



Нацыянальна свядомы штучны інтэлект

Ю.Адамаў, EPAM, Minsk

23 жніўня 2014

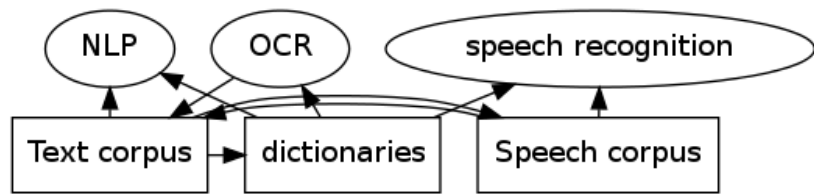


Што маецца на ўвазе

Think IBM Watson

Машына, якая можа размаўляць з чалавекам і самастойна збіраць інфармацыю з чалавечых тэкстаў

- OCR
- Аналіз тэкстаў і збор звестак з іх (NLP)
- Пазнаванне чалавечай гаворкі
 - TTS



- OCR
 - cuneiform
 - tesseract
- NLP
 - aot?
 - корпусы
 - tomita-parser
- Гаворка
 - CMUSphinx, pocketsphinx
 - Julius
 - ...

Корпусы тэкстаў

- Ёсць некалькі корпусаў (bnkorporus.info, github.com/poritski/YABC, ...)
- Ёсць пытанні з (C)
- Ідэі
 - crowdsourcing: opencorpora.org engine
 - Грамадскі набытак: 50 год Шмат класікаў ужо...

Корпусы гаворкі

- Праект voxforge (nothing/140h)
- Праект librivoх (1 зборнік вершаў)