

Second DIHARD Challenge License Agreement

In the remainder of this document the term **User** refers to:

_____ (Individual name)

and the term **User's Research Group** refers to:

_____ (University, Institute or Company name)

_____ (Specific department, area, organizational unit)

_____ (Geographic site where material will be used)

User agrees on behalf of **User's Research Group**, to receive media (CD-ROM, DVD, hard drive, web download, etc.), named below under "Corpora and/or Data Received," containing images, speech, text and/or video data from the Linguistic Data Consortium (LDC). Among the data provided is the Seedlings HomeBank Corpus (the "SEEDLingS corpus"). **User** agrees to use the material received under this agreement (the "Data") only for purposes of the Second DIHARD Challenge at the geographic site indicated above.

User and **User's Research Group** may include limited excerpts from the Data in presentations, articles, reports and other documents describing the results of work performed in the Second DIHARD Challenge, with the exception of audio clips from the SEEDLingS corpus, which require permission from Erika Bergelson on a per-file basis. Most but not all of the included audio clips in the SEEDLingS corpus are cleared for public-facing research use. Contact elika.bergelson@duke.edu to verify whether a clip in question may be played or used for presentations, articles, reports and other documents.

User further agrees not to otherwise publish, retransmit, disclose, display, copy, reproduce or redistribute the Data to others outside of **User's Research Group**.

If **User** fails to submit system results by the date established for such results in the Second DIHARD Challenge, **User** agrees to delete the Data from any computer or media onto which it has been copied and to return all media to LDC **on or before July 1, 2019**.

User may keep the Data if (a) there is a follow-on DIHARD Challenge and **User** has signed an agreement on behalf of **User's research group** to participate in that challenge; or (b) **User** executes on behalf of **User's research group** a subsequent LDC license that covers the Data.

User shall provide appropriate acknowledgement in all presentations, articles, reports and other documents describing the Data and/or the results of work performed in the Second DIHARD Challenge as follows: "Data used herein was taken from the DIHARD corpora (Ryant et al. 2018) and from the SEEDLingS corpus (Bergelson 2016)."

User shall provide appropriate citation in all presentations, articles, reports and other documents describing the Data and/or the results of work performed in the Second DIHARD Challenge as follows:

Bergelson (2016). Bergelson Seedlings HomeBank Corpus. doi: 10.21415/T5PK6D.

Ryant et al. (2019). DIHARD Corpus, Linguistic Data Consortium.

User shall direct to Erika Bergelson (elika.bergelson@duke.edu) any queries, questions or concerns about the Seedlings HomeBank Corpus.

User agrees that user's name and **User's Research Group** or affiliation will be communicated to the developer of the Seedlings HomeBank Corpus.

User designates _____ as the data contact person under this agreement. The data contact person shall receive all material distributed by LDC under this agreement. The data contact person is responsible for distributing such material to **User's Research Group**.

User shall send a signed copy of this agreement to LDC: (1) by e-mail to ldc@ldc.upenn.edu or (2) by facsimile, Attention: Membership Office, fax number (+1) 215-573-2175.

All Data is provided "as is" and neither the Linguistic Data Consortium nor the University of Pennsylvania warrant the accuracy, completeness, currentness, merchantability or fitness for a particular purpose of the Data. In no event will the Linguistic Data Consortium or the University of Pennsylvania be liable for any loss or injury caused in whole or in part by its negligence or contingencies beyond its control in procuring, compiling, interpreting, editing, writing, reporting or delivering the Data, or for any errors, omissions or inaccuracies in the Data, regardless of how caused. In no event will the Linguistic Data Consortium or the University of Pennsylvania be liable to **User, User's Research Group**, or to any other individual or organization for any decision made or action taken in reliance upon any part of the Data or for any consequential, direct, special or similar damages, even if advised of the possibility of such damages.

Corpora and/or Data Received

LDC2019E31 Second DIHARD Challenge Development Data

LDC2019E32 Second DIHARD Challenge Evaluation Data

Organization: _____

Name: _____

Mailing address: _____

Phone: _____

Fax: _____

Email: _____

Signature: _____

Date: _____

For LDC:

Christopher Cieri
Executive Director