

---

# Τεχνολογίες Φωνής στο Ηλεκτρονικό Εμπόριο και τις Υπηρεσίες Διαμεταγωγής

---

Νάσος Κατσαμάνης (nkatsam@cs.ntua.gr)  
**Εθνικό Μετσόβιο Πολυτεχνείο**  
**Σχολή Ηλεκτρολόγων Μηχανικών**  
**Ομάδα Όρασης Υπολογιστών, Επικοινωνίας Λόγου**  
**και Επεξεργασίας Σημάτων**  
**<http://cvsp.cs.ntua.gr>**

# Εισαγωγή

- E-Commerce → V-Commerce
- Σύνθεση Φωνής, Αναγνώριση Φωνής, Αναγνώριση Ομιλητή
- Προοπτικές και δυνατότητες εφαρμογής των τεχνολογιών φωνής
- Επεξεργασία Φυσικής Γλώσσας και Διαλογικά Συστήματα
- Πλεονεκτήματα: Ευχρηστία, Προσβασιμότητα

# Περιεχόμενα

- Ηλεκτρονικό Εμπόριο, Υπηρεσίες Διαμεταγωγής
  - Τι είναι; Ιστορία
  - Πλεονεκτήματα
  - Περιορισμοί
- Συστήματα IVR (Interactive Voice Response Systems)
  - Πλεονεκτήματα
- Εξελίξεις στο χώρο των τεχνολογιών Φωνής
- Ολοκλήρωση και Εφαρμογή
  - Διαλογικά Συστήματα, VoiceXML
- Σενάρια Εφαρμογής Τεχνολογιών Φωνής στις Ηλεκτρονικές Συναλλαγές
- Πλεονεκτήματα, Περιορισμοί, Συμπεράσματα

---

# Ηλεκτρονικό Εμπόριο και Υπηρεσίες Διαμεταγωγής

- Μέσο: Δίκτυα Υπολογιστών
  - Intranet, Extranet, Internet
  - Transaction Services
- Συναλλαγές , υπηρεσίες:
  - Αγορές, Πωλήσεις αγαθών (μεταξύ εταιριών, μεταξύ καταναλωτών-εταιριών)
  - Διαχείριση Κεφαλαίων ( e-banking, e-investment)
  - Υπηρεσίες Εκπαίδευσης ( e-education )
  - Υπηρεσίες Πληροφόρησης, Ψυχαγωγίας ( e-news, games )
  - Διοικητικές Υπηρεσίες ( e-government )

---

# Ηλεκτρονικό εμπόριο και Υπηρεσίες Διαμεταγωγής, Ιστορία

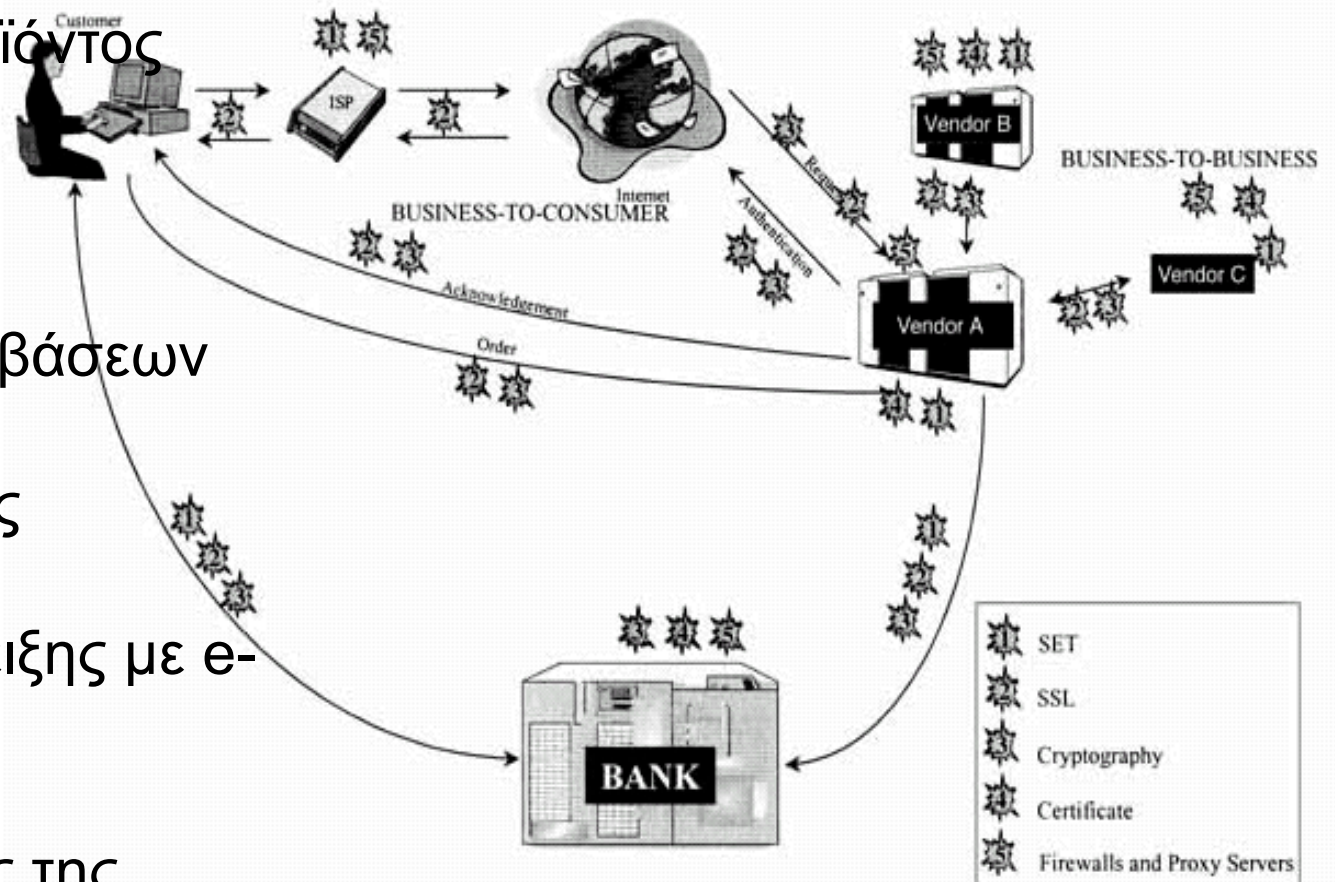
- Ιστορία:

- Electronic Data Interchange (1984),
- CompuServe (1992),
- Netscape (SSL, 1994),
- Amazon.com, eBay.com (1995)
- Φορολογικές ελαφρύνσεις μέχρι 2005 (ΗΠΑ, 2000)

- \$12,2 δισ. πωλήσεις το 2003

# Ηλεκτρονικό Εμπόριο και Υπηρεσίες Διαμεταγωγής, Σενάριο

- Επιλογή του προϊόντος
- Πληρωμή
- Πιστοποίηση  
Αυθεντικότητας
- Ενημέρωση των βάσεων  
της εταιρίας
- Δρομολόγηση της  
παραγγελίας
- Αποστολή απόδειξης με e-  
mail
- Δυνατότητα  
παρακολούθησης της  
παραγγελίας



Πηγή: Information Systems Audit and Control Association, <http://www.isaca.org/ecom.htm>

# Ηλεκτρονικές Συναλλαγές, Πλεονεκτήματα, Περιορισμοί

- Πλεονεκτήματα:
  - Άνεση, Ταχύτητα, Ευκολία
  - 24 ώρες το 24ωρο, 7 μέρες την εβδομάδα ( 24 x 7 )
  - Περιορισμένο κόστος, Μικρότερη Τιμή
  - Παγκοσμιοποιημένη Αγορά
  - Παγκοσμιοποιημένες Υπηρεσίες (π.χ. Εκπαίδευση)
- Περιορισμοί
  - Σχετικά μικρή προσβασιμότητα στο internet, έλλειψη εξοικείωσης
  - Ελλιπής Διαδικτυακή υποδομή, περιοριστικές συνδέσεις internet, λίγες online εταιρίες, online υπηρεσίες ( περιορισμένο e-government στην Ελλάδα )
  - Άτομα Ηλικιωμένα ή με ειδικές ανάγκες αποκλεισμένα
  - Ασφάλεια, μόνιμη ανησυχία
  - Προϊόντα μη κατάλληλα για online αγορές ( π.χ. ρούχα )
- ✓ ~~Εξαιρούνται χώρες όπως ΗΠΑ, Σουηδία, Νορβηγία, Καναδάς !~~

# Διαδραστικά Συστήματα με Φωνή

- Κλασσική Εναλλακτική: Call Centers
  - Πρόσβαση μέσω τηλεφώνου
  - Ακριβή λύση, περιορισμένο ωράριο
- Interactive Voice Response (IVR) Systems
  - Πρόσβαση μέσω τηλεφώνου
  - Αυτόματα συστήματα, δεν χρειάζονται άνθρωποι-χειριστές
  - Υποστηρικτικά των Call Centers για διαχείριση μεγάλης κίνησης
  - Συνεχές ωράριο
- Σημαντικές Αδυναμίες
  - Δύσχρηστα μενού
  - Περιοριστικό το πληκτρολόγιο (12 πλήκτρα)
- Σύγχρονη εναλλακτική: Κινητά τελευταίας τεχνολογίας, PDA (Personal Digital Assistant), WAP (Wireless Application Protocol)
  - Περιορισμένες δυνατότητες πλοήγησης, ασφάλεια κινητών επικοινωνιών (;), μικροσκοπικά πληκτρολόγια



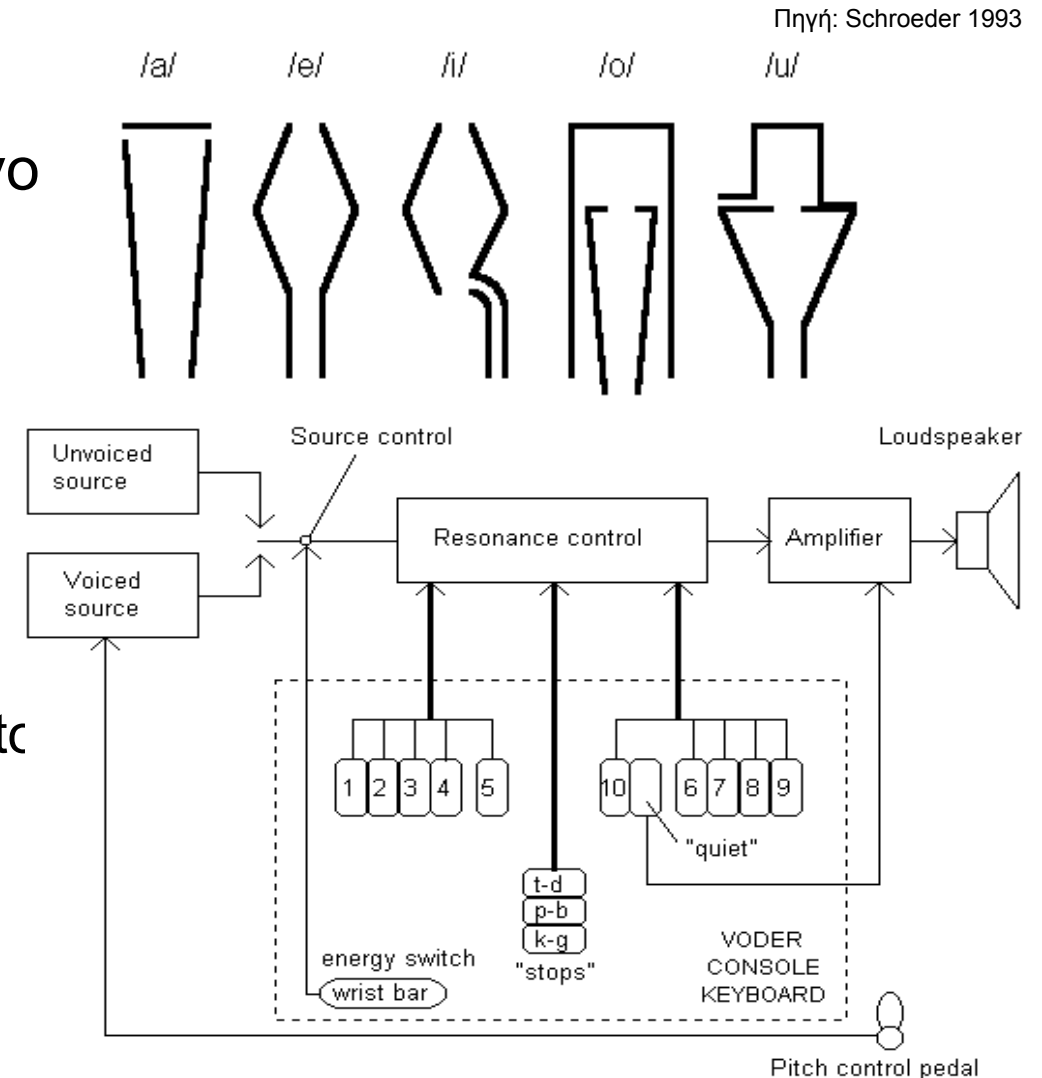
---

# Τεχνολογίες Φωνής, Αναγνώριση Φωνής

- Αυτόματη Αναγνώριση Φωνής (Automatic Speech Recognition)
  - Αναγνώριση Μεμονωμένων Αριθμών (Bell Labs, 1952)
  - 2001: Οδύσσεια του Διαστήματος (1968)
  - Τεχνολογία Αναγνώρισης Μεμονωμένων Λέξεων (Δεκαετία 1970)
  - Αναγνώριση Συνεχούς Λόγου, ανεξαρτήτως ομιλητή
  - Υπαγόρευση Κειμένου σε Υπολογιστή

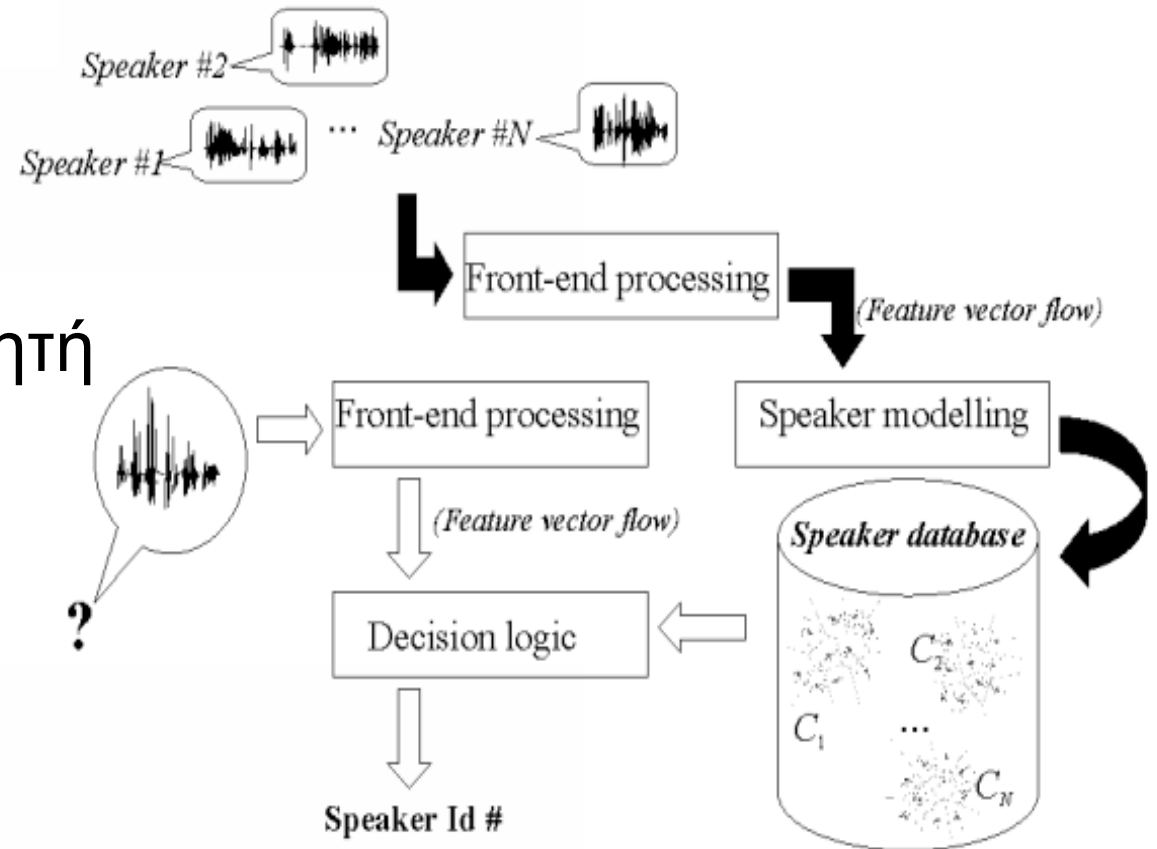
# Τεχνολογίες Φωνής, Σύνθεση Φωνής

- Σύνθεση Φωνής από Κείμενο (Text-to-Speech Synthesis)
  - Σύνθεση Φωνηέντων με χρήση Μηχανικών Ταλαντωτών, προσομοίωση φωνητικής οδού (Kratzenstein, 1779)
  - Ηλεκτρικοί Συνθέτες Φωνής (Voice Operating Demonstration, Dudley 1939)
  - Σύνθεση Φωνής με χρήση βάσεων δεδομένων
  - Ο υπολογιστής μιλάει



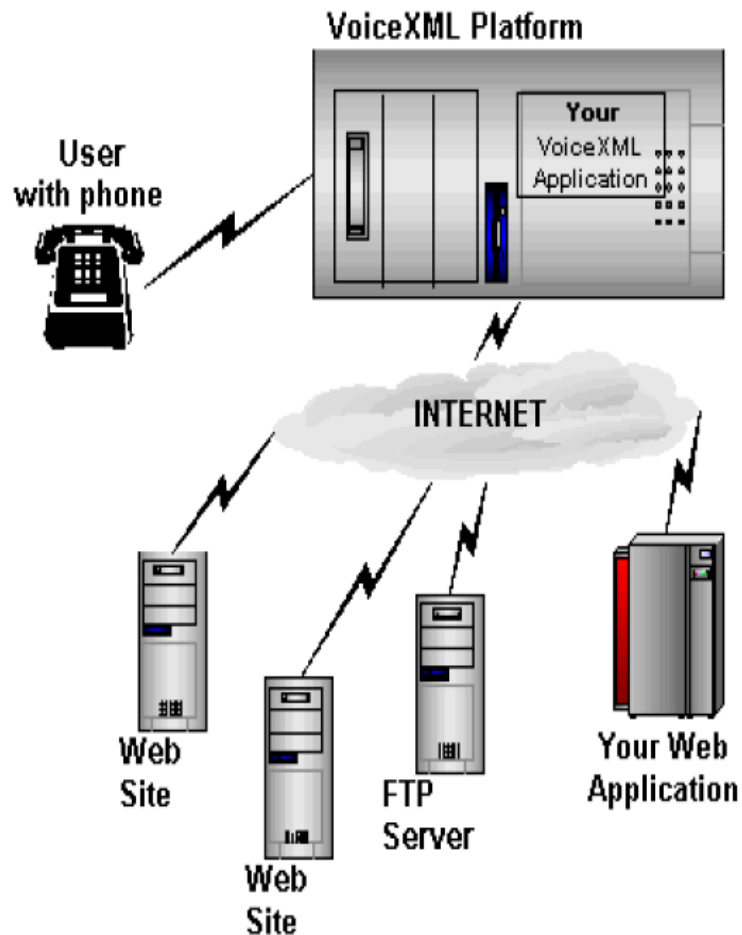
# Τεχνολογίες Φωνής, Αναγνώριση Ομιλητή

- Αναγνώριση Ταυτότητας Ομιλητή
- Εξακρίβωση Αυθεντικότητας Ομιλητή
- Βιομετρική μέθοδος
- Συνδυάζεται και με άλλες μεθόδους συνήθως



Πηγή <http://www.cs.joensuu.fi/pages/tkinnu/research/>

# Ολοκλήρωση και Εφαρμογή



- VoiceXML (Voice Extensible Markup Language)
  - Ιδρυτές: AT&T, Motorola, Lucent Technologies, IBM
  - VoiceXML 2.0 γλώσσα πρότυπο για διαδραστικά συστήματα διαλόγων με φωνή ( W3C )
- Παροχείς φωνητικών υπηρεσιών
- Φτηνό Υλικό
- Πλαίσιο Ανάπτυξης των Εφαρμογών (Application Development Framework)
- Επεξεργασία Φυσικής Γλώσσας και Διαλογικά Συστήματα

```
<?xml version='1.0'?>
```

```
<vxml version="2.0">
```

```
<menu id="Simple_Example">
```

```
<prompt> <audio>Welcome to my voice mail. Say Mail to leave me a voice mail and Operator to return back to our operator.</audio> </prompt>
```

```
<choice next="voicemail.vxml"> Mail </choice>
```

```
<choice next="operator.vxml"> Operator </choice>
```

```
<default><reprompt/></default>
```

```
</menu>
```

```
</vxml>
```

Πηγή: <http://cafe.bevocal.com/docs/tutorial/tutorial.pdf>

# Ηλεκτρονικές Συναλλαγές με Φωνή, Σενάρια

- Φωνητική Πλοήγηση σε ιστοσελίδες
  - Προϋπόθεση: Σωστή, Απλή Σχεδίαση, Γρήγορο Internet
- Συναλλαγές «εν θερμώ»
  - Απλά ένα μικρόφωνο σε κάθε συσκευή, π.χ. ραδιόφωνο
  - π.χ. αγορά εισιτηρίων για συναυλία
- Ψηφιακό Πορτοφόλι
  - Αναγνώριση Ομιλητή
- Ευκολία Χρήσης, Φιλικό περιβάλλον

# Πλεονεκτήματα, Περιορισμοί

## ■ Πλεονεκτήματα

### □ Για όλους:

- Ολοκλήρωση των τεχνολογιών, ευκολία χρήσης, συνεχής εξυπηρέτηση
- Γραφειοκρατία ;

### □ Για άτομα με ειδικές ανάγκες, απλά με τη χρήση της φωνής τους:

- Αγορά , επιλογή αγαθών (π.χ. τρόφιμα, [www.tesco.com](http://www.tesco.com)), δυνατότητα αποστολής στο σπίτι
- Πρόσβαση σε πολλές υπηρεσίες που σήμερα χρειάζονται αρκετή εξοικείωση με τους υπολογιστές ( διοικητικές υπηρεσίες, ενημέρωση, ψυχαγωγία )
- Δυνατότητες απασχόλησης

## ■ Περιορισμοί

- Θόρυβος, Αναγνώριση Συνεχούς Λόγου ;
- Φυσικότητα, Καταληπτότητα, Σύνθεση Φωνής ;
- Ασφάλεια, Αναγνώριση Ομιλητή ;
- Ελεύθεροι Διάλογοι, Επεξεργασία Φυσικής Γλώσσας ;
- Σχεδίαση κατάλληλων διεπαφών

## Συμπεράσματα

- Ανθρώπινο Ηλεκτρονικό Εμπόριο, ανθρώπινο Διαδίκτυο με την εφαρμογή των τεχνολογιών φωνής
- Πρόσβαση από παντού ( Σταθερά Τηλέφωνα, Υπολογιστές, Κινητά ) κάθε στιγμή
- Οφέλη για όλους (καταναλωτές, πολίτες, εταιρίες)
- Μεγάλη Δυναμική
- Φωνητική πύλη COSMOTE, Minoan Flying Dolphins, VoiceWeb (JUPITER από MIT στις ΗΠΑ)
- Προοπτικές για το μέλλον
  - Ακόμα οι τεχνολογίες φωνής δεν έχουν τελειοποιηθεί
  - Ελλάδα, το κράτος και η κοινωνία δεν έχουν επενδύσει όσο θα έπρεπε