

The National Institute of Informatics Data Set Terms of Use

Article 1. Purpose

These regulations have been provided as the Terms of Use for the Data Set (hereinafter “the Data Set”) of The National Institute of Informatics (hereinafter “the Institute”).

Article 2. Definition of Terms

The definitions of the following terms shall apply to the same terms appearing elsewhere in these regulations:

1. The Test Collection: a set of data built by the Institute used for evaluating text processing systems for information retrieval, question answering, and automatic summarization;
2. The Data Set: the document data base, task lists and correct data which appear in Appendix 1 and which together comprise the Test Collection;
3. User: a person who conducts research using the Data Set;
4. Joint User: a person who conducts research or assists in research using the Data Set in direct collaboration with a user.

Article 3. Users

Persons who may qualify as users of the Test Collection shall belong to one of the following items:

1. Employees of universities, colleges, and junior colleges;
2. Employees of national and public research institutes and corporations whose purpose is research or research assistance;
3. Employees who belong to private corporations and whose principal duty is research;
4. Other parties deemed by the NTCIR Test Collection Application Review Committee of the National Institute of Informatics (hereinafter “the Committee”) as being appropriate.

Article 4. Joint Users

Persons who may qualify as joint users of the Test Collection shall belong to one of the following items:

1. Persons who belong to the same organization as the Users and who are conducting joint research with them;
2. University graduate students and research students under the direct tutelage of a User.

Article 5. Applying to Use the Data Set

Persons who wish to use the Data Set shall request the approval of the Committee by submitting a prescribed User Application.

Article 6. Approval to Use the Data Set

The Committee shall approve applications mentioned in the preceding Article which are deemed appropriate and shall make the Data Set available to such applicants.

Article 7. Duration of Use of the Data Set

Users and Joint Users (hereinafter “Users etc.”) may use the Data Set from the date of approval mentioned in the preceding article until and inclusive of the final day of that fiscal year. Unless there is a notification of a change in or cancellation of the terms of approval on the final day of the fiscal year, use of the Data Set may be continued for a further fiscal year thereafter, and subsequently in increments of one year.

Article 8. Compliance with the Terms of Use

Users etc. shall comply with the following terms of use:

1. Users etc. will use the Data Set only for research purposes and/or assisting with research;
2. Users etc. will not use the Data Set for profit-making purposes;
3. Users etc. will not sell, lend, publish, distribute, or make transmittable to a third party the Data Set, copies of the Data Set, or processed data which would enable a third party to reproduce the Data set.;
4. Users etc. will not violate copyrights of the Data Set and related data when using the Data Set and when presenting research outcomes obtained by using it;
5. Users etc. will report to the Committee on the research activities report by use for The National Institute of Informatics Date Set.
6. Users etc. will clearly mention in any paper or thesis presenting the research outcomes obtained by using the Data Set that the Data Set was used. In addition to the bibliographic information (the titles of the publications, volume, page numbers, and names of publishers and dates), to submit one (1) offprint or (1) copy to the Committee;

7. Users etc. will receive the prior consent of the Committee in writing before making public any information data obtained by using the Data Set;
8. Users etc. will not use evaluation results obtained through use of the Data Set for advertising products or for profit-making purposes including advertising, and/or defamatory or slanderous purposes;
9. Users etc. will not violate user conditions established by the copyright holders;
10. Users etc. will uphold any other terms established separately by the Committee.

Article 9. Compliance of Joint Users

Users shall make joint users aware of these Terms of Use and require such users to comply with the regulations.

Article 10. Disputes between Users and Third Parties

The Institute shall not be responsible for any disputes arising between Users etc. and third parties relating to the use of the Test Collection.

Article 11. Revocation of User Privileges

The Committee may revoke or suspend the user privileges of Users who have violated the provisions of Article 8. Upon revocation of user privileges, such User(s) shall promptly return the Data Set and delete all data obtained by copying or processing the aforementioned data from any computer or media onto which it has been copied.

Article 12. Request for Submission of a Report

The Committee may request a report from Users concerning results and progress of the use of the Test Collection. Users shall promptly prepare a report in response to such requests.

Article 13. Notification of Changes in User Status

Users shall promptly notify the Committee for any of the following reasons:

1. When the user discontinues use of the Data Set;
2. When there is a change in details noted on the User Application.

Article 14. Additional Terms of Use

In addition to terms set forth in these regulations, other terms necessary for the implementation of these regulations shall be provided separately.

Supplementary Provisions
Enacted on March 24, 2003
Revised on May 29, 2012

These regulations shall become effective as of March 1 2014

Appendix 1: The list of data set which is available in accordance with the Terms of Use:

NTCIR-2 SUMM: Summarization Data
NTCIR-3 CLIR: Topics and Relevance assessment Data
NTCIR-3 QA: Questions and Answers Data
NTCIR-3 Web: Topics and Relevance assessment Data
NTCIR-4 CLIR: Topics and Relevance assessment Data
NTCIR-4 QA: Questions and Answers Data
NTCIR-4 SUMM: Summarization Data
NTCIR-4 Web: Topics and Relevance assessment Data
NTCIR-5 CLIR: Topics and Relevance assessment Data
NTCIR-5 CLQA: Questions and Answers Data
NTCIR-5 QA: Questions and Answers Data
NTCIR-5 Web: Topics and Relevance assessment Data
NTCIR-6 CLIR: Topics and Relevance assessment Data
NTCIR-6 CLQA: Questions and Answers Data
NTCIR-6 QA: Questions and Answers Data
NTCIR-7 ACLIA: Topics and Relevance assessment Data • Questions and Answers Data
NTCIR-8 ACLIA: Topics and Relevance assessment Data • Questions and Answers Data
NTCIR-8 GeoTime: Topics and Relevance assessment Data
NTCIR-9 Crosslink: Relevance assessment Data
NTCIR-9 INTENT: Queries and Relevance assessment Data
NTCIR-9 1CLICK: Queries and Relevance assessment Data
NTCIR-9 SpokenDoc: Queries and Relevance assessment Data
NTCIR-10 Crosslink: Queries and Relevance assessment Data
NTCIR-10 INTENT: Queries and Relevance assessment Data
NTCIR-10 1CLICK: Queries and Relevance assessment Data
NTCIR-10 MATH: Queries and Relevance assessment Data
NTCIR-10 SpokenDoc: Queries and Relevance assessment Data