



D1.19. Diseminare

Aceste rezultate au fost obținute prin finanțare în cadrul Programului PN-III Proiecte complexe realizate în consorții CDI, derulat cu sprijinul MEN – UEFISCDI,
Cod: PN-III-P1-1.2-PCCDI-2017-0818, Contract Nr. 73 PCCDI/2018:
“SINTERO: Tehnologii de realizare a interfețelor om-mașină pentru sinteza text-vorbire cu expresivitate”

© 2018-2020 – SINTERO

Acest document este proprietatea organizațiilor participante în proiect și nu poate fi reprodus, distribuit sau diseminat către terți, fără acordul prealabil al autorilor.

Denumirea organizației participante în proiect	Acronim organizație	Tip organizație	Rolul organizației în proiect (Coordonator/partener)
Institutul de Cercetări Pentru Inteligență Artificială “Mihai Drăgănescu”	ICIA	UNI	CO
Universitatea Tehnică din Cluj-Napoca	UTCN	UNI	P1
Universitatea Politehnică din București	UPB	UNI	P2
Universitatea "Alexandru Ioan Cuza" din Iași	UAIC	UNI	P3

Date de identificare proiect

Număr contract:	PN-III-P1-1.2-PCCDI-2017-0818, Nr. 73 PCCDI/2018
Acronim / titlu:	„SINTERO: Tehnologii de realizare a interfețelor om-mașină pentru sinteza text-vorbire cu expresivitate”
Titlu livrabil:	D1.19. Diseminare
Termen:	Noiembrie 2018
Editor:	Mircea Giurgiu (Universitatea Tehnică din Cluj-Napoca)
Adresa de eMail editor:	Mircea.Giurgiu@com.utcluj.ro
Autori, în ordine alfabetică:	Mircea Giurgiu, Adriana Stan
Ofițer de proiect:	Cristian STROE

Rezumat:

În acest livrabil se prezintă articolele științifice susținute și publicate de către partenerul UTCN la conferințe internaționale în anul 2018. Au fost identificate la începutul etapei conferințele de interes din domeniul aferent cercetărilor. Deoarece proiectul este doar la început, o mică parte din rezultate au fost diseminate, așa cum se va prezenta în continuare. Vom include lista articolelor și un rezumat cu principalele rezultate. Accesul la publicații (cf. 30.11.2018) este asigurat pe pagina web a proiectului în secțiunea dedicată.

De asemenea, este disponibilă și pagina web a sub-proiectului SINTERO (<http://speech.utcluj.ro/sintero/>), pagina cu demonstrații online de voci sintetice cu expresivitate (http://speech.utcluj.ro/sintero/prosody_examples), împreună cu paginile web ale celorlalte proiecte din cadrul proiectului complex ReTeRom.

Cuprins

1. Introducere	4
2. Identificarea posibilităților de publicare pe anul 2018 și realizări	4
3. Publicații științifice în anul 2018	4
4. Pagina web a proiectului SINTERO	5
5. Alte pagini web cu demonstrarea online a modului de control a prozodiei	6
6. Concluzii	6

1. Introducere

Acest livrabil prezintă o sinteză a articolelor publicate în anul 2018, pagina web a proiectului și o serie de demonstratoare online pentru controlul automat al expresivității vocilor sintetizate prin tehnici de modelare statistică. Acestea prezintă un punct de plecare pentru comparații cu rezultate viitoare în proiect. De exemplu, în prezent sunt în cercetare tehnici de adaptare a stilului de vorbire și a expresivității pe bază de structuri neuronale artificiale.

2. Identificarea posibilităților de publicare pe anul 2018 și realizări

Conform cu strategia de diseminare inclusă în formularul de aplicație, pentru fiecare an calendaristic s-au identificat posibilitățile de diseminare și de publicare de articole la conferințe științifice sau în jurnale.

Fiind la început, pentru anul 2018 a fost propusă participarea la conferința *CONSILR 2018, The 13th Edition of the International Conference on Linguistic Resources and Tools for Processing Romanian Language, Iasi, 22-23 Noiembrie 2018*, <https://profs.info.uaic.ro/~consilr/>

3. Publicații științifice în anul 2018

Autori	A Stan, M Giurgiu
Titlu	„A comparison between traditional machine learning approaches and Deep Neural Networks (DNN) for text processing in Romanian”
Ref.	The 13th Edition of the International Conference on Linguistic Resources and Tools for Processing Romanian Language, Iasi, 22-23 Noiembrie 2018, https://profs.info.uaic.ro/~consilr/
Rezumat	<i>This paper presents a comparison between traditional machine learning approaches (decision trees and multilayer perceptron) and the latest trend in artificial intelligence, deep neural networks for three separate tasks of text processing in Romanian. The tasks we examine are: lexical stress assignment, syllabification and phonetic transcription. The evaluation is performed on large manually transcribed lexicons and uses simple input features derived strictly from the orthographic form of the words. Results show that, depending on the task, the performance of each of the algorithms can vary, and that in some limited cases, the decision trees can outperform the deep neural networks.</i>

Stress assignment evaluation in terms of accuracy and F-measure

Algorithm	Features	With letter type		Without letter type	
		Accuracy	F-measure	Accuracy	F-measure
Decision Trees	WIN3	0.97	0.92	0.96	0.92
	WIN5	0.97	0.93	0.97	0.93
	WIN7	0.97	0.93	0.97	0.93
	Entire word	0.92	0.92	0.91	0.91
MLP	WIN3	0.88	0.72	0.84	0.56
	WIN5	0.92	0.79	0.87	0.65
	WIN7	0.92	0.82	0.87	0.66
	Entire word	0.79	0.74	0.52	0.46
DNNs	WIN3	0.96	0.92	0.96	0.90
	WIN5	0.98	0.95	0.97	0.95
	WIN7	0.98	0.95	0.98	0.95
	Entire word	0.96	0.96	0.95	0.95

Syllabification evaluation in terms of accuracy and F-measure

Algorithm	Features	With letter type		Without letter type	
		Accuracy	F-measure	Accuracy	F-measure
Decision Trees	WIN3	0.99	0.99	0.99	0.99
	WIN5	0.99	0.99	0.99	0.98
	WIN7	0.99	0.99	0.99	0.98
	Entire word	0.86	0.86	0.67	0.67
MLP	WIN3	0.96	0.93	0.78	0.67
	WIN5	0.96	0.95	0.81	0.70
	WIN7	0.97	0.96	0.83	0.74
	Entire word	0.65	0.63	0.30	0.21
DNNs	WIN3	0.99	0.99	0.99	0.99
	WIN5	0.99	0.99	0.99	0.99
	WIN7	0.99	0.99	0.99	0.99
	Entire word	0.98	0.97	0.97	0.97

Phonetic transcription - problematic letters - evaluation in terms of accuracy and F-measure

Algorithm	Features	With letter type		Without letter type	
		Accuracy	F-measure	Accuracy	F-measure
Decision Trees	WIN3	0.9882	0.9881	0.9859	0.9858
	WIN5	0.9877	0.9875	0.9837	0.9837
	WIN7	0.9868	0.9868	0.9825	0.9824
MLP	WIN3	0.86	0.84	0.80	0.73
	WIN5	0.90	0.88	0.86	0.80
	WIN7	0.93	0.92	0.87	0.83
DNNs	WIN3	0.99	0.99	0.98	0.98
	WIN5	0.99	0.99	0.99	0.99
	WIN7	0.99	0.99	0.99	0.99

4. Pagina web a proiectului SINTERO

EN DESCRIERE CONSORTIU ECHIPA REZULTATE CONTACT

SINTERO: Tehnologii de realizare a interfețelor om-mașină pentru sinteza text-vorbire cu expresivitate

Proiect finanțat de Ministerul Cercetării și Inovării, Program PN-III-P1-1.2-PCCDI, nr. 73/2018, durata: 2018-2020

Proiect component al proiectului complex

RETEROM

Proiecte paralele

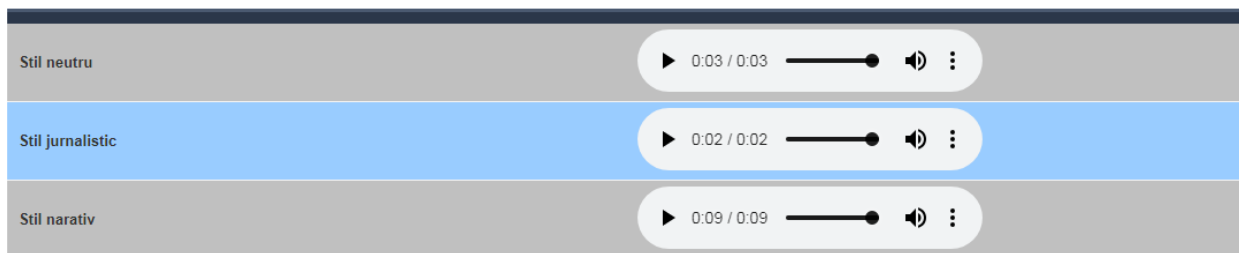
COBILIRO TEPROLIN TADARAV

5. Alte pagini web cu demonstrarea online a modului de control a prozodiei

În pagina http://speech.utcluj.ro/sintero/prosody_examples se găsesc mostre de semnal sintetizat prin modificarea prozodiei, după cum urmează:

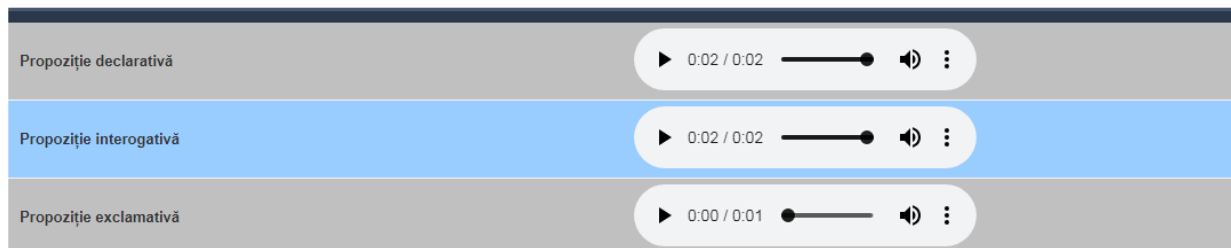
- 3 noi voci sintetice pentru 3 stiluri de vorbire studiate în livrabilul D1.16 (stilul neutru, stilul de prezentator de știri, stilul narativ din audio book)
- modificarea prozodiei, corespunzător tipului propoziției declarativă, interogativă, exclamativă
- adaptarea duratei și/sau a frecvenței fundamentale a vocii neutre la noul stil de exprimare
- adaptarea vocii neutre la stilul și expresivitatea celor 3 stiluri studiate folosind tehnica CSMAPLR.

Exemple voci naturale - stiluri de exprimare



Stil neutru	▶ 0:03 / 0:03
Stil jurnalistic	▶ 0:02 / 0:02
Stil narativ	▶ 0:09 / 0:09

Exemple voci sintetizate - modificarea tipului propoziției



Propoziție declarativă	▶ 0:02 / 0:02
Propoziție interogativă	▶ 0:02 / 0:02
Propoziție exclamativă	▶ 0:00 / 0:01

Exemple voci sintetizate - adaptarea vorbitorului la stilul jurnalistic



Voce stil neutru	▶ 0:07 / 0:07
Voce stil jurnalistic	▶ 0:06 / 0:06
Adaptarea duratei vocii neutre la stilul jurnalistic	▶ 0:05 / 0:05
Adaptarea frecvenței fundamentale a vocii neutre la stilul jurnalistic	▶ 0:07 / 0:07
Adaptarea duratei și F0 vocii neutre la stilul jurnalistic	▶ 0:05 / 0:05

6. Concluzii

Accesul la publicațiile elaborate în anul 2018 este asigurat la adresa <http://speech.utcluj.ro/sintero/rezultate/>. Pagina web are un conținut dinamic, adaptat cu realizările din proiect, astfel că pentru această raportare se pot accesa și mostre cu semnal sintetic generat de un modul de control automat al expresivității (http://speech.utcluj.ro/sintero/prosody_examples).