

This Data Usage Agreement ("Agreement") between Mohamed Bin Zayed University of Artificial Intelligence ("MBZUAI") and the User governs the royalty-free limited, revocable, non-transferable and non-exclusive license granted by MBZUAI to the User to access and use the ArVoice dataset ("Dataset") made available by the ArVoice Project, as described in the publication *ArVoice: A Multi-Speaker Dataset for Arabic Speech Synthesis* (proceedings of Interspeech, 2025)

1. Definitions

- **Dataset:** All original audio recordings made available by the ArVoice Project (Parts 1 and 2 as described in the publication).
- **User:** Any individual or entity that requests or uses the Dataset and signs this Agreement.
- **Project Team:** The creators and maintainers of the ArVoice dataset.

2. Permitted Uses

The Dataset may be used as authorized under this Agreement **exclusively for non-commercial research and educational purposes**, including but not limited to:

- Training, evaluating, or benchmarking machine learning models.
- Academic publications and presentations; small samples from Dataset can be used as examples.
- Developing open-source tools and models.

Users may distribute models or systems trained using the Dataset, provided that:

- No original Dataset audio files are included.
- Proper attribution is given (see Section 6).
- The synthesized voices do not match the original Dataset voices.
- The distributed models or systems are released under a strictly non-commercial license.

3. Prohibited Uses

Users shall not:

- Use the Dataset for any commercial purpose without explicit, written permission from the Project Team.
- Attempt to re-identify any individual in the Dataset.
- Redistribute the raw Dataset or any part thereof.
- Redistribute synthesized voices that match the original voices in the Dataset.
- Use the Dataset in any way that could harm, exploit, or misrepresent the voices or identities of the individuals included.

4. Data Security and Confidentiality

User shall not disclose, release, sell, rent, lease, loan, or otherwise grant access to the Dataset to any third party and must implement reasonable and appropriate safeguards to protect the Dataset from unauthorized access or use. This includes:

- Storing the Dataset on secure, access-controlled systems.

- Not sharing Dataset access credentials with unauthorized parties.
- Deleting the Dataset upon completion of the intended research or upon termination of this Agreement.

5. Data Protection Laws

The User must at all times comply with all applicable laws, rules, and regulations, including data protection laws.

5. Publication Requirements

Any publication or public disclosure based on the Dataset must:

- Clearly state the use of the ArVoice Dataset.
- Include a citation to the official ArVoice paper.

6. Attribution

Users must credit the ArVoice Project in any work that uses the Dataset. Acknowledgment should include:

- The name "ArVoice" Dataset.
- The publication citation.
- A link to the dataset homepage or official publication.

7. Term & Termination

This Agreement is effective at the date of the last signature, until terminated at the discretion of MBZUAI by providing a written notice to the User. It will terminate automatically if the User breaches any of its terms. Upon termination, the User must:

- Cease all use of the Dataset.
- Delete all copies of the Dataset from their systems.
- Decommission and delete all copies of models/systems trained on the Dataset.

8. Indemnity:

The User will defend, indemnify and hold MBZUAI harmless from and against any liability, loss, damage, cost and expense (including without limitation reasonable attorneys' fees) suffered as a result of any breach of this Agreement.

9. Contact Information

For questions regarding this Agreement or the Dataset:

Email: hanan.aldarmaki@mbzuai.ac.ae

Website: <https://h-aldarmaki.github.io/arvoice/>

10. Governing Law and Jurisdiction

This Agreement shall in all respects be governed by and construed in accordance with the laws of the United Arab Emirates. Any dispute or difference of any kind between the Parties in connection with or arising out of this Agreement or the breach, termination or validity hereof shall be finally settled in the Courts of Abu Dhabi Global Market.

9. User Information and Signature

Name: _____

Title: _____

Institution/Affiliation: _____

Email Address: _____

Date: _____

Signature: _____