Project title:	"Multimodal multilingual human-machine speech communication"
Project Acronym:	AI-SPEAK
Deliverable index:	D1.2c
Version:	2.0

DISSEMINATION PLAN

of the Project "Multimodal multilingual human-machine speech communication" (AI-SPEAK), as initially established during the series of meetings that took place in January 2024 including the Project kick-off meeting, and finalized by the end of WP1 (Preparatory activities) in M4.

This Dissemination plan represents an extension of the dissemination plan given in the Project proposal, and results from a more detailed analysis of the channels through which the dissemination of AI-SPEAK Project results can be effected.

The dissemination activities at the project are within the scope of work package 6, and covers the planning and timing of publications and demonstration activities. The following table gives the original description of WP6 as given in the Project proposal:

Work package number	6	Work package title	Dissemination
Lead SRO's acronym	FTNUNS		
WP Coordinator - team member's ID	TM1		
Team members' IDs	PI, TM1-TM8		
Objectives To increase the visibility of the Project and provide open access to its results; to strengthen international scientific collaboration, ensuring impact beyond the Project timeframe.			
Description of work (where appropriate, broken down into sub-activities), and role of the team members Most dissemination activities, with the exception of the Project website creation (D1.1), will be based on a detailed dissemination plan (D1.2). Besides the organization of major dissemination events (Sub-activity 6.1; PI and TM1) – the 25th edition of the international conference SPECOM in the first year (months 1-7) and a workshop for regional IT companies (months 24-25) – dissemination activities will also include the preparation of scientific publications for open access journals, participation at scientific conferences, appearances in electronic and printed media and on social networks, as well as preparation of appropriate marketing material, including brochures, leaflets and roll-up banners (Sub-activity 6.2). Work package will be led by Prof. Vlado Delić (TM1).			
Deliverables of the work package (brief description and month of delivery)			

** Papers published in journals and conference proceedings during the Project (the open datasets that will result from the Project will also contribute to Project visibility, but they are listed as deliverables of WP2).*

A key element in the development of the innovative dimensions of the Project are the results of scientific research in the field of audio-visual speech recognition and synthesis. It is our general intention, that according to the good practices in academia and in line with the principles of open research, the researchers should document their work and disseminate the results of the Project principally in the form of conference and journal papers related to both speech technology and their application within human-machine interaction. However, project results will be made public through other dissemination channels as well.

1. Dissemination channels

The initial plan proposes the **participation at renowned international conferences**, such as SPECOM, INTERSPEECH and/or ICASSP, as well as the presentation of results at conferences (co)organized by the Faculty of Technical Sciences. As members of programme committees for some of these conferences (Milan Sečujski (PI) and Branislav Popović (TM2): SPECOM; Vlado Delić (TM1): SPECOM and INTERSPEECH), they will ensure the visibility of the project results presented in appropriate sessions of these conferences. **Publication in open access journals** is also planned and appropriate funds for this have been allocated in the Project budget.

Regular project meetings, which will occur once in every 6 months, were also intended to serve as dissemination events, aimed at presenting the results of the Project to the interested parties from the fields of academia and industry as well as popularization of the use of high-quality speech technology. Depending on the availability of Project members and interested parties, some of the Project meetings may be shifted to the virtual domain. Alternatively, social networks and online media will also be used as alternative dissemination channels.

Most of the aforementioned events are well covered by the media and each appearance in **electronic and printed media** as well as presentation of results with significant market potential (wider application of speech technology) and/or a clear social impact (preservation of the language as well as the applications for the disabled), in our experience, triggers a series of appearances in other electronic and printed media, as well as on social networks, which will be intelligently exploited and stimulated. Participation at the International Fair of Technics and Technical Achievements in Belgrade is also planned (most likely in 2026, after the integration of Project results into a practical environment of automatic speech recognition or text-to-speech systems).

The **Project website** has been created at the beginning of the Project: https://www.ktios.ftn.uns.ac.rs/ai-speak/AI-SPEAK_sr.html. Its public section will be used for Project presentation and dissemination, as well as repository of relevant scientific papers that the members of the project team have published before or during the Project, as well as the implementations of the algorithms (computer code) developed within the project. The private section of the Project website will be used as a platform for communication between team members, including a repository of relevant scientific papers published by others, as well

as a repository of computer code shared between team members. The private section of the Project website, with its evidence of project activities, milestones and deliverables, will facilitate Project management and reporting.

The practical results of the project will be disseminated through their application within automatic speech recognition and text-to-speech systems, improving their performances. Namely, the core of the Project team is the Speech Technology (ST) group at the Department of Power, Electronic and Telecommunication Engineering at FTS-UNS, which has long-standing cooperation with the company "AlfaNum" from Novi Sad, a spin-off from FTS-UNS and a regional leader in the development of speech technology applications for Serbian and several other languages. As the proposed improvement of the flexibility and adaptability of speech technology will be evaluated and demonstrated by implementation into a digital voice assistant for Serbian, "AlfaNum", as a company with good strategic position on the market, with a number of contacts in the region and beyond, will also play an important role in Project dissemination as well as the transfer of its research and development results into industry.

Open data is an increasingly popular dissemination channel, and the multimodal multilingual (audio-visual) speech corpora that will be collected within the project (WP2) is expected to be useful to researchers worldwide. Namely, in order to establish a good starting point for the proposed experiments based on audio-visual speech recognition and synthesis, a multimodal multilingual speech corpus, referred to as **AI-SPEAK speech corpus** will be created. The corpus will contain recordings of speech in both Serbian and English from 25 adult speakers of both genders, together with video recordings of the movements of their lips. The intended quantity of speech data per speaker is 10 minutes, although it is possible that eventually more data will be obtained. The corpus will be recorded in strictly controlled conditions, in the IAC Mini anechoic chamber obtained through the Erasmus+ project SENVIBE. The speakers will deliver a fixed number of utterances in both Serbian and English, including spoken digits, names of letters, simple commands, as well as a number of short sentences. There will be one set of sentences which will be the same for all users, and another one which will be different for each speaker, in order to maximize the phonetic coverage and the usability of the corpus for all intended research directions. This corpus will be stored on the private section of the Project web platform, accessed according to the well-established practice, and moved to the public section of the Project web platform and become openly available to the research community once any information that may reveal the participant's identity (such as easily recognizable facial features) is blurred out or otherwise removed, and parts of recordings outside the intended scope of the corpus are destroyed. To ensure its visibility, at the end of the Project this corpus will also be uploaded onto the UCI Machine Learning Repository (a repository of freely available datasets for machine learning tasks), and it will be properly referenced in all publications as one of the very few publicly available multimodal (audio-visual) speech corpora with bilingual speakers. The second corpus that will be collected and processed on the Project will be similar in content, it will also be bilingual (Serbian and English), but to allow the research in line with the tendency of state-of-the-art machine learning systems to focus on exploiting huge quantities of inferior quality data, it will be based on existing videos published on the Internet. This corpus, referred to as the **Internet speech corpus**. This corpus will be stored in the private section of the Project web platform, but it will not be made openly available to the research community, since although the information related to the participant's identity will be removed, the original video would likely still be available somewhere on the Internet, and we cannot exclude the possibility that the speaker could be identified based on the original video, which would be likely to contain personal data. The speakers will be anonymized in this corpus as

well, and only the information on their gender will be retained.

2. Timing of dissemination activities

PI of the Project, according to his role of the leader of WP7 (Management) together with TM1 who is the leader of WP6 (Dissemination), will keep track of the publications of Project results and ensure that they are made public, in line with the principles of open data and open research.

The Project website and its News section will serve as a continuous platform for the dissemination of Project results. From the Project website it will also be possible to download all scientific papers that have been published within the Project, as soon as they have been accepted for publication at an international conference or in an international journal.

Several conferences that have been found to be particularly suitable for the dissemination of project results (owing to their renown, and/or to the membership of particular Project team members in their Programme committees) are listed in the following table:

Name	Character	Expected date	Venue	Note
SPECOM 2024	international	November 2024	Belgrade, Serbia	PI, TM1 and TM2 are PC members
SPECOM 2024	international	October 2025	Belgrade, Serbia	PI, TM1 and TM2 are PC members
YU INFO 2024	regional	March 2024	Kopaonik, Serbia	
YU INFO 2025	regional	March 2025	Kopaonik, Serbia	
YU INFO 2026	regional	March 2026	Kopaonik, Serbia	
INTERSPEECH 2025	international	August 2025	Rotterdam, Netherlands	
INTERSPEECH 2026	international	Sept.-Oct. 2026	Sydney, Australia	
ICASSP 2025	international	April 2025	Hyderabad, India	
ICASSP 2026	international	May 2026	Barcelona, Spain	

Since the Covid-19 pandemic and a growing demand of mitigating CO2 emissions with air travel, participating in conferences online has become more prevalent. Taking this into account, remote venues are also considered.

Several international journals have also been selected as likely candidates for publication of the Project results, owing to the specific scope of a journal, or to previous positive experiences of team members, most of which have long-standing experience in publication in scientific journals:

Name	ISSN	Domain	IF (2019) and M category	Note
Knowledge-Based Systems	0950-7051	Computer Science, AI	8.8 (M21a)	Previous publications by TM2
Expert Systems with Applications	0957-4174	Computer Science, AI	8.5 (M21a)	
Applied Intelligence	0924-669X	Computer Science, AI	5.3 (M22)	Previous publications by PI, TM1, TM2
International Journal of Interactive Multimedia and Artificial Intelligence	1989-1660	Computer Science, AI	3.6 (M22)	
Journal of Universal Computer Science	0948-695X	Computer Science, Software Engineering	1 (M23)	Previous publications by TM1