



LINGUISTIC DATA CONSORTIUM

3600 Market St., Suite 810, University of Pennsylvania, Philadelphia, PA 19104-2653 USA • Tel: (215) 898-0464 • Fax: (215) 573-2175 • email: ldc@ldc.upenn.edu

LDC User Agreement for Non-Members.

This User Agreement is provided by the Linguistic Data Consortium as a condition of accepting the databases named or described herein.

In the remainder of this document the term User refers to Ceri Malamud (Individual name) of Public.Resource.org (Affiliation),

and the term User's research group refers to: Public.Resource.org (Specific department or area within company, if appropriate, or University, Institute or Company/name).

This Agreement describes the terms between User/User's Research Group and Linguistic Data Consortium (LDC), in which User will receive material, as specified below, from LDC. The terms of this Agreement supercede any previous Membership Agreement for the Corpora received in Exhibit A.

Under this agreement User will receive one or more CD-ROM discs, DVDs, electronic files or other media as appropriate, named below under "CORPORA RECEIVED" (Exhibit A), containing speech, video and/or text data. User agrees to use the material received under this agreement only for non-commercial linguistic education and research purposes. In the event that User's use of LDC Corpora results in the development of a commercial product, User must join the LDC as a Commercial Member and pay all applicable fees prior to release of said commercial product. Unless explicitly permitted herein, User shall have no right to copy, redistribute, transmit, publish or otherwise use the LDC Databases for any other purpose and User further agrees not to disclose, copy, or re-distribute the material to others outside of User's research group.

USER ACKNOWLEDGES AND AGREES THAT "CORPORA RECEIVED" ARE PROVIDED ON AN "AS-IS" BASIS AND THAT LDC, ITS HOST INSTITUTION THE UNIVERSITY OF PENNSYLVANIA, AND ITS DATA PROVIDERS AND CORPUS AUTHORS MAKE NO REPRESENTATIONS OR WARRANTIES OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR CONFORMITY WITH WHATEVER DOCUMENTATION IS PROVIDED. IN NO EVENT SHALL LDC, ITS HOST INSTITUTION, DATA PROVIDORS OR CORPUS AUTHORS BE LIABLE FOR SPECIAL, DIRECT, INDIRECT, CONSEQUENTIAL, PUNITIVE, INCIDENTAL OR OTHER DAMAGES, LOSSES, COSTS, CHARGES, CLAIMS, DEMANDS, FEES OR EXPENSES OF ANY NATURE OR KIND ARISING IN ANY WAY FROM THE FURNISHING OF OR USER'S USE OF THE CORPORA RECEIVED.

For Public.Resource.org (User Affiliation)

Signature [Signature] Date 10/19/07

Name Ceri Malamud Title President

For LDC

Chris Cieri, Executive Director

EXHIBIT A

CORPORA RECEIVED (Include Title and Catalog Number)

1 JURIS LDC 98T32

2 _____

3 _____

4 _____

5 _____



JURIS

Item Name: JURIS

Authors: Alexandra Canavan and Paul Morgovsky

LDC Catalog No.: LDC98T32

ISBN: 1-58563-135-3

Data Type: text

Data Source(s): varied

Application(s): information retrieval

Language(s): English

Language ID(s): ENG

Distribution: 2 CD

Membership Year(s): 1998

Non-member Fee: US\$750.00

Reduced-License Fee: US\$375.00

Extra-Copy Fee: US\$300.00

Non-member License: yes

Online documentation: yes

Licensing Instructions: Subscription Members, Standard Members, Non-M

Citation: Alexandra Canavan and Paul Morgovsky

1998

JURIS

Linguistic Data Consortium, Philadelphia

Introduction

The text data contained on this two CD-ROM set represent a release of the JI (Justice Department Retrieval and Inquiry System) data collection that has been available to the Linguistic Data Consortium (LDC) by the U.S. Department of Justice. The time span of the text ranges from the 1700s to the early 1990s.

Data

There are 1,664 individual text files in the corpus, 1011 on the first CD-ROM and 653 on the second. The original archive consisted of 219 files ranging between 1 MB and nearly 70 MB in size. In order to make the data more accessible for use, we chose to divide the larger files into pieces, such that the average file size is about 2 MB when uncompressed (the largest uncompressed file size is about 70 MB). Divisions of the files were done at document boundaries, so all files contain whole documents.

There are a total of 694,667 document units in the corpus and these can be cat

to some extent with regard to their content. The following is a partial list of categories and their descriptions drawn from JURIS documentation contained in the corpus. The terminology and organization of categories are those used in the JURIS documentation:

- Case Law
- Executive Order
- Regulations
- Federal Register
- Statutory Law
- Administrative Law
- International Agreements
- Freedom of Information Act and related documents
- Indian Law
- Tax Law
- Brief

Updates

There are no updates at this time.

[About LDC](#) | [Members](#) | [Catalog](#) | [Projects](#) | [Papers](#) | [LDC Online](#) | [Search / Help](#) | [Contact Us](#) | [UPenn](#) | [Home](#) | [Obtaining Data](#) | [Creating Data](#) | [Using Data](#) | [Privacy Policy](#)

Contact: ldc@ldc.upenn.edu

(c) 1992-2007 [Linguistic Data Consortium](#), [University of Pennsylvania](#). All Rights Reserved.