

Agent Asistent Inteligent
Prelucrarea și generarea vorbirii

Mihalea Matei
Conf. Dr. Ing. Andrei Olaru
04.07.2017

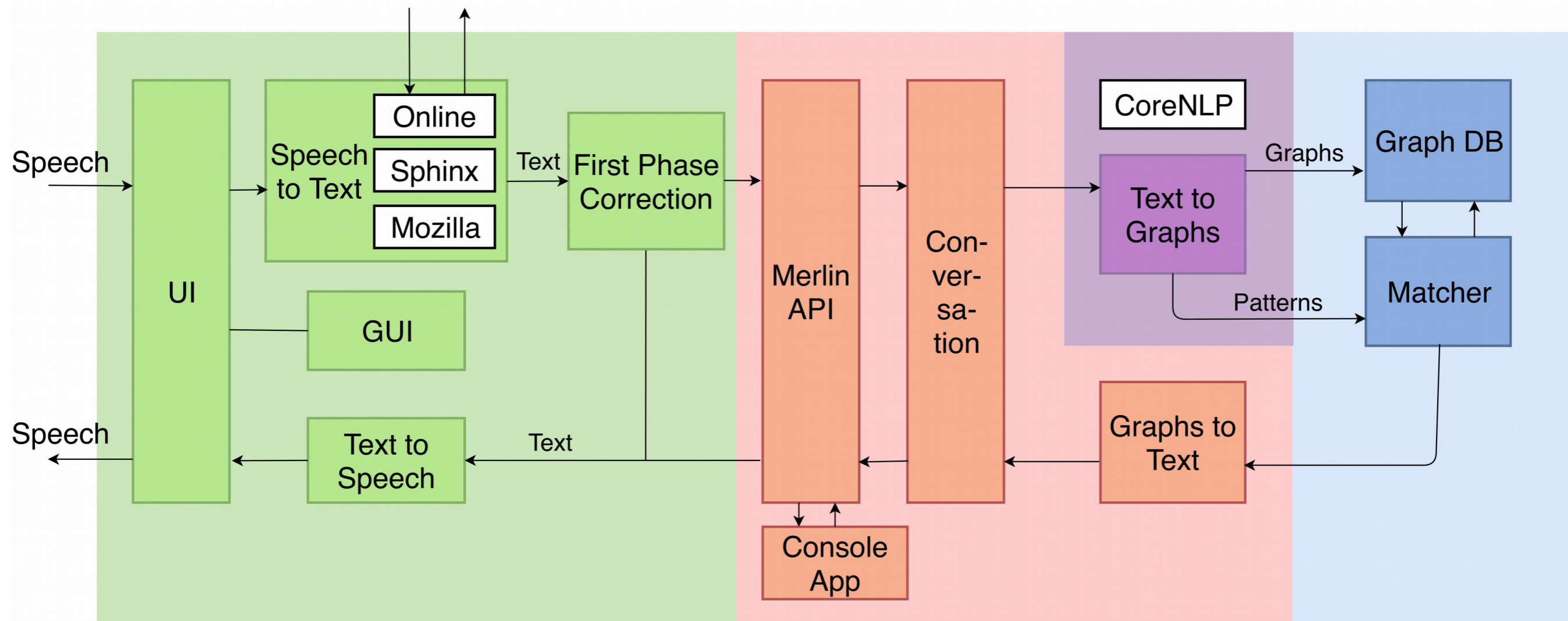
University Politehnica of
Bucharest



Introducere

- parte a proiectului Agent Asistent Inteligent
- are ca obiective implementarea unui asistent inteligent:
 - local
 - open-source

Arhitectura proiectului



Recunoașterea vocii

- Sisteme cloud
 - Bazate pe efectuarea de deep-learning pe date masive
 - Performanțe foarte bune
- Sisteme locale
 - Permit utilizatorilor antrenarea și folosirea propriilor modele
 - Dimensiunea datelor - limitată

Recunoașterea vocii – State of the Art

- Sisteme cloud
 - Google Cloud Speech API
 - Microsoft Bing Voice Recognition
 - IBM Speech to Text
- Sisteme locale
 - Mozilla DeepSpeech
 - CMU Sphinx
 - Kaldi

Recunoașterea vocii în cadrul proiectului

- Sphinx4
- Mozilla DeepSpeech
- Google Cloud Speech API

Înregistrarea comenzilor utilizatorului

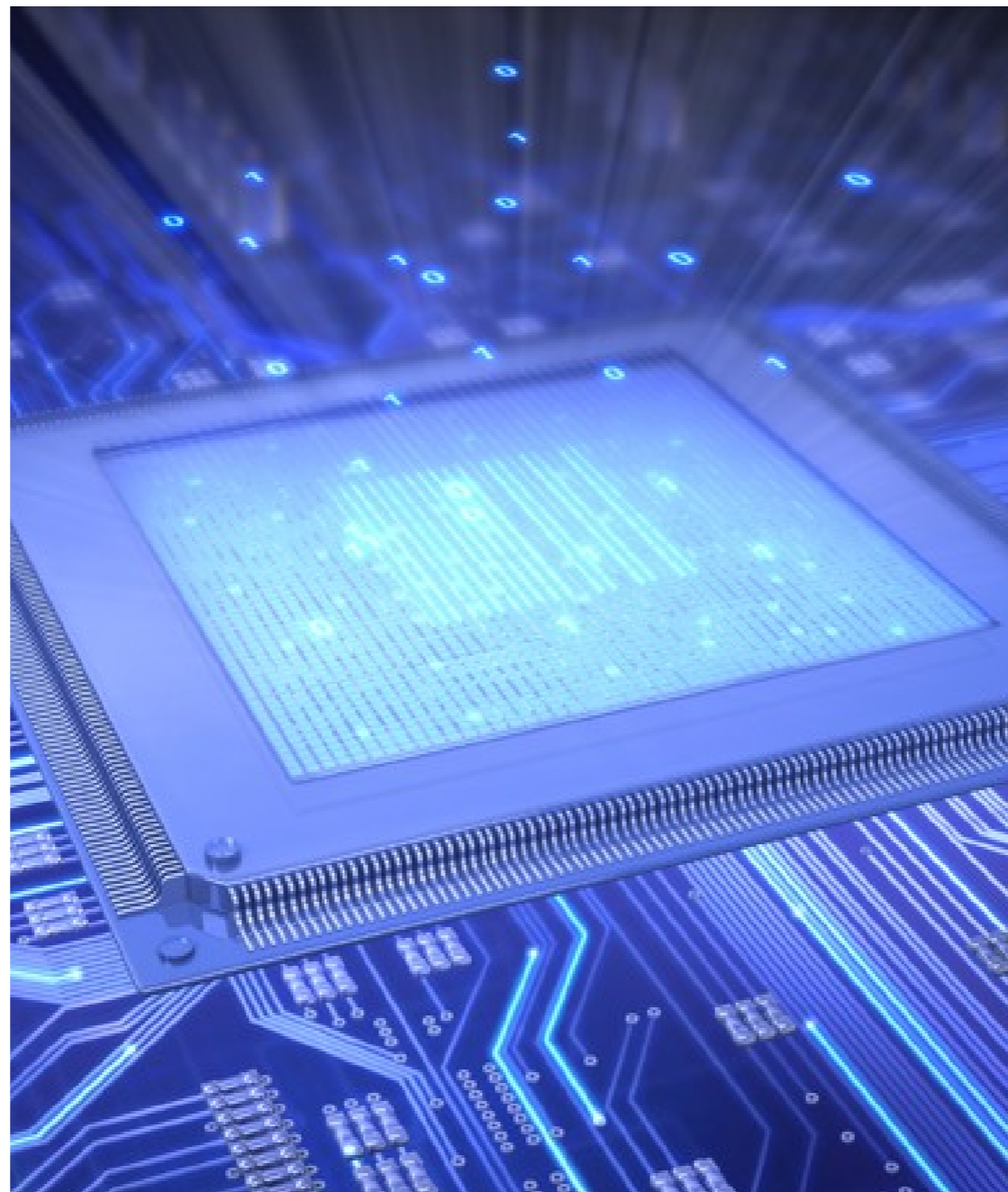
Am ales folosirea fișierelor:

- ✓ Debugging
 - ✓ Persistență
 - ✓ Corectarea erorilor
- x Viteză de execuție

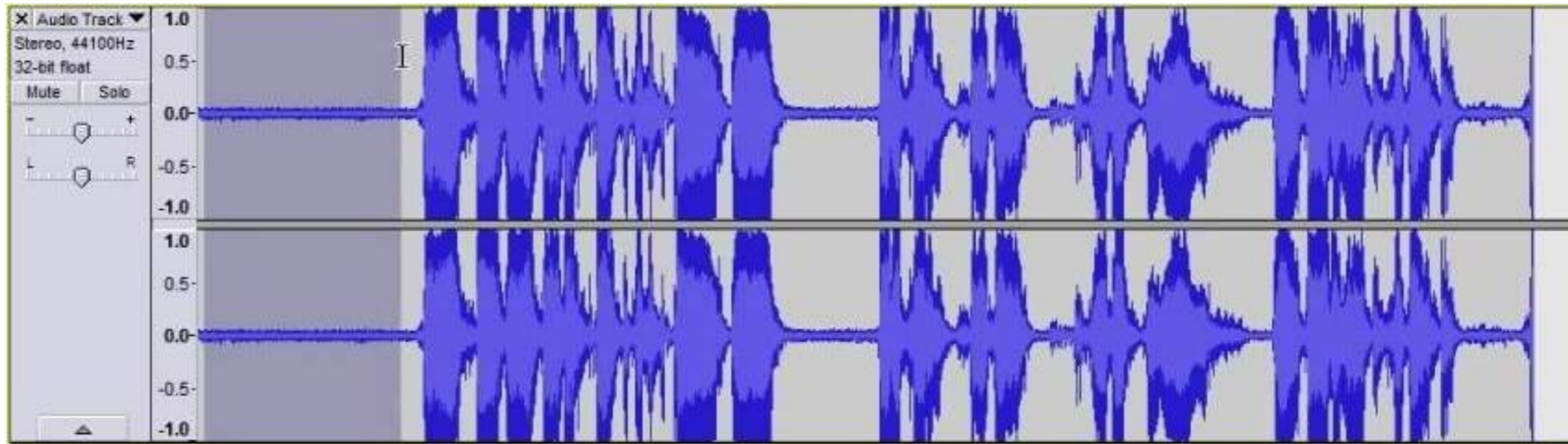
Modalități de înregistrare:

- Folosind un microfon
- Înregistrarea de pe placa de sunet

Corectarea erorilor



Eliminarea zgomotului



- Scop
 - Performanța aplicației în orice condiții
- Funcționare
 - 1) Se creează profilul zgomotului
 - 2) Pe baza profilului, se elimină zgomotul pentru tot fișierul

GUI



1. Buton pentru începerea înregistrării
2. Selectează durata înregistrării
3. Selectează modalitatea de traducere / înregistrare fișier
4. Bară de status

Resultate - 1

Resultate - 2

Thank you for your attention.



**Computer Science
& Engineering
Department**

matei.mihalea@cs.pub.ro
cs.pub.ro