ακόμη, προβάλλει θέματα που καλύπτουν σφαιρικά τον τομέα της υγείας, παρουσιάζοντας μια on-line εγκυκλοπαίδεια υγείας, καθώς και άλλες ενότητες που έχουν σχέση με διατροφή, άσκηση, εναλλακτικές θεραπείες κ.α.

Τα χαρακτηριστικά της φωνητικής πύλης Γραμμή για την υγεία είναι τα εξής:

- Χρησιμοποιείται ένας ενιαίος πενταψήφιος αριθμός σε όλη την Ελλάδα με αστική χρονοχρέωση. Ο αριθμός αυτός είναι ο **1535**.
- Λειτουργεί 24 ώρες τη μέρα, 365 μέρες το χρόνο.
- Λειτουργεί σύμφωνα με τους κανόνες που ορίζει η Αρχή Προστασίας Προσωπικών Δεδομένων.
- Διαθέτει σύστημα ασφάλειας για τη προστασία της διακίνησης της πληροφορίας.
- Διαθέτει τηλεφωνικό κέντρο για την εξυπηρέτηση των πελατών, για τη διεκπεραίωση συναλλαγών σε περίπτωση που ο χρήστης δεν μπορεί να εξυπηρετηθεί από το αυτόματο σύστημα, με αριθμό κλήσης **801 801 1535**.

## 8.1 Υπηρεσίες

Παρέχεται ένα πλήθος υπηρεσιών από το σύστημα γραμμή για την υγεία:

- Γίνεται εξακρίβωση των στοιχείων του ασθενή και του γιατρού κάθε φορά που αυτοί εισέρχονται στο σύστημα.
- Γίνεται καταχώρηση των στοιχείων νέων ασθενών
- Ο ασθενής μπορεί να αναζητήσει τα διαθέσιμα ραντεβού και να επιλέξει το νοσοκομείο, τον ιατρό ή τη κλινική που τον ενδιαφέρει.
- Επίσης ο ασθενής μπορεί να επιλέξει την ημέρα και την ώρα του ιατρικού ραντεβού

- Υπάρχει δυνατότητα κατοχύρωσης , ενημέρωσης ,ακύρωσης και μετάθεσης κάποιου ραντεβού
- Γίνεται ενημέρωση των ιατρών για τα απογευματινά ραντεβού που έχουν κλειστεί

## 8.2 Τεχνική περιγραφή

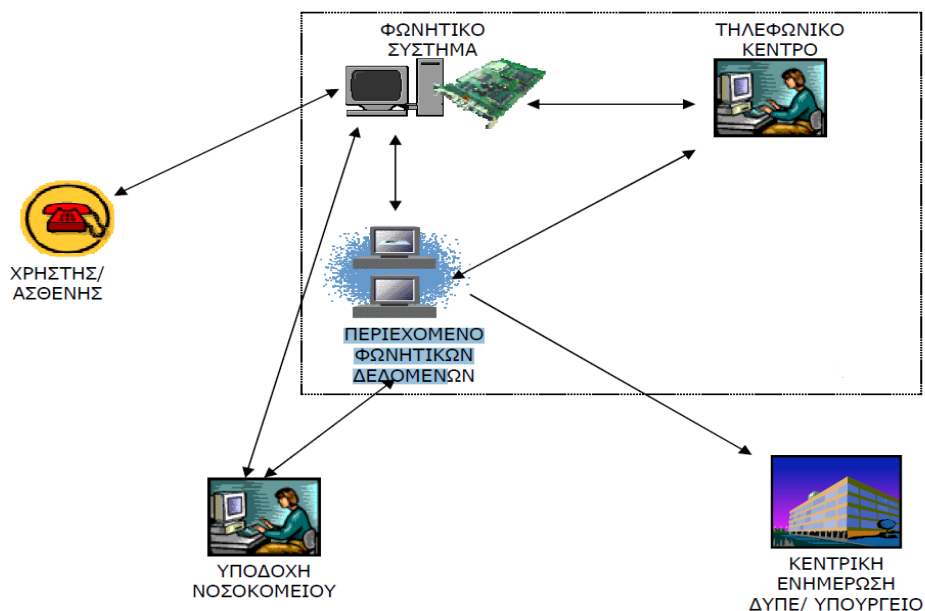
**ΧΡΗΣΤΗΣ/ΑΣΘΕΝΗΣ:** Τηλεφωνική κλήση διεκπεραίωσης / ενημέρωσης ιατρικών ραντεβού

**ΦΩΝΗΤΙΚΟ ΣΥΣΤΗΜΑ:** Διαδραστικός Αποκριτής Φωνής – Υπηρεσία Φωνής

**ΤΗΛΕΦΩΝΙΚΟ ΚΕΝΤΡΟ:** Εξυπηρέτηση κοινού από ανθρώπινο δυναμικό

**ΚΕΝΤΡΙΚΗ ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ:** Καταχώρηση διαθεσιμότητας προσωπικού, καταγραφή προσέλευσης, κλεισμένα ιατρικά ραντεβού

**ΚΕΝΤΡΙΚΗ ΕΝΗΜΕΡΩΣΗ :** Άντληση στοιχείων για εξαγωγή στατιστικών



Εικόνα 42 Αρχιτεκτονική φωνητικής πύλης 1535 Γραμμή για την Υγεία

### 8.3 Λειτουργία

- Ο χρήστης / ασθενής καλεί από το τηλέφωνο του τον αριθμό 1535 και συνομιλεί με το αυτόματο σύστημα.
- Αν ο χρήστης είναι νέος τότε θα πρέπει να συμπληρώσει μια φωνητική φόρμα για τη καταχώριση του στο σύστημα . Από τη στιγμή που θα γίνει η σωστή καταχώριση στη βάση δεδομένων τότε θα πρέπει να περάσει ένα χρονικό διάστημα τεσσάρων εργάσιμων ωρών μέχρι να γίνει η πραγματοποίηση της ενημέρωσης του συστήματος φωνητικής αναγνώρισης.
- Αν ο χρήστης είναι ήδη εγγεγραμμένος τότε είτε λέγοντας το ονοματεπώνυμο του και την ημερομηνία γέννησης του είτε πληκτρολογώντας το κωδικό που του είχε δοθεί από το σύστημα κατά την καταχώρηση του εισέρχεται στο σύστημα. Αν υπάρχουν δύο ή και περισσότεροι χρήστες στο σύστημα με τα ίδια στοιχεία τότε η κλήση προωθείται στο τηλεφωνικό κέντρο.
- Ο χρήστης αρχικά επιλέγει ανάμεσα στη τροποποίηση προηγούμενου ραντεβού και στη δημιουργία νέου ραντεβού
- Σε περίπτωση που επιθυμεί να τροποποιήσει ένα παλιό ραντεβού τότε προωθείται στο τηλεφωνικό κέντρο εξυπηρέτησης.
- Για τη δημιουργία νέου ραντεβού επιλέγει το είδος του ραντεβού που θέλει πρωινό ή απογευματινό, στη συνέχεια επιλέγει το νοσοκομείο. Έπειτα αναφέρει το όνομα ή την ειδικότητα του ιατρού, τη κλινική ή την εργαστηριακή εξέταση που επιθυμεί να κλείσει ραντεβού. Τέλος επιλέγει την ώρα και την ημερομηνία που τον βολεύει.
- Η συναλλαγή διεκπεραιώνεται με την εισαγωγή των στοιχείων του χρήστη.
- Οποτεδήποτε ο χρήστης μπορεί να πει «θέλω βοήθεια» για να πληροφορηθεί σχετικά με τη χρήση του συστήματος.
- Επίσης όταν πει «έξοδος» έχει τη δυνατότητα να εγκαταλείψει την υπηρεσία.
- Λέγοντας «κύριο μενού» έχει τη δυνατότητα να επιστρέψει στο σημείο εισόδου του στο σύστημα.

- Λέγοντας «προώθηση» ο χρήστης προωθείται αυτόματα στο τηλεφωνικό κέντρο εξυπηρέτησης
- Στις περιπτώσεις που ο χρήστης δε μπορεί α εξυπηρετηθεί ούτε από το τηλεφωνικό κέντρο τότε προωθείται στο νοσοκομείο για άμεση εξυπηρέτηση.

#### 8.4 Παράδειγμα

Ακολουθεί Παράδειγμα κλεισίματος ραντεβού με την χρήση της φωνητικής πύλης 1535

- **Φων. Πύλη:** Καλώς ήλθατε στη γραμμή για την υγεία 1535.  
Οποιαδήποτε στιγμή μπορείτε να πείτε θέλω βοήθεια.

- **Φων. Πύλη:** Σας ενημερώνουμε ότι η χρέωση είναι αστική όταν καλείτε από σταθερό τηλέφωνο του ΟΤΕ.  
Για τη δική σας ασφάλεια η κλήση μπορεί να ηχογραφηθεί. Αν θέλετε να αλλάξετε, να μεταθέσετε ή να επιβεβαιώσετε κάποιο παλιό σας ραντεβού , πείτε τροποποίηση, ενώ αν θέλετε να κλείσετε κάποιο νέο ραντεβού, πείτε νέο ραντεβού.

- **Πολίτης:** Νέο ραντεβού

- **Φων. Πύλη:** Πείτε πρωινά ραντεβού αν ενδιαφέρεστε για πρωινά ραντεβού, ή πείτε απογευματινά ραντεβού αν ενδιαφέρεστε για απογευματινά ραντεβού με χρέωση .

- **Πολίτης:** Απογευματινό ραντεβού

- **Φων. Πύλη:** Σας υπενθυμίζω ότι τα απογευματινά ραντεβού επί πληρωμή

- **Φων. Πύλη:** Πείτε το όνομα του νοσοκομείου που σας ενδιαφέρει.

- **Πολίτης:** Παίδων Αγία Σοφία

- **Φων. Πύλη:** Στο νοσοκομείο εξετάζονται παιδιά μέχρι και 14 ετών +  
Πείτε το όνομα ή την ειδικότητα του γιατρού που σας ενδιαφέρει ή το είδος της εργαστηριακής εξέτασης που σας ενδιαφέρει

- **Πολίτης:** Καρδιολογικό

- **Φων. Πύλη:** Πρόκειται για επανεξέταση ή πρώτη εξέταση;

- **Πολίτης:** Πρώτη εξέταση

- **Φων. Πύλη:** Θέλετε να κλείσετε ραντεβού για πρώτη εξέταση με καρδιολόγο στο Νοσοκομείο Παίδων Αγία Σοφία. Πείτε μου ναι ή όχι;

- Πολίτης: *Ναι*

- Φων. Πύλη: *Το πρώτο διαθέσιμο ραντεβού είναι με καρδιολόγο είναι με την ιατρό Αθηνά Βέκιου Σαμαρά την Τρίτη 25 Μαΐου στις 6:30 το απόγευμα. Το κόστος εξέτασης είναι 75 ευρώ. Μήνυμα: Η καρδιολογική εξέταση στοιχίζει 75 ευρώ, ενδέχεται να χρειαστεί και υπερηχογράφημα καρδιάς η τιμή του οποίου είναι 73 ευρώ εκτός αν έχετε ασφάλεια δημοσίου όποτε θα προσκομίσετε το βιβλιάριο ασθενοείας σας. Να το κλείσω; Πείτε μου ναι ή όχι;*

- Πολίτης: *Ναι*

- Φων. Πύλη: *Επαναλαμβάνω το μήνυμα: Η καρδιολογική εξέταση στοιχίζει 75 ευρώ, ενδέχεται να χρειαστεί και υπερηχογράφημα καρδιάς η τιμή του οποίου είναι 73 ευρώ εκτός αν έχετε ασφάλεια δημοσίου οπότε θα προσκομίσετε το βιβλιάριο ασθενοείας σας. Θέλετε να κλείσετε ραντεβού την Τρίτη 25 Μαΐου στις 6:30 το απόγευμα. Πείτε μου ναι ή όχι;*

- Πολίτης: *Ναι*

- Φων. Πύλη: *Για να καταχωρηθεί το ραντεβού σας πρέπει να μου δώσετε τα στοιχεία του ασθενή. Θα πρέπει να σας ενημερώσω ότι στην περίπτωση που τα στοιχεία αυτά είναι ασαφή ή ελλιπή υπάρχει πιθανότητα να μην ισχύει το ραντεβού σας. Πείτε το μικρό όνομα του ασθενή.*

- Πολίτης: *Γιώργος*

- Φων. Πύλη: *Πείτε το επώνυμο του ασθενή*

- Πολίτης: *Παπανικολάου*

- Φων. Πύλη: *Πείτε το πατρώνυμο του ασθενή*

- Πολίτης: *Λευτέρης*

- Φων. Πύλη: *Πείτε το έτος γέννησης του ασθενή*

- Πολίτης: *2001*

- Φων. Πύλη: *Τέλος, για να μπορεί το νοσοκομείο να σας ενημερώσει για οποιαδήποτε αλλαγή του ραντεβού σας πείτε ένα πρωινό τηλέφωνο επικοινωνίας*

- Πολίτης: *2103456789*

- Φων. Πύλη: *Παρακαλώ περιμένετε. Επαναλαμβάνω τα στοιχεία του ασθενή: Όνομα Γιώργος, επώνυμο Παπανικολάου, πατρώνυμο Λευτέρης, έτος γέννησης*

*2001 και τηλέφωνο 2103456789 .*

*Είναι σωστά τα στοιχεία που ακούσατε; Πείτε μου ναι η όχι.*

- Πολίτης: *Ναι*

- Φων. Πύλη: *Παρακαλούμε να είστε στην ώρα σας. Θα σας έχει κρατηθεί θέση στο parking του νοσοκομείου. Το ραντεβού σας καταχωρήθηκε. Το κόστος εξέτασης είναι 75 ευρώ εκτός αν ενδιαφέρεστε για υπερηχογράφημα καρδιάς οπότε το κόστος είναι 73.37 ευρώ.*

*Μήνυμα: Η καρδιολογική εξέταση στοιχίζει 75 ευρώ, ενδέχεται να χρειαστεί και υπερηχογράφημα καρδιάς η τιμή του οποίου είναι 73 ευρώ εκτός αν έχετε ασφάλεια δημοσίου οπότε θα προσκομίσετε το βιβλιάριο ασθενείας σας. Θέλετε να βγείτε από την υπηρεσία;*

- Πολίτης: *Ναι*

- Φων. Πύλη: *Ευχαριστούμε πολύ, γεια σας.*

## **Επίλογος**

Η υπηρεσία Γραμμή για την Υγεία 1535 βοηθά σημαντικά στην αποσυμφόρηση της λειτουργίας των νοσοκομειακών ιατρείων. Παρέχει υψηλής ποιότητας υπηρεσίες διαχείρισης νοσοκομειακών ραντεβού στους πολίτες / ασθενείς και τηλεφωνική υποστήριξη πάνω σε 12ωρη βάση (08.00-20.00) από εξειδικευμένο προσωπικό. Οι ασθενείς μπορούν εύκολα και άμεσα να κλείσουν ραντεβού, καθώς υπάρχει μόνο ένας ενιαίος τετραψήφιος αριθμός κλήσης πανελλαδικά για όλα τα νοσοκομεία, με αστική χρονοχρέωση, ενώ προσφέρεται άμεση εξυπηρέτηση των πολιτών / ασθενών και διαθεσιμότητα 24 ώρες την ημέρα 7 ημέρες την εβδομάδα, ακόμα και σε αργίες. Αναβαθμίζονται οι παρεχόμενες υπηρεσίες υγείας. Επίσης βγαίνουν αξιόπιστα στατιστικά τόσο σε επίπεδο νοσοκομείου όσο και σε επίπεδο περιφέρειας υγείας, μέσω της παρακολούθησης της εξέλιξης της κίνησης ανά Μονάδα Υγείας/ Νοσοκομείο.



## Βιβλιογραφία

### Ελληνικές πηγές

1. Αδαμόπουλος, Δ. *Τηλεματική – Τηλεπληροφορική και Δικτυκεντρικά Πληροφοριακά Συστήματα*.
2. Δημήτριος, Μ. *Φωνητικές Πύλες: αρωγός στην ανάπτυξη ψηφιακών εφαρμογών*.
3. Διγαλάκης, Β. *Σημειώσεις μαθήματος Εισαγωγή στην επεξεργασία φωνής*.
4. Μιχαλέτου, Ε. (2008). *Παραμετροποίηση σηματος ομιλίας για αναγνωρισησυναισθηματος ομιλητη. Πανεπιστήμιο Πατρών*
5. Μπλέτσας, Α. (1998). *Ελληνικό σύστημα αυτόματης εκφώνησης κειμένου*. Θεσσαλονίκη: ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ.
6. Μπούρας, Χ. Ι. *Τηλεματική και νέες υπηρεσίες*. ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ ,ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ.
7. Πασχόπουλος Αρσένης, Χ. Β. (2003). *Κινητό Εμπόριο (Mobile commerce)*.
8. Στεργιοπούλου, Σ. (2010). *Συστήματα Μετατροπής Κειμένου σε Ομιλία (TTS Systems)*. Θεσσαλονίκη: ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ.
9. Τσινάκος, Α. *Τηλεματική - Βασικές Αρχές και Έννοιες*. ΑΤΕΙ ΚΑΒΑΛΑΣ ΤΜΗΜΑ ΒΙΟΜΗΧΑΝΙΚΗΣ ΠΛΗΡΟΦΟΡΙΚΗΣ.

### Ξένες πηγές

10. Burkhardt, F. *Requirements for an Emotional Markup Language with Voice Portal Applications in Mind - a Position Paper*. Germany: Deutsche Telekom Laboratories.
11. Burkhardt, F., Ajmera, J., Englert, R., Stegmann, J., & Burleson, W. *Detecting Anger in Automated Voice Portal Dialogs*. Berlin, Germany: T-Systems Enterprise Services GmbH, Deutsche Telekom Laboratories.
12. Burkhardt, F., Ballegooy, M. v., Englert, R., & Huber, R. *An Emotion-Aware Voice Portal*. Berlin / Erlangen, Germany: T-Systems International GmbH, Deutsche Telekom Laboratories, Sympalog Voice Solutions GmbH.
13. Burkhardt, F., Markus, v. B., Engelbrecht, K.-P., Polzehl, T., & Stegmann, J. *Emotion Detection in Dialog Systems: Applications, Strategies and Challenges*. Berlin, Germany: Deutsche Telekom Laboratories.

14. Burkhardt, F., Polzehl, T., Stegmann, J., & Metze, F. *DETECTING REAL LIFE ANGER*. Berlin, Germany: Deutsche Telekom Laboratories / \*Technische Universität Berlin.
15. Clark, D. (n.d.). *Speech Recognition: The Wireless Interface Revolution*. *TECHNOLOGY NEWS*.
16. CORKREY, R., & PARKINSON, L. (2002). *Interactive voice response: Review of studies 1989–2000*. Australia: University of Newcastle, New South Wales.
17. Cristel, M., & Hauptmann, A. (2002). *Voice XML and Speech Applications*.
18. D'Alessandro, N. *An Introduction to Text-to-Speech Synthesis*.
19. Davis, C. (2001). *Automatic Speech Recognition and Access*.
20. *Dialogos Voice Platform*. (2007). Dialogos Speech Communication Systems S.A.
21. Dutoit, T. (n.d.). *A Short Introduction to Text-to-Speech Synthesis*. Ανάρτηση από [http://tcts.fpms.ac.be/synthesis/introtts\\_old.html](http://tcts.fpms.ac.be/synthesis/introtts_old.html)
22. Dutoit, T. *High-quality text-to-speech synthesis an overview*. Belgium: Faculte Polytechnique de Mons, TCTS Lab.
23. Erlanger, L. (2009). *The why and how of voice portals*.
24. F. Burkhardt, J. A. *Detecting Anger in Automated Voice Portal Dialogs*. Berlin, Germany: T-Systems Enterprise Services GmbH, Deutsche Telekom Laboratories.
25. Fagel, S. *AUDIOVISUAL SPEECH: ANALYSIS, SYNTHESIS, PERCEPTION, AND RECOGNITION*. Berlin : University of Technology.
26. Gallivan, P., Hong, Q., Jordan, L., Li, E., Mathew, G., Mulyani, Y., και συν. (2002). *VoiceXML Absentee System*. Pace University.
27. GOLDMAN, E. L., PANTTAJA, E., WOJCIKOWSKI, A., & BRAUDES, R. (2001). *Voice Portals—Where Theory Meets Practice*. INTERNATIONAL JOURNAL OF SPEECH TECHNOLOGY.
28. Holmes, J. N., Holmes, W. J., & Garner, P. N. *USING FORMANT FREQUENCIES IN SPEECH RECOGNITION*. Speech Technology Consultant.
29. Juang, B., & Rabiner, L. R. (2004). *Automatic Speech Recognition – A Brief History of the Technology Development*. Atlanta: Georgia Institute of Technology.

30. Kloos, C. D. (3, September 2004). *Arturo Azcorra-Salona Telematics: A Distinctive Discipline*. Madrid: Spain International Journal of Computer & Information Science.
31. LEE, F.-M., LI, L.-H., & HUANG, R.-Y. *Recognizing Low/High Anger in Speech for Call Centers*. Taiwan: Department of Information Management Chaoyang University of Technology.
32. Levin, E. *Designing the Voice User Interface for Automated Directory Assistance*. City University of New York.
33. Meng, Y. (2004). *Speech Recognition on DSP: Algorithm Optimization and Performance Analysis*. Hong Kong: The Chinese University of Hong Kong.
34. Neiberg, D., & Elenius, K. *Automatic Recognition of Anger in Spontaneous Speech*. Stockholm, Sweden: Centre for Speech Technology, CSC, KTH (Royal Institute of technology).
35. Oliver, D. *Prosody in text-to-speech synthesis*. Saarbrücken, Germany: Institute of Phonetics, Saarland University,.
36. Panagos, E., & Delis, A. (2005). *Selective Replication for Content Management Environments*.
37. Potamianos, G., & Garg, A. (2003). *Recent Advances in the Automatic Recognition of Audio-Visual Speech*.
38. Quinn, K. (2001). *Voice Portal Customer Presentation for The Dialogues Spotlight Colloquium*.
39. Radhakrisitnan, V. (2006). *VOICE PORTAL*. Kolenchery: College of Engineering.
40. Rudžionis, A., & Ratkevičius, k. (2004). *Voice Based Internet Services*.
41. Schröder, M., Pirker, H., Lamolle, M., Burkhardt, F., Christian, P., & Zovato, E. *Representing emotions and related states in technological systems*.
42. Schroeter, J. *Text to-Speech (TTS) Synthesis*.
43. Sharma, D. R. (2006). *Text-to-speech synthesis for punjabi language*. PATIALA: COMPUTER SCIENCE AND ENGINEERING DEPARTMENT.
44. SHINN, P., SHOMPHE, M., LEWIS, M., CAREY, K., & KIM, D. (2004). *Designing Language Models for Voice Portal Applications*. INTERNATIONAL JOURNAL OF SPEECH TECHNOLOGY.
45. Sidorova, J. (2007). *Speech Emotion Recognition*. Universitat Pompeu Fabra.

46. Simoneau, A. (2001). *How voice portal technology is advancing the evolution in communications.*
47. Sproat, R., Taylor, P., Tanenblatt, M., & Isard, A. *A MARKUP LANGUAGE FOR TEXT-TO-SPEECH SYNTHESIS.* Murray Hill: Bell Laboratories, Lucent Technologies.
48. Strik, H. *COMPUTER-BASED LEARNING: WHAT AUTOMATIC SPEECH RECOGNITION HAS TO OFFER.*
49. Tadlock, J. *The Voice Activated Services Revolution.*
50. Takeiki Aizono, D. E. *New Information Technologies to Achieve Next-generation.*
51. Tatham, M. (1985). *An Integrated Knowledge Base for Speech Synthesis and Automatic Speech Recognition.* London.
52. Tatham, M., & Morton, K. (2005). *DEVELOPMENTS IN SPEECH SYNTHESIS.* England.
53. *Telematics Past, Present and Future.* (2008).
54. *The Next Big Opportunity on the Web - An Introduction to Voice Portals.*
55. Umesh, S. (2010). *Automatic Speech Recognition – Research and Standards.* Department of Electrical Engineering Indian Institute of Technology Madras.
56. Vogt, T., And, E., & Bee, N. *EmoVoice | A framework for online recognition of emotions from voice.* Augsburg, Germany.
57. Vogt, T., Elisabeth, A., & Johannes, W. *Automatic Recognition of Emotions from Speech: A Review of the Literature and Recommendations for Practical Realisation.*
58. *Voice Portal First Solution.* (2009). Avaya Solution Interoperability Lab.
59. *Voice Portals White Paper.* (2001). Sonera Plaza Ltd Medialab.
60. *VoiceXML Tutorial.*
61. Winther, M. (2001). *The Rise of Voice Portals.*
62. Yacoub, S., Simske, S., Lin, X., & Burns, J. (2003). *Recognition of Emotions in Interactive Voice Response Systems.*
63. Yasushi, Y., Hitoshi, I., & Kazuhiro, W. (2003). *Technologies for voice portal platform.*

### **Επιλεγμένες Πηγές από το Διαδίκτυο**

- <http://searchnetworking.techtarget.com> (What is Telematics?)
- <http://www.serresbiz.com/method-www/attach/Thlematiki.htm>
- [www.ASAsShop.org](http://www.ASAsShop.org) (Telematics Past, Present and Future)
- <http://el.science.wikia.com/wiki/Τηλεματική>
- <http://searchmobilecomputing.techtarget.com/definition/voice-portal>
- [http://www.techworld.com.au/article/278949/why\\_how\\_voice\\_portals?fp=4&fpid=23](http://www.techworld.com.au/article/278949/why_how_voice_portals?fp=4&fpid=23) (The why and how of voice portals)
- <http://cslu.cse.ogi.edu/asr/> (Automatic Speech Recognition)
- [http://en.wikipedia.org/wiki/Speech\\_recognition](http://en.wikipedia.org/wiki/Speech_recognition)
- [http://en.wikipedia.org/wiki/Interactive\\_voice\\_response](http://en.wikipedia.org/wiki/Interactive_voice_response)
- [http://en.wikipedia.org/wiki/Speech\\_synthesis](http://en.wikipedia.org/wiki/Speech_synthesis)
- [www.telecom.tuc.gr/courses/speech/vxml\\_lectures/d02\\_voice\\_uis.pdf](http://www.telecom.tuc.gr/courses/speech/vxml_lectures/d02_voice_uis.pdf)
- <http://www.telecom.tuc.gr/>
- [www.voicexml.gr/Greek/files/banking.pps](http://www.voicexml.gr/Greek/files/banking.pps)
- <http://el.science.wikia.com/wiki/Τηλεματική>.
- <http://searchmobilecomputing.techtarget.com/definition/voice-portal> (voice portal (vortal))
- <http://www.seminarprojects.com/Thread-automatic-speech-recognition-full-report?pid=11809#pid11809>
- [www.neccare.com](http://www.neccare.com) (IVR White Paper)
- <http://www.voxeo.com/library/ivr.jsp>
- [http://cis.mantis.gr/solutions\\_05.htm](http://cis.mantis.gr/solutions_05.htm)