



ARM-1: System automatycznej konwersji mowy polskiej na tekst

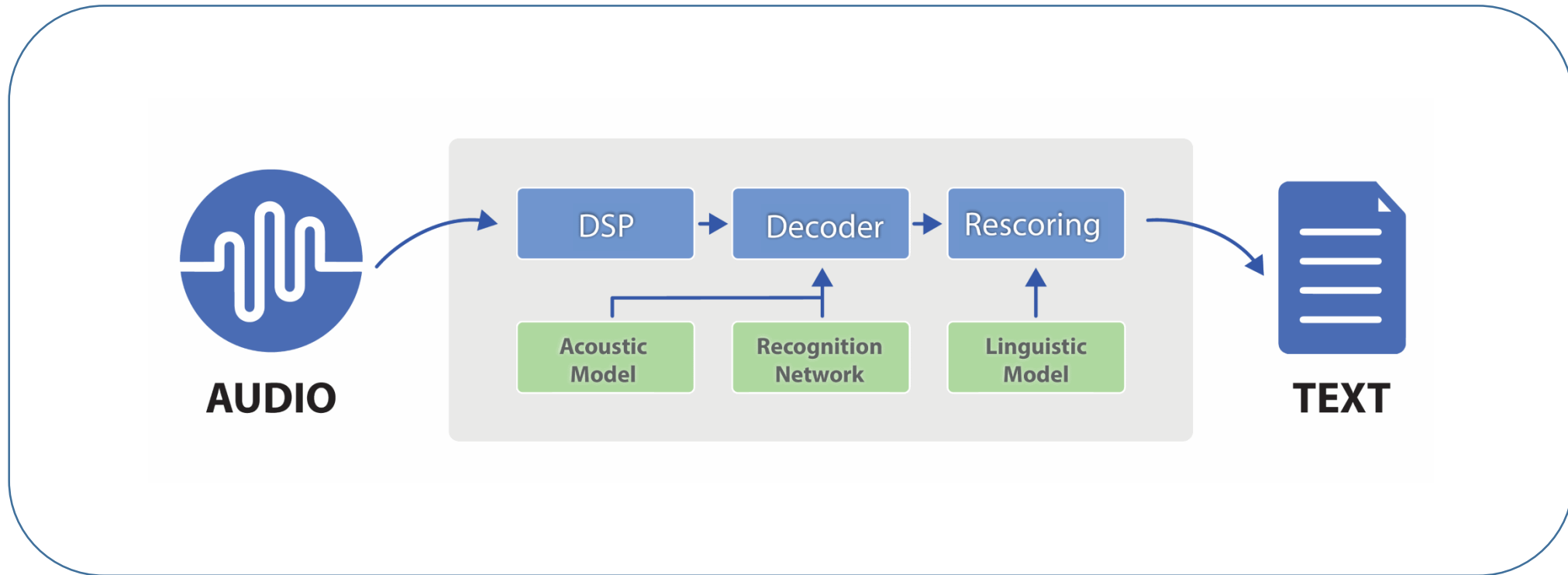
Prezentujący:
mgr. inż. Jerzy Jamroży



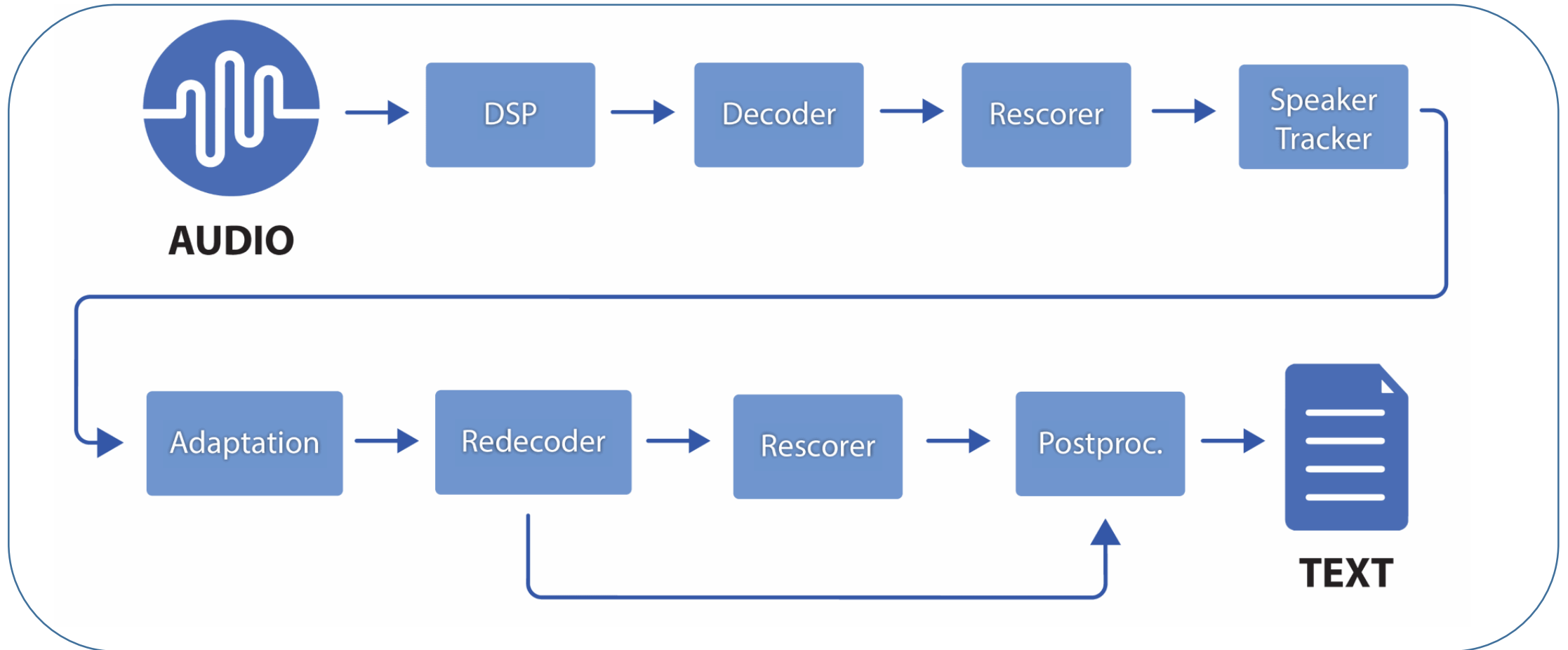
- **ARM-1 powstał w ramach projektu rozwojowy nr DOBR/0008/R/ID1/2013/03** współfinansowany ze środków NCBR, w oparciu o wcześniejsze badania i wyniki własne
- **Kluczowe założenia techniczne:**
 - Implementacja oparta o ukryte modele markowa (HMM) oraz modele gaussa (Gaussian Mixture Model)
 - 3-gramowy model językowy
- **Czas realizacji:** kwiecień 2013 r. – wrzesień 2016 r.
- Po zakończeniu projektu rozszerzenie o możliwość modelowania mówców oraz automatyczną, nienadzorowaną adaptację on-line



Silnik ARM-1



Silnik ARM-1 z nienadzorowaną adaptacją akustyczną



Akustyczna baza danych użyta do uczenia modelu akustycznego:

- mowa dyktowana: 205h
- wypowiedzi sejmowe: 58h
- programy radiowe: 20h
- programy telewizyjne: 40h
- dialogi telefoniczne: 40h
- nagrania sądowe

Nagrania zostały opisane na poziomie fonemów



Uczenie modelu akustycznego:

- wyznaczenie rozkładów parametrów audio dla stanów triphone'owych
- klastering w oparciu o 60 parametrów kontekstowych fonemów

Korpus tekstowy

- gazety
- czasopisma
- przemówienia sejmowe
- ogólnie dostępne zasoby Internetu

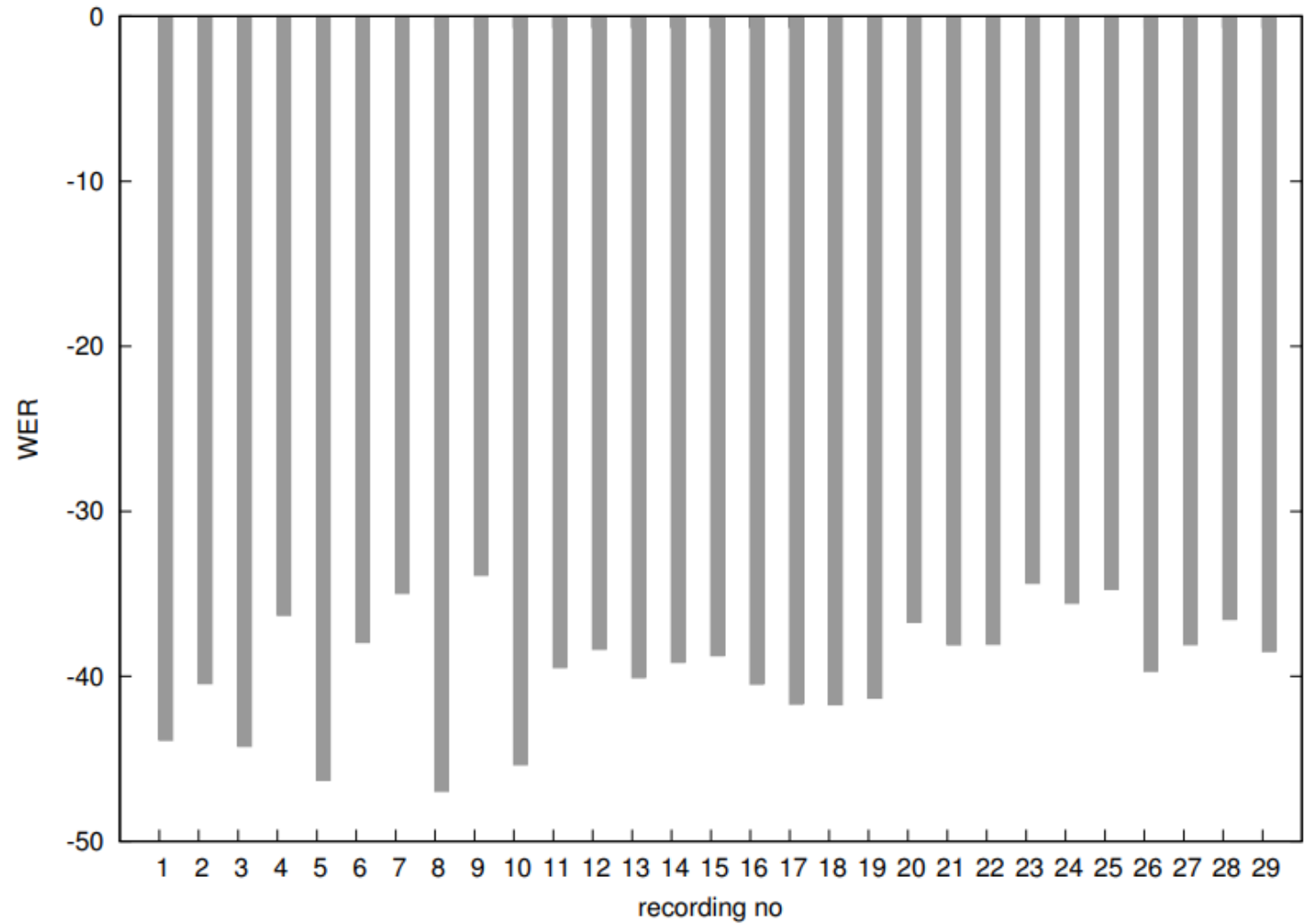
Razem ok. 20GB tekstu



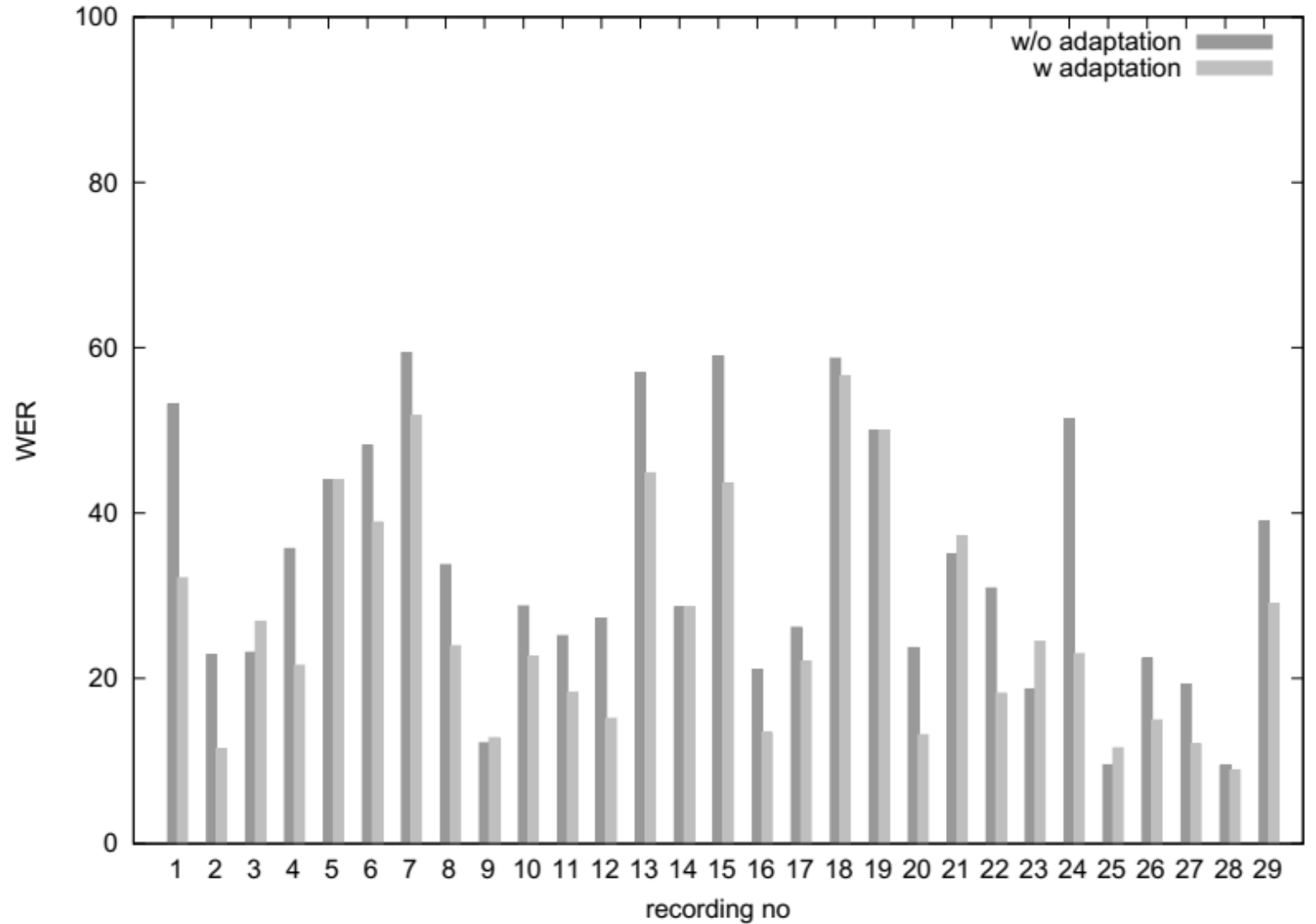
Uczenie modelu językowego

- model N-gram trzeciego stopnia
- Statystyki:
 - 280k unigramów
 - 74M bigramów
 - 60M trigramów

- Wyniki na nagraniach konkursowych



- Różnica pomiędzy oryginalną wersją silnika a wersją z adaptacją nienadzorowaną





Dziękuję za uwagę!

Poznańskie Centrum Superkomputerowo-Sieciowe

Dział nowych interfejsów z użytkownikiem

ul. Jana Pawła II 10, 61-139 Poznań

e-mail: arm@man.poznan.pl

www.pcass.pl, www.speechlabs.pl
