



Εθνικόν και Καποδιστριακόν  
Πανεπιστήμιον Αθηνών

Τμήμα Πληροφορικής και Τηλεπικοινωνιών

# Επεξεργασία Ομιλίας και Φυσικής Γλώσσας

Ενότητα 1: Εισαγωγή

Γεώργιος Κουρουπέτρογλου

[koupe@di.uoa.gr](mailto:koupe@di.uoa.gr)



# ΔΙΑΠΡΟΣΩΠΙΚΗ ΕΠΙΚΟΙΝΩΝΙΑ ΣΕ ΠΡΑΓΜΑΤΙΚΟ ΧΡΟΝΟ



A) εκ του πλησίον  
ή  
πρόσωπο με πρόσωπο

ΟΜΙΛΙΑ

---



B) σε απόσταση  
δίκτυο (π.χ. τηλεφωνικό)

ΟΜΙΛΙΑ

# ΔΙΑΠΡΟΣΩΠΙΚΗ ΕΠΙΚΟΙΝΩΝΙΑ ΣΕ ΜΗ ΠΡΑΓΜΑΤΙΚΟ ΧΡΟΝΟ



A) εκ του πλησίον  
σημειώματα

---



B) σε απόσταση  
επιστολές (δίκτυο  
ταχυδρομείου)

# ΕΠΙΚΟΙΝΩΝΙΑ ΑΝΘΡΩΠΟΥ – ΜΗΧΑΝΗΣ

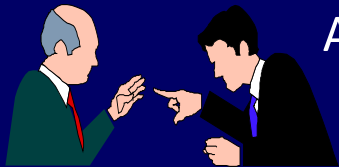


συνήθεις συσκευές εισόδου – εξόδου  
και κατάλληλο user – interface

**ΠΡΟΣΒΑΣΗ ΣΤΟ ΠΕΡΙΕΧΟΜΕΝΟ (και του WWW)**

# ΤΕΧΝΟΛΟΓΙΑ ΕΠΙΚΟΙΝΩΝΙΑΣ ΜΕ ΟΜΙΛΙΑ

ΔΙΑΠΡΟΣΩΠΙΚΗ ΕΠΙΚΟΙΝΩΝΙΑ ΣΕ ΠΡΑΓΜΑΤΙΚΟ ΧΡΟΝΟ (1/3)



A) εκ του πλησίον

**ΟΜΙΛΙΑ**

πρόβλημα παραγωγής ομιλίας 1.2% του πληθυσμού

πρόβλημα αποδοχής ομιλίας 2.7% του πληθυσμού

*Βοηθήματα παραγωγής ομιλίας*

*Βοηθήματα ομιλίας για κωφούς*



B) σε απόσταση

**ΟΜΙΛΙΑ**



δίκτυο (π.χ. ψηφιακής τηλεφωνίας, ISDN, κινητή τηλεφωνία, Voice over IP)

ψηφιοποίηση  
ομιλίας



συμπίεση  
ομιλίας




μετάδοση  
ομιλίας



αναπαραγωγή  
ομιλίας

# ΤΕΧΝΟΛΟΓΙΑ ΕΠΙΚΟΙΝΩΝΙΑΣ ΜΕ ΟΜΙΛΙΑ

## ΔΙΑΠΡΟΣΩΠΙΚΗ ΕΠΙΚΟΙΝΩΝΙΑ ΣΕ ΠΡΑΓΜΑΤΙΚΟ ΧΡΟΝΟ (2/3)

σε απόσταση:  voice mailing  
voice messaging



## ΕΠΙΚΟΙΝΩΝΙΑ ΑΝΘΡΩΠΟΥ - ΜΗΧΑΝΗΣ

ομιλία (αναγνώριση)

ομιλία (σύνθεση)



κοινό τηλεφωνικό δίκτυο

αναγνώριση



ΟΜΙΛΙΑ

σύνθεση



# ΤΕΧΝΟΛΟΓΙΑ ΕΠΙΚΟΙΝΩΝΙΑΣ ΜΕ ΟΜΙΛΙΑ

ΔΙΑΠΡΟΣΩΠΙΚΗ ΕΠΙΚΟΙΝΩΝΙΑ ΣΕ ΠΡΑΓΜΑΤΙΚΟ ΧΡΟΝΟ (3/3)

## ΕΠΙΚΟΙΝΩΝΙΑ ΑΝΘΡΩΠΟΥ – ΜΗΧΑΝΗΣ

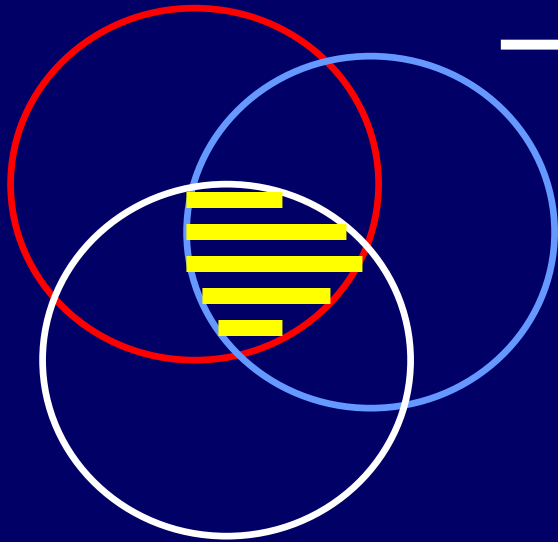
Πρόσβαση στο περιεχόμενο

- Άτομα χωρίς όραση: 0.1%
- Άτομα με μειωμένη όραση: 1,5%
- Άτομα με αναπηρία στα άνω άκρα: (ποσοστά στο γενικό πληθυσμό)
  - Δεν μπορούν να χρησιμοποιήσουν τα δάχτυλα ή τον καρπό: 0.3 %
  - Δεν μπορούν να χρησιμοποιήσουν έναν βραχίονα: 0.1 %
  - Μειωμένη ισχύς: 2.8 %
  - Μειωμένος συντονισμός: 1.4 %
- *Μετατροπή κειμένου σε συνθετική ομιλία.*
- *VoiceBrowser.*
- *χρήση s/w με φωνητικές εντολές.*
- *έλεγχος περιβάλλοντος με φωνητικές εντολές (μέσω κατάλληλου h/w).*



# ΣΗΜΑΝΤΙΚΟΤΗΤΑ ΟΛΟΚΗΡΩΣΗΣ ΤΕΧΝΟΛΟΓΙΑΣ ΟΜΙΛΙΑΣ

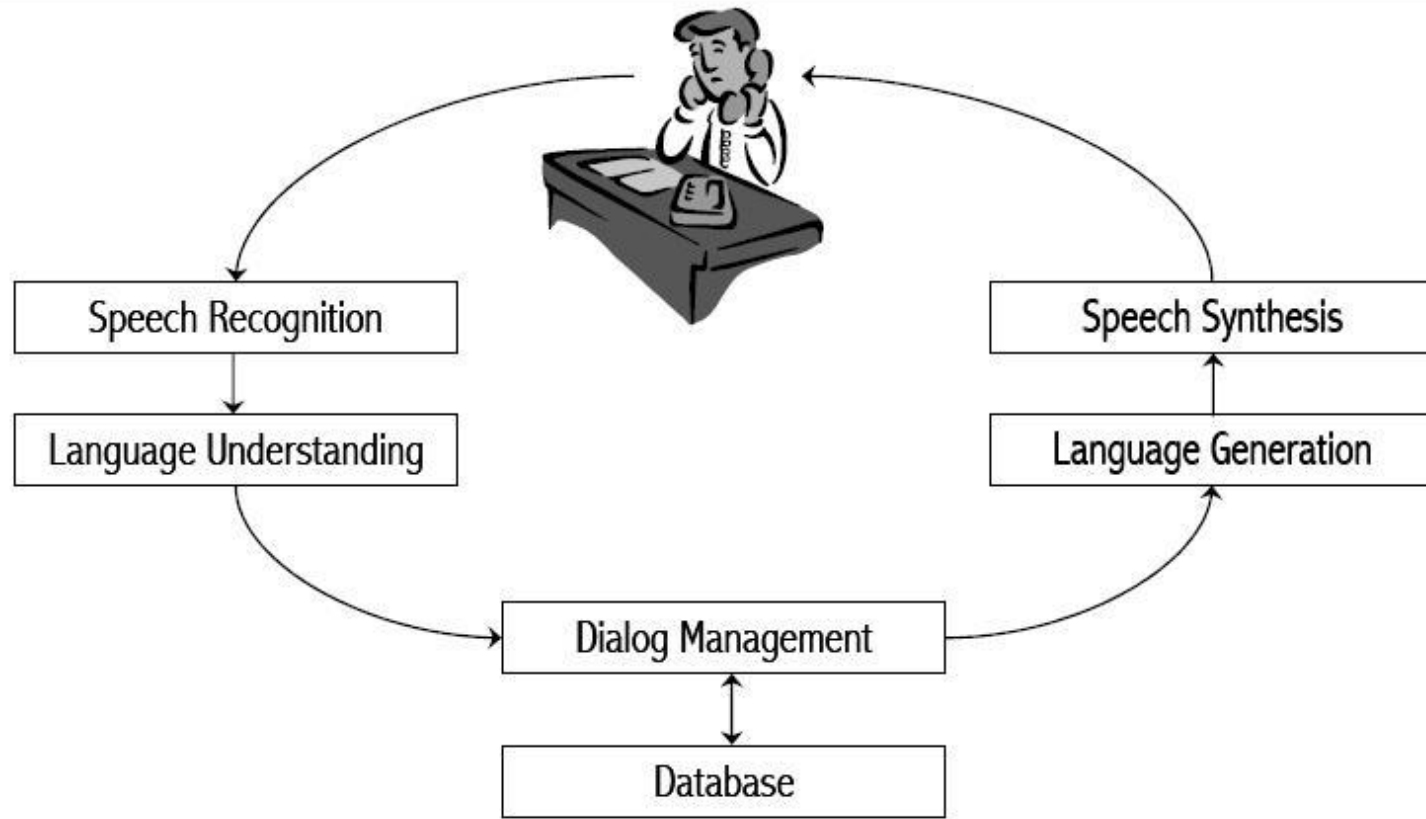
- διαπροσωπική επικοινωνία
- πρόσβαση στον Η/Υ & στις πληροφορίες
- έλεγχος περιβάλλοντος





# Σύστημα Φωνητικού Διαλόγου

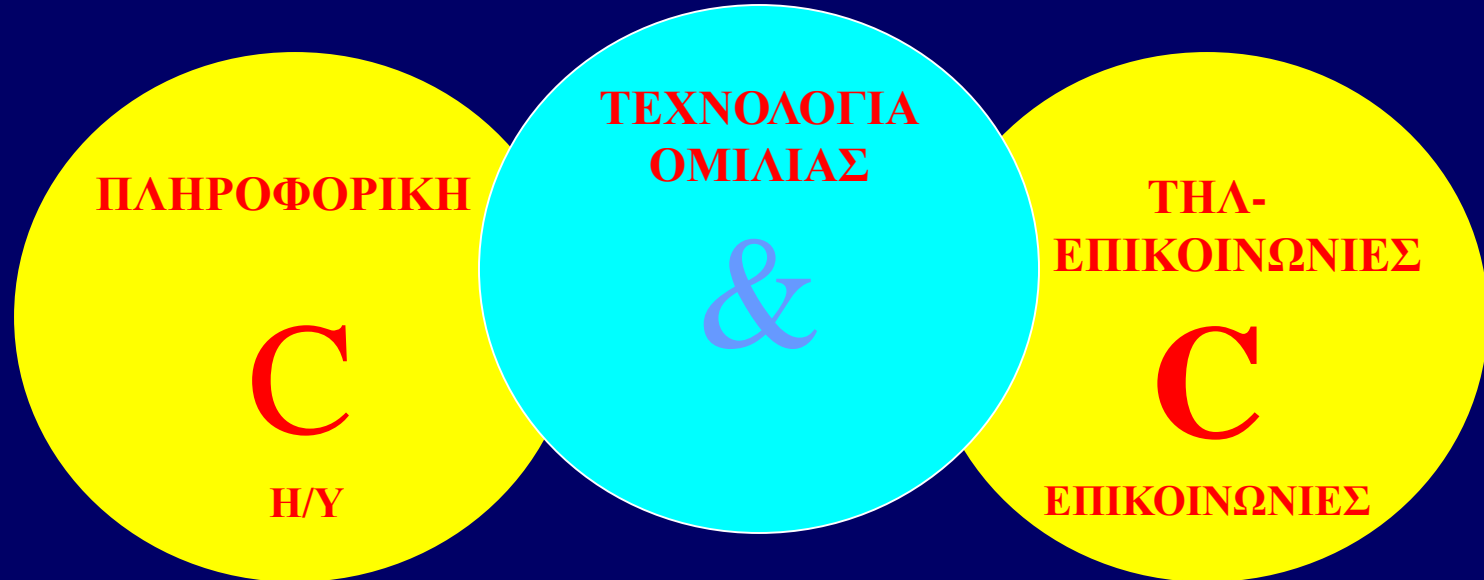
## The Architecture of an SLDS



# Τι περιέχει ένα Σύστημα Προφορικού/ Φωνητικού Διαλόγου;

- **Αναγνώριση ομιλίας** – αναλύει το εισερχόμενο ηχητικό σήμα ομιλίας ώστε να εξάγει γλωσσικές μονάδες όπως λέξεις ή φωνήματα.
- **Κατανόηση Γλώσσας** – καθορίζει το νόημα των εισερχόμενων.
- **Διαχείριση διαλόγου** – ελέγχει τη ροή της συζήτησης, διατηρώντας την ιστορία και το πλαίσιο, κατευθύνοντας την πορεία της, έχοντας πρόσβαση στη βάση δεδομένων, και μορφοποιώντας τις αποκρίσεις.
- **Βάση δεδομένων** – φυλάσσει τις πληροφορίες που παρέχουν το περιεχόμενο του διαλόγου.
- **Παραγωγή Γλώσσας** – θέτει τις αποκρίσεις ως λέξεις.
- **Σύνθεση Ομιλίας** – παράγει το εξαγόμενο ηχητικό σήμα ομιλίας.

# ΤΕΧΝΟΛΟΓΙΑ ΟΜΙΛΙΑΣ



# 1<sup>ο</sup> Σενάριο Προφορικού Διαλόγου με Υπολογιστές (1/2)

Το ξυπνητήρι χτυπά στις 06.30 π.μ. Η Στέλλα στριφογυρίζει λίγο, ανασηκώνεται και ψιθυρίζει «αναβολή». Δέκα λεπτά αργότερα, το ξυπνητήρι ξαναχτυπά. Η Στέλλα ξυπνάει και λέει «κλείσε».

Τεντώνεται, χασμουριέται, και σηκώνεται από το κρεβάτι. Γυρίζει το κεφάλι προς το μικρό κουτί δίπλα στο κρεβάτι και λέει: «Πώς είναι ο καιρός σήμερα;». «Στην Αθήνα σήμερα, συννεφιά αρχικά με λίγο ήλιο αργότερα, με θερμοκρασία στους 12 βαθμούς το πρωί με αύξηση στους 16 το απόγευμα», έρχεται η απάντηση από το κουτί. «Αύριο;» «Η πρόβλεψη για την Τρίτη είναι ελαφριά βροχή το πρωί και δυνατή βροχή και ίσως καταιγίδες από τα Δυτικά αργότερα το απόγευμα». «Και στη Βαρκελώνη;» «Η πρόβλεψη του καιρού για τη Βαρκελώνη είναι λιακάδα, με θερμοκρασίες από 16 βαθμούς το πρωί έως 25 το απόγευμα». «Και την Κόστα Μπράβα;» «Συγγνώμη», απαντά το κουτί, «δε γνωρίζω το μέρος που ρωτάτε. Παρακαλώ δοκιμάστε να ρωτήσετε για κάποια πόλη ή χώρα. Για παράδειγμα, ποιες πόλεις γνωρίζετε στην Ισπανία;» «Δεν πειράζει», απαντά η Στέλλα με μια ελαφριά έκφραση ενόχλησης ενώ κατευθύνεται στο μπάνιο.

# 1<sup>ο</sup> Σενάριο Προφορικού Διαλόγου με Υπολογιστές (2/2)

Έπειτα από ένα ντούζ, η Στέλλα μπαίνει στην κουζίνα, πιάνει μια φέτα ψωμί και το βάζει στην τρoστιέρα. «Όχι τόσο ξεροψημένο αυτή τη φορά». Πηγαίνει στο ψυγείο, βγάζει ένα κουτί γάλα και παρατηρεί ότι είναι σχεδόν άδειο. «Μην ξεχάσεις να παραγγείλεις ένα κουτί γάλα», λέει στο ψυγείο. «Αργότερα περιμένεις κάποιους φίλους για ζεστή σοκολάτα, ίσως θα έπρεπε να παραγγείλω δύο κουτιά», απαντά το ψυγείο. «Εντάξει», συμφωνεί η Στέλλα.

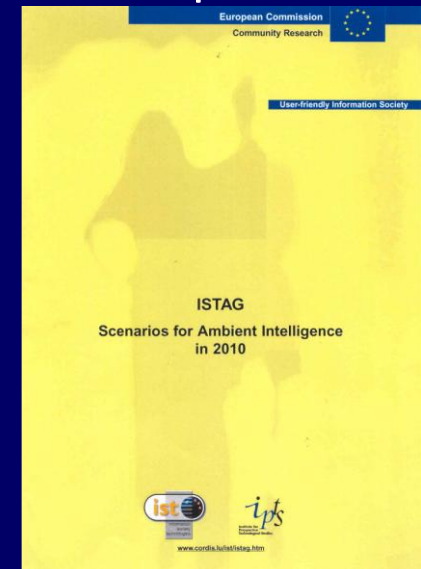
Έπειτα από το πρωινό η Στέλλα αρχίζει να σκέφτεται τις δουλειές της ημέρας. Πατά ένα κούμπι στη μικρή συσκευή που βρίσκεται στον καρπό του χεριού της που δείχνει την ώρα. Η συσκευή την καλημερίζει και τη ρωτά τι υπηρεσίες θα ήθελε. «Τι συναντήσεις έχω για σήμερα;», ρωτά η Στέλλα. «Έχεις μια συνάντηση με τον Παύλο στις 10, και μετά μια συνάντηση με την Άννα στις 11», απαντά η συσκευή. «Τηλεφώνησε την Άννα», λέει η Στέλλα, και η συσκευή αρχίζει να τηλεφωνεί. Η Στέλλα συνομιλεί με την Άννα για την αλλαγή της ώρας της συνάντησής τους. Λέει στη συσκευή να ορίσει καινούρια ώρα στο ημερολόγιο και να ελέγξει το email της. Όταν τελειώνει λέει «Ευχαριστώ, αυτά για την ώρα» στη συσκευή στον καρπό της και ετοιμάζεται να φύγει.



# 2° Σενάριο Προφορικού Διαλόγου με Υπολογιστές

**Maria and P-com: ISTAG** (*Information Society Technologies Advisory Group*)  
Scenarios for Ambient Intelligence 2010,  
European Commission, 2001

<ftp://ftp.cordis.europa.eu/pub/ist/docs/istagscenarios2010.pdf>



# Τρόποι αλληλεπίδραση Ανθρώπου – Υπολογιστή με Ομιλία

- Ο χρήστης φορά μικρόφωνο και ακουστικά συνδεδεμένα στον Η/Υ.
- Μέσω τηλεφώνου: σύνδεση με απομακρυσμένο Η/Υ (αόρατος στον χρήστη).
- Ενσωματωμένα συστήματα φωνής (π.χ. ηλεκτρικές συσκευές, ρούχα).
- Ενσωματωμένα μικρόφωνα & μεγάφωνα στον περιβάλλοντα χώρο που επικοινωνούν ασύρματα (π.χ. *Bluetooth*) με Η/Υ (πανταχόθεν & διεισδυτικοί υπολογιστές) – δίκτυα αισθητήρων (μικροφώνων).
- Talking avatars / conversational avatars.



This is Mary, one of the talking and listening characters with whom you can interact when using our software on your computer



Imagine a flat screen display and wireless microphone.





# Τι είναι ένα Σύστημα Προφορικού Διαλόγου (ΣΠΔ);

- είναι ένα υπολογιστικό σύστημα στο οποίο μπορείς να μιλήσεις για να εκτελέσει κάποια εντολή.
- Τα ΣΠΔ είναι συνήθως δύο ειδών :
  - **Συστήματα παροχής πληροφοριών:** σου παρέχουν πληροφορίες ως απάντηση σε επερώτηση, όπως είναι μια αίτηση για πληροφορίες σχετικά με ένα χρονοδιάγραμμα ή πληροφορίες σχετικά με τον καιρό.
  - **Συστήματα βάσει συναλλαγών:** σου επιτρέπουν να αναλάβεις κάποιου είδους συναλλαγή κράτησης, όπως το να αγοράζεις ή να πουλάς μετοχές, ή μία θέση σε αεροπλάνο.

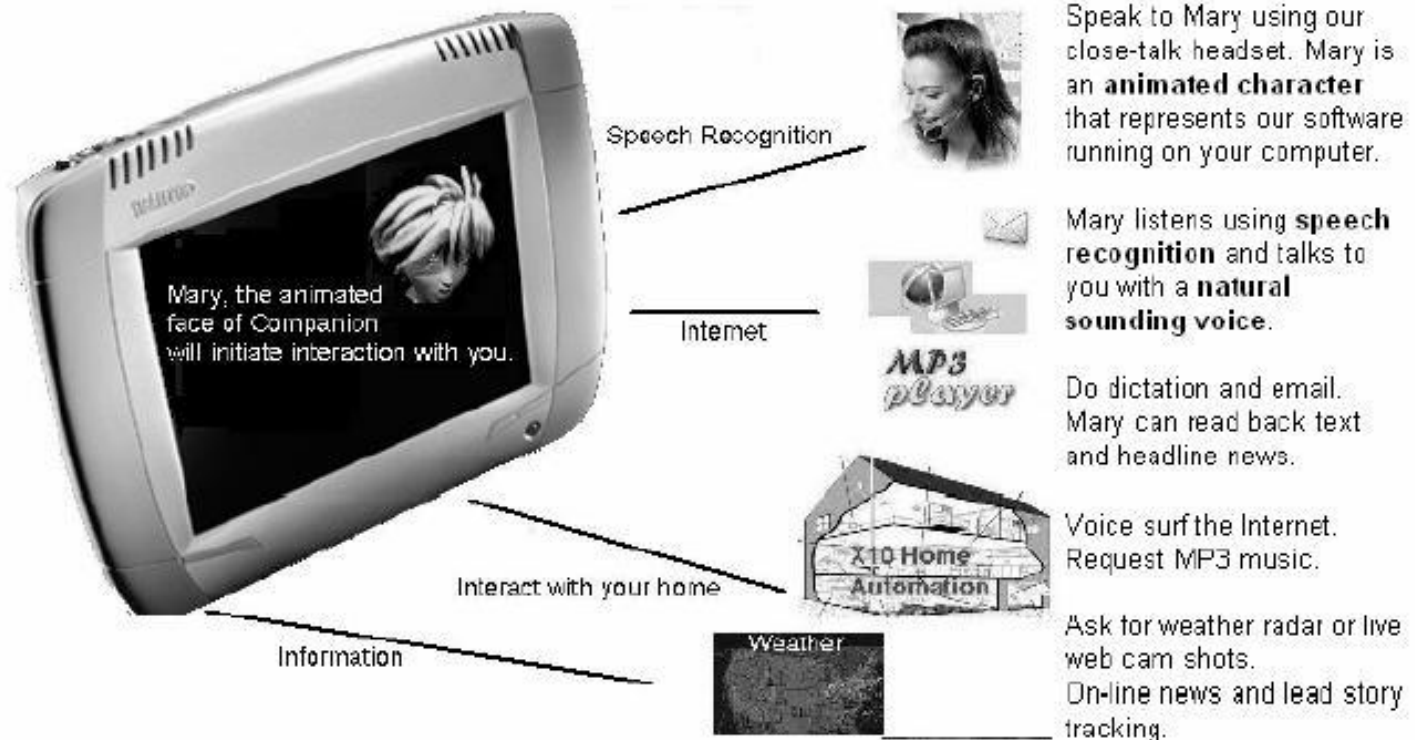
# Δύο Κύριες Εφαρμογές της Τεχνολογίας Αναγνώρισης Ομιλίας

- Στον προσωπικό Η/Υ:
  - εξαρτημένη από τον ομιλητή
  - εκτεταμένο λεξιλόγιο (δεκάδες χιλιάδες λέξεις)
  - εργασίες υπαγόρευσης
- Στην Τηλεφωνία ή Βασισμένα στην Τηλεφωνία:
  - ανεξαρτήτως ομιλητή
  - σχετικά περιορισμένο λεξιλόγιο (εκατοντάδες λέξεις)
  - αμφίδρομες εργασίες συναλλαγών

# Βασικές Χρήσεις αναγνώρισης ομιλίας σε προσωπικό Η/Υ

- Φωνητική Υπαγόρευση.
- Φωνητική αλληλογραφία.
- Φωνητικές εντολές.
- Φωνητική Πλοήγηση.
- Μετάφραση προφορικού λόγου (σε πραγματικό χρόνο) σε άλλη γλώσσα.

# Cutting-edge Technology: Talking Desktop



Voice Web Browser

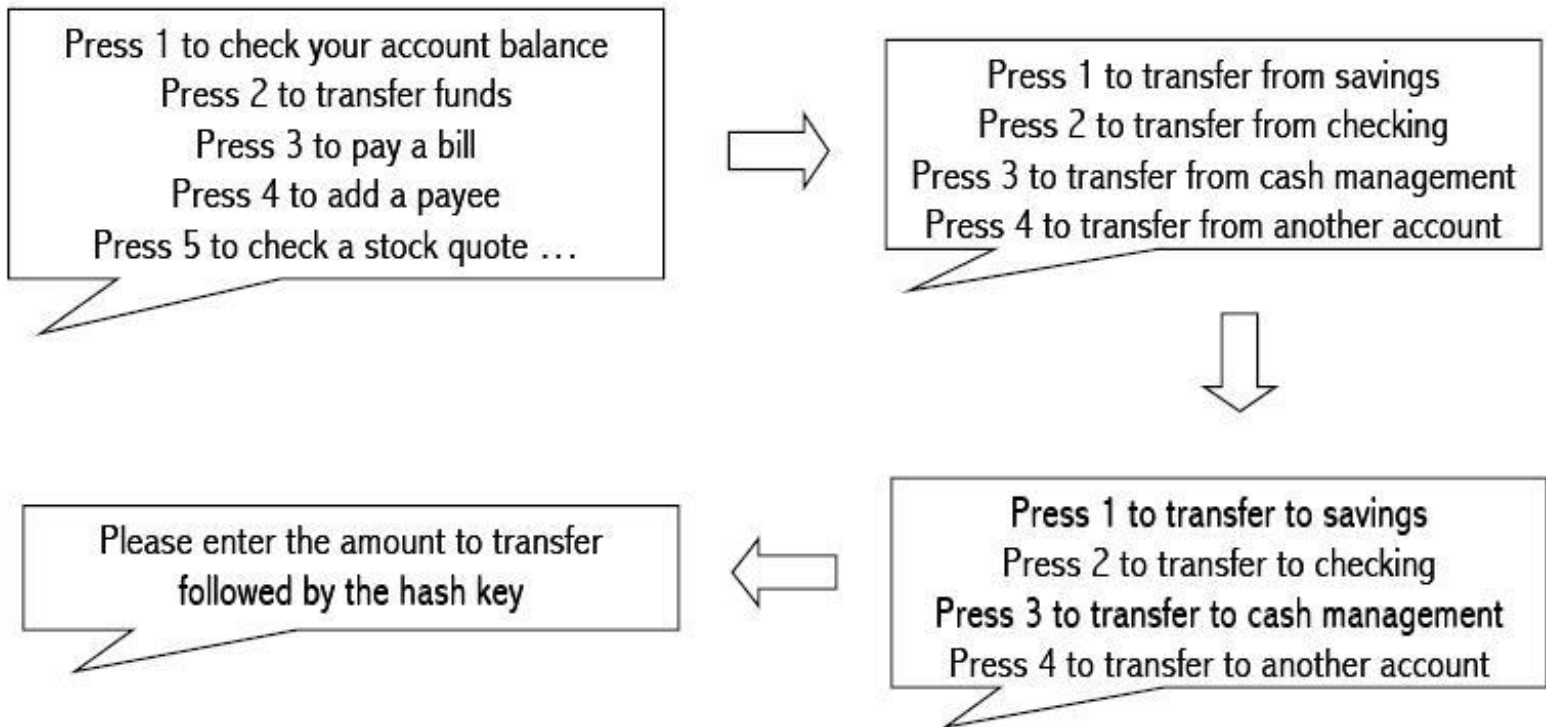


# Παραδείγματα Χρήσεων Τηλεφωνικών Συστημάτων Προφορικού Διαλόγου – Συστημάτων Προφορικού Διαλόγου Βασισμένων στην Τηλεφωνία

- Απομακρυσμένες τραπεζικές συναλλαγές.
- Ταξιδιωτικές κρατήσεις.
- Κρατήσεις σε εξωτερικά ιατρεία νοσοκομείων.
- Συναλλαγές μετοχών.
- Τηλε-στοιχήματα.
- Υπηρεσίες καταλόγου.
- κλήσεις για Taxi.
- Παραγγελίες φαγητού.


# Traditional Interactive Voice Response Systems

---



# Speech-Enabled Interaction

---



Transfer 500 dollars from  
savings to checking next  
Wednesday after 3pm

~25 seconds via speech — as compared with two minutes via touch tone

# Συνδέοντας το Διαδίκτυο με το Τηλέφωνο

- Ο παγκόσμιος ιστός εξαπλώνεται πέρα από τους προσωπικούς υπολογιστές.
- Τα τηλέφωνα βρίσκονται παντού.
  - Πολύ περισσότερα τηλέφωνα, παρά υπολογιστές.
  - Περισσότερα ασύρματα τηλέφωνα, παρά υπολογιστές, ειδικά στην Ευρώπη.
- Τα ασύρματα τηλέφωνα είναι εξαιρετικά μεταφέρσιμα.
  - Υποστήριξη ενδιαφερουσών υπηρεσιών βάσει τοποθεσίας (location based services).
- Δεν μπορείς να χειριστείς έναν υπολογιστή ενώ οδηγείς.
- Άμεση ενεργοποίηση.





# Το ενδιαφέρον στρέφεται:

- Αξιοποίηση του Διαδικτύου για ανάπτυξη εφαρμογών και παράδοση υπηρεσιών με χρήση φωνητικού διαλόγου
  - Τηλέφωνο αντί για προσωπικό υπολογιστή.
  - VoiceXML αντί για HTML.
  - Ένα φωνητικό πρόγραμμα περιήγησης (Voice Browser) αντί για ένα τυπικό πρόγραμμα περιήγησης.

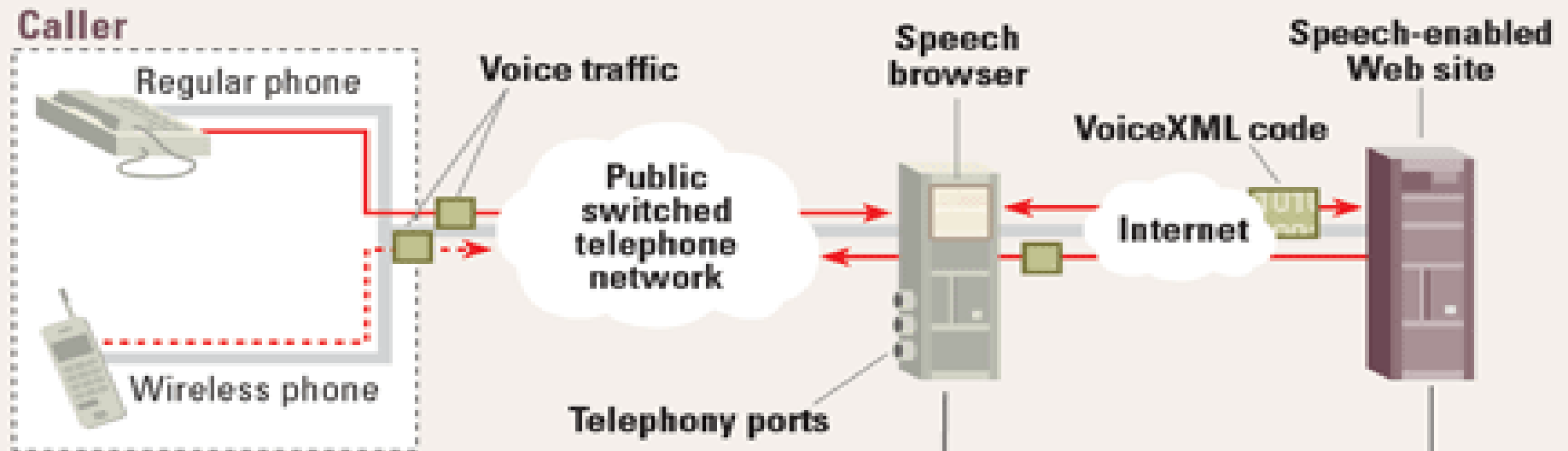
# Κοινοπραξία Παγκόσμιου Ιστού (W3C) και φωνή

- Φόρουμ για την VoiceXML  
[www.voicexml.org](http://www.voicexml.org)
- Ομάδα εργασίας της συνεργασίας W3C για  
Φωνητικά Προγράμματα Περιήγησης  
[www.w3.org/Voice](http://www.w3.org/Voice)

# HOW IT WORKS

## VoiceXML

VoiceXML allows application developers to create new speech-based Web content.

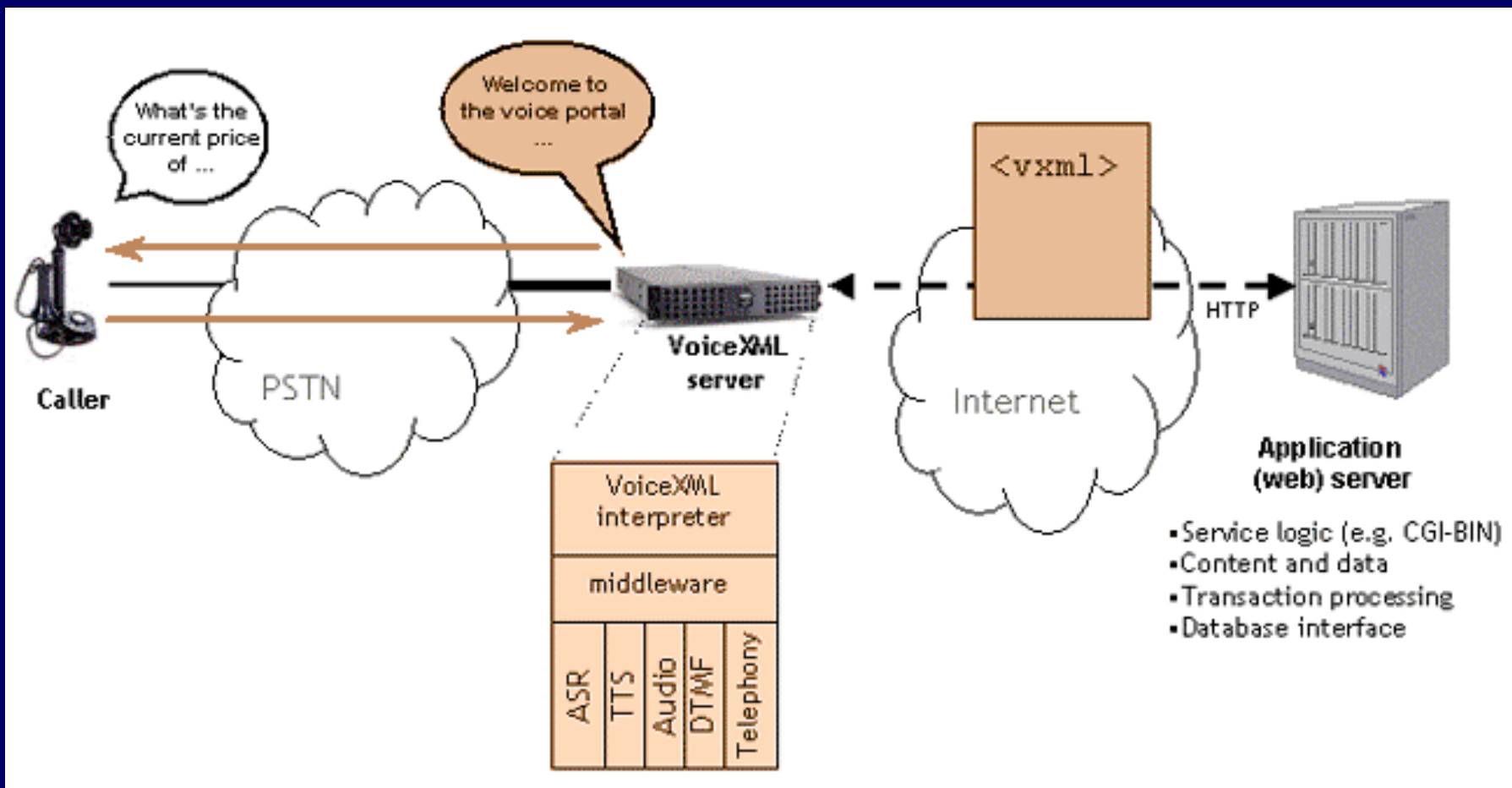


**1** A caller can use any voice-enabled device to access a speech-enabled Web site.

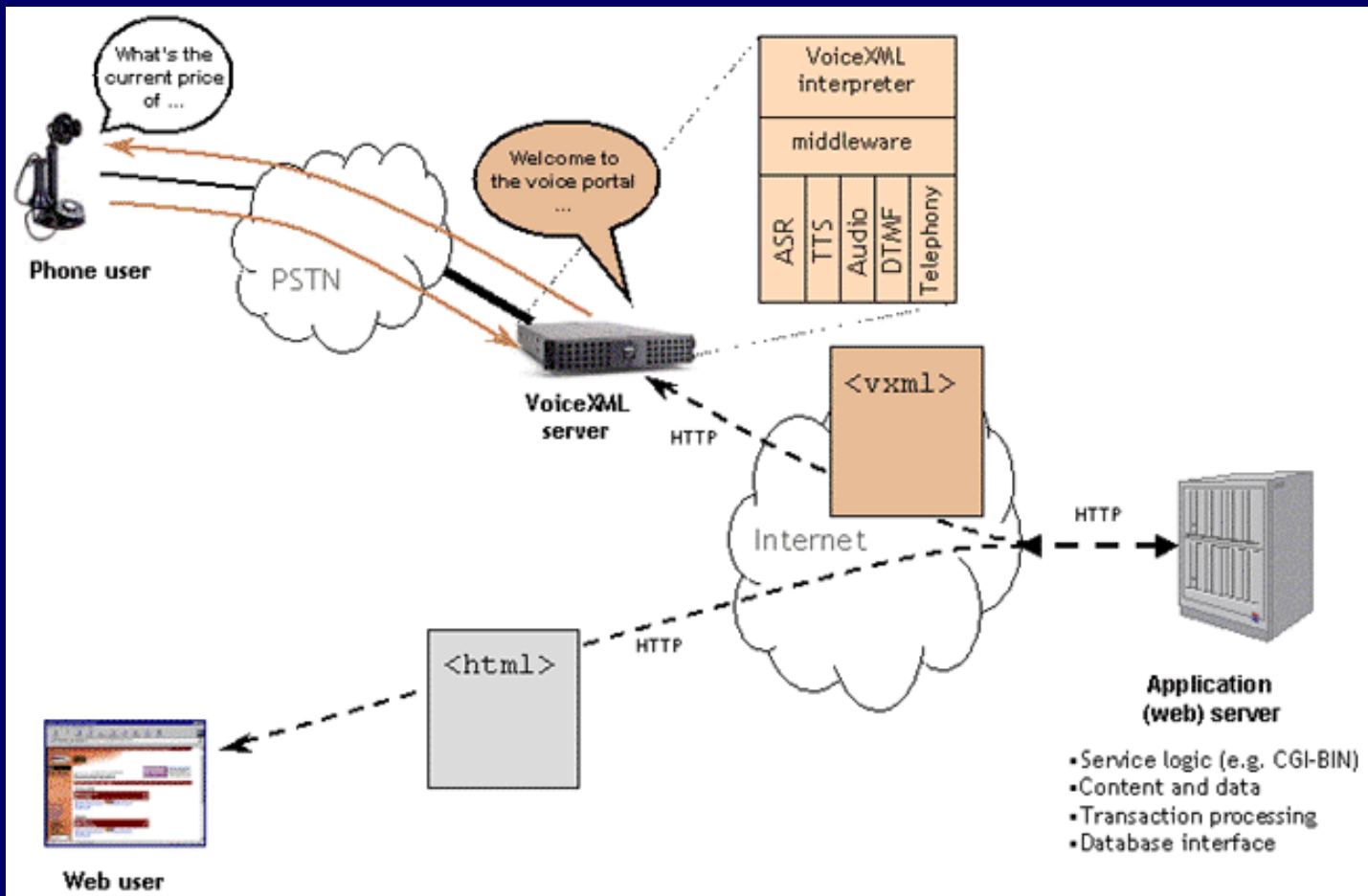
**2** A speech browser is the gateway between a voice call and an Internet connection, interpreting VoiceXML code and managing dialog between the caller and VoiceXML content at Web site.

**3** A server delivers information back to the caller in audio format.

# Αρχιτεκτονική μιας Φωνητικά Ενεργής Διαδικτυακής Εφαρμογής



# Σχέση μεταξύ μιας παραδοσιακής Διαδικτυακής Εφαρμογής και μιας Φωνητικά Ενεργοποιημένης Διαδικτυακής Εφαρμογής



# ΠΡΟΤΥΠΑ – ΤΥΠΟΠΟΙΗΣΕΙΣ

- SSIL (Speech Synthesizer Interface Library).
- SAM (Synthesizer Access Manager).
- **Microsoft Speech API ή SAPI.**
- **JSML (Java Speech Markup Language) < XML.**
- **VoiceXML (Voice Markup Language).**
- MeXE (Mobile Executive Environment for Mobile e – Commerce).
- V–Commerce = Voice-Commerce.

W3C: Voice Browser Activity Working Group.

ETSI: DSR (Distributed Speech Recognition).



# Έλεγχος Συσκευών με Φωνή (1/2)

- Λειτουργία Command-and-control.
- Από εφαρμογές λογισμικού ως τις οικιακές ηλεκτρικές συσκευές.
- Κατάλληλα και για περιπτώσεις «απασχολημένων χεριών» π.χ. αυτοκίνητο, χρήση ποντικιού.
- Χρήσιμη για άτομα με αναπηρίες που δεν μπορούν να χειριστούν τις συσκευές ή το λογισμικό.
- Πολλές φορές με μικρό λεξιλόγιο (10-50 λέξεις), άλλες φορές (π.χ. Πιλοτήριο αεροπλάνου) ~5000 λέξεις.

# Έλεγχος Συσκευών με Φωνή (2/2)

- Σε ενσωματωμένα συστήματα 50-100 φράσεις εξαρτώμενες από τον ομιλητή και 10-20 σε φράσεις ανεξάρτητες του χρήστη.
- Συνήθως άμεση αντιστοίχιση στη σημασιολογία των λέξεων ή φράσεων.
- Πιο σύνθετες φράσεις, όπως «άνοιξε την πόρτα του γκαράζ στις 7:00».
- Αλλά: π.χ. «κάθε Σάββατο και Κυριακή άνοιγε τα φώτα του καθιστικού στις 18:00 για 4 ώρες» πολλές εναλλακτικές λέξεις, πολλοί πιθανοί συνδυασμοί ενεργειών > ανάγκη χρήσης γραμματικής η οποία προσδιορίζει με πιο αφηρημένο τρόπο τις επιτρεπόμενες λέξεις, φράσεις και τους συνδυασμούς τους.
- Διαφορετικοί τρόποι έκφρασης για το ίδιο πράγμα: «άναψε τα φώτα στις 7», «στις 7 άναψε τα φώτα», «τα φώτα άναψέ τα στις 7».



# Εισαγωγή Δεδομένων με Φωνή

- Σε εφαρμογή λογισμικού (π.χ. Συμπλήρωση πεδίων φόρμας, τιμολογίων, αλλαγή ραντεβού).
- Μικρός αριθμός συνηθισμένων λέξεων.
- Ίσως μεγάλος αριθμός λέξεων εξαρτώμενων από την εφαρμογή (π.χ. Ένα σύστημα αναφορών τροχαίων δυστυχημάτων απαιτεί όλα τα ονόματα των δρόμων και των τοποθεσιών μιας πόλης).

# Λήψη Πληροφοριών με Φωνή

- Αρκετές εφαρμογές σε λειτουργία.
- Χρήση σχεδόν φυσικής ομιλίας.
- MIT: σύστημα πληροφοριών καιρού Jupiter + κρατήσεων θέσεων Pegasus (4000 πτήσεις) + Voyager για καθοδήγηση πορείας και κυκλοφορίας στην Boston (1977-1999: 30.000 κλήσεις, 80% επιτυχής αναγνώριση για νέους χρήστες).
- Μεγάλο μέγεθος λεξιλογίου.

# Φωνητική Υπαγόρευση

- Σε πραγματικό χρόνο ή από προεκφωνημένη ομιλία.
- Επιτρέπουν συνδυασμό φωνητικών εντολών για μορφοποίηση του κειμένου.
- Δεν υπάρχει διάλογος με τον χρήστη.
- Μπορεί όμως να ενσωματωθεί σε εφαρμογές (π.χ. Mailer).
- Πολύ μεγάλο λεξικό ή και εξειδικευμένο: μικρό λεξικό > καλύτερη ακουστική αναγνώριση (κυρίως σε ακουστικά ίδιες λέξεις), πολύ μεγάλο λεξικό > μικρότερα λάθη εισαγωγής λανθασμένων λέξεων.
- Μεγάλο πρόβλημα οι άγνωστες λέξεις που αντικαθίστανται από κάποια του υπάρχοντος λεξικού.



# Φωνητικός Προσωπικός Βοηθός

## Speech (Voice) enabled Virtual Assistant

- Χειρίζεται έναν αριθμό αυτόματων λειτουργιών: βρίσκει και καλεί τηλεφωνικούς αριθμούς, λαμβάνει e-mails, συμβουλεύεται και ενημερώνει ημερολόγιο συναντήσεων ή γεγονότων.
- Π.χ. «κάλεσε τον Γιάννη στο γραφείο», «Ραντεβού της επόμενης Δευτέρας από τις 5 ως τις 7 το απόγευμα», «Διάβασέ μου το e-mail από τον Πρόεδρο».

# Φωνητικά Συστήματα

## Προηγμένου Διαλόγου (1/2)

- Περιπτώσεις που ο χρήστης εκτελεί πολλές εξαρτώμενες ενέργειες.

Π.χ. Πληροφορίες και κρατήσεις πτήσεων, ενοικίαση αυτοκινήτου, κράτηση ξενοδοχείου.

- Επιτρέπουν τον χρήστη να μετακινείται μεταξύ των ενεργειών και το σύστημα θυμάται τι έχει συζητηθεί (π.χ. Να αποφασίσει για ενοικίαση αυτοκινήτου εξαρτώμενη από την τοποθεσία του ξενοδοχείου)

- Προσπαθεί να αντιμετωπίζει τα σφάλματα και τις παρανοήσεις

Παράδειγμα:

X: Που υπάρχει λιακάδα στην Καραϊβική;

Y: Συγνώμη, δεν είμαι σίγουρος τι είπες

X: Που υπάρχει λιακάδα στη Καραϊβική;

Y: Συγνώμη, μπορεί να μη γνωρίζω τη πόλη που ενδιαφέρεσαι. Προσπάθησε να ρωτήσεις για την χώρα της ή την ήπειρο.



# Φωνητικά Συστήματα

## Προηγμένου Διαλόγου (2/2)

- Προσθήκη «λογικής κρίσης» (reasoning) στο σύστημα.

Παράδειγμα:

Χ: Χρειάζεται να πάμε την ασθενή στο Ιπποκράτειο στην Αθήνα *(υπάρχει αντίστοιχο και στη Θεσ/νίκη)*

Υ: ΟΚ

Χ: Τι οχήματα είναι διαθέσιμα;

Υ: Υπάρχουν ασθενοφόρα στην Καρδίτσα και στον Βόλο.

Χ: Προτιμώ αυτό του Βόλου.

Υ: Γνωρίζεις ότι ο νότιος δρόμος Βόλου-Εθνικής είναι μπλοκαρισμένος λόγω έργων;



TRY

- ... changing Ramona's Personality
- ... changing Ramona's Accent
- ... asking her about well-know people, books, events
- ... moving your cursor around in her window
- ... earlier versions of Ramona

- RAMONA: φωτορεαλιστική προσωποποίηση πράκτορα συζήτησης.

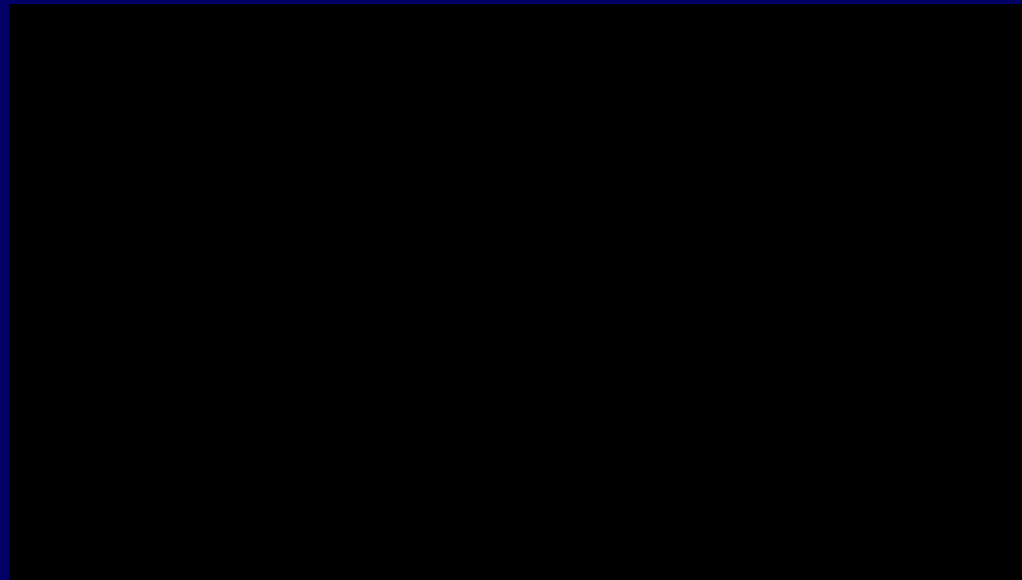
<http://www.kurzweilai.net/Ramona4.2/ramona.html>

# Maske-bot: A talking video humanoid robot

November 8, 2011

<http://www.youtube.com/watch?v=oFp1hpH25oI>

<http://www.kurzweilai.net/mask-bot-a-talking-video-humanoid-robot>





# Η πρόοδος των Τεχνολογιών Φωνής

είναι αποτέλεσμα ενός αριθμού παραγόντων:

- Τεχνολογικοί πρόοδοι στο hardware των υπολογιστών..
- Ανάπτυξη ειδικού λογισμικού και αλγορίθμων.
- Υποδομών για το VoiceWeb.
- Εμπορικής Ώθησης.

# ΑΝΑΠΗΡΙΑ

- *μια κατάσταση της ζωής ενός ατόμου η οποία χαρακτηρίζεται από διάφορου βαθμού δυσκολίες στην εκτέλεση των καθημερινών ανθρώπινων λειτουργιών και δραστηριοτήτων* (συμβατικός ορισμός).
- αποτέλεσμα οργανικών ή περιβαλλοντολογικών αιτιών, που δημιουργούν ένα σύνολο εμποδίων σε σημαντικές περιοχές της ζωής, όπως η αυτοεξυπηρέτηση, η απασχόληση, η εκπαίδευση, η ψυχαγωγία και η γενικότερη κοινωνική συμμετοχή.

# Ταξινόμηση Περιορισμών των Λειτουργιών ενός Ατόμου με τον Χρόνο Εμφάνισης

- Εκ γενετής (συγγενείς)
- Από ασθένεια ή ατύχημα (επίκτητοι)
- Με ηλικία (25% > ηλικιωμένοι από 2020)
- Προσωρινά
- Περιπτωσιακά



# Παραδείγματα Περιπτωσιακής Απώλειας Λειτουργιών (1/2)

- Λειτουργία **χωρίς όραση**: απασχολημένα μάτια (π.χ. οδήγηση), σκοτάδι.
- Λειτουργία **με χαμηλή όραση**: μικρή οθόνη, περιβάλλον με καπνό, ξεχασμένα γυαλιά.
- Λειτουργία **χωρίς ακοή**: πολύ θορυβώδες περιβάλλον, απασχολημένη ακοή, επιβαλλόμενη ησυχία (βιβλιοθήκες, συναντήσεις).
- Λειτουργία **με χαμηλή ακοή**: θορυβώδες περιβάλλον.

# Παραδείγματα Περιπτωσιακής Απώλειας Λειτουργιών (2/2)

- Λειτουργία **χωρίς ανάγνωση**: άλλη γλώσσα, επισκέπτες/τουρίστες, ξεχασμένα γυαλιά.
- Λειτουργία **χωρίς περιορισμένη νοητική ικανότητα**: πανικόβλητος, επηρεασμένος από αλκοόλ.
- Λειτουργία **με μειωμένη ικανότητα χεριών**: ζωηρό όχημα, διαστημικό όχημα.

# Αντιμετώπιση Αναπηρίας

- Αλλαγή των ικανοτήτων του ατόμου
  - Χειρουργικά.
  - με αποκατάσταση ή με εκπαίδευση.
  - Προσωπικές Υποστηρικτικές Τεχνολογίες (Π-ΥΤ).
- Εγκατάσταση προσαρμογών στο περιβάλλον
  - Υποστηρικτικές Τεχνολογίες Προσαρμογών (ΥΤ-Π).
- Αλλαγή τρόπου σχεδιασμού πραγμάτων
  - Όστε να είναι χρησιμοποιήσιμα από περισσότερους.
  - Καθολική Σχεδίαση (ΚΣ) / Σχεδίαση για όλους.

# Παράδειγμα: Άτομα με Προβλήματα Όρασης στις Αυτόματες Ταμειακές Μηχανές (ATM)

	Τυφλοί	Χαμηλής όρασης
Εντοπισμός ATM	0	0
Πρόσβαση στις ATM		
Ανάγνωση οδηγιών χρήσης	0	0
Εισαγωγή κάρτας	0	0
Ανάγνωση οθόνης	0	0
Χρήση πληκτρολογίου	0	0
Χρήση οθόνης αφής	0	0
Ανάκτηση χρημάτων		
Ανάγνωση απόδειξης		0
Ανάκτηση κάρτας	0	

	Ελάχιστα προβλήματα
	Μερικά προβλήματα
	Πολλά προβλήματα
0	Διαθέσιμη τεχνολογία ομιλίας για αντιμετώπιση του προβλήματος



# Παράδειγμα Ο τυφλός μαθητής κατά την ενταξιακή εκπαιδευτική διαδικασία αντιμετωπίζει προβλήματα: (1/2)

- Πρόσβασης στο έντυπο εκπαιδευτικό υλικό.
- Πρόσβασης στο βοηθητικό έντυπο εκπαιδευτικό υλικό (εγκυκλοπαίδειες, περιοδικά, εφημερίδες, κλπ).
- Πρόσβασης στον πίνακα της τάξης.
- Πρόσβασης στο υλικό που προβάλλεται στην τάξη με άλλα μέσα, όπως overhead projector, data projector).
- Συγγραφής σημειώσεων στην τάξη.



# Παράδειγμα Ο τυφλός μαθητής κατά την ενταξιακή εκπαιδευτική διαδικασία αντιμετωπίζει προβλήματα: (2/2)

- Συγγραφής εργασιών που του ανατίθενται.
- Συμμετοχής στις γραπτές δοκιμασίες.
- Πρόσβασης με τη χρήση του ηλεκτρονικού υπολογιστή (Η/Υ):
  - στο εκπαιδευτικό υλικό που υπάρχει: α) σε ηλεκτρονική μορφή, β) σε ηλεκτρονικές βιβλιοθήκες και γ) στο διαδίκτυο,
  - σε εκπαιδευτικές εφαρμογές.

Οι Τεχνολογίες Ομιλίας προσφέρουν μερικές ή ολικές λύσεις σε όλες τις παραπάνω περιπτώσεις

# Προσβασιμότητα (accessibility)

- Η δυνατότητα του χρήστη να αλληλεπιδρά με φυσικό τρόπο με το προϊόν ή την υπηρεσία.
  - π.χ. να το φτάνει, να έχει αρκετή ισχύ να το μετακινεί, κλπ.

# Προσβασιμότητα

Σχεδίαση και Ανάπτυξη εφαρμογών και υπηρεσιών  
πληροφορικής

έτσι ώστε να μπορούν

να τα χρησιμοποιούν αποτελεσματικά

περισσότεροι άνθρωποι

σε περισσότερες περιστάσεις ή πλαίσια χρήση

# περισσότεροι άνθρωποι (1/4)

## Άτομα με Αναπηρία:

- **Αισθητηριακή**
  - Τύφλωση ή χαμηλή όραση ή αχρωματοψία.
  - Κώφωση ή βαρηκοΐα.
- **Κινητική**
  - Επιδεξιότητα χεριών.
  - Τέντωμα και φτάσιμο.
  - Μετακίνηση.

# περισσότεροι άνθρωποι (2/4)

## Άτομα με Αναπηρία :

- Γνωσιακή/νευρολογική
  - Δυσλεξία.
  - Γλωσσική/επικοινωνία.
  - Ελλειμματική προσοχή.
  - Μνήμη.
  - Κατανόηση.
- Πολυαναπηρίες

# περισσότεροι άνθρωποι (3/4)



Άτομα με Αναπηρία: > 10 % πληθυσμού

# περισσότεροι άνθρωποι (4/4)

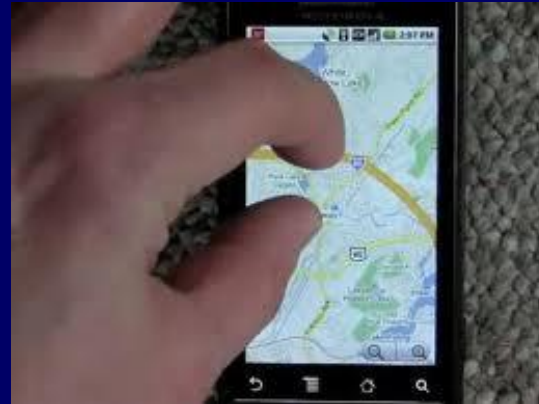
- Περιπτωσιακή ή περιστασιακή ανικανότητα
- Άτομα χωρίς γλωσσική ευχέρεια.
- Χρήστες παλαιότερης τεχνολογίας.
- Χρήστες νέων συσκευών πληροφορικής.
- Νέοι ή μη συχνοί χρήστες πληροφορικής.
- Χρήστες κινητών τηλεφώνων ή tablets.
- .....
- Ηλικιωμένοι.

# περισσότερες περιστάσεις ή πλαίσια χρήσης (1/2)





# περισσότερες περιστάσεις ή πλαίσια χρήσης (2/2)



# Ποικιλία πλαισίων χρήσης (1/2)

περιπτώσεις όπου οι χρήστες:

- Ίσως δεν μπορούν εύκολα ή και καθόλου να δουν, να ακούσουν, να χειριστούν ή να επεξεργαστούν κάποιες μορφές πληροφορίας.
- Ίσως έχουν δυσκολία να διαβάσουν ή να κατανοήσουν κείμενα.
- Ίσως δεν μπορούν να χειριστούν το πληκτρολόγιο ή το ποντίκι.
- Ίσως διαθέτουν μικρή οθόνη, ή οθόνη χωρίς γραφικά ή χαμηλής ταχύτητας σύνδεση στο διαδίκτυο.

# Ποικιλία πλαισίων χρήσης (2/2)

περιπτώσεις όπου οι χρήστες:

- Ίσως δεν μιλούν ή δεν καταλαβαίνουν με ευχέρεια τη γλώσσα στην οποία είναι γραμμένη μια πληροφορία.
- Ίσως βρίσκονται σε μια κατάσταση όπου τα μάτια τους ή και τα χέρια τους ή και τα αυτιά τους είναι απασχολημένα σε άλλες κύριες δραστηριότητες (π.χ. ενώ οδηγούν, όταν εργάζονται σε ένα θορυβώδες περιβάλλον).
- Ίσως διαθέτουν μια παλαιότερη έκδοση ενός φυλλομετρητή, ή έναν εντελώς διαφορετικό φυλλομετρητή, ή έναν φωνητικό φυλλομετρητή ή ένα διαφορετικό λειτουργικό σύστημα.

«Το 57 % των χρηστών Ηλεκτρονικών υπολογιστών ηλικίας 18 ως 64 ετών άμεσα ή έμμεσα ωφελούνται από τις τεχνολογίες προσβασιμότητας λόγω των δυσκολιών και των ανικανοτήτων στη χρήση των υπολογιστών»

Μελέτη της Forrester Research, Inc. για λογαριασμό της Microsoft

# Σχεδίαση για Όλους (Design for All – D4All)

- ισοδύναμος με τους όρους:
  - Καθολική Σχεδίαση και
  - Σχεδίαση Ενσωμάτωσης (inclusive design) *Ηνωμένο Βασίλειο*
- Χρησιμοποιείται περισσότερο στην Ευρώπη στον τομέα της Κοινωνίας της Πληροφορίας.

## Σχεδίαση για Όλους είναι:

- Η ενσυνείδητη και συστηματική προσπάθεια της εκ των προτέρων εφαρμογής αρχών, μεθόδων και εργαλείων με σκοπό την ανάπτυξη προϊόντων και υπηρεσιών Πληροφορικής και Τηλεπικοινωνιών που είναι προσβάσιμα και χρηστικά από όλους τους πολίτες και επομένως η αποφυγή της ανάγκης εκ των υστέρων προσαρμογών ή εξειδικευμένης σχεδίασης.

# ΒΑΣΙΚΑ ΘΕΜΑΤΑ ΕΠΕΞΕΡΓΑΣΙΑ ΟΜΙΛΙΑΣ

- Ψηφιοποίηση ομιλίας
- Αναγνώριση ομιλίας / ομιλητή
- Σύνθεση ομιλίας
- Κωδικοποίηση ομιλίας

# ΕΠΕΞΕΡΓΑΣΙΑ ΦΩΝΗΣ

“Επεξεργασία φωνής είναι ένας τρόπος χρήσης της τεχνολογίας ομιλίας ή των τεχνολογιών επεξεργασίας ομιλίας για την επίτευξη της διάδοσης και ανάκλησης των πληροφοριών.”

Η επεξεργασία φωνής αποτελείται από μια ομάδα τεχνολογιών που συνδυάζονται με διάφορους τρόπους για την επίτευξη συγκεκριμένων λειτουργιών. Συχνά αλληλεπιδρά με άλλα περιβάλλοντα, όπως τράπεζες δεδομένων και τηλεφωνικά συστήματα σαν μέρος ενός μεγαλύτερου τηλεπικοινωνιακού συστήματος διαχείρισης πληροφοριών.



# European Network of Excellence in Human Language Technologies

- <http://www.elsnet.org/>
- Roadmap: <http://elsnet.dfki.de/>







# Language Resources



# Language Processing



# Language Usage

technology systems tools info

2001

2002

2003

2004

2005

2006

2007

2008

2009

2010

2002

2003

2004

2005

2006

2007

2008

2009

2010

2001

speech synthesis  
controlled languages  
ontologies for many domains  
moderate complexity dialogs  
machine learning  
annotation of pragmatic content  
handwriting recognition  
prosodic formant speech synthesis  
semi-automatic ontology construction  
top-level ontologies  
standardized ontological encoding  
semantic enrichment for websites  
generic annotation schemes  
broadcast transcription  
domain based ontological lexicons  
formant speech synthesis  
continuous speech recognition  
unsupervised machine learning  
OS speech recognition

adequate NL generation  
speech/text transcodes  
pragmatic text recognition  
semantic analysis of free text  
syntactic analysis for ill-formed input  
continuous speech understanding  
large coverage grammars  
NL generation verbalizes information  
databases question answering  
context sensitive summarization  
fast domain adaption  
semantic analysis of text  
analysis for ill-formed input  
analysis of uncertain input  
speech summarization  
portable dialogue systems  
grammar-based generation  
multi-document summarization  
syntactic analysis adaption tools  
minimized domain adaption effort  
intelligent presentation generation  
FAQ translation systems  
spoken sentence translation  
free linguistic generation  
summarization systems  
intelligent QA systems  
concept-to-speech  
prosodic formant speech  
speaker verification  
superficial semantic processing  
information assimilation  
summarization system  
markup of discourse  
rapid domain porting  
portable translation systems  
multilingual authoring  
cross-language text retrieval  
simple QA systems  
reduced annotation effort  
standardized interactivity coding  
syntactic analysis

reasonable QA systems  
medium conversational systems  
fully natural input/output  
translator's workbenches  
meeting summarization  
mixed initiative dialogue  
fully natural animated characters  
spoken translation systems  
meeting transcription systems  
personal assistants  
medium complexity dialogue systems  
vocabulary based conversational systems  
spoken dialogue systems for cars  
meta-communication for dialogue systems  
vocabulary based character dialogue  
spoken car navigation  
spoken dialogue systems  
task-oriented spoken dialogue  
correct database presentation  
integration of spoken communication  
high-quality auto responder  
pro-active dialogue  
recognition-based tutors  
written dialogues systems  
task-oriented interpretation  
spoken dialogue apps

# Κύριες Εταιρείες Τεχνολογιών Φωνής

- SCANSOFT [www.scansoft.com](http://www.scansoft.com)
- NUANCE [www.nuance.com](http://www.nuance.com)
- IBM <http://www-01.ibm.com/software/voice/>
- BEVOCAL [www.bevocal.com](http://www.bevocal.com)  
<http://www.nuance.com/care/solutions/ondemand/cafe/>  
Café VoiceXML development environment
- VOXPILOT [www.voxpilot.com](http://www.voxpilot.com)  
VocaBase a VoiceXML application server
- PlumVoice <http://www.plumvoice.com/developers>  
free VoiceXML test bed



# Κύριες Ελληνικές Εταιρείες

- DIALOGOS [www.speech.gr](http://www.speech.gr)
- OMILIA [www.omilia.com](http://www.omilia.com)
- VOICEWEB [www.voiceweb.gr](http://www.voiceweb.gr)
- e-RHETOR <http://www.e-rhetor.gr/>

# Έργο LibAccessATHENA

Υποστήριξη ΑμεΑ χρηστών βιβλιοθηκών ΕΚΠΑ

<http://speech.di.uoa.gr/libaccess/>



# Προσανατολισμός & καθοδήγηση στον δομημένο χώρο - πληροφόρηση σημάνσεων



Ερευνητικό Έργο ΜΝΗΣΙΚΛΗΣ:

Προηγμένες Καθολικές Υπηρεσίες Θέσης σε Εσωτερικούς Χώρους

<http://speech.di.uoa.gr/mnisiklis/>

# Ερευνητικό Έργο **ΟΜΗΡΟΣ**:

Μεθοδολογική Προσέγγιση για τη Σχεδίαση Ακουστικής-Απτικής Αλληλεπίδρασης στη Μη-Οπτική Διεπαφή Χρήστη (user interface) με Έμφαση στην Προσβασιμότητα Ατόμων με Απώλεια Όρασης

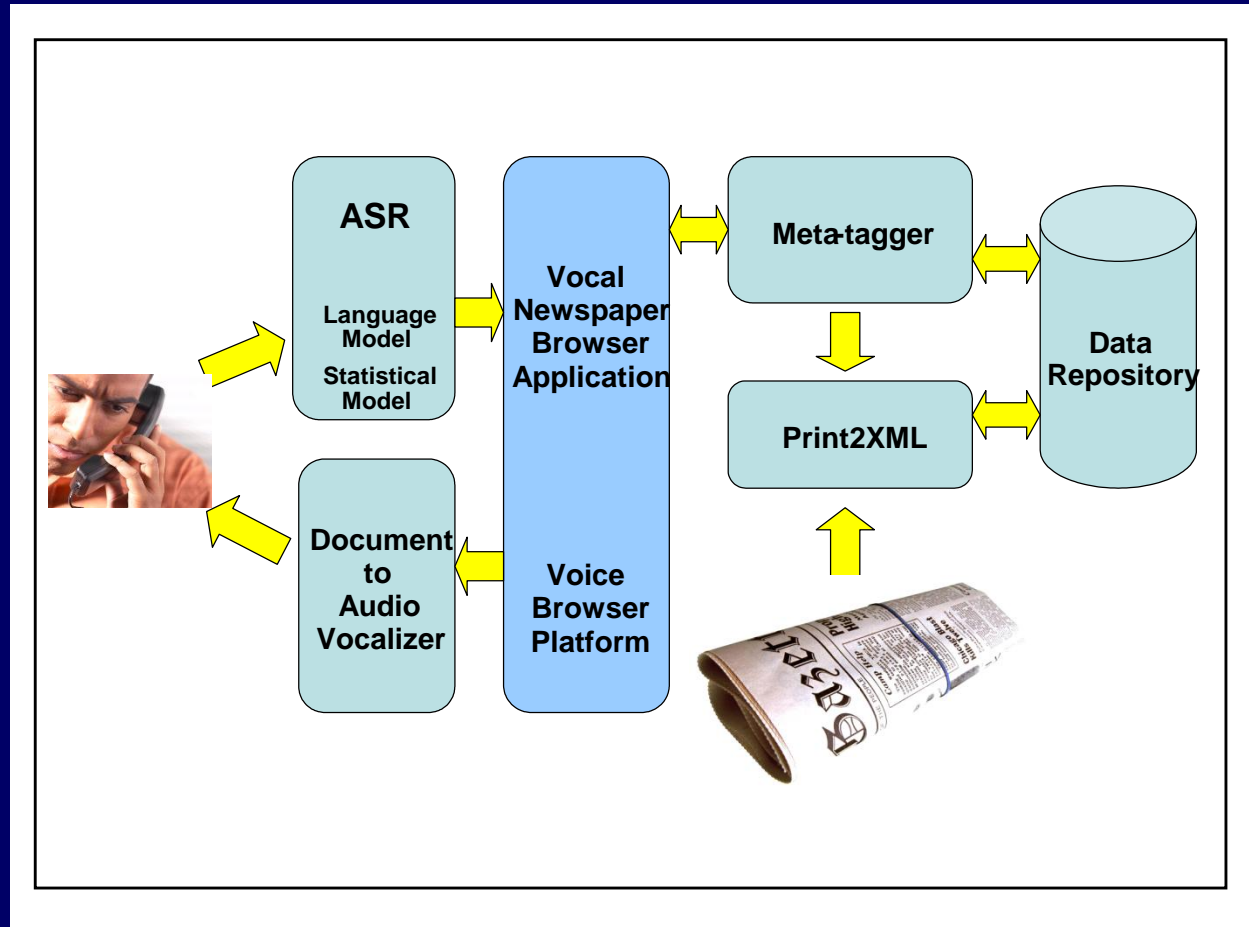


<http://speech.di.uoa.gr/homer/>



# Ερευνητικό έργο ΡΗΤΩΡ

Συστηματική προσέγγιση διεπαφής φωνητικού διαλόγου με ακουστικοποίηση της μεταπληροφορίας και οπτικής δομής εγγράφων με σκοπό την καθολική πρόσβαση σε έντυπο περιεχόμενο"



# ΔΗΜΟΣΘΕΝΗΣ:

## Σύστημα μετατροπής κειμένου σε συνθετική ομιλία

<http://demosthenes.di.uoa.gr/>

- Τεχνολογία συνθετικής φωνής
  - Φυσική ομιλία
  - Προσωδικά χαρακτηριστικά
- Χωρίς προ-ηχογράφηση
  - Εκφωνεί οποιοδήποτε κείμενο
  - Παραγωγή ομιλίας σε πραγματικό χρόνο
- Ταυτόχρονη υποστήριξη πολλών γλωσσών
  - Multilingual & polyglot



Ο Δημοσθένης είναι ένα πολυγλωσσικό σύστημα μετατροπής κειμένου σε συνθετική ομιλία, που αναπτύχθηκε στο Τμήμα Πληροφορικής του Πανεπιστημίου Αθηνών. Supports the English, but also μιλά εξίσου καλά και ελληνικά

ΔΗΜΟΣΘΕΝΗΣ 2005





# ερευνητικό έργο M-PIRO

<http://www.ltg.ed.ac.uk/mpiro/>



*Figure 2: The 3D collection of the M-PIRO virtual reality prototype: spoken descriptions are generated for each selected object.*

# Ερευνητικό έργο ΑΙΝΕΙΑΣ: Ανάπτυξη Ευέλικτων Συστημάτων Εναλλακτικής & Επαυξητικής Διαπροσωπικής Επικοινωνίας



# Συλλογή Δωρεάν Λογισμικού ΑΘΗΝΑ

στοχεύει να ενημερώσει αλλά και να παρέχει στα Άτομα με Αναπηρία και τους επαγγελματίες λύσεις Υποστηρικτικών Τεχνολογιών Πληροφορικής που δεν έχουν κόστος (Open Source ή Freeware)

<http://access.uoa.gr/ATHENA/>



Εθνικό & Καποδιστριακό Πανεπιστήμιο Αθηνών  
Τμήμα Πληροφορικής και Τηλεπικοινωνιών  
Εργαστήριο Φωνής και Προσβασιμότητας

ΑΘΗΝΑ: Συλλογή Δωρεάν Λογισμικού Υποστηρικτικών Τεχνολογιών για ΑμεΑ

Αρχική σελίδα Όλα τα βοηθήματα Συντελεστές Επικοινωνία Αναζήτηση:  Αναζήτηση

Εμφάνιση όλων των εφαρμογών  
[Προβολή](#)

Παρουσίαση κατά Κατηγορία

- [DAISY \(4\)](#)
- [Ανάγνωση Οθόνης \(6\)](#)
- [Αναγνώριση Ομιλίας \(3\)](#)
- [Αντίθεση/Χρώματα \(9\)](#)
- [Αριθμομηχανές \(2\)](#)
- [Δείκτες Ποντικού \(8\)](#)
- [Διευκόλυνση Κλικ \(15\)](#)
- [Εικονικά Πληκτρολόγια \(3\)](#)
- [Έλεγχος με το Κεφάλι \(3\)](#)
- [Εναλλακτική Επικοινωνία \(2\)](#)
- [Κείμενο σε Ομιλία \(10\)](#)
- [Μεγέθυνση Οθόνης \(14\)](#)
- [Μετατροπή σε Braille \(3\)](#)
- [Πλοήγηση στο Διαδίκτυο \(1\)](#)
- [Προσβασιμότητα Κειμένων \(13\)](#)
- [Προσομοίωση Ποντικού \(5\)](#)
- [Πρόβλεψη Λέξεων \(1\)](#)
- [Συζήτηση \(2\)](#)
- [Συντομοίσεις Πληκτρών \(8\)](#)
- [Φωνητικό Ταχυδρομείο \(1\)](#)
- [Ψηφιακά Ρολόγια \(2\)](#)

Ομάδα φωνής και προσβασιμότητας > [Ελεύθερα διαθέσιμα βοηθήματα](#) > Αρχική

**Η Συλλογή Δωρεάν Λογισμικού ΑΘΗΝΑ** στοχεύει να ενημερώσει αλλά και να παρέχει στα Άτομα με Αναπηρία και τους επαγγελματίες λύσεις **Υποστηρικτικών Τεχνολογιών Πληροφορικής** που δεν έχουν κόστος (Open Source ή Freeware).

Συγκεκριμένα στη συλλογή ΑΘΗΝΑ παρουσιάζονται και προσφέρονται με ένα οργανωμένο και συστηματικό τρόπο τα βοηθήματα Υποστηρικτικών Τεχνολογιών που επιλέχθηκαν ύστερα από έρευνα στο Διαδίκτυο και δοκιμάστηκαν στο [Εργαστήριο Φωνής και Προσβασιμότητας](#) του Πανεπιστημίου Αθηνών. Για κάθε Ελεύθερο διαθέσιμο Λογισμικό δίνεται μια τυποποιημένη σύντομη περιγραφή με τις περισσότερο χρήσιμες πληροφορίες όπως: όνομα προϊόντος, έκδοση, κατασκευαστής, κατηγορία προϊόντος, σχετιζόμενες αναπηρίες, περιγραφή εφαρμογής, λειτουργικό σύστημα, διαδικασία εγκατάστασης, ρυθμίσεις, και τρόπος λήψης του λογισμικού.

Τύφλωση Κινητική αναπηρία Χαμηλή Όραση Απόλυση Ακοής Αναπηρία Λόγου Δυσλεξία

**Μπορείτε να πλοηγηθείτε στην σελίδα της Συλλογής ΑΘΗΝΑ με τους εξής τρεις τρόπους:**

**"Παρουσίαση κατά Αναπηρία"** Επιλέγοντας τη συγκεκριμένη κατηγορία αναπηρίας που επιθυμείτε εμφανίζονται όλα τα βοηθήματα που περιλαμβάνονται σε αυτήν.

**"Παρουσίαση κατά Βοήθημα"** Επιλέγοντας μια από τις κατηγορίες βοηθημάτων εμφανίζονται όλα τα βοηθήματα που ανήκουν σε αυτή την κατηγορία.

Τέλος, επιλέγοντας **"Εμφάνιση Όλων"** μπορείτε να δείτε όλα τα βοηθήματα.

Οι Τεχνολογίες Πληροφορικής προσφέρουν σήμερα σημαντικές δυνατότητες που διευκολύνουν την καθημερινή ζωή και την κοινωνική ένταξη των Ατόμων με Αναπηρία (ΑμεΑ) και των ηλικιωμένων. Οι λύσεις αυτές που ονομάζονται Υποστηρικτικές Τεχνολογίες, αποτελούνται από ειδικό εξοπλισμό και λογισμικό για Η/Υ και παρουσιάζονται με συστηματικό τρόπο στην ιστοσελίδα <http://e-bility.gr>.



# Περιεχόμενο του μαθήματος (1/2)

- Βασικά χαρακτηριστικά σημάτων ομιλίας.
- Μηχανισμοί και πρότυπα παραγωγής ομιλίας.
- Ακοή και αντίληψη ομιλίας.
- Μέθοδοι ψηφιακής ανάλυσης σημάτων ομιλίας.
- Η μέθοδος της γραμμικής πρόβλεψης.
- Ψηφιακή κωδικοποίηση ομιλίας.

# Περιεχόμενο του μαθήματος (2/2)

- Μέθοδοι σύνθεσης ομιλίας.
- Μετατροπή κειμένου σε ομιλία.
- Μέθοδοι αναγνώρισης ομιλίας.
- Η ομιλία στην επικοινωνία ανθρώπου - μηχανής.
- Εφαρμογές στα συστήματα πληροφορικής και επικοινωνιών.

# [www.di.uoa.gr/speech/com](http://www.di.uoa.gr/speech/com)

- Εργαστηριακός εξοπλισμός
- Βιβλιογραφία
- **Ορολογία**
- Αρκτικόλεξα
- Διεθνές φωνητικό Αλφάβητο
- Ενδιαφέρουσες Ιστοσελίδες: Μάθηση μέσω internet  
Επιστημονικές ενώσεις  
Πανεπιστήμια  
Περιοδικά  
Λογισμικό  
Συστήματα Επίδειξης

XUEDONG HUANG | ALEX ACERO | HSIAD-WUEN HON

# S P O K E N

## LANGUAGE PROCESSING

*A Guide to Theory, Algorithm, and System Development*



Foreword by Dr. Raj Reddy  
Carnegie Mellon University



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

*Μαθήματα*  
**Επεξεργασίας Ομιλίας**

Γεωργίου Θ. Κουρουπέτρογλου

Αθήνα 2004



# Εργαστήριο

Θέμα 1: Φασματική ανάλυση. Συναρτήσεις παραθύρου. Ψηφιακά φίλτρα. Ανάλυση σε Χρόνο-Συχνότητα (Φασματογράφημα).

Θέμα 2: Φασματογράφημα στενής και ευρείας ζώνης, ενός σήματος ομιλίας. Προέμφαση της ομιλίας. Παράμετροι ομιλίας (Ενέργεια, Pitch, Formants, LPC).

Θέμα 3: Dynamic Time Warping (DTW). Hidden Markov Models (HMM).

Θέμα 4: Αναγνώριση ομιλίας.

Θέμα 5: Σύνθεση ομιλίας.

Εργαλεία:

- Cool edit 2000
- SFSwin (Speech Filing System)
- HVite
- PRAAT



<http://eclass.di.uoa.gr/>

## **ΕΠ19 - Επεξεργασία Ομιλίας και Φυσικής Γλώσσας**

- Βαθμολογία: εξετάσεις 70%
- Εργαστήριο: 30% (οι ασκήσεις παραδίδονται σε έντυπη και ηλεκτρονική μορφή)

# Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αθηνών» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Σημειώματα

# Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση 1.0.

# Σημείωμα Αναφοράς

Copyright Εθνικών και Καποδιστριακών Πανεπιστήμιον Αθηνών, Γεώργιος Κουρουπέτρογλου 2015. «Επεξεργασία ομιλίας και φυσικής γλώσσας. Εισαγωγή». Έκδοση: 1.0. Αθήνα 2015. Διαθέσιμο από τη δικτυακή διεύθυνση: <http://opencourses.uoa.gr/courses/DI36/>.

# Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση Παρόμοια Διανομή 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Ως Μη Εμπορική ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.

# Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.



# Σημείωμα Χρήσης Έργων Τρίτων

- "Η δομή και οργάνωση της παρουσίασης, καθώς και το υπόλοιπο περιεχόμενο, αποτελούν πνευματική ιδιοκτησία της συγγραφέως και του Πανεπιστημίου Αθηνών και διατίθενται με άδεια Creative Commons Αναφορά Μη Εμπορική Χρήση Παρόμοια Διανομή Έκδοση 4.0 ή μεταγενέστερη.
- Οι φωτογραφίες που περιέχονται στην παρουσίαση αποτελούν πνευματική ιδιοκτησία τρίτων. Απαγορεύεται η αναπαραγωγή, αναδημοσίευση και διάθεσή τους στο κοινό με οποιονδήποτε τρόπο χωρίς τη λήψη άδειας από τους δικαιούχους. "