AVSPOOF DATASET END USER LICENSE AGREEMENT (07/2012)

1. **Scope**
This End User License Agreement (hereinafter: “License”) is granted by Idiap Research Institute, Rue Marconi 19, BP 592, CH 1920 Martigny, Switzerland (hereinafter: “Licensor”) to the End User. This License is an Open Source License.

2. **Material under License**
The material under License (hereinafter: “Corpus”) is an Idiap Research Institute database described in Schedule 1 composed of audio and video disclosed on a voluntary basis by the concerned persons. The database contains the actual data as well as any derivative work, products or services based on all or part of the data. All data contained within the Corpus have been collected and processed in accordance with the laws applicable in Switzerland.

3. **Copyright**
The Corpus is the sole property of the Licensor and is protected by copyright. The Licensor reserves all rights to use in any way and distribute the Corpus. The Corpus shall remain the exclusive property of the Licensor. The End User acquires no ownership, rights or title of any kind in all or any parts with regard to the Corpus.

4. **License**
Subject to prior identification of the End User and signature of this License, the Corpus is freely available to the benefit of the End User. The Licensor grants to the End User the right to use the Corpus, for its own internal and non-commercial use and for the purpose of scientific research only.

   The End User shall respond for any infringement of the present License by one of their subsidiaries and/or student’s. The End User may only disclose, give access, and/or transfer the rights related to the Corpus to subsidiaries and/or students, under the following conditions:

   1) a copy of the present License has already been transferred to them;
   2) the subsidiaries/students have fully read and understood all terms and conditions of the present License;
   3) the access to the Corpus is granted under the close supervision of the End User;
   4) the access to the Corpus is granted under the sole responsibility of the End User.

   The Licensor grants to the End User the right to reproduce temporarily, to adapt, arrange and modify by any means the Corpus. The Licensor grants to the End User the right to rework and build upon the Corpus, or any component thereof, as necessary or desirable for research or technology development activity and create derivative products or services for the End User's own internal research and development. The End User is permitted to make a copy of the Corpus for archiving only. This License gives no right of any kind to the End User over the Corpus. The License is deemed non-exclusive and non-transferable to third parties.

5. **Access**
The End User may only use the Corpus after this License has been signed and returned to the Licensor. The End User must return the signed and dated License by email, in PDF format to the Licensor at the following address: data-manager@idiap.ch. The Corpus is accessible online at the following address: https://www.idiap.ch/dataset/avspoof.

6. **Copyleft and Open Science**
The Licensor grants to the End User the License defined under section 2 under the sole condition that any derivative work, product, services, scientific developments of any kind or any improvement or modification of any kind based on all or part of the Corpus shall be left freely available to other End Users. A copy of any modification of the Corpus shall be submitted to the Licensor at the following address: data-manager@idiap.ch. The End User shall ensure that, in the event of a modification of the Corpus, the substance of the Material is not altered and that, the facility still operates, and performs whatever part of its purpose remains meaningful. Any violation of this clause will give rise to immediate legal prosecution.
7. Distribution
The End User shall not, without prior authorization of Licensor, transfer in any way, permanently or temporarily, distribute or broadcast all or part of the Corpus to third parties. The End User shall send all requests for the distribution of the Corpus to the Licensor at the following address: data-manager@idiap.ch.

8. Research
Research includes all type of scientific research, irrespective of the object under scrutiny, aimed at achieving a progress in science. This dataset aims to assess the vulnerability of speaker verification systems

9. Commercial use
Any commercial use of the Corpus is strictly prohibited. Commercial use of the Corpus includes, but is not limited to:
- Proving the efficiency of commercial systems;
- Testing commercial systems;
- Using screenshots of subjects from the Corpus in advertisements;
- Selling data or making any commercial use of the Corpus;
- Broadcasting data from the Corpus.

Any violation of this clause will give rise to immediate legal prosecution by the Licensor. Any damages and/or unfair enrichment or the like of the End User due to the breach of the License shall be immediately restituted to the Licensor together with the derivative works, products and services based on all or part of the Corpus.

10. Publications
The End User shall reference the Corpus, or results obtained with it, in publications. Publications include, but are not limited to:
- Research papers,
- Articles,
- Presentations for conferences or educational purposes.

Small portions of the Corpus cannot be used in any publication. All publications that report on research that use the Corpus will acknowledge this as follows: "(Portions of) the research in this paper used the AVSpoof Dataset made available by the Idiap Research Institute, Martigny, Switzerland."

and will cite the following paper: Serife Kucur Ergunay, Elie Khoury, Alexandros Lazaridis, Sebastien Marcel. "On the vulnerability of Speaker Verification to Realistic Voice Spoofing*, BTAS 2015.

11. Illegal or criminal use of the Corpus
Any illegal or criminal use of the Corpus by the End User is strictly prohibited.

12. Legal Disclaimer
The Corpus is granted without any warranty. Licensor shall not be held responsible for any damage (physical, financial or otherwise) caused by the use of the Corpus. Licensor shall not be held responsible of any illegal or criminal use of the Corpus by the End User.

13. Jurisdiction
This License is subject to and interpreted in accordance with Swiss Law. Any claim arising on the basis of this License shall exclusively be submitted to the Courts of Martigny, Switzerland, subject to appeal to the Courts of Valais and to the Federal Court.
14. **Amendments**
The Licensor is allowed to amend this License at any time without prior consent of the End User. The End User shall be informed about changes and given the choice to opt out of this License within 10 days to the Licensor at the following address: data-manager@idiap.ch. Without any notice within 10 days and provided the amendment is not substantial, the amendment to the License will be fully applicable to the End User.

15. **Warranties**
The End User warrants that they are authorized signatory, adult and not legally forbidden to enter into this License. The End User warrants that they have read and understood all elements contained herein and that the signature apposed hereunder is the result of a fully aware decision.

By signing this License, the End User engages to strictly respect the conditions set forth herein and to respect all the laws applicable in Switzerland in relation to data and personality protection with regard to the data contained within the Corpus collected and processed by the Licensor:

Please also notice that students are not eligible, therefore the request must be made and signed by their supervisors:

End-User (Authorized Signatory):

<table>
<thead>
<tr>
<th>Organization Name (Institute, University, etc):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization Address:</td>
</tr>
<tr>
<td>Full Name (Block Letters):</td>
</tr>
<tr>
<td>Title / Function:</td>
</tr>
<tr>
<td>Email:</td>
</tr>
<tr>
<td>Date:</td>
</tr>
<tr>
<td>Signature:</td>
</tr>
</tbody>
</table>
**Description of the Corpus:**

The 'AVspoof' database is intended to provide a stable, non-biased spoofing attacks in order for researchers to test both their speaker verification systems and anti-spoofing algorithms.

The data acquisition process lasted approximately two months with 44 people, each participating in four different sessions configured in different environmental setups. Each recording session involves prepared (read) speech, pass-phrases and free speech recorded with three devices: one laptop with high-quality microphone and two mobile phones (iPhone and Samsung). These data can be found in the folder 'genuine'.

After the collection of the genuine data, 10 types of attacks were generated and that can be found in 'attacks/

* 4 Replay attacks:
  - replay_phone1
  - replay_phone2
  - replay_laptop
  - replay_laptop_HQ_speaker

* 3 Speech synthesis attacks:
  - speech_synthesis_logical_access
  - speech_synthesis_physical_access
  - speech_synthesis_physical_access_HQ_speaker

* 3 Voice conversion attacks:
  - voice_conversion_logical_access
  - voice_conversion_physical_access
  - voice_conversion_physical_access_HQ_speaker

Detailed statistics on the databases and baseline results using both ISV and i-vector system can be found in our publication: